WAR WITHOUT OVERSIGHT;
CHALLENGES TO THE DEPLOYMENT OF AUTONOMOUS WEAPON SYSTEMS

BY
PATRICK WILLIAM WALKER

A Thesis Submitted for the degree of
PhD in Modern War Studies
to the School of Humanities
in the University of Buckingham

JULY 2019

FINAL 1/7/2019 (DISTRIBUTION/READER'S VERSION)
ID. 1303207

Contents

ABSTRACT

Autonomous Weapon Systems (AWS) are defined as robotic weapons that have the ability to sense and act unilaterally depending on how they are programmed. Such human-out-of-the-loop platforms will be capable of selecting targets and delivering lethality without any human interaction. This weapon technology may still be in its infancy, but both semi-autonomous and other pre-cursor systems are already in service. There are several drivers to a move from merely automatic weapons to fully autonomous weapons which are able to engage a target based solely upon algorithm-based decision-making. This requires material step-change in both hardware and software and, once deployed, posits a significant change in how humans wage war. But complex technical difficulties must first be overcome if this new independent and self-learning weapon category can legally be deployed on the battlefield. AWS also pose basic statutory, moral and ethical challenges.

This thesis identifies the manifest complexity involved in fielding a weapon that can operate without human oversight while still retaining value as a battlefield asset. Its key research question therefore concerns the practical and technical feasibility of removing supervision from lethal engagements. The subject's importance is that several well-tried concepts that have long comprised battlecraft may no longer be fit for purpose. In particular, legal and other obstacles challenge such weapons remaining compliant under Laws of Armed Conflict. Technical challenges, moreover, include the setting of weapon values and goals, the anchoring of the weapon's internal representations as well as management of its utility functions, its learning functions and other key operational routines. While the recent development pace in these technologies may appear extraordinary, fundamental fault lines endure. The thesis also notes the inter-dependent and highly coupled nature of the routines that are envisaged for AWS operation, in particular ramifications arising from its machine learning spine, in order to demonstrate how detrimental are these compromises to AWS deployment models. In highlighting AWS deployment challenges, the analysis draws on broad primary and secondary sources to conclude that Meaningful Human Control (MHC) should be a statutory requirement in all violent engagements.

Abbreviations

AAR	After Action Review
AAV	Autonomous Aerial Vehicle
ACB	Advanced Capability Build
ADP	Army Doctrine Publication
AF	Activation Function
AI	Artificial Intelligence
AIKE	Artificial Intelligence and Knowledge Engineering
ACL	Autonomous Control Level
AGI	Artificial General Intelligence
ANNS	Artificial Neural Networks
APS	Active Protection Systems
ATAS	Automatic Target Acquisition System
ATD	Automatic Target Detection
AUVSI	Association for Unmanned Vehicle Systems International
AWS	Autonomous Weapon System
AVD	Aim Verification Device
BACS	Bayesian Application to Cognitive Systems
C3	Consultation, Command and Control
CACE	Change Anything Changes Everything
CARACaS	Control Architecture for Robotic Agent Command and Sensing
CCD	Coherent Change Direction
CCIP	Continuously Computed Impact Point
CCW	Convention of Certain Conventional Weapons
CHAGR	Centre for Historical Analysis and Conflict Research
CM	Continuous Mission
COA	Course Of Action
COG	Centre of Gravity
COTS	Commercial Off The Shelf
CQB	Close Quarter Battle
CSKR	Campaign to Stop Killer Robots
CVT	Controlled Variable Time
DARPA	Defense Advanced Research Projects Agency
DL	Deep Learning
DNN	Deep Neural Networks
DOD	US Department of Defense
DOF	Degrees of Freedom
ECW	Electronic Counter Weapons
EMP	Electro Magnetic Pulse
ERM	Empirical Risk Minimisation
ETL	Extraction, Transfer and Loading
FCS	Future Combat System
GAO	US Government Accountability Office
GCSP	Geneva Centre for Security Policy
GGE	Governmental Group of Experts (CCW)
GMS	Gun Management System

GPS	Global Positioning System
GPU	Graphics Processor Unit
HDR	High Dynamic Range
HFTB	Human Factors Test Bed
HLMI	Human Level Machine Intelligence
HRW	Human Rights Watch
HVEC	High Efficiency Video Coding
IAC	International Armed Conflict
ICJ	International Court of Justice
ICRC	International Committee of the Red Cross
IFF	Identification, Friend or Foe
IHL	International Humanitarian Law
IHRL	International Human Rights law
IOC	Initial Operational Capability
IPRAW	International Panel for the Regulation of Autonomous Weapons
ISR	Intelligence, Surveillance and Reconnaissance
LAWS	Lethal Autonomous Weapon Systems
LCS	Loader Control System
LEP	Life Extension Programme (procurement)
LOBL	Lock On Before Launch
LTDP	Long Term Defence Planning
LIDAR	Light Detection and Ranging
LOAC	Laws Of Armed Combat
MDB	Multi Domain Battle
MHC	Meaningful Human Control
MLP	Multi-layered Perceptron (neural network configuration)
NGO	Non-Governmental Organisation
NLU	Natural Language Understanding
NSA	Non-State Actors
OODA	Observe, Orient, Decide, Act
PED	Processing, Exploitation and Dissemination
PNN	Probabilistic Neural Network (configuration)
PNT	Positioning, Navigation and Timing
RAS	Remote Automated System
RMA	Revolution in Military Affairs
RMAS	Royal Military College Sandhurst
SGD	Stochastic Gradient Descent
SIPRI	Stockholm International Peace Research Institute
SOFM	Self-Organising Features Mapping
TAS	Tracking Adjunct System
TOS	Third Offset Strategy
UAS	Unmanned Aerial System
UAV	Unmanned Airbourne Vehicle
UGS	Unmanned Ground System
USARL	United States Army Research Laboratory
V&V	Validation and Verification
WBE	Whole Brain Emulation

Statement of originality

I hereby declare that my thesis entitled ***War Without Oversight: Challenges to the Deployment of Autonomous Weapons*** is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text, and is not substantially the same as any that I have submitted, or, am concurrently submitting for a degree or diploma or other qualification at the University of Buckingham or any other University or similar institution.

Signature: Date: 1 July 2019

Acknowledgements

Several people have helped me while writing this thesis. Professor Lloyd Clark has now been my academic supervisor on two occasions and I am grateful for his guidance. I look forward to a third outing. I have Steve Goose, Director of NGO Human Rights Watch's Arms Division, to thank for my interest in the subject. Without his and his colleagues' initial work on *Killer Robots*, the world would still be asleep on the challenge that removing weapon supervision poses to society. Several thought leaders working on the subject in the Third Sector have been generous with their time including Richard Moyes (Article 36), Professor Noel Sharkey (University of Sheffield's Robotics department) and Mary Wareham (Human Rights Watch and Campaign to Stop Killer Robots). I am grateful to John Borrie (Programme Lead, UNIDIR), Professor Philip Sabin (War Studies, King's College, London), General Sir Andrew Sharpe (Centre for Historical Analysis and Conflict Research, Royal Military Academy) and Major-General Patrick Cordingley (Retd., Commander, 7th Armoured Brigade, Gulf War, 1991) for their input. I would like to thank Elena Webb and Professor Stefan Hawlin (Department of English, University of Buckingham) for their help in this thesis' formatting and MHRA compliance.

I am also grateful to those individuals who kindly read drafts of my thesis prior to its final submission including Steve Goose, Major-General Patrick Cordingley, Richard Moyes, Mary Wareham, Richard Elliott (Retd., 4th/7th Dragoon Guards), Dr. Hamish Mykura (EVP, Programming and Development, National Geographic International), Dr. Matthias Strohn (University of Buckingham and Center for Historical Analysis and Conflict Research, Royal Military Academy), Dr. Hongbo Du (School of Computing, Buckingham University), Delina Goxho (Open Society Foundation) and Chris Woods (AirWars).

Given the fast-moving nature of this topic, it should be noted that November 2018 represented the cut-off date for this thesis' research.

Styling

This thesis follows formatting conventions set out in *MHRA Style Guide*, Third Edition (London: Modern Humanities Research Association, 2013) pp. 1-104.

1. Introduction

Machines have long served as instruments of war but it has traditionally been the relevant commander who has decided *how* such weapons are employed. Evolution of technology, however, has the potential to change that reality and the purpose of this thesis is to analyse the widespread implications of this development. The work therefore considers challenges and consequences arising from the deployment[1] of Autonomous Weapon Systems (AWS) that may be capable of executing lethal engagements without human oversight. For the purposes of this introduction, an autonomous weapon is an armament in which the identification and selection of human targets and the initiation of violent force are carried out under machine control. Lethal capacities are thus delegated by the weapon system to its sub-components in ways that preclude deliberative and accountable human intervention.[2] The AWS can be described schematically as a weapon with sensors, algorithms and effectors that can include stationery as well as mobile robotic components.[3] Data collected by its sensors is processed computationally to enable independent detection, tracking and classification of objects. Target recognition can then be achieved by comparing this sensed data remotely or, more likely in a communications-denied environment, with target types contained in that weapon's database or perception library.[4] Finally, the system includes a weapon to engage selected targets.[5]

[1] These timeframes are broadly adopted as at 2019 in considering AWS deployment. Such phasing clearly moves year to year as evidenced by UNIDIR, 'Framing Discussion on the Weaponization of Increasingly Autonomous Technologies', *United Nations Convention for Certain Conventional Weapons*, (2014), generally <http://www.unidir.ch/files/publications/pdfs/framing-discussions-on-the-weaponization-of-increasingly-autonomous-technologies-en-606.pdf>. The definition of 'near-term', however, is not set in stone. See: Gabi Siboni and Yoni Eshpar, 'Dilemmas in the use of autonomous weapons', *Strategic Assessment*, The Institute for National Security Studies, 14, 4 (2014) <http://www.inss.org.il/wp-content/uploads/systemfiles/Dilemmas%20in%20the%20Use%20of%20Autonomous%20Weapons.pdf>. Timelines are also covered by the Brookings Institute in its 2009 discussion with PW Singer ('Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century') and General James Mattis, (2012), generally <https://www.brookings.edu/wp-content/uploads/2012/04/20090126_wired.pdf> and by Chapter 2 (*Context*), specifically: 2.4 ('*Defence planning'*). For the purposes of this thesis, *near-term* relates to the period up to 2025 and *medium-term* relates to 2025-2040.

[2] Lucy Suchman, 'Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy', in R. Geiss, 'Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security', *German Federal Foreign Office* (Germany: Berlin, 11 April 2017), pp. 273-283.

[3] This thesis focuses on mobile rather than static weapon platforms. In considering relevant models for the deployment of compliant AWS, Chapter 4 (*Deployment*) reviews the close correlation between a machine's ability to change its position and that machine's *non-defensive* weapon tasking in what is a wide scale of likely weapon configuration and assignment. It is this continuum that makes precision problematic in AWS definition except in terms of that platform's ability to remain compliant once human supervision has been removed. See: US Department of Defense, *Summer Study on Autonomy*, 11, (Defense Science Board, 2016) <https://www.hsdl.org/?abstract&did=794641> [accessed 13 February 2018]. The DoD's document conceptualizes technologies that are key to the development of autonomous systems in terms of 'sense, think/decide, act, team'.

[4] ICRC, 'Autonomous weapons systems: Technical, Military, Legal and Humanitarian aspects', *Experts meeting, CCW*, 64 (Geneva, Switzerland, 2014) <https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014> [accessed 17 November 2017].

[5] This definition accords with that adopted by UN Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions. See: Christof Heyns, 'UN Document A/HRC/23/47, Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions, United Nations' (*Human Rights Council, 23rd Session*, Agenda item 3, 27 May 2013) <https://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A.HRC.23.47.Add.5_ENG.pdf>. AWS is here defined as a 'robotic weapon system that once activated, can select and engage targets without further intervention by a human operator'.

Why is this subject so important? AWS posit becoming the 'third revolution in warfare'.[6] Once developed, autonomous weapons 'will permit armed conflict to be fought at a disruptive scale and to a super-fast timetable'.[7] A lethal engagement that is undertaken by an unsupervised weapon is also the manifestation of handing over the decision to kill to a computer. The past two decades have seen a broad range of technological advances that have now made this a practicable possibility and, in so doing, AWS deployment might render obsolete several precepts of current battlecraft.[8] The removal of meaningful human control from targeting sequences also questions how best to frame battlefield command, battlefield control and leadership which may abruptly no longer be fit for purpose. Indeed, this thesis' context is that current debate on AWS is heavily influenced by the apparent certainty of AWS' deployment.[9] As also noted by Kalmanovitz, AWS' deployment contributes 'to the already stark asymmetries of so-called riskless warfare' in which risks are increasingly passed across to civilian populations in an opponent's territory.[10] The matter's importance is also heightened by the absence of any international agency that is authorised to test such weapons, to control AWS risks and ensure global protection of civilian interests.

Several barriers, however, exist to the removal of supervision from weapon systems. A key purpose of this thesis is to highlight sources of uncertainty that might impact upon the decision to remove supervision in such weapons. This requires a wide set of tests in order to resolve four uncertainties. How might this new technology change current combat methods? After all, 'Fire and forget' weaponry (and the issues that they raise) are not new.[11] Second, is it possible to shoehorn these technology developments into existing engagement rules in a manner that still safeguards compliance with legal frameworks already in place?[12] As noted by Kalmazovitz, international criminal law holds that commanders who fail to avoid non-negligible risks to civilians and other protected persons can themselves be liable for negligence or recklessness.[13] Third, what weighting should be applied in this review to *contextual* components that are a part of the battlefield's processes such as command, political and social vectors? Finally to this point, defence planning that considers the removal of human supervision in lethal engagement must involve more than usual guesswork. It is not yet clear how such weaponry will be deployed on the

[6] See, generally: BBC, 'Killer Robots: Experts warn of Third Revolution in Warfare', *BBC website*, (2017) <http://www.bbc.co.uk/news/technology-40995835> [accessed 12 November 2017].

[7] BBC, 'Killer Robots: Experts warn of Third Revolution in Warfare', generally.

[8] This conditional is discussed in detail in Chapter 3 (*Drivers*) and *Chapter 4 (Deployment)*.

[9] Kenneth Anderson and others, 'Adapting the Law of Armed Conflict to Autonomous Weapon Systems', *International Law Studies,* 90, 386, (US: Stockton Center for the Study of International Law, 2014), 389-390 <http://www.dtic.mil/dtic/tr/fulltext/u2/a613290.pdf>.

[10] Pablo Kalmanovitz, *Judgement, liability and the risks of riskless warfare*, *Autonomous Weapons Systems: law, ethics, policy,* (UK: Cambridge University Press, 2016), p. 158.

[11] Economist Magazine, 'Trying to Restrain the Robots', para 3 of 29, (19 January 2019) <https://www.economist.com/briefing/2019/01/19/autonomous-weapons-and-the-new-laws-of-war> [accessed 2 February 2019].

[12] Detailed analysis of legal, ethical and operational obstacles to the removal of human oversight in lethal engagements is undertaken in Chapter 5 (*Obstacles*).

[13] Kalmanovitz, pp. 156-157. Various commentators including Neha Jain and Geoffrey Corn extend this form of liability ('role responsibility') to the use of AWS. The issue is discussed in Chapters 4 (*Deployment*). Chapter 5 (*Obstacles*) also discusses definitions and responsibilities of commander, politician and the procurement executive (see below: AWS' *Delivery Cohort*).

battlefield. Similarly, it is not clear what further technical advances await nor what priority and resources will be made available to AWS deployment.[14] Faced with the deployment of independent weapons, defence planning becomes considerably more challenging given the speculation that must be made concerning AWS' shape and capability. It is therefore unavoidable to make uncomfortable assumptions about what is likely and what is unlikely in this field. Arguments around whether weapon independence technically constitutes a revolution in military affairs (RMA) are thus prompted by the foundational, irreversible changes in battlefield processes posited by autonomous processes. This relationship, moreover, is complicated by the *cumulative* effects of such challenges and, often, by those challenges' unanticipated secondary effects. An appeal of unmanned assets is also their promise of generating mass but this too has unexpected behavioural consequences. It will drive up the ratio of AI-driven systems (both physical and virtual) to soldiers in uniform, leading over time to 'proportionately fewer points of consciousness within the [whole] system' and, consequently, to an increasingly untested framework for battlefield oversight.[15] In this vein, it is the independent behaviour of these machines (in what, notes Barrons, will be technologically complex and uncooperative weapon systems) that must cumulatively impact battlecraft efficiences, subsequent 'allocation of human bandwidth' as well as the ability of AWS' Delivery Cohort's to deliver on its tasks.[16] AWS deployment will at the very least require 'frictionful adjustment'[17] across working practices, skills, training and command.

In December 2016, the Fifth Review Conference of the Convention of Certain Conventional Weapons (CCW)[18] agreed to formalise discussions that first began in 2013 and 'explore and agree on possible recommendations on options related to emerging technologies in the area of LAWS, in the context of the objectives and purposes of the Convention, taking into account all proposals – past, present and future'.[19] This was a breakthrough development. The first meeting of the CCW

[14] MC Haas, 'Autonomous Weapon Systems: The Military's smartest toys?' *The National Interest*, (2014) <http://nationalinterest.org/feature/autonomous-weapon-systems-the-militarys-smartest-toys-11708> [accessed 13 July 17].

[15] Ministry of Defence, 'Human-Machine Teaming*', UK MOD, Joint Concept Note 1/18*, (2018), p. 44 <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/709359/20180517-concepts_uk_human_machine_teaming_jcn_1_18.pdf>.

[16] General Sir Richard Barrons, Commander Joint Forces Command (Retd.) in conversation with the author, 23 June 2016. For the purposes of this thesis, the term *Delivery Cohort* is used as a device to convey the parties involved in delivering the deployment of AWS and will include, inter alia, the following taskings: neurophysiologists to coordinate AWS networks, psychologists to coordinate learning and cognition, biologists for adaption strategies, engineers for control routines, logisticians, roboticists, electrical specialists, behaviorists, politicians, NGOs, sociologists, lawyers, company directors, weaponists, military tacticians, manufacturers, professionals involved in miniaturization, simulation, configuration, coding, power supply and modularity, specialists in sensors, in distributed and decentralized routines, ethicists, specialists in tooling and calibration.

[17] Ministry of Defence, 'Human-Machine Teaming', generally.

[18] The CCW is the UN-body that is tasked to 'prohibit or restrict further the use of certain conventional weapons in order to promote disarmament' and the 'codification and progressive development of the rules of international law applicable in armed conflict'. See: 'Preamble*,* 1980 Convention on Prohibitions on the Use of Certain Conventional Weapons Which May Be Deemed Excessively Injurious or to have Indiscriminate Effects'*, United Nations Treaty Collections,* 22495, (2 December 1983) <https://treaties.un.org/doc/Treaties/1983/12/19831202%2001-19%20AM/XXVI-2-revised.pdf>.

[19] Mary Wareham, Director, HRW Arms Division, in conversation with the author, December 2018. The wording is taken from the 'Final Report of the 2016 Informal Meeting of Experts on Lethal Autonomous Weapons Systems, Genev, *UN Documents Publishing*, (10 June 2016) <https://www.unog.ch/80256EDD006B8954/(httpAssets)/DDC13B243BA863E6C1257FDB00380A88/$file/Report LAWS_2016_AdvancedVersion.pdf>. Convention here relates to the Geneva Convention. A detailed analysis on International Humanitarian Law (IHL), International Human Rights Law (IHRL), the tipping point between these two

Group of Governmental Experts (GGE) in November 2017 took place after three informal CCW meetings during 2014-2016 and underlines the global importance now attached to limiting the introduction of what may be a materially different means of applying force. The context, after all, is that signatories to the CCW generally agree that '*any* use of force, including through [AWS], must strictly comply with international law and, in times of armed conflict, with IHL'.[20]

1.1 Thesis historiography

The structure of this thesis is driven by available historiography. The subject has therefore been approached using several broad sources of written research.[21] First, the premise of removing supervision from lethal engagements is examined through the prism of *fundamental* research that has been published on the nuts-and-bolts architecture of artificial intelligence (AI). What might theoretically be possible in the broad space of machine robotics and how readily might this translate to unsupervised weapons? An assumption here is borrowed from Hammond that the core methodology of AI has not changed materially over the past quarter century.[22] As also noted by Guszcza and Maddirala, AI agents are, in several ways, similar today to years' past.[23] This may not appear to be an intuitive assumption. Hammond therefore highlights that recent[24] improvements in AI, albeit narrow AI[25], have occurred not because any methodological discontinuity has surfaced to disrupt practices but because 'necessary computational capacity, raw volumes of data, and processing speed [is] now available so that the technology can shine'.[26] This assumption requires further analysis as it forms the basis of this thesis' technical review. While the central premise of network training has existed for more than thirty years, early AI efforts were based on a relatively tiny universe of just a few thousand examples being applied to poorly differentiated problem sets. Today this same technique remains the default 'but [is] now applied to hundreds of billions of examples and run on machines with specialized chips that allow them to learn from these examples much faster'.[27] Exactly the same dynamic holds true, notes

frameworks and their role in the Law of Armed Combat (LOAC) can be found in Chapter 5 (*Obstacles*), specifically 5.1 ('*Geneva Convention and Laws of Armed Combat'*).

[20] United Nations Office at Geneva, 'Possible challenges two international humanitarian law due to increasing degrees of autonomy*',* (2016), *NGO Article 36* website, <http://www.article36.org/wp-content/uploads/2015/04/Article-36-remarks-CCW-150415-IHL.pdf>.

[21] A writing methodology for this thesis is included later in this chapter. See, generally, section 1.7 ('*Statement of methods'*) and subsequent sections setting out the thesis' broad treatment of data, sources and argument formulation.

[22] K Hammond, 'Why Artificial Intelligence is succeeding: Then and Now'*, Computerworld*, *Artificial intelligence Today and Tomorrow*, (2015), para. 4 <http://www.computerworld.com/article/2982482/emerging-technology/why-artificial-intelligence-is-succeeding-then-and-now.html> [accessed 1 March 2017].

[23] J Guszcza and N Maddirala, 'Minds and Machines: The Art of Forecasting in the Age of Artificial Intelligence'*, Deloitte University Press, Deloitte Review*, 19 (2016), paras. 3-4 of 27 <https://dupress.deloitte.com/dup-us-en/deloitte-review/issue-19/art-of-forecasting-human-in-the-loop-machine-learning.html> [accessed 2 November 2016].

[24] Likely timelines for these developments are discussed throughout this thesis. As per footnote 1 to this chapter, *near-term* relates to the period up to 2025 and *medium-term* relates to 2025-2040. *Recent* here relates to developments witnessed over the preceding decade.

[25] AI relates here to non-sentient computer sequences that are focused on one narrow task or, perhaps, the combination of several such narrow techniques that are enhanced by access to massive data sets. Discussion on the important distinction between AI and AGI (Artificial General Intelligence) is set out in Chapter 6 (*Wetware*), specifically: 6.1 ('*Software versus intelligence'*).

[26] K Hammond, para. 9.

[27] Ibid.

Knight, for how an autonomous weapon might operate.[28] It was previously impossible. Today it 'might now be feasible'[29] leading to what Sabin refers to as a 'revolution of rising expectations'.[30] The point is that it is not just that algorithms have been improved; the disruption has been the massive datasets and fast chipsets available to them that now allow extraction of meaningful signal from often-noisy data.[31]

Conflicting context, however, suggests instead that AI has had a chequered past 'full of hype and disappointment'.[32] The field has witnessed bubbles involving expert systems, neural networks, hard and fuzzy logic models as well as a dependence upon (and then subsequent relegation of) complex statistics in order to enable machine reasoning. This thesis' technical analysis therefore relies on a second, broad and quite separate historiography of research discussing *current* best practice and developments in AI. These sources are global. They come from diverse academic institutions, defence establishments, practitioners and third sector parties.[33] This second historiography cohort is, furthermore, typically less than five years old and reflects current empirical work rather than what is theoretically possible in the science as it might relate to AWS deployment. A further purpose of this thesis then becomes the intermediation of this second body of work in order to generate inferences that are relevant to the deployment of autonomous weaponry. Examination of AI capabilities as they relate to this *battlefield* context is otherwise absent, as negligible historiography currently exists for this exercise. While research may be available on the matter's primary arguments, this thesis seeks to knit directly these fundamentals with specific issues that concern the removal of human supervision over battlefield weapons. A set of examples provides context. An unsupervised weapon clearly requires intricate routines to administer its onboard goals and values. It will require dynamically managed utility functions, the set of mathematical routines ranking alternative course of action according to their worth to the weapon system and its programming.[34] It will also require an anchoring mechanism to ensure the platform's actions do not stray inappropriately from its intended purpose. These

[28] W Knight, 'The US Military wants its Autonomous Machines to explain themselves', *MIT Technology Review*, (2017) <https://www.technologyreview.com/s/603795/the-us-military-wants-its-autonomous-machines-to-explain-themselves/> [accessed 2 July 2017].

[29] A Jhingran, 'Obsessing over Artificial Intelligence is the wrong way to think about the future*', Wired Magazine*, Business, (2016) <https://www.wired.com/2016/01/forget-ai-the-human-friendly-future-of-computing-is-already-here/> [accessed 29 June 2017].

[30] Professor Philip Sabin, Professor of Strategic Studies at KCL, in conversation with the author, 29 June 2017.

[31] The issue of data efficacy is a key theme in assessing compliant AWS deployment and is covered in Chapter 7 (*Software,* specifically *Sources of technical debt*) and Chapter 9 (*Hardware*), specifically: 9.1 *('Hardware and sensor fusion issues for AWS'*).

[32] Rodney Brooks, 'The Seven Deadly Sins of AI Prediction', *MIT Technology Review'*, (6 October 2017), paras. 3-4 and generally <https://www.technologyreview.com/s/609048/the-seven-deadly-sins-of-ai-predictions/> [accessed 12 August 2018].

[33] The thesis' bibliography is thematically divided into five discrete sources: ethical, historical, legal, operational and technical. A comprehensive historiography of on-line sources then comprises a sixth section to the bibliography. For detail on this thesis' referencing, see section 1.7 ('*Statement of methods'*).

[34] For a discussion on these topics, see: Chapter 8 (*Software*), specifically: 8.3 ('*Utility functions'*) and 8.5 ('*Anchoring and Goal setting issues'*). More generally, Chapter 6 (*Wetware*) isolates complexities arising from likely fundamental architectures in autonomous weapons. Chapter 8 (*Software*) identifies specific fault lines that might result from specific operational *routines* likely to underpin weapons-directing artificial intelligence. Chapter 9 (*Hardware*) then highlights difficulties stemming from the *physical* properties of such systems. The point of these three chapters is to demonstrate the *cumulative* technical complexity that underlies AWS deployment. Taken together, the thesis' review of coding and architecture pinpoints discrepancies that exist between *capabilities* that are feasible and *tasks* that are essential to these platforms' function.

capabilities may have robust theoretical foundations but involve magnitudes of complexity if actually to be deployed in a battlefield setting.

The thesis' first research question therefore emerges directly from extrapolating these two elements of historiography. Is the removal supervision from lethal engagements *technically* feasible? This is not straightforward given the breakneck development of technical competencies. An obvious danger for this thesis is being blindsided by the unforeseen.[35] Indeed, it is not new for academics in this field to posit scenarios that are based on amalgamations of existing technologies which suggest that can targets can be engaged without such supervision.[36] Technical progress, however, tends to be chaotic with technologies evolving continually rather than arriving fully formed. A 'certain' dead-end today is tomorrow's ubiquitous breakthrough solution. Nor is it necessarily clear which *grouping* of technologies might overturn how battlecraft may, in time, be undertaken.[37] This is an important observation. As noted by Scharre, 'an autonomous weapon need not to be very intelligent. It is simply a weapon that selects and engages targets on its own'.[38] In this vein, the pace of technical advance requires that assumptions continually be made on AWS' timelines and capabilities and, in considering AWS' feasibility, on the broad context that must frame such removal of supervision from battlefield weapons.

It is this knitting together of themes that also informs the thesis' third section of historiography. Assessing AWS' feasibility requires a detailed appreciation of existing constraints that together require human supervision in lethal engagements be retained. This third historiography therefore comprises the corpus of work that argues *against* the introduction of autonomy in weapon systems. This constituent covers existing supranational Laws Of Armed Combat (LOAC) and, on a more granular level, evidence relating to Rules of Engagement (ROE).[39] It includes analysis of moral, ethical, and economic arguments against the withdrawal of Meaningful Human Control (MHC) across engagements.[40] This third component also covers what emerges as the key role of context and situational awareness in the argument. Again, much of this secondary material must be extrapolated from existing historiography in order to comment specifically on the matter of oversight in lethal battlefield engagement. The thesis' fourth

---

[35] Caroline Crampton, 'Why is it so hard to predict the future of technology?', *New Statesman*, (2017) <http://www.newstatesman.com/culture/observations/2017/01/why-it-so-hard-predict-future-technology> [accessed 25 June 2017].

[36] Daniel Oberhaus, 'Watch 'Slaughterbot': A Warning about the Future of Killer Robots', *Motherboard*, (2017) <https://motherboard.vice.com/en_us/article/9kqmy5/slaughterbots-autonomous-weapons-future-of-life> [accessed 17 November 2017]. Professor Stuart Russell, Director of Computing at Berkeley University, concludes that 'this is not speculation. It is the result of integrating and miniaturizing technologies that we already have'. Russell's *Slaughterbox* provides a useful baseline at the time of this thesis' writing and informs a degree of its contextual anlaysis; see introduction to Chapter 2 (*Context*).

[37] Battlecraft is generally defined here as the skills and techniques of military combat in the sustained fight between organised armed forces.

[38] Paul Scharre, 'Presentation at the United Nations Convention of Certain Conventional Weapons', *Informal Meeting of Experts on Lethal Autonomous Weapons, Geneva*, (2015), p. 2 <https://www.unog.ch/80256EDD006B8954/(httpAssets)/98B8F054634E0C7EC1257E2F005759B0/$file/Scharre+presentation+text.pdf>.

[39] See: Chapter 5 (*Obstacles*), specifically, 5.1 ('*Geneva Conventions and the Laws of Armed Conflict*'). The section discusses the nature of constraints that might collude against compliant AWS deployment (political, social, command, environmental, ethical, accounting, physical, behavioural and proliferation constraints).

[40] MHC is a key concept (and conclusion) for this thesis and is discussed in detail in Chapter 10 (*Oversight*), specifically: 10.1 ('*Meaningful Human Control*') and Chapter 11 (*Conclusion*).

historiographical element is then comprised of sources that together combine to *promote* weapon autonomy. This component provides the counterfactual to the thesis' preceding three sections and covers issues such as perceived political dividends, procurement advantages, operational benefits and the role of dual-use technologies. It also factors for the attraction of force multiplication and the appeal of remote engagement. It reviews those arguments where machine involvement in lethal engagements might actually lead to *better* ethical performance in combat situations and where AI might make battlefields safer for humans, especially civilians.[41] It might also be expected to cover empirical deployment of unsupervised systems and how such systems might be incorporated into future battlecraft. Here, however, the corpus remains surprisingly scant as evaluation of AWS deployment from an empirically technical perspective barely exists. The thesis' fifth and final historiographical component then reviews possible solutions in this AWS debate and acts as a synthesis in considering best current practice in weapon supervision and the matter of MHC in lethal engagements.

Given this framework, discrete areas of inquiry emerge as a basis for the thesis' research questions, each related to the central question around AWS feasibility. First, how intractable are technical, operational and contextual impediments to the eventual removal of human supervision from lethal engagements? Second, what might interim and transitional autonomous systems look like? Third, how will States adopt such weapons and how will they be fitted into battlecraft? A fourth question then asks whether a role exists (and the shape of that role) for MHC to become an over-arching control mechanism and, conceivably, a statutory umbrella for the deployment of unsupervised weapons. A complication is that this exercise depends upon assessing context, both its relevance and its weighting in AWS deployment. As noted by Talwar, after all, autonomy is not in itself a solution to any problem.[42] Instead, the utility of autonomous capability on the battlefield is a function of the 'ecology of each mission's needs and operating environment. There is no value here without context'.[43] This is doubly relevant. As Ricks concludes, 'at the end of the [planning] process, there will be gaps in facts that necessarily have to be filled by rational assumptions'.[44] This may appear uncomfortable given that computing was little more than a fringe activity less than two generations ago. It underlines, however, the role of context in determining whether supervision should remain a battlefield prerequisite. As such, this analysis is deliberately undertaken from a behavioural perspective. This is foremost a humanities-based analysis of AWS feasibility. It considers, after all, the *concepts* that underlie AWS deployment rather than, say, specifics of AWS' coding. The methodology of the thesis is not to provide detailed line-by-line technical assessment of AWS routines but instead to offer a conceptual analysis of the determinants and possible consequences of removing oversight from weapon technology. Technical challenges are therefore deliberately framed in the perspective of behaviour and context. While a decision to use force may depend on a mix of legal, ethical and strategic matters,

[41] See: Chapter 3 (*Drivers*), specifically: 3.4 ('*Ethical Drivers*') and the accompanying analysis of R Arkin, *Governing lethal behaviour: Embedding ethics in a hybrid deliberate/reactive robot architecture*, (Atlanta, GA: Georgia Institute of Technology, 2007), generally.

[42] Rohit Talwar and others, 'Keeping the Human Touch; humans need a new mindset to function in a tech-dominated society', *Financial Times*, People's Technology section, (2017) p. 35.

[43] US Department of Defense, 'The Role of Autonomy in DoD Systems', *Task Force Report*, (2012), p. 21<https://fas.org/irp/agency/dod/dsb/autonomy.pdf>.

[44] See, generally: T Ricks, 'Staff Planning: It's all about examining assumptions and then re-examining them', *Foreign Policy.com*, (2016) <http://foreignpolicy.com/2016/01/12/staff-planning-its-all-about-examining-assumptions-and-then-re-examining-them/> [accessed 12 June 2017].

the conclusion in this broad case is that machines remain enduringly incapable of understanding the context of their actions.

The analysis is concerned by what is *feasible* around human supervision. To that end, the thesis is framed less by its analysis of drivers accelerating AWS adoption and more, in fact, by challenges that lurk in deployment models which posit how weapons with gradually less human supervision may be adopted. What is not in doubt is a general lessening of human involvement across combat activities, whether physically through increased adoption of unmanned weapon platforms or generally through wide introduction of teaming autonomous processes across combat assets. This notion drives the UK Ministry of Defence's Joint Concept note on *Human-Machine Teaming* where 'future force design must find the optimal mix of manned and unmanned platforms, and balance employment of human and machine cognition for various tasks'.[45] The challenge, however, is that the idealised 'centaur model' founded upon human-machine teaming breaks down irreparably once battlefield actions are required that are faster than the human operator can provide. That model similarly relies on sufficient (and empirically unlikely) communication being available between machine and human. The machine is on its own. One final element of context is relevant to this preamble. In considering these challenges, this analysis is limited to the current era of narrow AI; should artificial *general* intelligence and machine sentience ever reach battlespace, then both the assumptions and deductions of this thesis will no longer be valid.

The thesis' litmus test is therefore whole-weapon independence with the analysis assuming deployment of wide-task AWS; if a weapon can select its own target and engage that target without human involvement then the thesis' technical commentary remains intentionally agnostic to the *level* of that autonomy. While this helps future-proof the work's conclusions, the thesis must still be sure about its assumptions in order to identify *cumulative* ramifications arising from AWS deployment. An example, inferred separately from Lombardi and Jones, is relevant in order to demonstrate this layered nature of the issue.[46] There is, for instance, likely to be grey area between the deployment of autonomous battlefield decision aids and, in time, stand-alone weapons that are capable of independent action.[47] The appeal here of weapons that partner human soldiers in outwardly narrow tasks is also that they provide force multiplication and do this at a potentially lower cost in terms of casualties. The broad capabilities, however, that these systems must first integrate are fundamentally challenging around target identification, target

[45] Ministry of Defence, 'Human-Machine Teaming', p. 44.

[46] B Lombardi, 'Assumptions and Grand Strategy', *Defence Research and Development, Canadian Centre for Operations Research and Analysis,* (2011), p. 38 <http://ssi.armywarcollege.edu/pubs/parameters/articles/2011spring/lombardi.pdf>.

[47] S Jones, 'AI and Robots line up for Battlefield Service', *Financial Times*, (2016) <https://www.ft.com/content/02d4d586-78e9-11e6-97ae-647294649b28?mhq5j=e2> [accessed 12 February 2017]. A decision aid here relates to a 'colleague' weapon that provides varying levels of force multiplication. See, generally: Chapter 4 (*Deployment*), specifically: 4.3 ('*Machine and human teaming models'*) and 4.5 ('*Flexible autonomy*') that consider current practice and relevant emerging technologies in autonomous weapons.

selection and target engagement.[48] They must also then be integrated into a State's battlecraft.[49] Moreover, effective human-machine collaboration will require that all team members (humans and 'colleague' machines alike) share common goals, values and utility functions notwithstanding that those goals may be expressed in different frameworks and semantics.[50] Task specificity must then require complex and bespoke rules within each class of independent weapon that will also vary by mission and vary over time. It is the complexity that is created by these conditions that undermines the feasible and compliant deployment of armed autonomous weapons and provides the basis of this thesis' inquiry.

1.2 Introduction to key concepts

This introduction now undertakes two tasks. It provides an overview to the argument's central concepts in order better to prepare the reader for discussion on the challenges to AWS deployment. It also provides an appraisal of important themes that comprise the work's individual chapter headings in order to map the thesis' overall structure.[51] The key hypothesis is that the right of combatants[52] to choose their means and methods of warfare[53] *is not unlimited.*[54] This highlights a key principle. Humans should exercise control over combat operations but also, crucially, over *individual* attacks.[55] This, after all, is a basic tenet of international humanitarian law (IHL), a central component of the Law of Armed Conflict (LOAC) or Law of War that underpins much of this thesis' later analysis.[56] Yielding the decision to kill to a machine is the key issue that confronts this study. There are several underlying elements to this observation and, for the purposes of this introduction (and expanded in later chapters), it is helpful to touch on a handful

[48] It is useful here to signpost the structure of the thesis. Evidence and argument for statements made in this Introduction are set out in subsequent chapters as follows: Context (Chapter 2); drivers to adoption and deployment (Chapter 3); current practices and likely pathways to the removal of human supervision in engagements (Chapter 4); legal and other obstacles in front of such adoption (chapter 5); architectural (Chapters 6 and 7) and control issues (Chapter 8) that challenge AWS deployment;likely equipment deficiencies (Chapter 9, *Hardware*); the concept and role of Meaningful Human Control (MHC) in lethal engagements (Chapter 10, *Oversight*). The thesis' general methodology is set out in Section 1.7 ('*Statement of methods'*) and subsequent sections of Chapter 1 (*Introduction*).

[49] This thesis primarily considers the deployment of weapons autonomy from the perspective of a sophisticated, resourced State and, generally, one of the 196 signatories to the Geneva Convention. This assumption is valid as the Convention confers an obligation to comply with, inter alia, the Laws of Armed Combat and other responsibilities and commitments set out in Chapter 5 (*Obstacles*).

[50] See: Chapter 8 (*Software*), specifically: 8.5 (' *Anchoring and goal setting issues'*), 8.6 ('*Value setting issues'*) and 8.3 ('*Utility function'*).

[51] See, generally, Thesis' chapter headings: *Context, Drivers, Deployment, Obstacles, Wetware, Firmware, Software, Hardware* and *Oversight.*

[52] For the purposes of this dissertation, a combatant is defined under the 'Third Geneva Convention' (*Article 3, GCIII*, 1949) as an individual taking direct part in the hostilities of an armed conflict.

[53] 'The Protocol Additional to the Geneva Conventions' (August 1947) and relating to 'The Protection of Victims of International Armed Conflicts (Protocol I)', (8 June 1977) hereinafter referred to as *Additional Protocol I*; the document refers alternately to 'methods and/or means of warfare', 'means and methods of attack' and 'weapon, means or method of warfare', <https://ihl-databases.icrc.org/ihl/INTRO/470> [accessed 12 May 2017].

[54] This principle is variously stipulated in, for instance, 'Article 22' of the *Hague Regulations* (1907); See: 'Respecting the Laws and Customs of War on Land' and 'Article 35(1)' of *Additional Protocol I*, <https://ihl-databases.icrc.org/ihl/INTRO/470> [accessed 12 May 2017].

[55] As defined in US Department of Defense, 'DoD Dictionary of Military and Associated Terms', (2017) <http://www.dtic.mil/doctrine/new_pubs/dictionary.pdf>.

[56] International Review of the Red Cross, 'A guide to the legal review of new weapons, means and methods of warfare: Measures to implement Article 36 of Additional Protocol I of 1977', *International Committee of the Red Cross Geneva*, 88, 864, (2006), p. 931.

of these conflicting issues. First, replacing human soldiers by machines might have benefits in terms of reducing casualties for the machines' owner, reducing his costs and addressing certain technical challenges, but it also might lower the threshold for going to battle and risk starting an arms race in new technologies.[57] The Future of Life Institute notes that the way that humans resolve conflict shapes the society in which those humans live, the consequence being that machines which can make the decision about who and when to kill clearly has the ability to fundamentally change this.[58] It is the availability of such independent weapons to *any* political cohort that might then have broad constitutional ramifications. Suddenly, the decision of only a very few people (those in a position to authorise AWS deployment) may be required for a polity to drift into war, so reversing a trend (in the West) by further centralising decisions to go to war.[59] Unlike nuclear devices, after all, AWS technology requires no costly or hard-to-obtain materials. Finally to this point, analysis of complex machines that can kill without oversight is unexpectedly involved, unexpectedly imprecise and requires detailed overlay of legal, ethical, contextual and operational constraints.[60] It is for this reason that this thesis' humanities-based assessment can still examine AWS deployment from both a technical *and* behavioural perspective.[61]

In this vein, it is useful to review certain terms that will appear throughout the thesis.[62] Autonomous weapons were initially categorised by the US Department of Defense into three sub-types, as noted by NGO Human Rights Watch (HRW) in its November 2012 report *Losing Humanity* and defined according to the amount of human involvement in their actions.[63] Human-*in*-the-loop weapons comprise robots that can select targets but only deliver force with a human command.[64] Human-*on*-the-loop weapons can select targets and deliver force under the oversight

[57] For a study of AWS' perceived combat benefits see: Mike Guetlein, 'Lethal Autonomous Weapons; Ethical and Doctrinal Implications', *US Department of Joint Military Operations*, (2005), generally <http://www.dtic.mil/dtic/tr/fulltext/u2/a464896.pdf>.

[58] Future of Life Institute, 'Autonomous Weapons: An Open Letter from AI and Robotics Researchers', *Future of Life*, (2015) <https://futureoflife.org/open-letter-autonomous-weapons/> [accessed 6 April 2016]. It may, for instance, become political consensus that autonomous systems are ideal for certain tasks such assassinations, destabilizing neighbouring States and other polities, subduing particular populations or even selectively killing a particular ethnic group.

[59] Daniel Suarez, 'The kill decision shouldn't belong to a robot', *Ted.com*, (2013) <https://www.ted.com/talks/daniel_suarez_the_kill_decision_shouldn_t_belong_to_a_robot> [accessed 18 May 2017].

[60] For discussion on situational awareness, see: Chapter 2 (*Context*), specifically: 2.6 ('*The role of situational awareness and uncertainty*'). For analysis on its technical ramifications, see: Chapter 9 (*Hardware*), specifically 9.1 ('*Hardware and sensor fusion issues for AWS*').

[61] For a discussion on technical faultlines to AWS, see: Chapters 6 (*Wetware*), specifically 6.5 ('*Missing Pieces*'), 8 (*Software*), specifically 8.1 ('*Coding methodologies*') and 8.2 ('*Coding errors*') and 9 (*Hardware*), generally. For discussion on legal, ethical and other behavioural constraints on AWS deployment see: Chapter 5 (*Obstacles*).

[62] A discussion of relevant definitions also appears in Paddy Walker, 'Killer Robots? The Role of Autonomous Weapons on the modern battlefield', MA thesis, *Buckingham University*, (2013), p. 5 and pp. 16-19.

[63] Bonnie Docherty, 'Losing Humanity – the Case against Killer Robots', *Human Rights* Watch, (2012) <http://www.hrw.org/reports/2012/11/19/losing-humanity-0> [accessed 2 June 2015], p. 2. It should be noted that neither the CCW nor UN States-parties are currently considering concerns over possible use of fully autonomous weapons *outside* of armed conflict (for example, policing and law enforcement and border control). HRW has subsequently investigated concerns in this area in its May 2014 'Shaking the Foundations' Report; see: <https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots#> [accessed 13 May 2018].

[64] These definitions have been developed by Human Rights Watch. HRW's paper is extensively cited in this thesis. For a discussion and timeline on reports on AWS, see: HRW and Harvard Law School International Human Rights Law Clinic, 'Reviewing the Record: Reports on Killer Robots from Human Rights Watch an Harvard Law School International Human Rights Law Clinic, (2018), generally <http://hrp.law.harvard.edu/wp-

of a human operator who can override the robots' actions. HRW's third definition covers fully autonomous robots that are capable of selecting targets and delivering force without any human input or interaction. This thesis concerns this third category of autonomous, human-*out*-of-the-loop weapon systems although, as evidenced below, the original heuristic is no longer wholly helpful as it does little to identify elements of the argument relating to what is potentially risky, dangerous or prone to unintended consequences.[65] It is also important to understand the difference between machines that are *autonomous* (machines that are self-learning and therefore 'evolving', the focus of this thesis), *automated* (machines that are nevertheless complex and rules-based but are not able to learn through assimilated feedback) and *automatic* machines that are simply based on programmed thresholds. It is the level of human control and intervention (and, conversely, the degree of machine freedom) that is the key distinction for this thesis. It is also the *whereabouts* in a weapon system where autonomous function is to be found.[66] Simply referring to autonomy as a general weapon attribute is too imprecise and it is thus the task *nature* being undertaken autonomously at subsystem or function level that matters to this analysis. To this point, Sharkey notes that certain autonomous functions such as navigation may of course be quite uncontentious while others, such as targeting, present enduring difficulties.[67]

Three other classifications require discussion. While *semi*-autonomy involves human oversight in target selection[68], *supervised*-autonomy involves machine-selected targets with humans subsequently confirming any lethal engagement.[69] It is then human and communication shortcomings that might lead to *full* weapon autonomy whereby the machine is selecting and engaging targets without recourse to human supervision. Sartor and Omicini also point out the distinction between 'capability-independence', the weapon's ability to accomplish a task, and 'organisational independence', the ability of the weapon to achieve that task 'within the sociotechnical infrastructure as a whole'.[70] The gap between these two independencies is a key theme of this thesis and, in its technological context, the relevant type of autonomy to this thesis

content/uploads/2018/08/Killer_Robots_Handout.pdf>. The extent of NGO advocacy against AWS deployment is best evidenced by the broad composition of civil society organisations comprising coalitions such as the Campaign to Stop Killer Robots (incorporating, inter alia, HRW, Amnesty International, AAR Japan, ICRAC, Mines Action Canada, Pugwash Conferences on Science and World Affairs, PAX Netherlands, Nobel Women's Initiative, Article 36, SEHLAC Latin America and Women's League for Peace and Freedom). An analysis of civil society's AWS debate is provided by University of Sheffield's Research Excellent Framework: Impact Case Study (Ref 3B), 'Shaping International Policy and Stimulating International Public Debtate of Autonomous Weapon Systems', Sheffield University, 2014, <https://www.sheffield.ac.uk/polopoly_fs/1.434127!/file/policy_study.pdf>.

[65] Future of Life Institute, 'Autonomous weapons: an interview with the experts' (with Ariel Conn, Heather Roff and Peter Asaro, *www.futureoflife.org*, (2016), p. 3 <http://futureoflife.org/2016/11/30/transcript-automous-weapons-interview-experts/1> [accessed 5 January 2017].

[66] Chapter 10 (*Oversight*), specifically: 10.1 ('*Meaningful Human Control*').

[67] Noel Sharkey, 'Saying "No!" to Lethal Autonomous Targeting', *Journal of Military Ethics*, *Ethical and Emerging Military Technology*, 4, (2010), 369-383.

[68] Examples including homing munitions, UAV with GPS guided munitions, counter–rocket artillery and certain sensor-fused weapons.

[69] Examples for this sub-group include Aegis and Patriot missile defence; specific weapon systems and their classification are dealt with later in this thesis. See, generally: Chapter 4 (*Deployment*).

[70] Giovanni Sartor and Andrea Omicini, in N Bhuta and others (eds.), *Autonomous Weapons Systems: Law, Ethics, Policy*, (Cambridge: Cambridge University Press, 2016), p. 44.

may be *task* autonomy.[71] In this sense, the categorization of weapons into lethal, non-lethal or less-lethal is actually less helpful as it masks the fact that weapon effects are never solely a function of that weapon's design but also depend upon its use and the vulnerabilities of those affected by it.[72]

1.3 Timelines around capabilities

This thesis concerns the *future* battlefield.[73] Weapons capable of identifying, tracking and engaging incoming targets without supervision (thus far, broadly limited to a defensive role) already function without human engagement in their decision-making.[74] Generally stationary, such platforms are currently designed repeatedly to perform pre-programmed actions within tightly set parameters and time frames.[75] Such weapons currently perform in quite structured and controlled environments.[76] Notwithstanding a widespread 'revolution in expectations' around weapon development[77], broadly capable AWS deployment remains still a hypothetical construct[78] leading States and other polities to be wary of defining too clearly their responses to the technology, in particular on positions around statutory instruments that might ban weapon autonomy.[79] The context is that, while pre-cursor technologies are reviewed below, any broad genre of unsupervised weaponry remains in its infancy despite long-dated attempts to automate weapons.[80]

---

[71] As opposed, for instance, to *personal* autonomy which frequently dominates ethical analysis of autonomy. Accordingly, a weapon system is regarded here as autonomous if it is able to select and engage military targets without human intervention to carry out that task.

[72] Serious injury or mental trauma is recognised in ECtHR, 'Abdullah Yasa et al versus Turkey', App. No.44827/08, *European Court of Human Rights, Information Notes on the Court's Case*-law, 176, Judgement, (July 2014). As set out in Chapter 4 (Treatment amounting to instrument or degrading treatment concerns under both IHL and IHRL.

[73] For a discussion on likely timelines, see: Chapter 2 (*Context*), specifically: 2.4 ('*Defence planning'*).

[74] SIPRI analysis, introduction to Chapter 3 (*Drivers*).

[75] See: Chapter 4 (*Deployment*), specifically: 4.3 ('*Human and machine teaming models'*).

[76] Jurgen Attman and Frank Sauer, 'Autonomous Weapon Systems and Strategic Stability', *Survival*, 59, 5, (2017), p. 118.

[77] For a useful discussion on 'the battlefield of the future' see, generally: R Wood, 'The Technical Revolution in Military Affairs', (2010) <www.holtz.org/library/technology/technical_revolution_in_military_affairs> [accessed 2 March 16]. As Wood comments, 'innovation in weapons design is driven by the absolute need for survival and, thus, will always advance aggressively… The study of innovation is crucial since the technology of war interacts with the actual practice of fighting'. See also: R Rubenstein and others, *Practicing military anthropology: Beyond expectations and traditional boundaries* (VA: Sterling, Kumarian Press, 2013), generally.

[78] In particular, the task type and capability set available to AWS weapon classes. See, generally: Center for a New American Security, 'Autonomous Weapons and Human Control', *CNAS*, (2016), pp. 4-5 <https://www.files.ethz.ch/isn/196780/CNAS_Autonomous_Weapons_poster_FINAL%20(1).pdf>.

[79] The position of the UK Foreign Office is an example. See: John Templeton Stroud, UK UN delegation, *Fifth Review Conference, Convention for Conventional Weapons*, HRW/Article 36 side event, December 2016. UK's negotiating position within the CCW has instead been to point to both current absence and long-term unfeasibility of AWS to define its hands-off stance on AWS deployment policy.

[80] See, generally: Chapter 4 (*Deployment*).For a detailed review of pre-cursor weapon autonomy see also: Walker, p. 19. The earliest attempt at a powered unmanned aerial vehicle was A.M Lowe's 'aerial target' of 1916. See: Jane's 'Book of Remotely Piloted Vehicles', Collier Books, (1977) <https://books.google.co.uk/books/about/Jane_s_pocket_book_of_remotely_piloted_v.html?id=8o9TAAAAMAAJ> [accessed 28 May 2017]. Since the 1980s, technical advances have allowed engineers to contemplate bringing autonomy to battlefield weaponry. Transistors, for instance, have 'shrunk from the size of a fingernail to today's high-end microprocessor with its billion transistors'.

It is also useful for this introduction to provide definition around the timelines under discussion. The challenge that this creates is neatly framed by Clarke's aphorism from 1962: 'It is impossible to predict the future and all attempts to do so in any detail appear ludicrous within a few years'.[81] While this thesis' consideration of defence planning later refutes Clarke's pessimism[82], for its purposes 'near-term' is reckoned very generally to be within the six-or-so-year period to 2025.[83] 'Medium-term' relates to developments that may be expected to occur in the fifteen years thereafter to 2040. Opinions on the timetable for weapons-directing artificial intelligence are, however, 'as confident as they are diverse'.[84] By way of context, Bostrom estimates from recent surveys that human-level machine intelligence (HLMI) has a fifty per cent probability of arriving by 2040 and a ninety percent probability by 2075. Procurement timelines depend, however, upon the scope of weapon capabilities being envisaged and a more granular example is required. Sadowski, Chief Robot Scientist at the US DoD, suggests that autonomous machines for quite narrow 'hauling techniques for the military' may be deployable within ten years and more complicated convoy applications will be available 'sometime after'.[85] Even within this narrow vertical, he forecasts a time frame of not less than fifteen years before any 'widespread deployment of fully autonomous line-haul convoys'. Other parties are less conservative and evidence the contradictory nature of these development timelines. In an open letter published in July 2015, leading practitioners in the field of artificial intelligence warned that the underlying technology behind lethal autonomous systems would be feasible 'in years, not decades'.[86] In considering the likely capabilities that such AI-based systems might require, its signatories concluded that artificial intelligence technology has *already* reached a point where the deployment of such systems is already practically (if not legally) feasible.

A further challenge is also to define the *capabilities* that will comprise AWS.[87] To be autonomous, a weapon system must have the capability to select independently among different courses of action in order to accomplish goals that are based on its knowledge (both received and learned) and subsequently derived understanding of the world, itself and its immediate

---

[81] David Bawden, 'The nature of prediction and the information future: Arthur C. Clarke's Odyssey vision', *Aslib Proceedings,* 49, 3, (1997), pp. 57-60.

[82] See: Chapter 2 (*Context*), specifically: 2.4 ('*Defence planning*'). See also: Chapter 4 (*Deployment*), specifically: 4.2 ('*Planning tools'*).

[83] As set out in footnote 1, these timeframes are broadly adopted as at 2019 in considering AWS deployment. Such phasing clearly moves year to year as evidenced by UNIDIR, 'Framing Discussion on the Weaponization of Increasingly Autonomous Technologies', *United Nations Convention for Certain Conventional Weapons*, (2014), generally <http://www.unidir.ch/files/publications/pdfs/framing-discussions-on-the-weaponization-of-increasingly-autonomous-technologies-en-606.pdf>. See also: The Brookings Institution's 2009 discussion with PW Singer ('Wired for War: The Robotics Revolution and Conflict in the Twenty-first Century') and General James Mattis, (2012), generally <https://www.brookings.edu/wp-content/uploads/2012/04/20090126_wired.pdf>. Also; Major-General Patrick Cordingley (Commander, 7th Armoured Brigade, Gulf War, 1991), in conversation with the author, June 2016.

[84] Nick Bostrom, *Superintelligence; Paths, Dangers, Strategies,* (Oxford: Oxford University Press, 2014), p. 19.

[85] Breaking Defense, 'Interview with Bob Sadowski, US Army Chief Roboticist', *www.breakingdefense.com,* (2016) <http://breakingdefense.com> [accessed 12 October 2016].

[86] Future of Life Institute, (2015), generally.

[87] See: Chapters 8 (*Software*) and 9 (*Hardware*). The thesis concentrates upon *State* actors' institutional deployment of AWS rather than use of autonomous weapons by *non-state* parties. Much of the analysis is derived in sources from the United States; this arises from much of the thought leadership around the subject currently coming from America. Unless otherwise stipulated, it is intended that the arguments hold for any sovereign State involved in these platforms.

situation.[88] This requires a very broad list of competencies that is investigated in subsequent chapters.[89] Ansell notes that autonomous weapons must be anchored by a degree of reasoning that is based on data, an ability to act independently according to that data, a capacity to decide and an awareness of its surroundings.[90] Weapons, after all, which are merely automatic 'are controlled by an on/off switch'.[91] A ramification for this thesis is that unsupervised weapons will be characterised by 'a movement from today's 'mission execution' to tomorrow's 'mission performance', the difference being that the former simply executes a pre-programmed plan whereas autonomous performance involves mission outcomes that can vary even during a mission'.[92] Identifying the technical stretch that exists between such envisaged capabilities and machine functions that will likely be available becomes an important component of this analysis.[93]

1.4 AWS classification issues

Two further introductory classifications are relevant to frame this thesis. Weapons autonomy should be viewed as a sliding scale that extends, on the one hand, from weapons capable of narrowly defined supervised tasks to autonomous machines capable of complex thought processes.[94] Second, individual AWS components must work independently[95] but also act *together* in order to provide self-authorisation in an engagement.[96] Such capabilities are already under test. Since 2014, unmanned aircraft demonstrators such as the X-47B have been able to fly a mission with no involvement of a ground-based 'pilot'.[97] The tipping point between weapon automaticity and autonomy is therefore quite clear; it becomes the combination of platform and weapon that

[88] Department of Defense, 'Unmanned Systems Integrated Roadmap FY2013-2038', *Under-Secretary of Defense Acquisition, Technology and Logistics*, Reference 14-S-0553, Washington, p. 4, (November 2013) <http://www.dtic.mil/dtic/tr/fulltext/u2/a592015.pdf>.

[89] Chapter 8 (Software), specifically: 8.1 ('*Coding methodologies'*) and 8.4 ('*Software processing functions'*).

[90] '… and ideally all of these four characteristics together'; Dr Darren Ansell, University of Central Lancashire's School of Computing, in conversation with the author, Chatham House Conference, February 2014.

[91] Ibid.

[92] 'Unmanned Systems Integrated Roadmap FY 2013-2038', p. 67.

[93] See: Chapters 6 (Wetware), 7 (*Firmware*), 8 (*Software*) and 9 (*Hardware*). As detailed in this analysis, AWS mission performance will involve the unit's ability to integrate sensing, deep learning, perceiving, analysing, communicating, planning, decision making and executing in order to achieve such complex mission goals.

[94] See: Chapter 4 *(Deployment),* specifically: 4.3 ('*Machine and human teaming models*') and 4.5 ('*Flexible autonomy*'). See also: Royal Air Force Directorate of Defence Studies, 'Air Power – UAVs: The wider context', ed. Owen Barnes, (2015), p. 69 <https://www.scribd.com/document/52847466/Air-Power-UAVS-The-Wider-Context> [accessed 12 June 2017].

[95] See: 'Unmanned Systems Integrated Roadmap, FY2013-2038', *Department of Defense; '*Unmanned systems that have the option to operate autonomously are typically fully pre-programmed to perform defined actions repeatedly and independent of external influence or control', p. 66.

[96] Dr Ansell, School of Computing, University of Central Lancashire, in conversation with the author, Chatham House Conference on Autonomous Military Technologies, February 2014; Ansell's position highlights the distinction between 'full machine authority' versus machines that act unless revoked ('direct support'), machines that advise and act if authorised ('in support'), and machines that are merely 'advisory'. Again, see: Chapter 4 *(Deployment),* specifically: 4.3 ('*Machine and human teaming models'*).

[97] Jonathan Marcus, 'Robot Warriors: Lethal machines coming of age', *BBC magazine*, (2014) <http://www.bbc.co.uk/news/magazine-21576376> [accessed 19 June 2016]. Mary Wareham of HRW points out that the X-47B has since being repurposed into an air-born fuel tanker reconfiguring the internal payload originally intended to carry weapons.

creates a 'weapon that is lethal, a weapon that on its own can kill'[98] and one that can 'select and engage targets without further intervention'.[99] This, however, is only part of the equation as it relates to AWS deployment. Russell and Norvig rightfully promote a different perspective in characterising autonomy as 'an agent's capacity to learn what it can to compensate for partial or incorrect prior knowledge'.[100] Conversely, autonomy is absent should that agent simply rely for its operation upon the prior knowledge of its designer rather than on its own percepts.[101] The important distinction is identified by Hanon whereby the weapon has autonomous choice regarding target *selection* in the use of lethal force.[102]

1.5 Thesis structure

How then does the available historiography inform this thesis' high-level structure?[103] Its argument is built upon behavioural analysis (deployment's context, accelerators and obstacles) that is complimented by technical analysis (AWS' architectural, coding and processing challenges). A methods statement discussing this thesis' processes, in particular those procedures covering its treatment of primary source material, is included in this introductory chapter.[104] Indeed, the thesis' opening chapters seek to map such a basis against which to consider the fielding of AWS. On the one hand are politics, culture and society.[105] A quite separate view is provided by the

[98] Jody Williams, Human Rights Watch, recipient of Nobel Peace Prize for work on landmine ban, in conversation with author, CCW GGE meeting, Geneva, 24 November 2017.

[99] Christof Heyns, 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions', p. 8. This distinction informs the position of NGO Article 36. See: Article 36, 'Structuring debate on autonomous weapon systems', Memorandum for delegates to the *Convention on Certain Conventional Weapons* (CCW), Geneva (14 November 2013). Similarly, HRW's Mary Wareham points to the definition of weapons 'able to select and attack targets without any human intervention' being the broadly adopted working definition for AWS (in conversation with the author, June 2014); See also: Campaign to Stop Killer Robots, 'Urgent Action Needed to Ban Fullt Autonomous Weapons', *CSKR London*, (23 April 2013), generally <http://stopkillerrobots.org/wp-content/uploads/2013/04/KRC_LaunchStatement_23Apr2013.pdf>.

[100] SJ Russell and P Norvig, 'Artificial Intelligence: A modern approach', (Upper Saddle River, NJ: Prentice Hall, 2010) p. 39. The focus here is on cognitive capacity, especially the capability to obtain new knowledge through interaction with the agent's immediate environment. See: Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues*').

[101] Sartor and Omicini, p. 50.

[102] A popular analysis of the subject is provided by Leighton Hanon, 'Robots on the Battlefield – are we ready for them?', *American Institute of Aeronautics and Astronautics*, (2004), p. 7 <http://arc.aiaa.org/doi/abs/10.2514/6.2004-6409> [accessed 2 January 2014]. AWS technology is already subject to a defined set of recognised institutional measurements. These grey scales of autonomy range from Level 1 (simple remote guidance) to Level 10 (full autonomy). Autonomous Control Level (ACL) 6 will, for instance, allow multiple unmanned weapon systems to recognise multiple targets and allocate those targets between systems. A further example is useful. ACL 9 is intended to enable groups of automated systems to assess the battlefield, the number and the location of targets. ACL 9 incorporates AWS' analysis of targets' threat potential in order to allocate overall mission priorities. It even envisages the skipping between low and high value targets. For a popular analysis of the subject, see: *Drone360 Magazine*, p.62 <http://www.drone360mag.com> (April/May 2016) [accessed 16 February 2016]. It is noteworthy for this analysis that the US Military's Global Hawk AUV has autonomous take-off and landing, can self-determine destinations, adjust and set speeds, altitude, roll, pitch and yaw but only ranks at *2.5* on the ACL scale of 1-10.

[103] An overview is as follows: First, it looks to establish a contextual framework to frame the breadth of factors that influence the broader question of weapon control (Chapter 2, *Context*). It then identifies drivers influencing the removal of human oversight in lethal engagements (Chapter 3, *Drivers*). Material obstacles exist, however, for any material shift towards independent weapons. Chapters 5 through 9 (*Obstacles, Wetware, Software, Firmware,* and *Hardware*) seek first to identify and then to evaluate the significance of technical and other related faultlines that exist to AWS deployment. These chapters together consider AWS' technical feasibility. Chapter 10 (*Oversight*) then analyses the concept of MHC as a key pivot to the AWS debate.

[104] See, generally: Section 1.7 ('*Statement of methods*').

[105] See: Colin Gray, *Strategy and defence planning; meeting the challenge of uncertainty,* (Oxford: Oxford University Press, 2014). Also: Colin Gray, *The Future of Strategy,* (Polity Books, 2015) and *Another Bloody Century,* (Phoenix,

'military man's framework'[106] of technological promise and battlefield experience tempered by an understanding of what can *practically* be achieved.[107] Two issues arise from this contention. While Sabin suggests that every generation may think it is witnessing a technical discontinuity in military affairs[108], three fifths of the thesis will nevertheless be taken up reviewing *technical* faultlines that, on a cumulative basis and when assessed together, certainly impact upon compliant deployment of independent weapons. Second, the very complexity of AWS' deployment framework (juggling these factors' relative precedence in order to arrive at appropriate context) itself deserves attention, as it too must inform subsequent 'planning certainty'. Hammes borrows from Bismark to make the point: 'The Statesman is like a wayfarer in the forest who knows in which direction he is walking but not at what point he will emerge from the trees'.[109] Chapter Two's review of context thus informs much of the thesis' subsequent analysis and quite deliberately covers the broadest possible scope; ethical, legal, political, social and cultural components are each integers that vie for prominence with the technical throughout this deployment debate and each requires appropriate heft. The role of the 'human dimension'[110], for instance, has recurring weight in setting this thesis' assumptions. This is borne out by developments in recent military doctrine: The US Army Combined Arms Center identifies precisely such challenges arising from 'the rapid evolution of methods, the complex and dynamic mix of cultures, a broad range of actors and unprecedented proliferation of technology' and does so in order to optimize its 'most agile resource, *its people'*. [111] Only within this framework does the Center see its resources being able 'to thrive in the ambiguity and chaos of 2015'.[112] Tipping points exist, therefore, when one element (here, perhaps the human dimension) trumps other contextual components and, certainly, technical components within this thesis' broad argument; the assertion will actually be that good soldiers with lesser equipment will eventually outplay poor soldiers armed with latest technology.[113] It is always contextual issues that will fashion the deployment of weapons autonomy.

In this vein, this thesis considers *why* and to *whom* the deployment of AWS might appeal and the several drivers that press for autonomy in States' (and others') arsenals. Unsurprisingly, this is

2005). Gray is Director, Centre for Strategic Studies, Department of Politics and International Relations, Reading University. In particular, Chapter 2 (*Context*) seeks to repurpose several of Gray's arguments within the AWS debate.

[106] General Sir Richard Barrons, Commander Joint Forces Command (Retd.) in conversation with the author, 23 June 2016.

[107] Several sources consider this relationship. See, generally: J-C Ruano-Borbalan, 'Technology, Science and Society; Norms, Cultures and Institutions matter', *Journal of Innovation Economics and Management*, 1, 22, (2017), 3-8 <https://www.cairn.info/revue-journal-of-innovation-economics-2017-1-page-3.htm> [accessed 12 December 2016]

[108] Professor Philip Sabin, Professor of Strategic Studies at KCL, in conversation with the author, 29 June 2017.

[109] T Hammes, 'Assumptions – A Fatal Oversight', *Infinity Journal*, 1, (2010) <https://www.infinityjournal.com/article/1/Assumptions__A_Fatal_Oversight/> [accessed 2 July 2017].

[110] US Army, 'Army Human Dimension Strategy 2015; Building cohesive teams to win in a complex world', *ACAC*, (2015) p. 1 <http://usacac.army.mil/sites/default/files/publications/20150524_Human_Dimension_Strategy_vr_Signature_WM_1.pdf>.

[111] US Army Combined Arms Center, 'The Human Dimension White Paper; a framework for optimizing human performance', *USACAC*, (2014) p. 6 <http://usacac.army.mil/sites/default/files/documents/cact/HumanDimensionWhitePaper.pdf>.

[112] Ibid., pp. 6-11.

[113] Here, 'assertion', *not* assumption; see: J. Storr, *The Human Face of War,* 25, 7 (A&C Black, 2009) p. 200.

also an involved relationship. While recent growth in military robotics and unmanned systems[114] may have multiple drivers, empirical advantages to their deployment are more complex to identify[115], a phenomenon that is reviewed in Chapter Three (*Drivers*) through its analysis of technology creep and the rise of machine autonomy in other commercial sectors.[116] Given that such systems are unmanned and, as such, with a definable economic cost should they be lost, appropriate weighting can also be given to operational factors, the appeal of 'force multiplication' and the optionality that AWS deployment may provide commanders through adoption of riskier (and likely idiosyncratic) tactics.[117] Increasing autonomy in weapon systems also suggests quicker reaction to adversarial threats through an accelerated targeting-decision-cycle and the speeding up of data processing.[118] As noted by Russell, it promises better persistence and better endurance while also reducing humans' exposure to enemy fire.[119] A driver therefore stems from what is referred to above as a 'revolution in expectation'.[120] Here, weapons autonomy appears to be an inescapable development.[121] Indeed, the US Department of Defense's *Unmanned Systems Integrated Roadmap FY 2013-2038* plots an unambiguous course whereby 'the prevalence and uses of unmanned systems continues to grow at a *dramatic* pace'.[122] It has thus become a broad procurement assumption[123] that robotics are ideal for 'dull' missions[124] (long-duration undertakings with mundane tasks ill-suited for manned systems), 'dirty' missions (exposure to hazardous conditions) as well as deep (behind enemy lines) and 'dangerous' missions.

Chapter Three also reviews those theoretical arguments that promote AWS deployment as a means of raising ethical standards on the battlefield. The notion here is that machines promise to

---

[114] See, generally: Marketsandmarkets.com, 'Military Robots Market to be worth 21.11 US$ Billions by 2020', *Markets and Markets,* <http://www.marketsandmarkets.com/PressReleases/military-robots.asp> [accessed 17 July 2017].

[115] Robotic World blog, 'The use and advantages of military robots', *A Robotic World*, <http://minerrobot.weebly.com/the-use-and-advantages-of-military-robots.html> [accessed 17 January 2017].

[116] Throughout this work, the term *driver* is used to refer to accelerators, prompts and catalysts for particular further actions. Source: Dictionary.com <http://www.dictionary.com/browse/drive>: i) definition 25, 'vigorous onset or onward course towards a goal or objective' and ii) coincidentally, definition 26 'a strong military offensive'. For a pictorial review of recent combat robotics see: S Melendez, 'The Rise of the Robots: What the future holds for the world's armies*,* (2017), *Fast Company blog* <https://www.fastcompany.com/3069048/where-are-military-robots-headed> [accessed 4 July 2017].

[117] The term 'battlefield commander' does not relate in this work to any particular rank but is used throughout to convey a nomenclature of that superior who is controlling, usually, the in-theatre deployment of AWS. This may or may not be separate from other political authority in the decision process. It also relates to a level of accountability and legal responsibility. Force multiplication is discussed in detail in Chapter 3 (*Drivers*), specifically: 3.3 ('*Structural and procurement drivers'*) and 3.5 ('*Operational drivers*').

[118] See: Chapter 3 (*Drivers*). The very best human fighter pilot needs at least 0.3 seconds to respond to a simple stimulus and more than twice as long to make a choice between several possible responses. Chapter 3 (*Drivers*) similarly discusses the OODA Loop ('Observe, Orient, Decide and Act').

[119] S Russell, D Dewey and M Tegmark, 'Research priorities for robust and beneficial artificial intelligence'*, AI Magazine*, (2015), 104-114 <http://futureoflife.org/data/documents/research_priorities.pdf>.

[120] Professor Philip Sabin, Professor of Strategic Studies at KCL, in conversation with the author, 29 June 2017.

[121] Heather Roff, 'The Self-Fulfilling Prophesy of High-tech War'*, Duck of Minerva*, (2015) <http://duckofminerva.com/2015/12/the-self-fulfilling-prophecy-of-high-tech-war.html> [accessed 3 March 2017].

[122] Department of Defense, 'Unmanned Systems Integrated Roadmap FY 2013-2038', p. 20.

[123] Melendez, (2017), generally.

[124] The relevance of autonomous unmanned machines for dull, dirty, deep and dangerous missions is well covered by PJ Neal, 'From Unique Needs to Modular Platforms: The Future of Military Robotics'*, US Naval Institute*, (2010), pp. 1-7.

'remove' humans from that combat frontline. Regardless of training procedures, it has proved empirically unrealistic to assume that humans obey LOAC in moments of combat stress: In this case, Malhoney notes that soldiers 'are, at best, a variable tool in waging war'.[125] The corollary is therefore that AWS deployment may empirically be justified by recurring lapses in battlefield *jus in bello.*[126] Instead, a theoretical construct can be posited whereby autonomous weapons operate under a transparent, code-based 'ethical governor' that edits machine actions, goals and values *in advance* of lethal engagement. Further to this point, AWS may operate more appropriately than human soldiers if built without need for self-protection. Again, however, this is to mask difficulties if AWS deployment is to be reliable and compliant: Just as it is unclear how unsupervised machines can be shoehorned into battlefield operations, it is uncertain how removing oversight will, in its wide definition, affect performance.[127] Moreover, efforts to deploy battlefield autonomy deflect focus from the fact that changes in how missions are accomplished will result in wholly new consequences that must be understood and integrated into subsequent battlecraft.[128] The thesis' deployment analysis therefore considers several continua that exist between, at one end, the notion of self-learning and independent 'killer robots' that sentiently roam the battlefield to the more likely advent of task-specific, human-machine teaming that involves hardware with specific autonomous capabilities in specific applications.[129] A purpose of Chapter Four (*Deployment*) is thus to consider definitional conflicts that arise from the wide variety of static/mobile platforms and defensive/aggressive assignments which might comprise AWS deployment as well as their incorporation into subsequent battlecraft.

As noted by Cummings, autonomy is also likely to enable the execution of wholly *new* mission types[130], particularly in areas such as cyber and electronic warfare in which decision speed is critical to success.[131] Its adoption, however, is more likely to be characterised by incremental replacement of human oversight from an ever-increasing number of currently supervised battlefield tasks. That scope is illustrated by a set of lethal and non-lethal scenarios that is set out by the US Department of Defense's 2016 study of weapons autonomy.[132] These range from covertly-deployed networks of smart mines to stand-alone systems controlling the rapid-fire exchange of cyber weapons, from swarming autonomous machines intended to disrupt enemy operations to unmanned, sentient aircraft capable of adaptively jamming enemy positional, navigational and timing (PNT) capabilities.[133] Scope here also concerns the mechanisms for AWS

---

[125] Col S Malhoney, *Ethics Theory for the Military Professional*, 32, 3, (Air University Review, 1981), p. 55.

[126] *Jus in bello* et al is discussed in Chapter 5 (*Constraints*), specifically: 5.1 ('*The Geneva Convention and Laws of Armed Combat*').

[127] See, in particular: Chapter 4 (*Deployment*), specifically 4.7 ('*Operations and causes of failure'*).

[128] Here, consequences are understood primarily in their operational rather than any tactical or strategic meaning. Battlecraft is defined throughout as the skills and techniques that comprise military combat including procedures and the deployment of military assets.

[129] See, generally: Human Rights Watch, Arms Division, *Losing Humanity – the Case against Killer Robots,* (USA: Washington, 2012) <http://www.hrw.org/reports/2012/11/19/losing-humanity-0>. The organisation is widely recognised to have been the first NGO to highlight the issues posed by the deployment of autonomous weapons.

[130] A detailed tasking review is provided by ML Cummings, 'Artificial intelligence and the future of warfare' *The Royal Institute of International Affairs*, (2017) <https://www.chathamhouse.org/publication/artificial-intelligence-and-future-warfare> [accessed 2 April 2017].

[131] US Department of Defense, 'Summer Study on Autonomy'**,** *US Defense Science Board***,** p. 11.

[132] Ibid., p. 4.

[133] Ibid., generally.

overrule and interruption. Are such platforms to be command-executing whereby the weapon receives its order, carries it out and then pauses to await the next command or a sovereign-type system that has open-ended mandate to operate on the battlefield in pursuit of broad objectives?[134]

1.6 Introduction to AWS feasibility

It is this portfolio of challenges that occasions the thesis' overarching inquiry into the matter of technical feasibility and whether or not such deployment obstacles are in fact intractable. Chapters Five to Ten identify and unpick these challenges. First, chapter Five (*Obstacles*) identifies non-technical constraints, reviewing the legal framework into which such weapons must fit in order to highlight areas where compliance is likely to be problematic.[135] This thesis adopts ICRC's broad assumption that procuring parties (here, States rather than non-State players) *do* place weight on their weapons being LOAC compliant.[136] If, for instance, it is a legal condition to battlefield engagement[137] that there be unambiguous distinction between combatants and non-combatants[138], how then might AWS fit into existing frameworks that require such a finely nuanced calculation?[139] There are, after all, four such legal hurdles that comprise LOAC. Each requires satisfaction. Is each specific attack proportional? Is it militarily necessary? Has due process has demonstrably been undertaken to ensure the selected target is a *bone fide* combatant? Moreover, LOAC requires that appropriate action be taken to prevent unnecessary human suffering arising from that (and every) lethal engagement. The analysis will demonstrate that other 'soft' complexities exist. It is, as an example, very difficult for coding to capture the fluid nature of LOAC, in particular the imprecision that exists between International Humanitarian Law (IHL), International Human Rights Law (IHRL) and rules on targeting.[140] It will also be challenging to write compliant engagement routines given very different national interpretations that exist on the most basic components of formal rules of engagement.[141] Both ICRC and Palin separately note that it is the battlefield consequences of this complexity[142] that should dictate humans remain

---

[134] Nick Bostrom, *Superintelligence; Paths, dangers, strategies,* (Oxford: Oxford University Press, 2014) p. 148. See also: Chapter 8 (*Software*), specifically: 8.3 ('*Utility function'*), 8.5 ('*Anchoring and goal-setting issues'*) and 8.6 ('*Value setting issues'*).

[135] Specifically the legal, ethical, political, social and economic constraints to removing human supervision in lethal engagements.

[136] International Committee of the Red Cross, 'The Use of Armed Drones Must Comply with the Laws of Armed Combat', *ICRC*, (2013) <https://www.icrc.org/eng/resources/documents/interview/2013/05-10-drone-weapons-ihl.htm> [accessed 2 February 2016].

[137] A detailed discussion on the ramifications of IHL and IHRL on inter-State lethal engagements is undertaken in Chapter 5 (*Obstacles*), specifically 5.1 ('*The Geneva Convention and Laws of Armed Conflict*').

[138] Lucy Suchman and Jutta Weber, 'Human-Machine Autonomies' in *Autonomous Weapon Systems: Law, Ethics, Policy,* (Cambridge: Cambridge University Press, 2016), p. 1.

[139] For a detailed discussion on this nuance, see: JFR Boddens Hosang, 'Rules of Engagement; rules on the use of force as linchpin for the international law of military operations*', UvA-DARE*, (University of Amsterdam, 2017), pp. 59-86.

[140] See: Chapter 5 (*Obstacles*), specifically: 5.1 ('The *Geneva Convention and Laws of Armed Conflict'*).

[141] Throughout this thesis, UK ROE is taken from JSP 383, 'The Joint Service Manual of the Law of Armed Conflict', *Ministry of Defence*, (UK MOD Publications, 28 August 2013) <https://www.gov.uk/government/collections/jsp-383> [accessed 12 January 2017].

[142] International Red Cross, 'Handbook on International Rules Governing Military Operations', *ICRC*, (2013) <https://www.icrc.org/sites/default/files/topic/file_plus_list/0431-

intimately involved in the engagement sequence.[143] The point is that it is often not straightforward to determine *which* legal framework[144] on hostilities applies in each engagement.[145] The distinction highlights a general dislocation between, on one hand, the capabilities that should comprise a legally-compliant yet unsupervised weapon and, on the other hand, what might realistically be possible using a code-based framework to execute on those same required actions. It is this observation that reinforces the degree of priority given to the role of technology in this thesis' arguments.[146]

The distinction provides a helpful bridge to considering whether the whole construct of weapon independence is terminally undermined by these technical shortcomings. Forming a view on this relationship principally requires dissection of machine learning (ML) processes to determine how they can fit with operational imperatives but also with LOAC obligations.[147] As noted by Christensen, the challenge here is to make forecasts in a discipline where technology development is so fast-moving.[148] A key precept is highlighted by Benitez whereby all weapons and all weapon consequences must be *controllable* and, in AWS deployment, whether that control can still be maintained absent of supervision across the components that make up each lethal engagement.[149] Indeed, a broad corpus exists to evidence that this is a long-standing requirement for the moral acceptability, political legitimacy and general legality of organised violence.[150] A focus for this thesis is therefore to determine the balance that must exist, on the one hand, between envisaged tasks for the AWS and, on the other, between prospective technical capabilities and the practicalities of battlefield control as exercised by those engaged in AWS deployment (for the purposes of this thesis, hereinafter labelled the Delivery Cohort as defined below).[151] Several ramifications arise from this incongruity. Enabling a weapon to select targets makes it problematic for the relevant human commander both to predict and to understand every specific target, every *precise* engagement moment, the location where violence is administered and the environment within which violent effects are undertaken. This informs a conclusion that commanders deploying AWS are prima facie unable to control machine behaviours that are within

handbook_on_international_rules_governing_military_oprations.pdf>. The publication runs for 1,464 pages and covers general obligations during combat including targeting and command responsibilities.

143 R Palin, *Multinational Military Forces: Problems and Prospect,* (London: Adelphi Paper 294, Routledge, 2005) p. 34.

144 Be it International Humanitarian Law (IHL) or International Human Rights Law (IHRL).

145 Suchman, 'Situational awareness and adherence to the principle of distinction', p. 8.

146 US Department of Defense, 'The Role of Autonomy in DoD Systems', pp. 21-29.

147 See again: Chapters 6 (*Wetware*), 7 (*Firmware*) and 8 (*Software*).

148 For a broad discussion on current global autonomous developments (Antarctica, Ocean, Swarm technologies), see: H Christensen, 'On their own: Research on autonomous technology is developing increasingly sophisticated capability in air, marine and around robotic vehicles', *Georgia Tech Research Institute*, undated <https://gtri.gatech.edu/casestudy/autonomous-technology-research-developing-increasi> [accessed 2 June 2017].

149 Mike Benitez, 'It's About Time: The Pressing Need to Evolve the Kill Chain', *War on the Rocks*, (2017) <https://warontherocks.com/2017/05/its-about-time-the-pressing-need-to-evolve-the-kill-chain/> [accessed 29 October 2017].

150 The harmful effects of weapons must be foreseeable and must at all times be under the control of those who employ them. See: International Committee of the Red Cross, *Draft rules for the limitation of the dangers incurred by the civilian population in time of war,* Article 14, *ICRC*, (1956). See also: International Law Commission, *Articles on the responsibility of States for internationally wrongful acts,* Article 8 and 23 (1), UNGA Res 56/83, (2001).

151 For the purposes of this thesis, the term Delivery Cohort is a useful device used throughout to aggregate the several parties involved in implementing AWS adoption. See footnote 16 to this chapter for definition.

their responsibility.[152]

Goff and Brooks point out that an ability to make appropriate judgement may already be diluted as irregular warfare removes conventional designations from battle zones and competencies.[153] A different point (but supporting this notion of erosion and predictability) is made by Santor and Omicini who note that even when a human operator is tasked with authorising force (here, 'pushing the button' and the release of lethal force against an identified and considered target), 'important aspects of the decisional process that leads to selection and engagement [will already have] been delegated to automated system[s]'.[154] Indeed, the twin cores of machine autonomy and human involvement in the targeting loop are not a zero sum such that increasing the one results in a corresponding decrease of the other. As suggesed by SIPRI, this would otherwise become a virtuous circle.[155] This, however, is oveshadowed by the greater hitch to ensuring predictability whereby AWS' instructions must remain prescriptive specifications that cannot properly be defined by current coding methodologies[156], particularly around compliant selection of targets.[157] Given that such function must be based entirely upon code-based representations of the machine's immediate environment (there can, by definition, be no human tuning of its processes), any unforeseen change[158] to that environment or uncorreographed operation outside of that environment 'will necessarily lead to unpredictability in its functioning'.[159] Machine predictability therefore arises as a key priority to commanders and the Delivery Cohort.[160] Asaro and others conclude, moreover, 'as the behaviour of automated systems becomes more complex, and more dependent on imports from environmental sensors and external data sources, the less predictable [AWS] become'.[161] As an adjunct, this thesis must also consider the 'temporal' aspects of AWS deployment and the consequences of warfare's automation narrowing the timeframe for consequential situational assessment.[162] In reflecting

---

[152] Chapter 5 (*Obstacles*), specifically: 5.6 ('*Behavioural constraints'*).

[153] Kendall Gott and Michael Brooks, *Warfare in the Age of Non-State Actors*, (Kansas: Combat Studies Institute Press, 2007), pp. 209-230 and 225-342.

[154] Sartor and Omicini, p. 61.

[155] Stockholm International Peace Research Institute, *'*Implementation of Article 36 Weapon Reviews in light of increasing autonomy in weapon systems*', SIPRI*, (2015) <https://www.sipri.org/media/press-release/2015/implementing-article-36-weapon-reviews-light-increasing-autonomy-weapon-systems> [accessed 25 April 2017].

[156] Lucy Suchman and J Weber, 'Human-machine Autonomies' in N Bhuta et al (eds), *Autonomous Weapons Systems: Law, Ethics, Policy* (Cambridge University Press, 2016), p. 85.

[157] Suchman, *'*Situational awareness and adherence to the principle of distinction*'*, p. 9.

[158] L Weiss, 'Autonomous Weapons in the Fog of War'*, IEEE Spectrum*, (2012) <http://spectrum.ieee.org/robotics/military-robots/autonomous-robots-in-the-fog-of-war> [accessed 6 July 2017].

[159] Paul Scharre, 'Robotics on the Battlefield, Part II: The Coming Swarm'*, CNAS*, (2014) <https://s3.amazonaws.com/files.cnas.org/documents/cnas_TheComingSwarm_Scharre.pdf.

[160] General Gary Luck and others, 'Joint Operations: Insights and Best Practices'*, Joint Warfare Center, US Joint Forces Command,* (2008), p. 21 <http://www.au.af.mil/au/awc/awcgate/jfcom/joint_ops_insights_july_2008.pdf>.

[161] Peter Asaro, *Cybernetics and autonomous weapons: reflections and responses,* XXXIII, 3, (Italy: Paragigmi: Rivista di critica filosofica, 2015), pp. 83-107.

[162] John Horgon, 'The undiscovered mind: Exploring the world of artificial intelligence' in *When Machines outsmart humans*, (Nick Bostrom blog, Futures, Vol 35:7, 2000) <https://nickbostrom.com/2050/outsmart.html>; Kassan similarly concludes on AI machines that 'we don't know what to build, much less how to build it'. Horgan is also pessimistic; 'the neuroscience appears to be making anti-progress; the more information we acquire, the less we seem to know'.

upon predictable outcomes (and given that the definition between surveillance and targeting practices continues to blur), even the boundaries between weapon types is not always clear as weapon sub-components become less distinct and even geographically distributed.[163]

A key purpose of this thesis is to query the depth and persistence of such flaws. For this, the analysis generally assumes the deployment of wide-task AWS with broad autonomous capabilities. What emerges is an imprecise set of component relationships, a brittleness arising from a broad portfolio of sub-systems and the creation of 'technical debt' that stems from AWS' complexity.[164] Furthermore, these consequences are exacerbated by AWS' foundational requirement for dynamic re-basing of its routines including anchoring, goal and value setting as well as the updating of both utility and planning functions.[165] These may appear discrete issues of calibration, verification and testing but this would be to ignore an unending requirement (and, notes Marchant, the associated difficulty)[166] for AWS to undertake moment-by-moment tuning absent of third party involvement.[167] Calibration in this case relates not to maximizing performance but to the verifiable measure of maintaining LOAC compliance.

In doing so, it is necessary to question the likely fit between AWS deployment and battlefield operations including, inter alia, a review of metrics such as adopted rules of engagement and command.[168] A purpose of this section is therefore to evidence a fundamental change in practices that will be needed as AWS are deployed. Indeed, several well-tried concepts that have long comprised battlecraft may no longer be fit for purpose. Chapter Ten (*Oversight*) considers the key and enduring role of humans in lethal engagement and whether MHC might provide an on-going control mechanism and statutory umbrella for AWS deployment. The matter has been an agenda item in the United Nations' Convention for Conventional Weapons since 2014 and, while progress in that body continues to be patchy, the historiography for this thesis' final section is nevertheless developing quickly.[169] Roff and Moyes, for instance, identify the overarching significance of *selection* over targets to be the critical control function of a weapon.[170] In this vein, the concept of MHC encapsulates the 'when, where and how weapons are used; what or whom they are used

[163] Katharine Hall Kinderrater, 'The Emergence of Lethal Surveillance: Watching and Killing in the History of Drone Technology', *Security Dialogue*, 47, 3, (25 January 2016), 223-238.

[164] Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt'*).

[165] Chapter 8 (*Software*), specifically: 8.3 ('*Utility function'*) and 8.5 ('*Anchoring and goal setting issues'*).

[166] 'During 2009, US drones sent back… more than 24 years' worth of video footage'. Here, the *Economist Magazine* predicted that 2011 would produce 30 times as much information'.

[167] See, generally: G Marchant and others, 'International Governance of Autonomous Military Robots', *Columbia Science and Technology Law Review*, XII, (2011), p. 274 <http://stlr.org/download/volumes/volume12/marchant.pdf>. See also: Chapter 10 (Oversight), specifically: 10.3 ('*Validation and testing*').

[168] Chapter 10 (*Oversight*), specifically: 10.1 ('*Meaningful Human Control'*). See also: Chapter 5 (*Obstacles*), specifically; 5.6 ('*Behavioural constraints'*).

[169] HRW's Mary Wareham notes that there were no discussions on killer robots at the CCW until November 2013 when the annual meeting agreed to add the subject to the agenda. In October 2012, HRW was to co-found the Campaign to Stop Killer Robots.

[170] Heather Roff and Richard Moyes, 'Meaningful Human Control, Artificial Intelligence and Autonomous Weapons', *UN Convention for Conventional Weapons*, (2016), generally. See also: International Committee of the Red Cross, 'Statement of the International Committee of the Red Cross', *ICRC,* Geneva, (2015).

against; and the effects of their use'.[171] In considering MHC's role, the emphasis in this thesis is specifically on human targets; that is, the identification of humans or human-inhabited objects (buildings, vehicles) as lawful targets for engagement.[172] The thesis reviews (and adopts) Roff and Moyes' position that MHC 'is best considered as operating at three different layers: *ante bellum*, *in bello* and *post bellum'*.[173] The distinction is important as it suggests that MHC must imbue *all* phases of battlecraft. The ramification is that *each* of these factors informs and then shapes each *other* constituent of the control debate relating to 'the design, acquisition and use of tools of violence'. MHC is therefore not a single test in the engagement sequence. Instead, it should be an overarching benchmark that provides a framework to violence which is applicable right down to the level of *individual* direct attacks. Similarly, its obligations require the operator and his command chain (however that may then be comprised) to evaluate the expected outcome of using that specific weapon in that specific context with this obligation existing for each and every lethal engagement.[174] Moreover, notes Asaro, the AWS is a machine deployed for a certain purpose: In itself it is devoid of agency or intentionality and, on this basis, it is particularly appropriate first to consider the broad role of context as one component of challenge to AWS deployment.[175]

1.7 Statement of methods

The purpose here is to detail this thesis' methods, in particular its handling of primary sources which comprise a key part of the work's original contribution to its subject matter. The chapter's balance therefore reviews evidence-gathering, sampling and data collection methods as well as the ethics framework underpinning the thesis' writing as a social science research project. The thesis has primarily been undertaken using *qualitative* methods. This requires amplification. For the purposes herewith, qualitative methods involve 'description of *kinds* of characteristics without exclusive recourse to terms of measurements or amounts'.[176] They are multi-method in their focus involving an interpretative, naturalistic approach to the subject matter. They also provide a study structure 'within natural settings, attempting to make sense of and interpret phenomena in terms of the meanings of people bring to them'.[177] The thesis' extensive attention upon the role of *context* in removing weapon supervision is a consequence of this approach which is underpinned by Tashakkori and Teddlie's 'belief in the value-ladeness of inquiry, a belief in the theory-ladeness of facts but also belief that reality is multiple and constructed'.[178] By way of balance, this thesis' structure is also driven by the contention of King, Keohane and Verba that such differences (between quantitative and quantitative approaches) 'are mainly ones of

[171] Article 36, 'Killing by machine; key issues understanding for meaningful human control', 2015, generally.

[172] Suchman, 'Situational awareness and adherence to the principle of distinction', p. 4. The thesis is therefore less concerned with *defensive* weapon systems that operate on the basis of unambiguous signals from another (unmanned or uninhabited) device that comprises an imminent threat.

[173] Roff and Moyes, p. 3.

[174] Indeed, there is a personal legal requirement for the human pulling that trigger to investigate, understand and then act upon that balance of probabilities.

[175] Peter Asaro, *Determinism, Machine Agency and Responsibility,* 2, (Italy: Politica & Societa, 2014), pp. 265-292.

[176] R Murray Thomas, *Blending Qualitative and Quantitative Research Methods in Thesis and Dissertations*, (Corwin Press, Sage, California, 2003), p. ix.

[177] Ibid., p.1.

[178] Abbas Tashakkori and Charles Teddlie, *Mixed Mixed Methodology: Combining Qualitative and Quantitative Approaches*, Applied Social Research Methods Series, Volume 46, (Thousand Oaks, Sage publications, 1998), p. 13.

style and specific technique' and that the best research should combine features of each paradigm.[179]

For this purpose, the thesis' bibliography is divided into five sections. The first four sections set out book and journal references underpinning the thesis' ethical, historical, legal, operational and technical analysis. The bibliography's more extensive section then lists on-line sources and articles.[180] Such historiographical evidence points to the comprehensive secondary research that is widely available on, for instance, the fundamentals of AI and what then is theoretically possible from such technologies in removing supervision from lethal engagements. The guiding purpose to the methods listed below (case study, personal experience, interview, observational, historical and interactional inputs) has therefore been to deduce, refine and then opine from this reference bank. In this vein, it is possible to compose and prove a theory around the challenges to the deployment of autonomous weapons. Interviews in this space were then directed at proving assumptions that AI's core methodologies have not changed materially over the past quarter-century.[181] Arguments here are not particularly controversial and interviewees were happy to corroborate the broad premise. References similarly abound discussing current best practice, laboratory and theoretical developments in capabilities and those specific competences (with their underlying technologies) required to enable such removal. Research, however, is largely absent on assessing the *feasibility* of such purpose.

In addressing this task, the thesis' qualitative methods deliberately borrow from a portfolio of research models. An example here is Creswell's *Grounded Theory Approach*.[182] The dissertation's research questions (the matter of technical and operational feasibility, the matter of interim models for removing weapon supervision, States' adoption of weapon autonomy, the persistence of flaws in these models and, finally, the likely fit between fielding autonomous weapons and battlefield operations) were therefore created in order to generate a comprehensive theory on AWS deployment. Existing models were deliberately set aside in order to allow a substantive argument to emerge (in this case, the multi-facetted and cumulative challenge arising from such weapons' technical debt, contextual limitations and enduringly poor predictability). The approach also focuses on how interests and parties are affected by this new theory. Under 'grounded theory', research is derived from data acquired through fieldwork, interviews, observations and the broadest possible documentary sources but is also characterised by *further* data collection arising from new concepts and arguments as they arise during the writing process. Here, arguments and data points were loosely coded in order to align them with relevant battlecraft characteristics, conditions and consequences in order to build and emerging story line for the overall research. Finally to this point, the resulting theory (here, meaningful human control in the kill chain and the enduring difficulty of ceding control to an

[179] Gary King, Robert Keohane and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research*, (Princeton University Publications, 1994), pp. 5-7.

[180] The thesis cites more than 850 such on-line sources. Some 180 journals underpin the thesis' ethical, historical, legal, operational and technical analysis.

[181] K Hammond, *Why Artificial Intelligence is Succeeding: Then and Now*, (Computerworld, Artificial Intelligence Today and Tomorrow, 2015), para 4 and generally.

[182] John Creswell, *Educational Research – Planning, Conducting and Evaluating Qualitative and Quantitative Research*, (Boston: Pearson Publishing, Chapter 13 'Grounded Theory Designs, 2012), pp. 422-499.

algorithm) can be reported in a narrative framework as a set of propositions.[183] As noted by Strauss and Corbin, an advantage of this approach is the method's *breadth* such that the resulting theory, if suitably well researched, can be 'abstract enough and include sufficient variation to make it applicable to a variety of contexts related to that phenomenon'.[184]

For the purposes of this thesis and given the foregoing approach, evidence is defined as an argument or assertion that is backed up by information in a wide range of forms. Here, therefore, evidence comprises a broad literature review, stakeholder consultation and the broadest possible meta-analysis. Cairney, for instance, notes that *psychology* frequently impacts upon decision-making and decisions are rarely made purely on the basis of scientific evidence.[185] The study's qualitative methods were therefore devised around a tailored set of research questions in order to provide a detailed understanding of deployment considerations, the likelihood of success arising from such deployment models as well as the forces impacting decisions for AWS' deployment. Such an understanding is critical given the role of technical development and its implementation to this matter, the expectation that such technical innovation creates in both public and policy-makers as well as the generally under-researched nature of the subject matter. Here, therefore, the author deliberately adopts the perspective of a social constructivist whereby 'reality is socially, culturally *and* historically constructed'.[186]

It also underlines the importance of this thesis' extensive use of primary sources. Original interview generally underpins this thesis' interpretivist paradigm 'which portrays a world in which reality is socially constructed, complex and ever-changing'.[187] In order to triangulate this evidence, the author has tested several military sources to ascertain that cohort's views on weapon independence. The thesis' footnotes reference ninety-five instances of data points derived from interviews. The thesis' key innovation therefore lies in an combinatory analysis of long-standing theory and current best practices with the empirics of operational requirements and common sense in order to provide original analysis of deployment challenges.[188] A useful characteristic of quantitative research, after all, is that targeted interviews are appropriate notwithstanding their small sample size given the method's description and analysis of a research subject without limiting the scope of that research or the nature of participants' responses.[189] Interviews split broadly between experts from legal, military, NGO, academic, practitioner (industry and procurement) and other relevant institutions (here, RUSI, Chatham House and Prowler.io Decision Summit). Interviews were either structured (the likes of Generals Cordingley and Sharp, Professors Sabin and Du, NGO experts Wareham, Goose, Moyes

---

[183] Ian Dey, *Grounding Grounded Theory: Guidelines for Qualitative Inquiry*, (Emerald Group Publishing, June 1999), pp. 1-2.

[184] Anselm Strauss and Juliet Corbin, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, (Sage Publishing, 1990), p. 23.

[185] Cairney, P, *The Politics of Evidence-based Policy Making*, (Palgrave Macmillan, 2016).

[186] Bloomberg LD and Volpe M, 'Completing your qualitative dissertation – a route map from beginning to end', (London: Sage Publications), 2012.

[187]

[188] Here, research strategy, methods, approach and data collection; the selection of the sample, Peter analysis, difficult considerations; the researcher limitations of the project.

[189] OP Atieno, *An Analysis of the Strength and Limitations of Qualitative and Quantitative Research Paradigms,* Problems of Education in the 21st Century, Masinde Muliro University, Kenya, 13, (2009), p.14.

and Sharkey), semi-structured (for example, Professors Clark, Asaro, Roff, Suchman, Cummins as well as industry experts such as Scharre and Borrie) or unstructured (ICRC's Maurer, General Barrons, Doctors West and Strohn). Primary material was also collected through impromptu interviews at symposia and conferences. Finally to this point, comments arising from earlier drafts invariably occasioned dialogue which is usually footnoted in subsequent versions as having taken place 'in conversation with the author'. The process followed for these several formats is discussed below.

In general, the thesis' interviewing methodology follows the protocols set out in NGO Human Rights Watch's formal 'Interview Manual'.[190] The publication identifies core principles, minimum standards and best practice in the conduct of fact-finding interviews. The protocols are intended to ensure 'factual, accurate [and] ethical' analysis that also 'aspires to innovation and learning'.[191] Under this umbrella, the author has chosen to use first-hand interviews as the major means of primary data collection as they allow original perspective that reflects participants' own notions of what are complex issues. As noted by Patton, 'qualitative interviewing begins with the assumption that the perspective of others is meaningful, knowable and able to be made explicit'.[192] There are, however, recognised limitations to such interviewing including researcher bias, power relations and transferability. As also noted by HRW's handbook, qualitative research has drawbacks as it generally does not involve random, representative samples.[193]

The format and ethics of the process have generally been governed by HRW protocols, in particular the NGO's internal memorandum '*Establishing and Writing about Broader Patterns in Human Rights Watch Research*'.[194] Given the importance of primary source material to this thesis' conclusions, all participating interviewees were therefore informed of the purpose of the research in advance. In all cases, the public, communal and non-invasive nature of the subject matter meant that interviews were freely and engagingly given. Participants all had a specialty in one or more aspects of the argument and, as above, the purpose of the thesis' qualitative research was to process their contributions into a coherent theory on the removal of weapon supervision. Under the ESRC's *Research Ethics Guidebook*, a thesis' ethical framework should comprise processes to ensure informed consent, confidentiality (as appropriate) of research respondents, their voluntary participation as well as the independence and impartiality of ensuing research.[195] Consideration was therefore given to whether the thesis' subject matter or

---

[190] Human Rights Watch, *HRW Interview Manual*, (February 2016; latest edition, but derived from earlier editions of the same), generally.

[191] Ibid., p.1. Matters of consent (pp. 34-40), question formation (pp. 44-46, 77-79), corroboration (pp. 53-54) and record-taking (pp. 89-93) are dealt with in the Manual's Part III.

[192] Patton MR, *Qualitative research and evaluation methods*. Third edition (London: Sage publications 2001), p. 341.

[193] HRW protocols, 'Establishing and writing about Broader Patterns in Human Rights Watch Research', (2016), p.1. Care must therefore be taken to avoid making broad claims based on common experiences of a small sample set. For this reason, the author has relied upon primary interviews but also secondary sources and thematic experience in order to inform the thesis' wider claims that is built upon evidence accumulated through individual cases and other information.

[194] Human Rights Watch, *Establishing and Writing about Broader Patterns In Human Rights Watch Research*, (2019), HRW Publications. See Sections 3, 'Establishing a Pattern in Writing' and 4, 'Weighing All Evidence'.

[195] Economic and Social Research Council, Institute of Education, University of London, the Research Ethics Guidebook: A Resource for Social Scientists, pp. 1-2.

researcher's background might influence the process' ensuing output. For the most part, the researcher was of a similar age and experience base in order to foster appropriate equilibrium. Finally to this point, the thesis research also accords with relevant guidelines set out in the BSA's 2017 *Statement of Ethical Practice*, specifically in relation to informed consent (and, as necessary, post-hoc consent), the use of a gatekeeper, acceptable practices for covert research (including selective rights of participants to review and edit) as well as the construction and storage of field notes.[196] The Association's *Annex* on digital research is particularly relevant in its discussion of 'situational ethics' in order to provide discretion, flexibility and innovation. The Annex correctly notes that the field of Internet research is 'dynamic and heterogeneous as reflected in the fact that as at the time of this writing no official guidance regarding Internet research ethics have been adopted at any national or international level'.[197]

[196] The British Sociological Association, Statement of Ethical Practice, 2017, https://www.britsoc.co.uk/media/24310/bsa_statement_of_ethical_practice.pdf, p.5 and generally.

[197] British Sociological Association, Ethical Guidelines and Related Resources for Digital Research Annex, 2017, <https://www.britsoc.co.uk/media/24309/bsa_statement_of_ethical_practice_annexe.pdf>, p. 8 and generally. In particular, the annex deals with crediting, consent and the treatment of Community research sources.

2. Context: The role of context in the removal of weapon supervision

While the primary purpose of this thesis is to review AWS feasibility, the aim of this second chapter is rather to provide appropriate context to the lessening of human supervision in lethal engagements. It is to place AWS within recent, relevant history in order to evaluate their deployment against a broad portfolio of social, political and technical norms. The chapter is a frame of reference against which this thesis' subsequent behavioural and technical analysis may be interpreted. Context here is an analysis of the circumstances that form the setting for such independence. The issue is to appraise the significance of weapon autonomy to battlefield practices and then to assess the degree to which erosion of human supervision amounts to a discontinuity in how war will be waged once AWS have been deployed.[1] In particular, the exercise should be a useful adjunct in reviewing catalysts that *promote* adoption of independent weapons, the subject of the following chapter.

An interesting starting point for this analysis is provided by Stuart Russell, professor of computer science at UC Berkeley, whose video-piece 'Slaughterbots' portrays a 'fictional near-future in which autonomous explosive-carrying microdrones are killing thousands of people around the world'.[2] Stuart's imaginary narrator quips that autonomy allows 'you to separate the bad guys from the good... watch the weapons make the decisions... take out your enemy virtually risk-free'.[3] Russell's portrayal is one of several radical pictures to emerge contemplating combat which has been transformed by the adoption of broad-task, wide-capability autonomous weapons.[4] In so doing, a plausible narrative emerges on how AWS might be integrated into States' battlecraft, here defined as the techniques and drills of military combat, that is relevant both to this chapter's analysis on context but also to the wider issue of AWS feasibility including sophisticated tactical uses (area denial, defence in depth, flank security, deep operations, establishment of an efficient kill-box, asymmetric operations) but also strategic uses (high profile engagements and other 'morale sapping actions').[5] Russell's analysis has several constituents that are relevant to this

[1] Discussion on Revolution in Military Affairs (RMA) follows below in this introduction to Chapter 2 (*Context*).

[2] Evan Ackerman, 'Lethal Microdrones, Dystopian Futures, and the Autonomous Weapon Debate', *IEEE Spectrum*, 15 November 2017, <https://spectrum.ieee.org/automaton/robotics/military-robots/lethal-microdrones-dystopian-futures-and-the-autonomous-weapons-debate> [accessed 16 November 2017]. The eight minute video is available at <https://www.youtube.com/watch?v=9CO6M2HsoIA> and <https://www.youtube.com/watch?v=ecClODh4zYk> [accessed 10 November 2017]. Russell's portrayal suggests AWS with multiple roles. It notes that AWS deployment might come with a high risk of failure but with little consequence being attached to that failure. It also posits material expansion to the scope and daring of AWS, the liklihood of an AWS arms race, a lower bar for future engagements and, finally to this point, fast cross-proliferation of lethal autonomy between military services.

[3] Stuart Russell, 'Slaughterbots', YouTube, 0.15 minutes/1.50 minutes/2.47 minutes, <https://www.youtube.com/watch?v=ecClODh4zYk> [accessed 2 November 2017]. Russell concludes the video in person by stating that 'this short film is more than speculation. It shows the results of miniturising technologies that we already have [and that] allowing machines to choose to kill humans will be devastating to our security and freedom'. Ibid., (7.15 and 7.34 minutes).

[4] See, generally, Stuart Russell, 'Take a Stand on AI Weapons', *Nature*, 521, 7553, 27 May 2015, <https://www.nature.com/news/robotics-ethics-of-artificial-intelligence-1.17611#russell> [accessed 6 November 2017]. Russell opines that 'the capabilities of autonomous weapons will be limited more by the laws of physics (constraints on range, speed and payload) than any deficiencies in the AI systems that control them…. One can expect platforms deployed in the millions, the agility and lethality of which will leave humans utterly defenceless'. An overview is usefully provided by Erik Sufge, 'What might a Killerbot Arms Race Look Like?', *Popular Science*, 28 May 2015, https://www.popsci.com/what-would-killerbot-arms-race-look/ [accessed 12 November 2017].

[5] Matt Bartlett, 'The AI Arms Rave in 2019', *Towards Data Science*, 28 January 2019, <https://towardsdatascience.com/the-ai-arms-race-in-2019-fdca07a086a7> [accessed 13 March 2019]. See also: James Vincent, 'China is worried an AI arms race could lead to accidental war', *The Verge*, 28 January 2019,

chapter's review of context. On one hand, it echoes arguments that are centuries old such as the suggestion of Bridge that AWS deployment will 'cause war on a vast scale'.[6] On the other, it seems to reject Wood's argument that 'war is cruelty and you *cannot refine it*'.[7] The gulf between these two positions (on the one hand, AWS' promise of targeted, refined violence versus the 'gritty and fundamental violence of war'[8]) explains much of the instability that follows technical changes throughout the practice of warfare.[9] It also explains the widely varying conclusions reached by parties trying to associate particular battlefield outcomes with particular combat technologies.[10] Such divergence, however, is to be expected given the much-expanded data sources available at the time of writing with which to undertake battlespace analysis.[11]

Russell's scenario underlines the degree of transformation that AWS deployment might posit for how war is prosecuted. Indeed, Wood highlights that States (and other parties) wishing to challenge battlefield status quo *must* embrace 'new types of warfare' in order to achieve genuine battlefield advantage.[12] AWS deployment is generally complicated by parties' smokescreening on account of the strategic benefits attached to surprise and, notes Cooke, by the time lag that exists between the commercial development of a technology and its subsequent rollout as battlefield capability.[13] AWS, moreover, have yet properly to be deployed and contextual analysis must therefore remain largely a matter of conjecture. As again noted by Wood in *Technical Revolution in Military Affairs*, 'only battle can provide hard evidence of a weapon's utility, forcing a chaotic discourse as stakeholders attempt to predict the outcome of their various strategies and options'.[14] The very notion of context therefore presents its own challenge in the assessment of AWS deployment. Not only are the consequences of removing supervision from weapons uncertain, but,

<https://www.theverge.com/2019/2/6/18213476/china-us-ai-arms-race-artificial-intelligence-automated-warfare-military-conflict> [accessed 13 March 2019].

[6] Mark Bridge, 'Killer robots 'will cause war on a vast scale'', *The Times*, 22 August 2017 <https://www.thetimes.co.uk/article/elon-musk-among-technology-experts-calling-for-ban-on-killer-robots-mustafa-suleyman-deepmind-ryan-gariepy-clearpath-robotics-mark-zuckerberg-facebook-stephen-hawking-noam-chomsky-5ndnblj0v> [accessed 6 November 2017].

[7] Richard Wood, 'The Technical Revolution in Military Affairs', *Holtz.org blog*, undated <http://holtz.org/Library/Technology/Technical%20Revolution%20in%20Military%20Affairs.htm> [accessed 3 November 2017]. This author's *italics* for emphasis.

[8] Brian Ferguson, 'Ten Points on War', *Social Analysis*, 52, 2, Berghahn Journals, (2008), 32-49 (p. 32).

[9] Ofer Fridman, 'Revolutions in Military Affairs that did not happen: a framework for analysis', *Comparative Strategy*, Volume 35, issue 5, (2016), pp. 388-406.

[10] For an overall analysis of offensive versus defensive operations, see: Jack Levy, 'The offensive/ defensive balance of military technology: A theoretical and historical analysis', *International Studies Quarterly*, 28, 2, (1984), pp. 219-220 and 221-222.

[11] Bill Holford, 'Big Data on the Battlefield', *IT ProPortal*, (1 February 2017) <https://www.itproportal.com/features/big-data-on-the-battlefield-an-introduction/> [accessed 12 October 2017].

[12] Wood, pp. 2-4. The discontinuity of AI in weapon processes is also highlighted by Jay Tuck whereby AI will be 'one thousand times smarter than we are, moving at speeds one hundred thousand times as fast as we can think and digesting information and data one million times quicker than we can'; Jay Tuck, 'Artifical Intelligence: It Will Kill US', TedX Hamburg Salon, 0.43 minutes, 31 January 2017, <https://www.youtube.com/watch?v=BrNs0M77Pd4> [accessed 10 October 2017]. An analysis of such 'new' process is undertaken in Chapter 3 (*Drivers*), specifically: 3.2 ('*Technology creep and dual-use technology trends'*) and 3.3 ('*Structural and procurement drivers'*). For discussion on evolving capabilities see also: introductions to Chapters 8 (*Software*) and 9 (*Hardware*). The matter of discontinuity in military affairs in discussed below. See here Gordon Cooke, 'The Future Battlefield', *US Army*, (16 July 2018) <https://www.army.mil/article/208553/the_future_battlefield> [accessed 12 January 2019].

[13] Wood, p. 3 and Cooke, generally.

[14] Wood, p. 6.

notes Halford, such consequences may be quite different for each affected party given 'what is a dynamic, shifting relationship between social and technical agents'.[15]

Notwithstanding such uncomfortable imprecision, it is this chapter's conclusion that context is central to understanding challenges posed by AWS deployment especially, notes Scharre, given the significance that removing weapon supervision may pose for the conduct of future military affairs.[16] Arguments, after all, appear throughout this thesis on how weapon autonomy will disrupt how States conduct their politics and their wars.[17] The contextual key is that such disruption arises from what is a *cumulative* set of circumstances which are melded from technical trends, from procurement trends and, crucially, from social trends. It also arises from the *number* of deployment models, weapon specifications, configurations and other exogenous deployment factors that combine to complicate such analysis. Notwithstanding this patchwork, it is context that becomes the common prism through which to weight these components that together comprise the deployment debate that is comprised of 'a complicated, multi-facetted set of interlocking processes'.[18] While later chapters[19] will demonstrate that the step-change of AWS deployment requires an '*assemblage* of technical competences'[20], it is a particular role of this chapter to evidence that these new battlefield models require significantly more than new hardware and lines of code if they are to be adopted.

In order to assess context's significance, this chapter is divided into three sections. An introduction to its significance is followed by analysis of context from the perspective of the defence planner. The chapter then considers the contextual function of ambiguity and 'situational awareness' in unsupervised engagements. Throughout, the substance of AWS' deployment challenge is complicated by difficulties establishing cause and effect. As evidenced in the following chapter, there are many and quite uncorrelated drivers to the removal of human supervision from weapon systems, each one promising discrete improvement in one or other processes such as better operational performance, reduced costs[21], broader combat options, force multiplication or advantageous expansion of the battlefield into new arenas. Given this mix, a contextual danger is that AWS deployment creates circularity between consequences (here, the consequences of

[15] Professor Susan Halford, Department of Sociology, President of British Sociological Association, University of Southampton, in conversation with the author, January 2018.

[16] Paul Scharre, 'Why we must not build automated weapons of war'*,* *Time Magazine*, 25 September 2017 <http://time.com/4948633/robots-artificial-intelligence-war/> [accessed 12 November 2017]. For a discussion of likely roles within general AWS deployment, see: Chapter 4 (*Deployment*), specifically: chapter introduction and 4.3 ('*Machine and human teaming models'*).

[17] Amitai Etzioni and others, 'Pros and Cons of Autonomous Weapon Systems'*,* *US Military Review*, May-June 2017 <http://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/> [accessed 13 November 2017]. This point is explored in Chapter 5 (*Obstacles*), specifically 5.8 ('*Ethical and Accountability constraints'*).

[18] Wood, p. 8.

[19] In particular, see: Chapters 8 (*Software*) and 9 (*Hardware*).

[20] Jai Galliott and Mianna Lots, *Super Soldier: ethical, legal and social implications,* (London: Routledge, Political Science, 2016), p. 14 and p. 17.

[21] David Francis, 'How a new Army of Robots can cut the Defense Budget'*, The Fiscal Times*, 2 April 2013 <http://www.thefiscaltimes.com/Articles/2013/04/02/How-a-New-Army-of-Robots-Can-Cut-the-Defense-Budget> [accessed 4 October 2017]. The article states that the Pentagon spends $850k per annum 'just to keep each soldier on the battlefield'. This compares to the cost of automatic weapons systems such as TALON ($230k), the SGR-Ai ($200k) and the 710 Warrior ($350k). Again, deployment of such systems is discussed in Chapter 5 (*Deployment*).

introducing autonomous capabilities onto the battlefield) and what then becomes 'new' context (new norms whereby the effects of AWS deployment materially alter what was that battlefield's 'preceding' context). Earlier deductions arising from the analysis of battlefield cause-and-effect may all at once no longer be reliable. In promising, for example, a reduction in the number of frontline personnel, AWS deployment is also facilitating removal of these soldiers from harm's way with its own set of quite separate contextual ramifications.[22]

Given also this thesis' focus on AWS' feasibility, the balance required of this chapter is to weigh the matter of technology in what is a much larger deployment equation. At one end of this argument, media reporting is near universal in judging near-term technology (here, the advent of weapon autonomy) as a game-changer in battlefield practices.[23] In this vein, the *Economist's* special report on *The Future of War* cites Hoffman, Fellow of the National Defence University, 'that these new technologies have the potential not just to change the character of war but even its supposedly immutable nature as a contest of wills'.[24] Hoffman's conclusion is that 'for the first time, the human factors that have defined success (sic) in war - will, fear, decision-making and even the human spark of genius - may be less evident'.[25] Hoffman is perhaps referring less to 'Clausevitzian Genius'[26] ('an individually outstanding gift of intelligence and temperament') but to more commonplace examples of front-line élan and individuality that historically have determined outcomes. Nevertheless, his deduction concurs with several of the deployment models reviewed in later chapters and relates not just to the battlefield (where, notes Clark, 'will is executed and contact made') but also to the role afforded to those deciding upon AWS' deployment.[27] The contextual inference is that pace of technical change will disturb the other integers that make up the basis of contextual analysis. Reporting from the February 2018 Pyeonchang Winter Games (although recorded before the Opening and not during it), Basset suggests that those watching the event's Opening Ceremony had just 'witnessed a sight never seen before, a record-setting 1,218 drones joined in mechanical murmuration'.[28] Translated to the battlefield, a capability such as swarming

---

[22] Helene Cooper, 'Air Force Plans Shift to Obtain High-Tech Weapon Systems', *New York Times*, 30 July 2014 <https://www.nytimes.com/2014/07/31/us/politics/air-force-calls-for-cheaper-quicker-weapons-development.html> [accessed 19 July 2017].

[23] Economist Magazine, 'Autonomous Weapons are a game-changer', *The Economist*, 25 January 2017 <https://www.economist.com/special-report/2018/01/25/autonomous-weapons-are-a-game-changer> [accessed 19 July 2017].

[24] Economist Magazine, 'The New battlegrounds', *The Economist*, January 2018, p. 13.

[25] Ibid., p. 14.

[26] Omar Mohamed, 'The Master Strategist: Clausewitzian Genius', *Real Clear Defense*, (27 November 2016), paras. 3,4 and 12 of 12 <https://www.realcleardefense.com/articles/2016/11/28/the_master_strategist_clausewitzian_genius_110387.html> [accessed 12 January 2019].

[27] Professor Lloyd Clark, thesis supervisor, in conversation with the author, June 2017. See also: Chapter 4 (*Deployment*) and Chapter 6 (*Wetware*), specifically: 6.3 ('*The Delivery Cohort*'). The Cohort is a useful artifice to describe this decision group and is used throughout this thesis. It likely encompasses, inter alia, the following constituents: neurophysiologists to coordinate AWS networks, psychologists to coordinate learning and cognition, biologists for adaption strategies, engineers for control routines, logisticians, roboticists, electrical specialists, behaviorists, politicians, NGOs, sociologists, lawyers, company directors, weaponists, military tacticians, manufacturers, professionals involved in miniaturization, simulation, configuration, coding, power supply and modularity, specialists in sensors, in distributed and decentralized routines, ethicists, specialists in tooling and calibration. For the purposes of this analysis, together this interest group is termed the *Delivery Cohort*.

[28] Brian Barrett, 'Inside the Olympics Opening Ceremony World-Record Drone Show', *Wired.com*, 9 February 2018 <https://www.wired.com/story/olympics-opening-ceremony-drone-show/> [accessed 10 February 2018]. For a disparaging record of the event, see: Rani Molla, 'Intel's Drone Light Show Never Got Off the Ground for the 2018 Winter

may have properly novel effects for battlefield practices. It may enable lethal machines to overwhelm defences[29], it may allow innovative coordination of previously disparate multi-agent forces and, in so doing, it may make obsolete current battlefield practices. Ilachinski's contention that autonomous swarms turn current battlefield practices on their head is at the same time calling into question the very framework with which to review the deployment of unsupervised weapon systems.[30] It is the role of context to provide measure to such developments.

2.1 Warfare's continuum of methods

Broad change (and, as below, the instability that it brings to defence planning) certainly justifies scrutiny into how battlefield practices will be affected by AWS deployment. As a first port, van Creveld's *Technology and War*[31] provides a stepping-stone for this exercise.[32] The background is that van Creveld posits a long continuum in warfare from practices that are characterised by the substitution of 'firepower mass for manpower mass' (van Creveld's age of *machines*) to practices that are defined instead by the integration of technology into complex combat networks (analogous to AWS, his age of s*ystems*). This is contextually relevant as AWS deployment sits neatly in this framework. While this may validate Hoffman's ideas around a new era of warfare (one that follows Russell's dystopian analysis of a battlefield dominated by lethal autonomous systems), it also provides context that links AWS technologies to a long line of earlier weapon innovations which have promised to upend the battlefield.[33] It is here, notes Wylie, where an argument can be made that history is punctuated with possible Revolutions in Military Affairs (RMA) of which AWS deployment is just the last in a list of examples.[34] Does, however, deploying a weapon system that, at root, is different simply by having no human in the loop constitute a wholly novel method of waging war? Mansoor notes, after all, that disruption in the past has usually been brought about by a much broader combination of technical breakthrough, organisational adaptation *and* doctrinal innovation.[35] Given this thesis' later finding that that autonomy cannot take over every battlefield task, Bousquet suggests too that AWS deployment may represent a break with past disruptions in

Olympic Opening Ceremony', *Recode*, (10 February 2018) <https://www.recode.net/2018/2/10/16998652/drones-guinness-world-record-pyeongchang-2018-winter-olympics> [accessed 12 January 2019].

[29] Kris Osborn, 'Swarming mini drones: Inside the Pentagon's plan to overwhelm Russian and Chinese air defences', *The Buzz, National Interest*, 10 May 2016 <http://nationalinterest.org/blog/the-buzz/swarming-mini-drones-inside-the-pentagons-plan-overwhelm-16135> [accessed 1 November 2017] paras. 3-6 of 21. Also, see: Chapter 4 (*Deployment*), specifically: 4.6 ('*Swarming models*').

[30] Andrew Ilachinski, 'AI, Robots and Swarms: Issues, Questions and Recommended Studies', *CAN Corporation*, January 2017 <https://www.cna.org/CNA_files/PDF/DRM-2017-U-014796-Final.pdf> pp. 6-7 and pp. 7-9.

[31] Martin van Creveld, *Technology and War: From 2000 BC to the Present Day,* (USA: Simon & Schuster, 2010), pp. 217-219, p. 311.

[32] See also: Maxim Worcester, 'Autonomous Warfare: A Revolution in Military Affairs', *ISPSW Strategy Series: Focus on Defence and International Security*, Issue 340, (April 2015), pp. 2-3 <https://www.files.ethz.ch/isn/190160/340_Worcester.pdf>.

[33] Promise here relates to the 'Revolution in expectations' discussed by Sabin. See: Chapter 6 (*Wetware*), specifically: 6.1 ('*Software versus intelligence*').

[34] Admiral JC Wylie, *Revolutions in Military Affairs*, (UK Essays, November 2013) <https://www.ukessays.com/dissertation/examples/history/revolution-in-military-affairs.php#citethis> [accessed 9 July 2017].

[35] Peter Mansoor, 'The Next Revolution in Military Affairs', *Strategika, Hoover Institute*, Issue 39, (15 March 2017), para. 1 of 4 <https://www.hoover.org/research/next-revolution-military-affairs>. In considering the ramifications of AWS deployment, Mansoor here discusses the relative development of the first firearms, the socket bayonet, the dreadnought battleship, carrier aviation and blitzkrieg.

the discontinuity it brings will occur 'without necessarily requiring wholesale organisational disruption in its prosecution'.[36] Bousquest similarly notes that an RMA can only be analysed within the *context* of that technology's 'wider socio-technical milieu' into which these battlefield changes are taking place.[37] This is not, however, to fix that an RMA can only be technically driven. The issue is that the wide universe of AWS deployment models (and the organisational developments that must accompany such models) undermines this exercise. By inference, however, Bousquest offers a way out: In considering links between AWS and an RMA, weapon deployment must be studied 'with relation to *broader* assemblages' that focus instead on a portfolio of weapon inputs such as 'its mode of production, the value attributed to it, its distribution in the social field and its employment'.[38] The key is that this chapter's analysis be based on *broad* context. None of the characteristics under review are either intrinsic or exclusive to the autonomous weapon. They are, unsurprisingly, contextual and arise directly from how the weapon is fielded and other quite human interventions. While AWS might be designed and then refined by its Delivery Cohort with particular uses in mind, it should be inferred from this thesis' technical analysis that these uses will rarely correlate exactly to the intentionality of their creators and, as noted again by Bousquest, are 'always liable to being repurposed as they enter into different assemblages'.[39]

Certain practical factors exist that bias the analysis of whether a new weapon system constitutes an RMA. Disproportionate attention to technology's role in war can, warns Backstrom, be allocated just to the *weaponry* component of deployment.[40] The challenge of technological determinism then surfaces where perceived changes in war's conduct are uncritically attributed to particular battlefield technologies. Conversely, insufficient contextual weight might be given to other influences such as logistics, tactics and operational changes (changes, for instance, to the tooth-to-nail ratio of combat troops to support staff) and the broad portfolio of effects that ensue.[41] Other factors may be marginalised when evaluating battlefield context. A platform's simple ubiquity may be ignored: The cumulative effect of the 'humble Kalashnikov'[42], nearly one quarter of the five hundred million firearms in current circulation, posits an entirely different set of outcomes (ubiquitous, easy to use and dependable[43]) than might be expected from expensively procured AWS

[36] Antoine Bousquet, 'A Revolution in Military Affairs? Changing technologies and Changing Practices of Warfare', *Technology and World Politics*, Routledge, (2017), pp. 2-3 <https://www.academia.edu/34469743/A_Revolution_in_Military_Affairs_Changing_Technologies_and_Changing_Practices_of_Warfare>.

[37] Bousquet, p. 2.

[38] Ibid., p. 2-4.

[39] Ibid., pp. 6-8. See also: Chapter 4 (Deployment). For a discussion on changing capabilities posited by AWS see also: Economist Magazine, 'Autonomous Weapons are a Game-changer', *Economist*, 25 January 2018 <https://www.economist.com/special-report/2018/01/25/autonomous-weapons-are-a-game-changer> [accessed 23 July 2018].

[40] Alan Backstrom and Ian Henderson, 'New Technology and Warfare', cit. *International Review of the Red Cross*, Volume 94, Number 886, (Summer 2012), pp. 483-485 ('*New Capabilities in Warfare*').

[41] Tamara Campbell and Carlos Velasco, 'An Analysis of the Tail to Tooth Ratio as a measure of Operational Readiness and Military Expenditure Efficiency', *Naval Postgraduate School*, (US: Monterey, December 2002), p. 5 and pp. 117-127 ('*Conclusions*') <http://www.dtic.mil/dtic/tr/fulltext/u2/a411171.pdf>.

[42] David Blair, 'AK 47 Kalashnikov: The Firearm that has killed more People than any Other', *The Telegraph*, 2 July 2015 <http://www.telegraph.co.uk/news/worldnews/northamerica/usa/11714558/AK-47-Kalashnikov-The-firearm-which-has-killed-more-people-than-any-other.html> [accessed 23 July 2017].

[43] WeaponsMan blog, '5 Reasons for the AK's Legendary Reliability', *WeaponsMan*, <http://weaponsman.com/?p=12534> [accessed 7 August 2018].

(scarce, unpredictable, prone to independent outcomes and reliant on challenging battlefield integration). In the same vein, specific operational effects may be skewed, again compromising the analysis' rigour: First reported in September 2006, the SGR-AI operates as an unsupervised robotic weapon in the DMZ between North and South Korea and, unsupervised, it offers identification and subsequent tracking of intruders, engaging them without the intervention of a human operator.[44] Various relevant observations arise. First, that platform's on/off switch illustrates an 'unstable slightness of difference'[45] between an AWS and a non-autonomous weapon and, second, the risk of the SGR-AI performing poorly (illegally?) is much reduced given its limited task of static targeting and the crucial environmental consideration that the DMZ is a thoroughly constrained area of operation.[46]

2.2 The role of context in AWS' argument

Context then becomes a key tool in order to provide confidence around this thesis' conclusions. It provides common analytical foundation, continuity and setting to what, after all, is inquiry into a 'future-orientated and convoluted phenomenon'.[47] What then constitutes context in this deployment debate and what significance should it be given in this inquiry? A starting point is the British Army's *Action Centred Leadership* and the explicit priority that it gives to the role of context in military operations.[48] The notion of '*Understanding Context'* is given unambiguous precedence by placing it on top of the Army's three leadership silos, '*Achieve the Task*', '*Build Teams'* and '*Develop Individuals'*.[49] Context is certainly the key factor in how the Army articulates how to conduct its mission and, within this framework, it is defined as 'the collection of circumstances that form the setting for an event in terms of which it can be fully understood'.[50] Context encompasses a broad group of conditions, some tangible and some intangible, that impact all components of military tasks including their definition, planning, communication, their execution, support and their evaluation. By way of balance, however, this is at odds with the prescription of Hoffman who attributes changes in war's character not to context and intangibles but directly instead to the deployment of specific battlefield technology (here, again, the adoption of weapon autonomy).[51] This posits a further dilemma. Can a weapon algorithm capture battlefield context in order to instigate confidence in its subsequent performance? Given the ambiguity under which AWS will be

[44] Robotzeitgeist, 'Robots answer battle call for South Korea', 2006 <http://robotzeitgeist.com/tag/sgr-a1, Robotics Zeitgeist, 2006> [accessed 11 December 2016]. For additional discussion on this and other precursor AWS, see: Chapter 4 (*Deployment*). For its deployment, see: Global Security, 'Samsung Techwin SGR-A1 Sentry Guard Robot', (September 2006) <https://www.globalsecurity.org/military/world/rok/sgr-a1.htm> [accessed 12 August 2018]. HRW's Mary Wareham questions whether the weapon remains operational.

[45] Professor Peter West, Computer Science Department, Cornell University, in conversation with the author, 12 December 2017.

[46] Guglielmo Tamburrini, 'On banning autonomous weapon systems: From deontological to wide consequentialist reasons', cit. Bhuta and others, *Autonomous Weapons Systems: law, ethics, policy,* p. 126. Any human detected in the prohibited area is classified as a legitimate target.

[47] Richard Moyes, Article 36, 'War without oversight; challenges to the deployment of autonomous weapons', Buckingham University Humanities Research Institute Seminar, 13 May 2018.

[48] Centre for Army Leadership, 'Army Leadership Doctrine', Edition 1 (UK: RMAS Camberley, 2016), p. 17.

[49] Ibid., Chapter 4 ('*What Leaders Do'*).

[50] Ibid., pp. 16-19.

[51] Economist Magazine, 'The New battlegrounds', p. 13.

operating, it is this dynamic that provides much of this thesis' contextual foundation as well as its behavioural and technical analysis set out in subsequent chapters.

The issue also has an inescapable temporal component with the matter of duration creating two separate challenges. The first relates to timeframes. In considering battlefield autonomy, the Economist's 2018 report on *The Future of War* limits its forecasts to the 'next twenty years or so, because beyond that the uncertainties become overwhelming'. In so doing, it still 'offer[s] its predictions with humility'.[52] Given battlefield developments since, say, the turn of the Millennium, that twenty years may appear a long timeframe but it is also one that must be tempered by context.[53] Latiff and others conclude, after all, that relevant timeframes should be much shorter.[54] The question of duration should also be framed by technical bottlenecks as well by a failure to distinguish between what may be short-term outcomes of experts' research programmes and other more ambitious (albeit distant) goals envisaged for AWS deployment.[55] A second 'temporal' issue relates to the use of context as a *tool*. Strachan, currently Professor of International Relations at St Andrews University, posits that 'history is the study of change'.[56] The student, suggests Strachan, should embrace Liddell Hart's observation that 'the past is a foreign place' where situations never exactly repeat and, instead, it is change that is the norm. In setting out this contextual framework, Strachan is therefore borrowing from historian Bloch whereby an examination of 'how and why yesterday differed from the day before… can reach conclusions which will enable it to foresee how tomorrow will differ from yesterday'.[57] Gray, Emeritus Professor of International Relations and Strategic Studies at Reading University, offers a different contextual view, one that is anchored in what he describes as the *continuum* of history. For Gray, the relevant context is that history repeats. There may always be an element of chaos that is present in strategic history[58] but, as he opens his work on the friction and uncertainty of future warfare, *Another Bloody Century,* 'historical perspective is the only protection against undue capture by the concerns and fashionable ideas of

[52] Economist Magazine, '*The New Battlegrounds*', pp. 3-4.

[53] Michael Marshall, 'Timeline: Weapons Technology', *New Scientist*, 7 July 2009 <https://www.newscientist.com/article/dn17423-timeline-weapons-technology/> [accessed 12 June 2018]. The first decade of this century saw the advent of national defence shields, active denial systems, high energy lasers, pulsed energy projectiles, neuroscience-based human enhancement and, in Metal Storm, a gun capable of firing several million rounds per minute. See: Chapter 1 (*Introduction*), specifically: 1.2 ('*Introduction to key Concepts'*). As defined in this thesis' introduction, near-term' is reckoned generally to be within the period 2025. 'Medium-term' then relates to developments that may be expected to occur in the fifteen years thereafter to 2040.

[54] Paul Scharre, 'Making Sense of Rapid Technical Change', *Center for a New American Security*, (17 July 2017), generally <https://www.cnas.org/publications/commentary/making-sense-of-rapid-technological-change>. See also: Robert Latiff, 'How Technological Advancements Will Shape the Future of the Battlefield', *Signature*, (13 October 2017), generally <https://www.signature-reads.com/2017/10/how-tech-advancements-will-shape-future-battlefield/> [accessed 13 January 2019].

[55] Tamburrini, cit. Nehal Bhuta and others, p. 122. As discussed in Chapter 3 (*Drivers*), consequentialist reasons to deploy AWS are often based on expectations (here, AWS being able to reduce casualties and collateral damage). Advancing this scenario contributes to unrealistic expectations about AWS compliant deployment.

[56] Professor Sir Hew Strachan, currently Professor of International Relations at St Andrews University, in conference with the author, 22 February 2017. Author's italics. Also, notes Strohn, this does not mean that such change is 'unforeseeable' but that the study of periods of change, such as the French Revolution, is really for the realm of the historian because this change 'manifests the fundamental issues': Dr Matthias Strohn, in conversation with the author, January 2019.

[57] Hew Strachan and Sibylle Scheipers (eds.), *The Changing Character of War,* (Oxford: Oxford University Press, 2011), p. 7. See also, generally: Marc Bloch, 'A Strange Defeat: A Statement of Evidence Written in 1940', (UK: Norton, 1999), generally.

[58] Colin Gray, *The Future of Strategy,* (UK: Polity Books, 2015), p. 1.

today'.[59] Divergent from Strachan, this definition of context also has its supporters (endorsed, by inference, in US Chief of Staff General Milley's 2017 presentation to the RUSI Land War Conference through his dictum on the 'arrogance of the present').[60] For Gray, there is an observable continuity to history.[61] His contention is not that nothing changes but rather that little changes in matters of profound importance. The contextual relevance to AWS deployment is that, for Gray, future warfare 'will be strategic history much as usual' regardless of such weapon development.[62] While the balance of this chapter disagrees with this contention, the relative positions of the two historians are nevertheless interesting precisely because they demonstrate the importance of context in framing the removal of weapon supervision. For this reason, notes Clark, it may be more apt to consider deployment in terms of war's enduring themes but now overlaid by the novel context that comes from weapon independence.[63] As developed in this thesis' technical review, while the force of human factors in combat practices may remain unaltered, AWS' context is complicated by, inter alia, their de facto independence and intrinsic unpredictability.

Given such divergence, a purpose to this chapter is to navigate between these two positions in order to determine how context might *practically* impact on AWS deployment. The inference is that context will play a determinant in shaping that deployment. This is true for a wide portfolio of reasons, the nub of this chapter, whether through restrictions brought about by popular pressure or political dealings, by third sector activism, by human ingenuity in defeating AWS' technologies, by economic factors whereby scarce resources are allocated elsewhere, by organisational inertia (or, indeed, friction) or by other as yet unidentified considerations. This, therefore, is the empirical role of context at work. Moreover, day-to-day nuances of AI-driven weaponry are rarely out of the public eye. In August 2017, for instance, the world's Press carried a coordinated statement from the world's top academics and industry leaders warning the United Nations about the dangers of autonomous weapons.[64] Alerting the UN to 'a third revolution in warfare', the letter warns that 'once this Pandora's box is opened it will be hard to close'.[65] On the other hand (and sitting apart on that continuum), General McMaster, erstwhile National Security Adviser to the US White House, reminds his readers in 2017 to be skeptical about ideas that divorce war from its political nature, in particular where those theories 'promise fast change and efficient victories through the application of advanced military technologies'.[66] The two positions may appear opposing end-points of the debate on context but this is to miss that the two statements are not actually mutually exclusive.

---

[59] Gray, *Another Bloody Century,* p. 13.

[60] General Mark Milley, speech to RUSI Land War Conference, 27 June 2017 <https://rusi.org/sites/default/files/20170627-rusi_lwc17-gen_milley.pdf> and in subsequent conversation with the author.

[61] See also: Lt Col Frank Hoffman, 'Thinking about future conflict'*, Marine Corps Gazette*, Volume 98, Issue 11, (November 2014), paras. 18-22 of 42 <http://pqasb.pqarchiver.com/mca-members/doc/1619980305.html?FMT=TG> [accessed 38 November 2017].

[62] Gray, *Another Bloody Century*, p. 14.

[63] Professor Lloyd Clark, in discussion with the author, 10 January 2019.

[64] Noah Kulwin, 'Elon Musk and over 100 AI Experts warn UN about Killer Robots'*, Vice News*, 21 August 2017 <https://news.vice.com/en_us/article/9kdgkz/elon-musk-and-over-100-ai-experts-warn-u-n-about-killer-robots> [accessed 30 June 2017].

[65] For the text of the open letter: Future of Life Institute, 'An Open Letter: Research Priorities for Robust and Beneficial Artificial Intelligence', undated <https://futureoflife.org/ai-open-letter/> [accessed 12 March 2018].

[66] Lt Gen HR McMaster, 'On the Study of War and Warfare'*, Modern War Institute*, (24 February 2017) <https://mwi.usma.edu/study-war-warfare/> [accessed 12 February 2018] para. 6 of 10.

For the purposes of this chapter, it is McMaster's evaluation that generally shapes this thesis whereby the *cumulative* challenges to AWS adoption will together construct challenging restriction on wholesale deployment of unsupervised weapons. Separately, however, this thesis' analysis recognises that *incremental* adoption of component autonomy will certainly come to dominate weapons procurement. The analysis thus follows Howard's aphorism that 'the roots of victory and defeat often have to be sought far from the battlefield'[67] and it is likely to be social, political and cultural factors that will trump technical considerations in the adoption of battlefield autonomy.[68] A further dynamic arises: While battlefield autonomy will certainly make significant contribution to near-term battlefield practices, a second inference is that such cause-and-effect will remain context-specific whereby AWS may make headlines in certain brilliant applications but may otherwise remain a marginal (albeit contributory) asset in broader settings and circumstances. This is investigated in Chapter Four (*Deployment*) and its analysis of likely procurement models.

Recognising that discontinuity cannot be confined simply to advances in weapon platforms, Drucker highlights that parties must broaden their definitions and scope of context's role in AWS deployment.[69] Rapidly developing social media models, widespread popular connectivity and the reach of other technological developments are, notes Staff, combining to create far-reaching cultural and societal disruptions which are 'taking place with unparalleled speed'.[70] The pervasiveness of computers is again a defining characteristic. Such contextual components have cumulative although imprecise effects on how disruptions evolve which are difficult to distinguish.[71] In the case of AWS adoption, a relevant source is provided by the British Army's *Leadership Doctrine* and its deliberate separation between the *nature* of conflict (an immutable, unchanging contest of will that involves uncertainty, chaos, chance and friction) and the *character* of conflict (which changes continually and is shaped primarily by politics and, crucially for the purposes of this thesis, the technology of the age).[72] Here, the deployment of weapon autonomy must be a matter of conflict's character (and the volatility in this condition) and not of conflict's fundamental nature. This separation (the nature of conflict and, pertinent to AWS, the character of conflict) may be a constructive device but, in considering context, in no way does it solve for the deployment challenges of removing weapon supervision. Instead, it suggests that defence planners' list of assumptions is becoming ever *less* sharp and, contextually, it is within this unpredictable mix that AWS deployment must be viewed.[73] The 'force' in this discussion is, after all, that same force that is being used against human beings and is not merely defensive action that is being taken against objects (incoming munitions or robots). The contextual suggestion is that such force can

[67] Lloyd Clark, 'Blitzkrieg: Myth, Reality and Hitler's lightning war – France 1940', (UK: Atlantic Books, 2016), p. 2.

[68] UNESCO, 'The Infernal Cycle of Armaments', *International Social Science Journal*, 28, 2, 252-254 <http://unesdoc.unesco.org/images/0001/000197/019707eo.pdf>.

[69] By inference: Peter Drucker, 'The Age of Discontinuity: Guidelines to our changing Society', ninth edition, (USA: Transaction Publishers, New Brunswick, 2011), pp. 4-10.

[70] E Staff, 'High Profile Panel Warns of unavoidable, far-reaching technical revolution', *The Education Post*, 7 July 2017 <http://educationpostonline.in/2017/07/07/high-profile-panel-warns-of-unavoidable-far-reaching-tech-revolution/> [accessed 2 August 2017].

[71] A Rachleff, 'What 'disrupt' really means', *Techcrunch.com*, 16 February 2013 <https://techcrunch.com/2013/02/16/the-truth-about-disruption/> [accessed 4 April 2017].

[72] Centre for Army Leadership, '*Army Leadership Doctrine',* Edition 1, (UK: RMAS Camberley, 2016), p. 10.

[73] See, generally: Gray, 'Strategy and Defence Planning: Meeting the Challenge of Uncertainty', p. 2. Gray's useful position is that 'defence planning is substantially guesswork, and it has to be such – educated guesswork one hopes, but still guesswork'.

only be undertaken as part of a process that is properly deliberative (one that involves human decision-making and meaningful human control in the initiation of violence) which, in common with Asaro, must empirically be shaped by both broad context *and* the laws of armed conflict.[74] While this too is discussed in later sections, the point is to demonstrate context's pervasive narrative in the matter of AWS deployment.

2.3 Context's behavioural significance

Given the largely intangible nature of context's constituents (as noted Cordingley and expressed here as the 'definition, planning, communication, execution, support and evaluation' of military tasks discussed above), how then might this collection of behavioural circumstances influence AWS deployment?[75] Imposition of force by one individual against his foe has long been an intensely personal affair. While human beings may have been increasingly remote from the point of violence, that human being has up to now taken the decision to take life.[76] Notwithstanding, therefore, that decision-makers and weapon operators alike have become increasingly remote from that point, Heyns notes that ethical and legal norms for such behaviours 'have developed over the millennia to determine when one human may use force against another, in peace and in war, and have assigned responsibility for violation of these norms'.[77] It is this broad set of conditions that still underpins relevant context with which to consider the role of that decision-maker and operator. The issue, however, is that AWS deployment may grossly destabilize this dynamic. Importantly, it is also a matter of *degree* as advanced military powers already undertake combat missions without using the full range of their technological capabilities.[78] Unsurprisingly, this argument has a verso; although outside the scope of this analysis, constraints to the removal of supervision (for instance, the statutory requirement for MHC) are empirically less relevant to non-State polities. In the case of irregular forces fielding independent weapons, 'the difference between an autonomous weapon and a drone is only four exploding bolts'[79] as the limitations of LOAC, ethics and precedent are trumped by non-State priorities of expediency, belief and culture.

Defining the *cultural* component of such context is even more complicated.[80] Culture, notes Karppi, is subject to rapid, inexplicable change, often occasioned by catalysts that are exogenous

[74] Peter Asaro, 'On banning autonomous weapons systems: human rights, automation and the dehumanisation of lethal decision making', *International review of the Red Cross*, 94, 86, (2012), p. 2 and pp. 8-17.

[75] Major-General Patrick Cordingley, Commander, 7th Armoured Brigade, Gulf War, 1991, in conversation with the author, May 2016. See also: introduction above to this chapter.

[76] As noted by Clark, that use of force should not necessarily conflate with loss of life. Professor Lloyd Clark, in conversation with the author, 10 January 2019.

[77] Christof Heyns, 'Autonomous Weapon Systems: living in a dignified life and dying a dignified death', cit. Bhuta and others, '*Autonomous Weapons Systems: law, ethics, policy,* p. 2.

[78] Hew Strachan, *The Direction of War – Contemporary strategy in historical perspective,* (Cambridge: Cambridge University Press, 2013), p. 167. Also: Major-General Patrick Cordingley, in conversation with the author, May 2016.

[79] Remote Control, 'Hostile drones: the use of drones by non-State actors against British targets', *Remote Control Project*, (The Oxford Research Group, 2013) <http://remotecontrolproject.org/hostile-drones-the-hostile-use-of-drones-by-non-state-actors-against-british-targets/> [accessed 2 January 2017].

[80] Tero Karppi and others, *Killer Robots as Cultural Techniques*, (International Journal of Cultural Studies, Sage Journals, October 2016), 1. Here, cultural components relate to the psychological connection between an individual's self and culture represented by the social behaviours and norms of that society.

and unexpected.[81] The contextual driver for AWS deployment is that cultural (and political) mindsets are created and broken up by disparate influences that are unstable and exhibit rapid change.[82] Furthermore, any one such influence will likely not correlate with another and may mask cross-border, cross-party and other affiliatory stimuli. Contextual components here include demographic considerations, gender considerations and, in certain geographies, the emergence of more individualistic cultures that further complicate pigeonholing.[83] Cultural influences, moreover, often defy easy definition. It is challenging to make strict contextual classifications given that individual elements tend, notes Gitelman, to split into numerous unstable sub-elements.[84] Given such volatility, it is unsurprising that ethical issues, legal constraints and technical expectation may combine with social and other human influences to inform how political decisions may be maneuvered on AWS deployment. Such trends are difficult to predict, they come in bunches and they interact erratically. To this extent at least (notwithstanding it being an uncomfortable generalization on warfare), this analysis conforms with Gray's conclusion that 'the course and outcome of war is shaped by many factors, not least of the human, the cultural, and the political, in addition to the possibilities opened by machines'.[85] Adamsky's *The Culture of Military Innovation* further evidences the importance of such context in his conclusion that, contrary to the USSR, post-WW2 US was still able to achieve technical development without developing doctrines that integrated such technology into the formal systems of military organisation.[86]

It is also this *pace* of cultural change that becomes a contextual driver in considering independent weapons.[87] Pace creates contextual difficulty around determining *which* and *when* each new norm (cultural, social, political) might next prevail.[88] Against this background, what is an apt example of cultural context?[89] Reaction to weapon systems, for instance, has often been erratic. Public disquiet against Spitfire aircraft flying over Southern England in 1937 turned very quickly to public adulation in advance of the Battle of Britain.[90] Its contextual relevance to AWS deployment is

[81] Michelle LeBaron, 'Culture and Conflict', *Beyond Intractability* blog, (July 2003) <http://www.beyondintractability.org/essay/culture_conflict> [accessed 2 August 2017].

[82] Ibid.

[83] Mark Rupert, *Support the Troops: Populist Militarism and the Cultural Reproduction of Imperial Power,* generally (USA: Maxwell School, Syracuse University, undated <http://faculty.maxwell.syr.edu/merupert/Populist%20Militarism.pdf>.

[84] L Gitelman, *Always Already New; Media, History and Data of Culture,* (USA: MIT Press, 2006), pp. 4-6.

[85] Gray, *Another Bloody Century,* p. 22.

[86] Dima Adamsky, *The Culture of Military Innovation: Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the US and Israel*, (US: Stanford University Press, 2010), pp. 31-44.

[87] C Pierreault, 'The Pace of Cultural Evolution', *PLOS.org*, (14 September 2012), <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0045150> [accessed 5 August 2017]. Peirreault investigates rates of change in human technology and their correlation to cultural and biological evolution. See also: D Barea and Y Silverstone, 'New Rules for Cultural Change', *Accenture Strategy*, (2016) <https://www.accenture.com/t20161216T040430__w__/us-en/_acnmedia/PDF-24/Accenture-Strategy-Workforce-Culture-Change-New.pdf>.

[88] For a useful analysis on cultural shift, see: David Brookes, 'When Cultures Shift', *New York Times*, op-ed, 17 October 2016 <https://www.nytimes.com/2015/04/17/opinion/david-brooks-when-cultures-shift.html?_r=0> [accessed 12 February 2017].

[89] See, generally: T. Ricks, 'The widening gap between Military and Society', *Atlantic Magazine*, July 1997 <https://www.theatlantic.com/magazine/archive/1997/07/the-widening-gap-between-military-and-society/306158/> [accessed 10 March 2017].

[90] Professor Lloyd Clark, thesis supervisor, in conversation with the author, December 2016.

that all regimes are eventually attentive to public opinion.[91] Whilst elements of a society might momentarily abhor autonomous weapons on ethical grounds, any subsequent undermining of national loyalties (a weakening, perhaps, within the willingness of its people to fight) might conversely lead others to advocate remote weaponry as a means to influence neighbours while avoiding casualties. Finally to this point, it should be reinforced that this thesis' commentary is based upon deployment of *State*-sponsored AWS. The emergence of AWS under entirely new fighting paradigms, perhaps by an as-yet undefined State system or, more obviously, by a non-State polity, remains broadly overlooked which, notes GCSP, will have quite different contextual ramifications.[92]

2.4 Defence planning

The effect then of context is to make the processes of the defence planner never more complicated.[93] In its accounting, defence-planning assumptions must be wide and in lockstep with current ethical, legal and technical norms.[94] It must incorporate political, environmental and mission analysis, the development of options, the review and then matching of capabilities and resources and, finally, the ensuing development of appropriate alternatives.[95] In this vein, context again is key. Planners must throughout have in mind an agreed future context for battlefield processes in order (and the purpose of this thesis) to establish whether that framework should accommodate compliant deployment of AWS. The complication is that such a framework must also feature *technical* assumptions if unsupervised weapons are to be realised (including, inter alia, their capabilities and tasking) while still reckoning for political precepts, strategic precepts as well as the effects of uncertainty, notes Barrons, that must arise from what are inevitable shortcomings in defence planning.[96] Given such complexity, it is context that provides baseline when planning for defence in a future that in large part cannot be understood.

---

[91] Martin Dimitrov, 'Tracking Public Opinion under Authoritarianism'*, Russian History*, 41, (2014), pp. 230-233 <http://www2.tulane.edu/liberal-arts/political-science/upload/dimitrov2014russian_history.pdf>.

[92] Geneva Centre for Security Policy (GCSP), 'Perils of Lethal Autonomous Weapon Proliferation: Preventing Non-State Acquisition', *GCFSP*, (2018), generally <https://www.gcsp.ch/News-Knowledge/Publications/Perils-of-Lethal-Autonomous-Weapons-Systems-Proliferation-Preventing-Non-State-Acquisition> [accessed 19 January 2019]. See also: W Wallach, *Towards a Ban on Lethal Autonomous Weapons: Surmounting the Obstacles,* 60. 5, (USA: Communications of the ACM, 2017) , p. 28.

[93] Dejan Stojkovic and Bjorn Robert Dahl, 'Methodology for Long-Term Defence Planning', *Norwegian Defence Research Establishment*, (February 2007) <https://www.ffi.no/no/Rapporter/07-00600.pdf> generally. See, also: Ministry of Defence and Foreign and Colonial Office, 'UK's International Defence Engagement Strategy', (UK MOD/FCO, 2017) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/596968/06032017_Def_Engag_Strat_2017DaSCREEN.pdf>. For reference on defence planning and methodologies, see: NATO office, 'Nato Defence Planning Process', updated 28 June 2018 <https://www.nato.int/cps/en/natohq/topics_49202.html> [accessed 15 January 2019].

[94] R Brooks, 'In defence of Killer Robots: Hold on there, technophobe hippies. When it comes to 'doing no harm', robots are a hell of a lot better than humans'*, Foreign Policy*, 18 May 2015 <http://foreignpolicy.com/2015/05/18/in-defense-of-killer-robots/> [accessed 6 August 2017].

[95] Dejan Stojkovic and Bjorn Robert Dahl, *Methodology for Long-Term Defence Planning*, p. 4.

[96] General Sir Richard Barrons, *Ditchley Foundation Panel,* at RUSI, 28 February 2017 and in subsequent conversation with the author. For discussion on defence planning shortcomings, see also: VS Shekhawat, 'Challenges in Defence Planning', *Institute for Defence Studies and Analyses*', (October 2006) <https://idsa.in/strategicanalysis/ChallengesinDefencePlanning_vsshekhawat_1006> [accessed 17 January 2019]. Also: Parliament of Finland, '*Long-term Challenges of Defence: Final Report of Parliamentary Assessment Group*', (May 2014), pp. 3-8 <https://www.eduskunta.fi/FI/tietoaeduskunnasta/julkaisut/Documents/ekj_5+2014.pdf>.

Just as weapon systems are not created in a vacuum, so defence planning requires context as it is unable to pilot itself 'given that the future never arrives and can never be known with certainty'.[97] Smith also notes that it does *not* follow simply because the feasibility of a technology is demonstrated that a 'military will want or be able to adopt it as there are numerous political, economic, operational, ethical and cultural factors that come into play'.[98] This narrative is important as all military polities practice some form of defence planning.[99] The skill of such planning is also the pursuit of a course of minimum regret.[100] Accordingly, the disruptive capabilities posited by weapon autonomy pose unusual challenges calling into question, as they do, first principles around States' politics, around fundamental law and ethics, the future of military command and battlefield organisation, allocation of resources as well as assumptions on leadership and combat assets. In considering these challenges, the balance of this thesis must be similarly beholden to context's significance in forming its conclusions.

Given that context is an inexact tool, the issue around planning analysis concerns the degree of exactness then required in order for it to be a *relevant* tool. The behavioural point is that an understanding of context and deployment assumptions only need be *broadly* correct to mitigate (as best as possible) outright error. Long-term defence planning (LTDP) and planners' subsequent procurement policy does not need to be 'in some absolute sense correct. Instead, it needs only to be correct enough'.[101] Its connection to deployment of independent weapons is quite specific given LTDP's key commission, notes Stojkovic and Dahn, must be to 'deter aggression and to ensure homeland defence'.[102] Given that defence planning marks out how polities should prepare for their future security, it also follows that there can be no objectively correct answer from the exercise. Context then provides the basis for a reasonable smell-test that should deliver long-lasting relevance.[103] Planners, like economists, remain in 'the dismal position'[104] of having certainty neither about the effects of output nor the worth of their defence choices given that both are bound up with accident and chance.[105] Planners, after all, are building uncertain models using uncertain data-points each that exacerbate overall planning uncertainty. It is against this contextual backdrop that the possible disruption of AWS is so telling: Militaries cannot be rearmed overnight while, in the

[97] Colin Gray, 'Defence planning, surprises and prediction', presentation to *Multiple Futures Conference, NATO's Allied Command Transformation*, (8 May 2009) <http://www.act.nato.int/images/stories/events/2009/mfp/mfp_surprise_prediction.pdf> p. 4.

[98] Ron Smith, *Military Economics: The interaction of Power and Money,* (UK: Palgrave Macmillan, 2009), p. 132.

[99] Source: *The NATO Defence Planning Process*. See: <http://www.nato.int/cps/en/natohq/topics_49202.htm> [accessed 25 May 2017]. NATO operates a formal Defence Planning Process comprised of five steps and resulting in a 'single, unified political guidance for defence planning setting out overall aims and objectives to be met by the Alliance and defining priorities and timelines for use by the planning domains'.

[100] Gray, *Another Bloody Century*, p. 42.

[101] Ibid., p. 47.

[102] D Stojkovic and B Dahn, 'Methodology for long-term Defence Policy', *Norwegian Defence Research Establishment*, p. 13, (28 February 2007) <http://www.ffi.no/no/Rapporter/07-00600.pdf>.

[103] Stojkovic and Dahn, p. 8.

[104] Adam Rogers, 'The Dismal Science remains Dismal, say Scientists', *Wired magazine, Science*, paras. 3 and 11 of 18, 14 November 2017 <https://www.wired.com/story/econ-statbias-study/> [accessed 9 January 2018].

[105] See: K Booth and N Wheeler, *The security dilemma: Fear, cooperation and trust in World Politics,* (UK: Basingstoke, Palgrave MacMillan, 2008). Also: Stojkovic and Dahn, pp. 15-16.

face of discontinuity, refits and periodic modernization as well as uncertain (and often extended) procurement timescales all at once become a newly inappropriate risk.[106]

Avoiding miscalculation also focuses defence planning on political consequences and, in particular, the economic constituent of such consequences. Foot points out that the costs of full warfare between the US and China might amount to some thirty per cent of China's Gross Domestic Product (GDP) and seven per cent of US GDP.[107] Studies unsurprisingly note that full State-upon-State war could cost both protagonists 'substantial military losses to bases, [military assets], significant political upheaval at home and abroad, and huge numbers of civilian deaths'.[108] Such disproportionate costs encourage instead a rising role for alternative means of warfare which, notes Kennon, includes 'political warfare' and the employment by defence planners of military, intelligence, diplomatic, financial, public relations, covert and other psychological means that 'fall short of conventional warfare'.[109] Chapter 4 (*Deployment*) also notes that it might encourage the deployment of certain hybrid and lighter touch deployment models involving independent weapons. Planning certainty around autonomy is, after all, hindered by the inability to conduct historical audit on recent battlefield precedents. Similarly, the further into the future that context is stretched, the less confident must be its predictions. As Gray concludes, 'many a reputation has been dented when vanity seduced its owner to venture a guess too far'.[110]

Finally to the defence planner's assessment of AWS feasibility, not all contextual drivers support AWS' deployment. In particular, this thesis' later analysis concurs with Lorber's position that technical weaponry exhibits very broad difficulties[111] which, if autonomy is to realise full potential, must incorporate a maintenance tail, innovative training, logistics, replenishment and repair (together with its coordinating staff work for each such component). Failure at any point in this long chain will be akin to failure in any one of its multipart technical components and, moreover, will likely cause similarly unacceptable degradation in the weapon's performance.[112] This accords with Drozdova that such planning must also be undertaken within the contextual

[106] For a discussion on procurement timetables and challenges, see: Martin Zapfe and Michael Haas, 'Arms Procurement: The Political-Military Framework', *CSS Analyses in Security Policy*, 181, (November 2015), generally <http://www.css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/CSSAnalyse181-EN.pdf>.

[107] Rosemary Foot, 'Constraints on Conflict in the Asia-Pacific: Balancing the 'War Ledger'', 66, 2, (UK: Political Science, 2014), p. 119.

[108] Seth Jones, 'Much 'Political Warfare' in our Future', *BreakingDefense*, 2 February 2018, paras. 6-13 of 19 <https://brekingdefense.com/2018/02/much-political-warfare-in-our-future> [accessed 6 February 2018].

[109] Seth Jones, 'The Return of Political Warfare', *Defense Outlook 2018*, February 2018, p. 3 <https://www.csis.org/analysis/return-political-warfare> [accessed 6 February 2018]. See also Anja Kaspersen and others, 'Ten Trends for the Future of Warfare', *World Economic Forum* (3 November 2016), Paragraphs 2 ('*Speed Kills*'), 3 ('*Fear and Uncertainty increases Risk*'), 6 ('*A Wider Cast of Players*'), 7 ('*The Grey Zone*') and 9 ('*Expanded Domain of Conflict*) <https://www.weforum.org/agenda/2016/11/the-4th-industrial-revolution-and-international-security/> [accessed 18 January 2019].

[110] Gray, *Another Bloody Century*, p. 41.

[111] Azriel Lorber, *Misguided Weapons: Technical Failures and Surprise on the Battlefield,* (USA: Potomac Books, 2002), pp. 8-10 ('*Other factors including common sense*'). See: Chapters 6 (*Wetware*), 7 (*Firmware*) and 8 (*Software*), specifically: 6.5 ('*Missing pieces*'), 7.1 ('*Sources of technical debt*') and 8.2 ('*Coding errors*').

[112] For a detailed analysis, see: Chapter 4 (*Deployment*), specifically: 4.7 ('*Operations and causes of failure*').

footing of general AWS vulnerability (in particular, to offsetting tactics which are themselves based on relatively 'low' and ubiquitous technologies).[113]

2.5 **Context's human angle**

The principal driver in AWS deployment remains, however, the role of the human in battlefield processes. Industrial and scientific advances have not yet sidelined the rating, soldier or pilot and, in AWS deployment, it is this thesis' suggestion that people will 'continue to matter most'.[114] Empirically, innovative methods and weaponry can only be effective if integrated with (and, notes Latiff, under the command of) troops who are both appropriately trained in its use and, crucially, motivated to win the fight. Motivation is a further contextual component in judging the feasibility of weapons without human supervision. It encompasses traits such as incentive, impetus, catalyst and drive. Regardless of weaponry, Bahnsen and Cone have highlighted that it is these characteristics that win fights although Clark significantly notes that this relationship is clearly altered by, first, fighting 'at distance' and, second, by fighting with assets that are themselves autonomous.[115] In this vein, robots may help humans in that fight but, in the final analysis, the context is that it is *human* resources that then seal the battle. A difficulty is that this relationship is imprecisely influenced by several competing conditions such as battle's proximity, by leadership and, note Grinker and Spiegel, by physiological and psychological mechanisms that, together, describe what is human context.[116] Nevertheless, it holds that the 'essence of the matter' (here, the violence of combat) all around the front-line soldier remains unchanged regardless of the weapon system at his or her disposal.[117] As noted in a recent US Army Handbook, '[h]owever much the tools of war may improve, only soldiers willing and able to endure war's hardships can exploit them'.[118] Taken together, therefore, such analysis starts to frame the human context with which to review AWS deployment, a dynamic that is based on human precedence over technology and, after all, moulded by the aphorism that it is 'usually much easier to predict technological change, even to understand how it should work, than it is to comprehend what it will *mean*'.[119] While such a framework may be based upon the competence of the fighting soldier and 'his skill in injuring people and damaging things', it does not, however, deny a role for battlefield autonomy.[120] Indeed, Wadwa and Johnson note that context can accommodate a relationship between man and weapon in an age where man may no longer be in charge of that weapon's effects.[121] This, of course, highlights an additional

---

[113] Katya Drozdova, *Low-tech threats in the Hi-tech Age: Subversive networks across ideologies, technologies and times,* (USA: University of Michigan Press, Analytical Perspectives on Politics Series, 2002), p. 1. Also: Major-General Patrick Cordingley, in conversation with the author, January 2019.

[114] Robert Latiff, *Future War: Preparing for the New Global Battlefield*, (USA: Knopf Doubleday Publishing, September 2017), pp. 3-16 ('*Introduction'*).

[115] John Bahnsen and Robert Cone, *Defining the American Warrior Leader*, (USA: Parameters, December 1990), pp. 24-28. Also, Professor Lloyd Clark, in conversation with the author, January 2019.

[116] Roy Grinker and John Spiegal, *Men Under Stress*, (US: Philadelphia Press, 1945), '*Abstract'*.

[117] Ibid., p. 25, here the 'utter inalienable violence of combat'.

[118] US Army, *Serving a Nation at War,* (USA: Washington DC, 2004), p. 8.

[119] Gray, '*Another Bloody Century'*, p. 39. *Italics* by the thesis' author for emphasis.

[120] Ibid., p. 45.

[121] Vivek Wadwa and Aaron Johnson, 'Robots could eventually replace soldiers in warfare. Is that a good thing*?'*, *Washington Post*, 5 October 2016 <https://www.washingtonpost.com/news/innovations/wp/2016/10/05/robots-could-eventually-replace-soldiers-in-warfare-is-that-a-good-thing/?noredirect=on&utm_term=.1a5862445221> [accessed 22 August 2018].

(even unexpected) conundrum: While humans remain responsible for AWS' design, its testing and validation, its coding, calibration and maintenance, the human is also abrogating himself from the moment when that weapon kills.

In terms of context, how then might AWS deployment posit a new relationship between human soldier and the battlefield? After all, change on a human level is not obvious: Warfare, point out Stojkovic and Dahn, remains a set of actions pivoting around human inputs and, in the State system at least, the principal contribution to defence planning remains the role of politics and political process.[122] The particularly human context to AWS deployment is that autonomous, automatic and manual means of violence are, essentially, all similar means of dispensing lethality, the means to cause death or damage.[123] Notwithstanding rogue human behaviour[124], in State systems of warfare it remains political (and therefore human) processes that occasion *use* of lethal force.[125] Furthermore, social and cultural dimensions reflecting the characteristics of those States' communities must similarly originate from entirely human interactions. The contextual corollary is that conflict will persist as a universally human activity long after AWS deployment and, in this sense, deployment of AWS is just one further ingredient to humans' ability to initiate violence for political purposes. While it may tinker with the equation of personal danger, AWS deployment does not change the unchanging *nature* of war. Nor does it refute that difficulties abound in such contextual tests. In the first place, while human activity will continue to dominate battlecraft, it is not possible to translate the context that arises from this conclusion into empirical AWS behaviour given, of course, fundamental (and enduring) challenges to weapon coding and, as discussed in this thesis' technical review, challenges arising from AWS' underlying machine-learning spine.[126] Second, not all wars are settled by violence and, as noted by Turitto, it is important that conclusions around context do not inappropriately underweight low-intensity and other non-lethal types of warfare such as non-State, economic, electronic and other psychological means of coercion.[127]

Other elements of human context will inform the processes of AWS deployment. An example is in the pace of AWS adoption where humans either embrace speed-of-change or put up resistance to transformation. Cantwell recounts a survey on attitudes towards unmanned weapons within the US Air Force which found that one third of pilots had 'wrapped up their professional identity so tightly around the act of flying' that they would rather leave the service than fly a remotely piloted

[122] Stojkovic and Dahn, p. 17.

[123] The distinction (death or damage) is deliberate and its legal consequences are taken up in Chapter 5 (*Obstacles*), specifically: 5.5 ('*Article 36 and LOAC-compliant weaponry*'), and Chapter 10 (*Oversight*), specifically: 10.1 ('*Meaningful Human Control*').

[124] Lonnie Harelson, 'The Principles of War: Valid Yesterday, Today and Tomorrow', *Joint Forces Staff College, Norfolk US,* (2005) <http://www.dtic.mil/dtic/tr/fulltext/u2/a436747.pdf> pp. 24-26 ('*The Future Has A Need For the Principles of War*'). Fighting wars is not a science and context must therefore factor in the unpredictable nature of human involvement (and weapon conformance) in its processes.

[125] For discussion on the use of force and relevant case studies, see: International Committee of the Red Cross, *The Use of Force in Armed Conflict: Interplay between the Conduct of Hostilities and Law Enforcement Paradigms*, pp. 13-43, (ICRC, November 2013).

[126] This is severally discussed in this thesis' technical analysis. See Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt*'), and Chapter 8 (*Software*), specifically: 8.1 ('*Coding methodologies*') and 8.2 ('*Coding errors*').

[127] James Turitto, 'Understanding Warfare in the Twenty First Century', *International Affairs Review*, Volume XVIII, Number 3, (Winter 2010) <http://www.iar-gwu.org/node/145> [accessed 8 February 2018].

aircraft.[128] Boulanin and Verbruggen note similar skepticism in other groups of military service personnel.[129] Suspicion around what appears 'new' arises, after all, from a long-held precept that 'every change in a military context is risk'.[130] This caution, moreover, is widely reinforced, not least by soldiers' lack of trust in unproven technology and attendant changes to their responsibilities[131], by organisational stasis, by frequent lack of coherent vision between military services and, notes Singer, an overarching focus on *current* operational priorities.[132] In this vein, AWS' inherent traits (composite and therefore complex systems, impenetrable software that operators cannot readily fix and, as later detailed, AWS' inherently unpredictable outcomes) are unlikely to inspire rank-and-file confidence in the event of technical lapses and, as noted in the US Army's *Army Equipping Strategy*, such technologies' likely requirement for a long unstable period of adoption.[133] Empirically, technical challenges are also difficult to refute once lodged in the soldier's mind and, as noted by Wheeler (and in this thesis' technical review), questions will persist whether weapon autonomy will perform as intended in complex battlefields situations.[134] In this sense, weapon autonomy 'is not a concrete visible object but instead a diffuse, remote set of capabilities hidden deep within a larger weapon system'[135] creating different views within different branches of military service about which AWS technologies are critical on the future battlefield. It is this imprecision that has also complicated the role of Civil Society in its efforts to contain such technologies. As noted by Roberts and Evanoff, 'the distinguishing features that make autonomous weapons enticing military investments also make placing restrictions on them difficult.'[136]

The precise impact of such biases may be debatable but are certainly shaped by human context, in particular the different operational realities faced by each combat branch.[137] Heuristically, AWS deployment will be governed in part by the priority given by each military

[128] Houston Cantwell, *Beyond butterflies: Predator and the evolution of unmanned aerial vehicles in Air Force culture,* (USA: School of Advanced Air and Space Studies: Maxwell Air Force Base, AL, 2007), pp. 81-85.

[129] Boulanin and Verbruggen, p. 72.

[130] Michel Goya, cit. Boulanin and Verbruggen, p. 71.

[131] Peter Singer, 'Tactical Generals: Leaders, Technology, and Perils', *Brookings Institute*, (7 July 2009), paras. 9-10, 14-15 and generally <https://www.brookings.edu/articles/tactical-generals-leaders-technology-and-the-perils/> [accessed 18 January 2019].

[132] Ibid. Also: Boulanin and Verbruggen, p. 71.

[133] US Army, 'The Army Equipping Strategy', *Army G-8*, (2009), pp. 4-5 <http://www.g8.army.mil/pdf/Army_Equipping_Strategy.pdf>.

[134] Scott Wheeler, 'Trusted autonomy: conceptual developments in technology foresight', *Defence Science and Technology Group Report*, (Australian Government, Department of Defence, Victoria, 2015), <http://www.dtic.mil/dtic/tr/fulltext/u2/a626723.pdf>.

[135] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[136] Megan Roberts and Kyle Evanoff, 'Can Civil Society Succeed in its Quest to Ban Killer Robots?', *World Politics Review,* 17 November 2017, https://www.worldpoliticsreview.com/articles/23636/can-civil-society-succeed-in-its-quest-to-ban-killer-robots [accessed 13 March 2019].

[137] A useful example is provided by US inter-service attitudes towards swarming and multi-vehicle control. While the US Air Force remains sceptical about the technology (and has no current interests in its development), the US Navy and Army are enthusiastic about its operational possibilities: See Chapter 4 (*Deployment*), specifically 4.6 ('*Swarming model'*) and, generally, John Arquilla and David Ronfeldt, *Swarming and the future of conflict,* (USA: Santa Monica: Rand Corporation, 2000), <https://www.rand.org/pubs/documented_briefings/DB311.html> [accessed 12 December 2017].

service to embedding those technologies that it deems most useful to its current practices.[138] Boulanin and Verbruggen note that these biases are exacerbated when budgets are tight and when acquisition of untried hardware (that may not deliver capabilities for several years) is seen as being less attractive than procuring readily available, proven technology which will be immediately accretive.[139] In this vein, context perhaps has a role in prompting States to adopt a more cautious and incremental path towards independent weapons. As an adjunct, Rosen notes that it is often not until the occurrence of major conflict that such adoption of new technologies properly accelerates.[140] A final contextual drag to deployment is identified by Harari and arises from the phenomenon of human knowledge (as a proxy here for battlefield technology) increasing at an unprecedented speed. In this case, he argues that new-found knowledge actually fosters paralysis: In an effort to understand what is happening, 'knowledge explosion' instead leads planners and the procurement executive into stasis given ever *less* ability to 'make sense of the present or forecast the future'.[141] Clark similarly notes that institutional belief that lessons have been learned from an earlier campaign or prior technology may reduce subsequent creativity and questioning.[142] By inference, this then becomes the next 'wide socio-technical milieu' discussed above in which context actually slows the pace of battlefield change.[143]

A further nuance is that military technology (inter alia, AWS' numbers, ranges, accuracy and costs) may appear easier to calculate and entrench than less definable assets such as training, morale, organisation, doctrine and the quality of leadership.[144] This contextual wrinkle may also prove an important factor in AWS deployment given that the defence planner eventually needs to decide the extent to which she agrees with the aphorism that 'historically, good men with poor ships are better than poor men with good ships'.[145] Contextually, the inescapable allocation issue is between the *relative* worth of man and machine. The verso, however, is stressed by Cordingley that any over-reliance on the technical can actually be undone by a portfolio of *context-driven* factors such as poor operational direction, ambiguous strategy, over-extended logistics and, in the case of AWS deployment, simple operation in a cluttered unfamiliar geography.[146] It is this line of analysis that points to it being the *use* made of that weapons technology rather than the technology itself which is then important to AWS context and that it is the human combatant rather than the machine combatant that remains the most important component to this piece.[147] On this basis,

---

[138] Army Technology blog, 'Prioritizing Procurement', *www.armytechnology.com*, (20 November 2016), generally <https://www.army-technology.com/features/feature45999/> [accessed 12 June 2018].

[139] Boulanin and Verbruggen, p. 72.

[140] S Rosen, *Winning the Next War,* (USA: Cornell University Press, Ithaca, NY, 1991), generally.

[141] Yuval Noah Harari, *Homo Deus*, (UK: Penguin Random House, 2016), p. 46.

[142] Professor Lloyd Clark, in conversation with the author, January 2019.

[143] Bousquet, p. 2. Although written in 1991, see also: Craig Moore and others, *Measuring Military Readiness and Sustainability*, (National Defense Research Institute, RAND Publishing, Santa Monica, 1991) pp. ix-xiv and generally.

[144] Gray, *Another Bloody Century*, p. 98.

[145] Admiral Alfred Thayer Mahan, *The influence of Sea Power upon the French Revolution and Empire 1793-1812,* Volume one, (USA: Boston, 1898), p. 102.

[146] Major-General Patrick Cordingley, Commander, 7th Armoured Brigade, Gulf War, 1991, in conversation with the author, January 2019. See also: Gray, *Another Bloody Century*, p. 121.

[147] Outside the scope of this analysis, historical examples might include Japanese aviation's technical superiority over American machinery in 1942 in the Pacific. Similarly, German operational and tactical skill had material impact in negating French and British land forces superiority in early 1940.

context around deployment should therefore afford greater weight to intangibles such as superior training, discipline and morale, effective leadership, greater numbers, intelligence and logistics. Although an over-simplification and written in 2004, the point is usefully laboured in the *US Army Handbook* by its exaggerated statement that '[h]owever much the tools of war may improve, only soldiers willing and able to endure war's hardships can exploit them'.[148]

This precept is backed up by Coker's analysis of Thucydides whereby it has often proved historically possible to compensate for technological disadvantage and where, in a world today where knowledge diffusion is extensive, there is sound argument to assume that technological advantage will be increasingly fleeting.[149] While weapon innovation may be a constant, it is also one where effects may empirically be blunted by emulation, by parallel discovery and adaption or by evasion. Indeed, notes Macias, newer and better devices 'are always just round the procurement corner' and there is therefore no resting point for the defence planner on the continuum of weapons development.[150] Finally to this point, each such novel technology can perform no better than the crew and personnel who must direct them (in the case of AWS, the set-up and monitoring of those systems). Inferring from Adams' *Future Warfare and the Decline in Human Decision Making* would suggest that AWS' calibration and its need for dynamic updating will mean that removing general human supervision only amplifies AWS' vulnerability to all of those contextual challenges identified above.[151] It is this condition that leads Cordingley to question whether AWS' Delivery Cohort will practically delegate key tasks (the security, perhaps, of a flank) to machines without human oversight.[152]

2.6 The role of situational awareness and uncertainty

In order to complete this chapter's analysis, its final section now considers context from the perspective of the engaging weapon. What contextual evaluation must reasonably be included in AWS' targeting and decision processes? Key to this is the notion of 'situational awareness' which Dorstal defines as 'understanding of the operational environment in *all of its dimensions* – political, cultural, economic, demographic as well as military factors'.[153] Situational awareness is also a catchall notion that should include both strategic and tactical awareness. Its importance is evidenced by Suchman in her conclusion that such capability is a pivotal component for weapon

[148] US Army, *Serving a Nation at War,* (USA: Washington DC, 2004), p. 8.

[149] Christopher Coker, 'Still "the human thing"? Technology, human agency and the future of war', *International Relations*, 32.1, (2018) <http://eprints.lse.ac.uk/87629/1/Coker_Human%20Thing.pdf> pp. 23-38. See also: Gray, *Another Bloody Century*, pp. 121-128. For a discussion of Boot's theory of technical nullification, see: Chapter 5 (*Obstacles*), specifically, 5.6 ('*Behavioural constraints'*).

[150] Amanda Macias, 'Weapons of the Future: Here's the New War Technologies Lockheed Martin is Pitching to the Pentagon', *CNBC*, (6 March 2018), generally <https://www.cnbc.com/2018/03/06/future-weapons-lockheed-martin-pitches-new-war-tech-to-pentagon.html> accessed 18 January 2019].

[151] Inferred from: Thomas Adams, *Future Warfare and the decline in human decision making,* 2, (Parameters, 2001), p. 12 <http://ssi.armywarcollege.edu/pubs/parameters/articles/2011winter/adams.pdf>.

[152] Major-General Patrick Cordingley, in conversation with the author, June 2017.

[153] Brad Dorstal, 'Enhancing situational understanding through the employment of unmanned aerial vehicles', *Centre for Army Lessons Learned*, (2001) <http://www.globalsecurity.org/military/library/report/call/call_01-18_ch6.htm> [accessed 6 March 2017]. Such awareness incorporates melding judgements that have been based upon, inter alia, classification of persons, escalation/de-escalation of force, definition and ramifications of protected places, battlefield responsibilities and permissions, appropriate selection of armaments as well as the three principles of distinction, military necessity and proportion.

adherence to International Humanitarian Law (IHL) as well as to 'any other form of legally accountable rules of conduct in armed conflict'.[154] The interpretative routines that define situational awareness cannot, moreover, be specified in law given, generally, the uncertainty of actions in situations of armed conflict. What is clear, however, is that such military rules require these same capabilities in *each* specific and 'actually occurring' situation. The purpose of this section is therefore to review the tenets of what comprises appropriate situational awareness in autonomous self-directing weapons. It is also to judge how they might impact AWS' contextual framework ahead of reviewing AWS' fundamental feasibility in this thesis' later technical sections.

Tyugu notes that multiple complications arise from AWS' having to precede lethal engagement with such an awareness check.[155] Coding for situational awareness is challenging precisely because 'future autonomy will need to process sensor data and inputs in order more effectively to create its own internal situation model to direct its decision-making'.[156] This contrasts with human senses that provide a continuous flow of stimuli. An adjunct challenge is then the mediation that is necessary between the weapon's battlefield inputs, each of which must compete to provide the host with an intuitive, comprehensive and moment-to-moment picture on its surrounding environment. In this respect, human perception is an effortless process. Human eyes, for instance, do not tell the brain what objects they see any more than a camera currently informs what objects it is capturing.[157] How, therefore, might situational awareness be undertaken by AWS? The challenge here is the supervision and adjudication of complex sensory signals as well as the management of their intensity. Empirically, such signals will be filtered according to threshold values as well as by circuits that are designed to determine 'match', 'mismatch' and 'novelty' relationships existing between these sensory and feedback signals and the AWS' initial representations.[158] This must be undertaken under 'right first time, every time' criteria[159], including context-based routines that match AWS' sensed data with both targeting parameters and required rules of engagement.[160] Technical narrative evidences these challenges. The self-directing weapon must generate a mismatch message if signal arrays do not correspond. Similarly, it must produce novelty signals whenever a sensory signal appears without a corresponding feedback signal. Even at a theoretical level, achieving situational awareness in machines is very difficult. The constraint is that such processes require broad feedback routines, created (and then processed) at each stage of this awareness routine with individual outputs being re-introduced, prioritised, filtered, weighted and,

---

[154] Suchman, *'*Situational awareness and adherence to the principle of distin*ction',* p. 1. See also: Chapter 5 (*Obstacles*), specifically: 5.1 ('*Geneva Convention and the Laws of Armed Conflict'*), generally.

[155] Enn Tyugu, 'Situational Awareness and control errors of cyber weapons'*, Cognitive Methods in Situational Awareness and Decision Support*, IEEE International Multi-disciplinary Conference, (February 2013), ('*Abstract'*).

[156] United Air Force, Office of the Chief Scientist, *Autonomous horizons - system autonomy in the air force; a path to the future*, (USA; USAF Publishing, June 2015), p. 33 <http://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf?timestamp=1435068339702> [accessed 19 August 2016].

[157] Pentti Haikonnen, *The cognitive approach to conscious machines,* (UK: Imprint academic, 2003), p. 42.

[158] For instance: Raytheon Patent US6952001B2, 'Integrity Bound Situational Awareness and Weapon Targeting', (2005) <https://patents.google.com/patent/US6952001B2/en> [accessed 19 February 2017]. The routines and their challenges are covered in detail in Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues'*).

[159] Meghan Han, 'Lethal Autonomous Weapons and Info-Wars: A Scientist's Warning', *Medium*, 6 July 2017, paras. 7, 12 and 21 of 27 <https://medium.com/@Synced/lethal-autonomous-weapons-info-wars-a-scientists-warning-cc95798bc302> [accessed 12 February 2018].

[160] UK Government, '*Rules of Engagement*', generally.

as necessary, acted upon (and all in keeping with machine goals and without distortion of the weapon's primary feedback). For compliant function, the weapon's controller must presumably predict (and search) for other contextual representations in order to affirm immediately prior battlefield readings.

This points to a further challenge: Coding for the separation of desired and undesired associations (in what, after all, is a chaotic battlefield environment[161]) presents an enduring bottleneck to enabling appropriate situational awareness. While its management is covered in this thesis' later technical analysis, awareness routines (for the purposes of weighting context's role in AWS deployment) must also enable suppression of 'inappropriate' sensory input in order to ensure that the weapon's momentary representation accurately reflects its current state as modified by new sensory input as well as output from subsequent feedback loops.[162] Situational awareness, furthermore, must account for *associative* meanings across very different types of sensor groups requiring, in turn, complex cross-connections between AWS' different sensory modules. Are the weapon's contextual associations to be pre-set and immutable or rules-based and varying, and within what framework might they then be improved through subsequent learning processes? Hew notes that it is how these issues are managed that will define AWS context from the perspective of the engaging weapon.[163] Such fundamentals therefore dictate both the weapon's overall system architecture but also the balance between integration and manipulation of sensed information and, contextually, the effectiveness of the weapon that can then be deployed.[164]

2.7 Contextual inputs for AWS operation

Divining appropriate context from its surroundings is similarly a hardware issue in AWS deployment. In this vein, Suchman points out that existing weapon sensors may be able to identify an object as a human but cannot currently make necessary discrimination among persons as required by the legal principle of distinction.[165] Furthermore, notes the World Economic Forum, zones of combat will be increasingly fluid, contested and defined by uncertain boundaries making it insufficient for AWS to rely solely upon prior internal models to define context and compliant operation.[166] This is a telling point. Machine sensors will presumably fail where combatants are disguised or in situations where an enemy employs obfuscation or false denial. Wilke notes the rise

[161] "Its grammar, indeed, may be its own, but not its logic" (Clausewitz, cit. Robert Cassity and Jacqueline Tame, 'The Wages of War Without Strategy: Beyond the Present – A Call to Clausewitz and to Conscience', *Strategy Bridge*, Abstract and generally, (23 August 2017) <https://thestrategybridge.org/the-bridge/2017/8/23/the-wages-of-war-without-strategy> [accessed 17 January 2019].

[162] Nathan Brennon, 'Coordinated Machine Learning and Decision Support for Situational Awareness', *Sandia National Laboratories, Report*, (September 2007), pp. 23-29 <http://prod.sandia.gov/techlib/access-control.cgi/2007/076058.pdf>.

[163] Patrick Chisan Hew, 'The Generation of Situational Awareness – A Near to Mid-Term Study', *Defence System Analysis Division*, (Australia: Australian Army Publishing, July 2006), pp. 6-8 ('*Technical Bottleneck Issues*') <http://www.dtic.mil/dtic/tr/fulltext/u2/a465252.pdf>.

[164] Haikonnen, p. 190.

[165] Suchman, 'Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy', p. 3.

[166] See, generally: Anja Kaspersen and others, 'Ten Trends for the Future of Warfare', *World Economic Forum* (3 November 2016), Paragraphs 2 ('*Speed Kills*'), 3 ('*Fear and Uncertainty increases Risk*'), 6 ('*A Wider Cast of Players*'), 7 ('*The Grey Zone*') and 9 ('*Expanded Domain of Conflict*'), <https://www.weforum.org/agenda/2016/11/the-4th-industrial-revolution-and-international-security/> [accessed 18 January 2019].

of the *unlawful combatant*, categorised in this instance as the 'illegitimate, non-innocent, suspicious civilian'.[167] Sophisticated technologies may include facial or gait recognition but remain reliant on pre-established databases where profiles remain inherently vulnerable to false positives, dissembling and inaccurate categorization. Such databases must also be updated dynamically in what is likely a communications-denied environment.[168] Currently posited methodology for AWS deployment, observes Schuppli, relies inappropriately upon third party intelligence gathering, assessment and military action, including the calculation of who can legally be killed, largely be performed by machines and 'based on an ever expanding database of aggregated information'.[169] An issue for AWS deployment is therefore that context is particularly complicated exactly because there can 'be neither an objectively correct answer nor one that is testable save by the verdict of future events'.[170]

As pointed out by Macauley, a final point of context for AWS deployment arises from the pell-mell pace of advance of technology across all of AWS' componentry.[171] Such pace of change questions what may eventually be properly disruptive and what is currently (at the point of decision) sustainable in these systems. In this vein, the *US Army Research Laboratory* (USARL) looks at the impact of weapon technology thirty years hence in order to identify what it judges will become key drivers in battlecraft.[172] Their focus is driven in large part by measures that seek to address current and, arguably, enduring shortcomings in the collection and processing of battlefield context. Evidencing its significance, the USARL focuses upon matters *disrupting* context-collection by adversaries including competences 'that could be used to deny, deceive, disrupt, degrade and compromise adversary information and information-related processes'.[173] Their principal prediction is for 'super-sensing and sense-making [enhanced] humans' rather than inanimate intelligent agents on the battlefield.[174] The contextual point here is that their forecast is *not* for independent AWS. Instead, therefore, context should be underpinned by what is expected in 2050 to be self-arranging and self-marshalling forces (incorporating material autonomous capabilities) but with 'fluid command and control' whereby human individuals, human-led teams and software agents will, as appropriate, 'self-organise, dynamically creating and modifying collaborative

[167] Christiane Wilke, 'Civilians, combatants and the history use of International law', *Critical Will: Law ad the Political*, 28 July 2014 <http://criticallegalthinking.com/2014/07/28/civilians-combatants-histories-international-law/> [accessed 5 May 2017].

[168] MC Haas and SC Fischer, 'The Evolution of Targeted Killing Practices: Autonomous Weapons, Future Conflicts and the International Order', *Contemporary Policy*, 38:2, (August 2017), p. 284 <https://www.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Haas&Fischer_2017_TargetedKillingPractices.pdf>.

[169] Susan Schuppli, 'Deadly Algorithms', *Continent*, Issue 4.4, p. 20-27 <http://www.continentcontinent.cc/indenx.php/continent/article/view/212> [accessed 7 May 2017].

[170] Gray, *Strategy and defence planning; Meeting the Challenge of Uncertainty,* p. 2.

[171] Thomas Macauley, 'The Future of Technology in Warfare: From AI Robots to VR Torture', *Techworld*, 13 January 2017 <https://www.techworld.com/security/future-of-technology-in-warfare-3652885/> [accessed 15 May 2017].

[172] US Army Research Laboratory, *Visualizing the tactical ground battlefield in the year 2050: workshop report,* (USA: ARL-SR-0327, June 2015). Here, the 'obtaining, collecting, organizing, fusing, storing and distributing relevant information' includes capabilities around 'command and control functions and processes including reasoning, influence, planning, decision-making and collaborating'.

[173] Ibid., p. 2.

[174] Ibid., p. 8.

processes'.[175] This broad narrative is contextually supported by the UK Ministry of Defence's recent *Future Operating Environment 2035*. In this case, an attempt is made to define likely characteristics of the future battlefield in an effort to inform the UK's on-going defence capability.[176] It is rather *because* of the uncertainty that such prediction entails that a broad grasp of context is so important. In the case of AWS deployment, the prediction for 2035's battlefield is that it will be 'congested, cluttered, connected and constrained' and, by inference, requiring MHC in weapon management.[177]

Chapter Two is therefore intended to provide a prism through which the complications of self-directing weapons can be judged. Broad context offers a framework through which to identify and then weight the several layers of challenge that might impact AWS deployment. Given the very wide 'art of the possible' (here, the dystopian scenario put forward by Stuart Russell), it is context that provides both narrative and boundaries in order to understand what appear as very quickly evolving sets of circumstances.[178] An example has been the role of the defence planner and her requirement to determine tipping points and what comprises certain disruption. A second prism is then called for by the importance of context *within* the weapon system. It is this step which then calls into focus whether independent weapons can both capture and then manage the relevant situational awareness that is a necessary prerequisite for compliant lethal engagement. Together, context is therefore a fundamental tool in this thesis' overall investigation. It is consistently relevant to understanding why (together, '*drivers'*), how (together, '*deployment'*) and why not (together, '*obstacles'*) AWS deployment may or may not be viable.

[175] Ibid., p. 11.

[176] UK Ministry of Defence, *Strategic Trends Programme, Future Operating Environment 2035,* (UK: Crown Copyright, August 2015), p. viii.

[177] UK Ministry of Defence, *Strategic Trends Programme,* p. 1.

[178] Susan Schuppli, 'Deadly Algorithms', p. 26.

3. Drivers: Factors accelerating the removal of weapon supervision

This thesis now looks to understand the several drivers that contribute, either singularly or in concert, to the removal of human supervision in lethal engagements.[1] Such drivers are rarely discrete and tend instead to overlap. They have the broadest genesis, different intensities, diverse trajectories and, as detailed below, may be political, social, ethical, moral, economic, operational, doctrinal or structural in their spur. Empirically, drivers accelerating the adoption of autonomy will share several characteristics. An important recognition, however, must be that parties remain in the early stages of developing robotic battlefield technologies[2] and, moreover, that fully autonomous weapons constitute an extreme of this robotics continuum.[3] In reviewing what is visible at States' policy level[4], the Stockholm International Peace Research Institute (SIPRI) concludes that 'there is not yet a declared arms race in autonomy'.[5] The broad narrative to this chapter is therefore that State positions on autonomy will 'develop and mature as they increasingly integrate robotic technologies into their arsenals'.[6] Furthermore, the notion of a continuum that exists between currently robotic battlefield assets and those same assets that increasingly operate without human supervision underlines much of this chapter's analysis.[7]

At its root, several push and pull factors exist that will influence State adoption, first, of robotics and, subsequently, of battlefield autonomy. These are framed by militaries' requirement to improve processes and stay ahead of likely adversaries' capabilities but also by public opinion and,

[1] For the purposes of this thesis, 'driver' is defined as 'a factor which causes a particular phenomenon to happen or develop'.

[2] Paul Scharre, 'The Coming Swarm: Robotics on the Battlefield', *Real Clear Defense*, (19 October 2014), generally <https://www.realcleardefense.com/articles/2014/10/20/the_coming_swarm_robotics_on_the_battlefield_107499.html> [accessed 8 December 2017].

[3] Khasha Ghaffarzadel, 'New Robotics and Drones, 2018-2038: Technologies, Forecasts, Players', *IDTechEX Reports*, (2018), generally <https://www.idtechex.com/research/reports/new-robotics-and-drones-2018-2038-technologies-forecasts-players-000584.asp> [accessed 20 January 2019]. Several continua (relevant to AWS deployment) are reviewed in this thesis including those in weapon tasking and assignment, in procurement practices and in defence planning. Continua also exist in battlefield practices (firepower mass to manpower mass to technology mass), in weapon capability and the move from automation to autonomy as well as the sequences within target acquisition, selection and dispatch. For a discussion on this relationship, see: Noel Sharkey, 'The 'Evitability' of Autonomous Robotic Warfare', *International Review of the Red Cross*, 94, 886, (Summer 2012), generally <https://www.icrc.org/eng/assets/files/review/2012/irrc-886-sharkey.pdf>.

[4] Stopkillerrobots.org, 'Country Policy Positions', *SKR*, (25 March 2016) <http://www.stopkillerrobots.org/wp-content/uploads/2015/03/KRC_CCWexperts_Countries_25Mar2015.pdf>.

[5] SIPRI is the Stockholm International Peace Research Institute undertaking research into conflict, armaments, arms control and disarmament. See: Boulanin and Verbruggen, '*Mapping the development of autonomy in weapon systems*'. The SIPRI analysis is based on the world's top ten arms producing nations measured by arms sales and includes the USA, Russia, China, France, Germany, Israel, South Korea, Japan, India and the UK. HRW's Mary Wareham points out that the first such database was compiled by ASU's Dr Heather Roff and Richard Moyes of Article 36, <https://futureoflife.org/wp-content/uploads/2017/01/Heather-Roff.pdf?x41605> [accessed 19 December 2017]. For the most current list of States calling for a ban on AWS, see: Campaign to Stop Killer Robots <https://www.stopkillerrobots.org/wp-content/uploads/2018/11/KKC_CountryViews22Nov2018.pdf>.

[6] Boulanin and Verbruggen, p. 58.

[7] Economist Magazine, 'Autonomous Weapons are a game-changer', *Economist*, 25 January 2017, generally <https://www.economist.com/special-report/2018/01/25/autonomous-weapons-are-a-game-changer> [accessed 12 March 2018].

notes Clark, by general defence debate.[8] Robotics, after all, and subsequent removal of human intervention from these processes posit several benefits and a purpose of this chapter is to review this potential. At its most basic, battlefield robotics can sidestep circadian shortcomings to human performance such as fatigue as well as often-inconvenient human requirements for upkeep and other support. Apthorp, for instance, notes that robots will likely have a main role in army logistics:[9] The payload of Boston Dynamics' 340-pound robotic BigDog is three times greater than that which can be hauled by the regular infantryman.[10] As important, however, are *push* factors in this dynamic, those contextual drivers inferred from Gillespie that include AWS' deployment's political, economic, social, technical, environmental and legal factors.[11] It is these components that provide the common thread to this chapter's analysis between, first, the adoption of robotic battlecraft and, thereafter, the migration of unmanned battlecraft towards autonomous processes. That migration is impacted by a swathe of quite uncorrelated influences such as the political landscape[12], the ramifications of an ageing population (with, perhaps, its decreasing tolerance for casualties[13]), the quickly evolving character of warfare that is evidenced in later chapters[14] as well, of course, as increased use of robotics in the commercial and domestic domains.[15] While the certain impact of these general drivers is hard to measure, it is this chapter's assumption that they will contribute cumulatively to adoption of autonomy in battlefield practices.

Given then that these factors are combinatory, structure is required within which to categorise the drivers that together facilitate adoption of AWS. Stanford futurologist Seba provides such a framework which, once repurposed, is useful to understanding removal of human supervision in lethal engagements.[16] Adapting then his model to AWS deployment, a disruptive weapon system should exhibit four characteristics. It should be possible to identify permanent and dramatic changes in the cost curves of technologies that comprise that weapon (here, the costs of the weapon as a function of the total quantity of such weapons produced). There should be dramatic increase in the weapon's capabilities brought about the convergence of these technologies. Such transformation, furthermore, should be accompanied by material innovation in the deployment models of these technologies as well as integration of the resulting weapon system into current practices. Finally, the disruption process should be subject to what Seba identifies as an S-curve

[8] Professor Lloyd Clark, in conversation with the author, January 2019, in particular 'the recent imperative of value based defence'.

[9] Claire Apthorp, 'Using Autonomy To Supply the 'Last Mile'', *Army Technology*, 25 June 2017 <http://www.army-technology.com/features/featureusing-autonomy-to-supply-the-last-mile-5852408/> [accessed 7 November 2017].

[10] See: Boston Dynamics, <https://www.bostondynamics.com/ls3> [accessed 2 March 2018].

[11] Andrew Gillespie*, Foundations of Economics; PESTEL Analysis of the Macro-Environment,* (Oxford: Oxford University Press, 10 March 2011) <https://www.kantakji.com/media/1610/ty3.pdf>.

[12] David Houle, *Entering the Shift Age: The End of the Information Age and the New Era of Transformation,* (USA: SourceBooks, 2012), generally.

[13] William Boettcher and Michael Cobb, 'Don't Let Them Die in Vain: Casualty Frames and Public Tolerance for Escalating Commitments in Iraq*', Sage Journal*, Volume 53, Issue 5, 13 July 2009, 677.

[14] Oxford Changing Character of War Centre, generally <http://www.ccw.ox.ac.uk/research/> [accessed 19 January 2019].

[15] International Federation of Robots, 'World Robotics 2017', *IFR*, Executive summary, 2017, pp. 15-17 <https://ifr.org/downloads/press/Executive_Summary_WR_2017_Industrial_Robots.pdf>.

[16] Tony Seba, 'Clean Disruption: Why conventional energy and transportation will be obsolete by 2030'*, Presentation to Swedbank*, (17 March 2016) <http://www.swedbank.no/idc/groups/public/@i/@sc/@all/@lci/documents/presentation/cid_1987411.pdf>.

adoption whereby a period of humdrum development is at once inflected by a tipping point in the product's adoption after which the system's implementation switches from linear to exponential.[17] To justify his analysis, Seba points to the recent cost curve disruptions in lithium batteries, in computing power, in LIDAR[18] as well as acceleration arising from innovative implementation models such as Open Source computing.[19] Seba's model supports this chapter's later analysis of drivers as well as, generally, this thesis' consideration of AWS' as a discontinuity in battlecraft.

The general procurement foundation for autonomous weaponry has been in place since the late 1990s.[20] More than two decades ago, researchers at UCLA and Hewlett-Packard had succeeded in building microscopic integrated circuits using single molecules as building blocks.[21] Heath, the UCLA professor leading that project, suggested at the time that a 'molecular computer with the processing power of one hundred conventional personal computers would be about the size of a grain of salt'.[22] The procurement implications of this remain significant to fielding independent weaponry including, notes Long, inexpensive capacity relative to previous peer hardware, ubiquitous supercomputing and 'almost unlimited memory capacity in devices so small that they are on the scale of insects'.[23] Doctrinal developments have been similarly significant. Already identified in the 1980s as one constituent of the framework for an 'automated battlefield', weapon autonomy has long been part of US military calculations.[24] In this way, US Defense Secretary Hugel's 2016 *Defense Innovation Initiative*[25] (otherwise referred to as *Third Offset Strategy* or TOS) is only the most recent initiative in a series of State initiatives seeking to use innovation in defence to overcome operational challenges.[26] Using predictions on what constitutes AWS deployment may

---

[17] Tony Seba, 'Clean Disruption' on *YouTube*, <https://www.youtube.com/watch?v=2b3ttqYDwF0> [accessed 12 June 2017].

[18] Ibid. Seba identifies a 14% improvement in Lithium battery performance per $1 from 1995 to 2010 and a 20% improvement thereafter; in 2000, one teraflop of processing power cost $46,000,000 and required housing in 150 square metres' space. In 2016, 2.3 teraflops cost just $59 and was available as a personal hard drive. By 2019, it is expected that 20 teraflops, the performance criterion for autonomous vehicles, will be available; the LIght Detecting And Radar (LIDAR) unit that cost $150,000 in 2012 currently costs just $250. This, notes Seba, is expected to fall to $90 in 2019.

[19] Source: OpenSource.com, 'What is open source computing?' <https://opensource.com/resources/what-open-source> [accessed 12 February 2018].

[20] Walker, *Killer Robots?,* pp. 26-28.

[21] David Rotman, 'Molecular computing', *MIT Technology Review*, (May 2000) <https://www.technologyreview.com/s/400728/molecular-computing/> [accessed 15 May 2017]. Also: John Markoff, 'Computer Scientists are Poised for Revolution on a Tiny Scale', *NY Times, Technology*, 1 November 1999 <https://archive.nytimes.com/www.nytimes.com/library/tech/99/10/biztech/articles/01nano.html> [accessed 15 May 2017].

[22] Adams, p. 4.

[23] Time Digital, '*Just right for Mini-Me: the Mini-Micro-PC*', *Time Digital blog*, 18 July 1999, cit. Adams, p. 4.

[24] Jeffrey Long, *The Evolution of US Army Doctrine: From Active Defense to the Airland Battle and Beyond*, (USA: US Army Command and General Staff College, Fort Leavenworth, Kansas, 1991), p. 122 and generally. Also: John Romjue and others, *Prepare the Army for War: A Historical Overview of the Army Training and Doctrine Command, 1973-1993*, (USA: TRADOC Historical Series, Office of the Command Historian, Virginia, 1993), pp. 44-48 <http://www.dtic.mil/dtic/tr/fulltext/u2/a267030.pdf>.

[25] See: Thomas Ridd, *Rise of the Machines: A Cybernetics history,* (USA: WW Norton, NY, 2016), generally.

[26] John Louth and Christian Moeling, 'Technological Innovation: The US Third Offset Strategy and the Future of Transatlantic Defense*', Armaments Industry European Research Group*, Policy Paper, (December 2016), pp. 3-6 <http://www.iris-france.org/wp-content/uploads/2016/12/ARES-Group-Policy-Paper-US-Third-Offset-Strategy-December2016.pdf>. See also: Cheryl Pellerin, 'Third Offset Bolsters America's Military Deterrence', *US Department of Defense*, (31 October 2016) <https://dod.defense.gov/News/Article/Article/991434/deputysecretary-third-offset-strategy-bolsters-americas-military-deterrence/> [accessed 7 July 2017]. Pellerin's analysis of the US 'Third Offset

therefore be one method of assessing its drivers but such an exercise is not clear-cut:[27] 2016's TOS was, after all, preceded by several quadrennial defence reviews with similarly heroic labels such as *Reconnaissance-Strike Complex*, *Transformation*, *Air and Sea Battle* and *Anti-Access/Area Denial*.[28] Practice, however, does not necessarily follow rhetoric and it is this disconnect that shapes how this chapter is organised.[29] First, it provides an analysis of the factors encouraging battlespace embrace of autonomy while reviewing the high-level ramifications of this move. It then reviews the notion of technology creep and the growing practice of dual-use in technologies behind machine (and therefore weapon) autonomy.[30] The chapter next considers structural and procurement drivers as well as constraints on their implementation before reviewing ethical, operational and then technical drivers to the introduction of autonomy into battlefield practices.

3.1 Current practice

A constructive starting point for an analysis on drivers comes from SIPRI's dataset which (current at the time of this thesis' writing) identifies three hundred and eighty-one different weapon systems that feature degrees of autonomy in some critical functions.[31] Crucially, however, this metric does not equate to three hundred and eighty-one *wholly* autonomous systems.[32] Two points arise. Bishop and Phillips' *Unmanning the Homeland* highlights what it identifies as a phenomenon of 'progressive' weapon automation.[33] Their inference supports the existence of a continuum from manned to autonomous weapons via robotics and automation. The progression from manned to robotic to autonomous main battle tanks may appear unsystematic but, notes Snow, it nevertheless provides a relevant precedent to States' adoption over time of wider unmanned battlefield

Strategy' highlights the role of AI (particularly machine learning) as a key component in improving the strengths and cost-effectiveness of its forces.

[27] Department of Defense, 'Unmanned Systems Integrated Roadmap FY2013-2038'. Here, DoD states that autonomy in unmanned systems will be critical to future conflicts 'that will be fought and won with technology'. The document highlights that 'the special feature of an autonomous system is its ability to be goal-directed in unpredictable situations'. It also sets out a 25-year vision for the development, production, test, training, operation and sustainment of unmanned systems technology across the Department of Defense. See also: Colin Roberts, 'Killer Robots: Moral Concerns versus Military Advantage', *The National Interest*, 3 November 2016, para. 1 of 11 <http://nationalinterest.org/blog/the-buzz/killer-robots-moral-concerns-vs-military-advantages-18277> [accessed 17 July 2017].

[28] Kathleen Hicks, 'What will replace the Third Offset? Lessons from Past Innovation Strategies', *Defense One*, 17 March 2017, para. 2 of 9 <http://www.defenseone.com/ideas/2017/03/what-will-replace-third-offset-lessons-past-innovation-strategies/136260/> [accessed 18 February 2018].

[29] John Gentry, *Doomed to Fail: America's Blind Faith in Military Technology,* (USA: Parameters, 32.4, 2002), pp. 88-90 <http://www.comw.org/rma/fulltext/0212gentry.pdf>.

[30] Deborah Shapley, 'Technology Creep and the Arms Race; A World of Absolute Accuracy', *Science Magazine*, Volume 201, Issue 4362, (29 September 1978), p. 1192 <http://science.sciencemag.org/content/201/4362/1192> [accessed 22 July 2017].

[31] See: Chapter 4 (*Deployment*), specifically analysis of Boulanin and Verbruggen, '*Mapping the Development of Autonomy in Weapon Systems*', pp. 57-84 ('*Drivers and obstacles*') <https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_1.pdf>.

[32] The point here is that such systems include an *element* of autonomous function in their operation. SIPRI has therefore identified unsupervised components in the weapon rather than that whole weapon being capable of autonomous function.

[33] Ryan Bishop and John Phillips, 'Unmanning the Homeland', *International Journal of Urban and Regional Research*, 26.3. (2002), 620-625.

systems.[34] In this vein, both Snow and Cummings argue that it is the pace of introduction as well as the breadth of robotic tasking that mask what is a profound disruption in procurement.[35] A further example is provided by the DoD's adoption of unmanned aircraft systems: By July 2013, the number of UAV already exceeded ten thousand units, demonstrating wide acceptance of battlefield automation.[36] Likewise, an inaugural requirement in 2013 by DARPA's *Robotics Challenge* for machines to undertake complex automated tasking (here, the negotiation of rough terrain, removal of debris and dealing with multipart maintenance tasks) has since matured into autonomy becoming a fundamental requisite in DARPA's subsequent challenges.[37]

It is this chapter's broad assumption that extensive implementation of automation in battlefield assets is broadly evident[38] which, notes Melendez, when taken together may posit a watershed in the framing of military technology.[39] How has this fulcrum been reached? As noted by Turnbull, it is the combination of broad-based innovation, much of it generated by academic, commercial and non-military parties, which have together proved important catalysts in accelerating weapon disruption.[40] This is not unexpected given that the development timelines for these technologies have been short.[41] By 1995 UT Arlington students had already demonstrated that an AAV was able to take off autonomously, locate and identify bio-hazardous material, map the location of each such barrel and return to its start point.[42] Three years later, a largely autonomous UAV no bigger than a model airplane successfully negotiated the Atlantic to land at a predetermined landing point in mainland Europe.[43] This short timeline is mirrored in the US by that State's own institutional roadmap. As early as 2001 (and motivated by cost, policy, opportunity and political imperatives), the US Senate Armed Services Committee had formalised its own aggressive schedule, albeit a guideline, towards autonomous and unmanned hardware by funding two central goals: Within ten years, one third of all deep-strike aircraft should be unmanned and, within fifteen years of that date, one third of all ground combat vehicles should operate without human beings on

[34] Shawn Snow, 'The US Army is Developing Autonomous Armoured Tanks', *Army Times*, 29 August 2017, paras. 3-4 and 6 <https://www.armytimes.com/news/your-army/2017/08/29/the-us-army-is-developing-autonomous-armored-vehicles/>) [accessed 9 December 2017].

[35] Shawn Snow, 'The US Army is Developing Autonomous Armoured Tanks', generally. See also, generally: Missy Cummings, 'Artificial Intelligence and the Future of Warfare', *Chatham House Research Papers*, (January 2017) <https://www.chathamhouse.org/sites/files/chathamhouse/publications/research/2017-01-26-artificial-intelligence-future-warfare-cummings-final.pdf>.

[36] Department of Defense, 'Unmanned Systems Integrated Roadmap FY 2013-2038', p. 5.

[37] DARPA website, 'DRC Trials 2013 Countdown: A Look at the Competition Course', *DARPA*, (16 December 2013), <https://www.darpa.mil/news-events/2013-12-16> [accessed 15 January 2018].

[38] Postnote, *Automation in Military Operations*, (UK: Houses of Parliament, Parliamentary Office of Science and Technology, Number 511, October 2015), generally.

[39] S Melendez, 'The Rise of the Robots: What the future holds for the world's armies', *FastCompany.com*, 12 June 2017, paras. 3,4 and 6 <https://www.fastcompany.com/3069048/where-are-military-robots-headed> [accessed 7 June 2018]

[40] Grant Turnbull, 'Off the Shelf: Re-thinking Innovation in the Military', *Army Technology*, 2 March 2014 <https://www.army-technology.com/features/featureinnovation-stagnation-re-thinking-innovation-in-the-military-4187511/> [accessed 18 January 2018].

[41] Boulanin and Verbruggen, pp. 85-89 ('*Relevant Innovations*').

[42] Arthur Reyes and others, 'Overview of the University of Texas and Arlington's Autonomous Vehicles Laboratory', *Department of Computer Science and Engineering, Technical Report CSE-2003-13*, (2013), generally. By way of subsequent narrative, in the four years to FY 2010, flight hours for UAS increased from 165,000 hours to more than 550,000 hours and the inventory from less than 3,000 to 6,500. See, here: T Adams, pp. 57-71.

[43] Arthur Reyes and others, p. 14.

board.[44] Such direction and what Antonova terms 'institutional purpose' remain crucial deployment drivers in the dilution of human oversight in military hardware.[45] In the case of the US DoD, moreover, the introduction of autonomy has also been aligned across services: a main tenet of its 2012 Task Force Report, *The Role of Autonomy in DoD Systems*, is to endorse AWS' *operational* benefits (itemized by domain, scenario and environment) and the 'capability surprise' that autonomy offers.[46] It has, notes Cordingley, 'been a perfect storm of innovation enhancing multiple capabilities that individually and cumulatively are key to battlefield autonomy'.[47]

In order to understand the enduring impact of AWS' drivers, it is also necessary to evaluate Collingridge's assessment that 'weapon automation is here to stay'.[48] Indeed, the inference is that weapon development (of whatever type) is unstoppable[49] and it is this broad range of battlefield possibilities promised by technology at the time of writing which reinforces the impact of weapon automation.[50] Acknowledging its widespread implementation, Scharre highlights that 'the same intelligence that allows self-driving cars to avoid pedestrians could allow future weapons that hunt and attack targets on their own'.[51] Just as Thanmoor notes that more than seventy States are currently involved in developing autonomous weaponry, some three thousand companies have demonstrably invested in the UAV procurement chain.[52] By 2017, the Las Vegas Drone Show already had more than 950 trade stands.[53] This, however, creates complications as a driver. Regulation of autonomous technology is generally perfunctory as it oversees, for instance, research efforts that are simultaneously adopting technologies for self-driving vehicles, for advanced nursing and eldercare and other verticals where machines are already taking actions with potentially lethal consequences.[54]

---

[44] Army Times, 'Gearing Up for Robot Wars', generally.

[45] Sam Jones, 'AI and Robots Line Up for Battlefield Service', *Financial* Times, 16 November 2016, para. 6 of 19 <https://www.ft.com/content/02d4d586-78e9-11e6-97ae-647294649b28> [accessed 10 January 2018]. See also: Albena Antonova, 'Institutional and Organisational Transformation in the Robotic Era', *IGI Global* , (August 2018), p. 3 ('*Introduction to Digital Transformations in Era 4.0*').

[46] US Department of Defense, 'The Role of Autonomy in DoD Systems', *US DoD Task Force Report*, 20301-3140, Section 2.0 (April 2012) <https://fas.org/irp/agency/dod/dsb/autonomy.pdf>. See also: US Department of Defence, 'DARPA and Army Select Contractors for future Combat Systems Programs', *OASD Public Affairs*, News Release 236-00, Washington, 9 May 2017.

[47] Major-General Patrick Cordingley, Commander, 7th Armoured Brigade, Gulf War, 1991, in conversation with the author, January 2019.

[48] David Collingridge, *The Social Control of Technology, (USA:* Francis Pinter, 1980), generally.

[49] David Majumdar, 'Introducing the 5 Deadliest Weapons of the US Military' *The National Interest*, (27 December 2018), generally <https://nationalinterest.org/blog/buzz/introducing-5-deadliest-weapons-us-military-39907> [accessed 12 January 2019].

[50] Jon Wallace, 'SciFi Eye: The Disturbing Future of Autonomous Weapons', *The Engineer*, 19 September 2017, para. 1 <https://www.theengineer.co.uk/autonomous-weapon-systems/> [accessed 23 February 2018]. Also: Boulanin and Verbruggen, pp. 85-89 ('*Relevant Innovations*').

[51] Paul Scharre, 'Why We Must Not Build Automated Weapons of War', *Time Magazine*, 25 September 2017, para. 2 <http://time.com/4948633/robots-artificial-intelligence-war/> [accessed 24 *February 2018].*

[52] Ishaan Thanroor, *'Should The World Kill Killer Robots Before It's Too Late?', Washington Post*, 12 May 2014 <http://www.washingtonpost.com/blogs/worldviews/wp/2014/05/12/should-the-world-kill-killer-robots-before-its-too-late/> [accessed 12 March 2018].

[53] See: Interdrone, <https://www.interdrone.com> [accessed 12 November 2017].

[54] Anderson and Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a ban won't work and How the Laws of War Can,* (USA: National Security and Law Essays, Hoover Institute, Stanford University, 2013), p. 2.

A further driver is that autonomy's fundamental technologies can be procured from several well-diversified sources.[55] While accessibility is just one factor, a finding from Seba's disruption model is nevertheless that rapid scaling-up can lead to a tipping point in a technology's general adoption[56] and that this feature in itself soon becomes a driver to that technology regardless of its underlying industry.[57] In the case of AWS, it may therefore represent a point in time when it becomes foolhardy for defence planning to ignore the technologies' potential.[58] There is usually a human angle to this driver. As noted by Norton, an unexpected consequence to weapon deployment (here, the use of AWS) will likely be a ramp-up in weapon-specific personnel responsible for realizing and then implementing these technologies.[59] It is such on-going human engagement which Hammond notes has the simple behavioural effect of further embedding those systems' adoption.[60] Norton estimates that in 2016 it took seventeen people just to fly an unmanned aircraft.[61] While these numbers are only indicative and will certainly change as technology evolves, human tasking here includes piloting, weapon control and calibration. In this case, the overall number of people involved in supporting that UAV unit may number some two hundred professionals coordinating planning and maintenance, launch and recovery, surveillance and, crucially, extensive ISP and PED efforts (Processing, Exploitation and Dissemination) arising from the weapon's data collection and other relevant activities.[62]

It is therefore the AWS' 'assemblage' of possible capabilities that becomes a central procurement driver.[63] Pairing battlefield priorities with pieces of technology that individually promise to solve those imperatives combine, in theory, to make out-of-the-loop weaponry an attractive proposition.[64] Unsurprisingly, this provides the narrative to commercial parties

[55] A Google search on 'DIY drone kit' returns 677,000 sites; see also: <http://www.buildyourowndrone.co.uk/> [accessed 17 November 2017].

[56] For analysis of Seba's concept, see footnotes 16-18 of this chapter.

[57] Peter Kestner, 'Encouraging Autonomy', *KPMG website*, (30 November 2017) <https://home.kpmg.com/xx/en/home/insights/2017/11/encouraging-autonomy.html> [accessed 12 November 2017].

[58] European Commission, 'Reflection Paper on the Future of European Defence', *Europe*: *Com 20170 315*, (7 June 2017), p. 10 <https://ec.europa.eu/commission/sites/beta-political/files/reflection-paper-defence_en.pdf>.

[59] Travis Norton, 'Staffing for Unmanned Aircraft Systems (UAS) Operations', *Institute for Defense Analyses (IDA)*, (June 2016), pp. 91-97 and generally <https://www.ida.org/idamedia/Corporate/Files/Publications/IDA_Documents/SFRD/2016/P-5253.ashx> [accessed 12 February 2019].

[60] Daniel Hammond, 'Autonomous Weapons and the Problem of State Accountability', *Chicago Journal of International Law*, Volume 15, Number 2, Article 8, (Winter 2015), 662-663 <https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1085&context=cjil> [accessed 5 May 2017]. Notwithstanding its origin in healthcare technology, see also: Oliver Mytton and others, 'Introducing New Technology Safely', *QSHC*, (2011), pp. 9-13 <https://qualitysafety.bmj.com/content/qhc/19/Suppl_2/i9.full.pdf>.

[61] Travis Norton, 'Staffing for Unmanned Aerial Systems (UAS)', pp. 43, 89 and 91. See also: David Hambling, *Swarm Troopers: How small drones will conquer the World,* (USA: Archangel Ink, 2015), p. 34.

[62] Hambling, p. 42.

[63] Jai Galliott and Mianna Lots, *Super Soldier: ethical, legal and social implications,* (USA: Routledge, Political Science, 3 March 2016), p. 14 and p. 17. See also: The US Patriot and Phalanx anti-missile system, the Israeli Iron Dome anti-missile system and South Korea's border denial system (see: *The Telegraph*, 13 July 2010 <https://www.telegraph.co.uk/news/worldnews/asia/southkorea/7887217/South-Korea-deploys-robot-capable-of-killing-intruders-along-border-with-North.html>) [accessed 8 August 2018].

[64] In order to define 'battlefield activities', a useful starting point herewith is still provided by US Army Field Manual, *Intelligence Preparation of the Battlefield*, Section 1 (USA: FM 34-130, July 1994)

advancing such solutions. BAE Systems' 2014 press release for its Taranis UAV accentuates its system's 'autonomous brains'[65], the pitch for their platform and a pre-cursor AWS, suggesting military advantage, cost benefit relative to manned equivalents and the capability of remote yet surgical engagement.[66] Just as it is technical potential which is the potent procurement driver, it is the fusion of such technologies (as suggested in Seba's disruption model) that has the ability both to create a tipping point in adoption as well as an institutional 'fear-of-missing-out'.[67] A further example provides context. From fifteen miles distance, the US Predator UAV can already capture images in which features just four inches across can be distinguished.[68] Subsequent introduction of CCD (Coherent Change Detection) allows that machine to note differences between the current scene under observation and one recorded previously enabling, in theory, identification of disturbances left by an IED on the side of the road.[69] Similarly, the US Predator's multi-spectral targeting system is made possible through the fusion of discrete technical advances:[70] A stabilized gimbal mount with two axes of rotation keeps the weapon's camera pointed in exactly the same direction regardless of the platform's motion. The point is to evidence the cumulative nature of such technologies.[71] The driver in this case is that broad developments across individual componentry can in combination suddenly affect materially how force can be deployed.

Technical innovation also occasions disruptive *refinement* to current deployment models. An example is provided by platform lethality. The US and UK's Reaper UAV can currently carry fourteen Hellfire missiles.[72] Alternatively, it can be deployed carrying just four such missiles but with a pair of laser guided five hundred-pound bombs.[73] In this way, the weapon's payload

<https://fas.org/irp/doddir/army/fm34-130.pdf>. The document covers battlefield environment, effects, threats, management as well as battle execution, space and time as it relates to AWS deployment.

[65] Chris Smith, 'What is Teranis? Everything you need to know about Britain's undetectable drone', *BT website*, (21 November 2017) <http://home.bt.com/tech-gadgets/future-tech/taranis-unmanned-aerial-vehicle-stealth-11364110510493> [accessed 12 March 2018].

[66] *Guia Marie Del Prado,* 'This drone is one of the most secretive weapons in the world', *Business Insider*, 29 September 2015, paras. 2-3 and 7 <http://uk.businessinsider.com/british-taranis-drone-first-autonomous-weapon-2015-9?r=US&IR=T> [accessed 9 September 2017].

[67] Peter Singer and August Cole, 'Humans Can't Escape Killer Robots but Humans can be Held Responsible for Them', *Vice News,* 15 April 2016 <https://news.vice.com/article/killer-robots-autonomous-weapons-systems-and-accountability> [accessed 8 February 2018].

[68] Air Force Technology, 'RQ-1/MQ-1/MQ-9 Reaper UAV', *AFT*, <http://www.airforce-technology.com/projects/predator-uav/> [accessed 4 March 2018].

[69] Mark Preiss and Nicholas Stacy, 'Coherent Change Detection: Theoretical Description and Experimental Results', *Australian Department of Defence*, DSTO-TR-1851, Edinburgh, (2006), p. iii. See also: Hambling, p. 41.

[70] Source: *Army Recognition*, <https://www.armyrecognition.com/us_american_unmanned_aerial_ground_vehicle_uk/mq-1_predator_unmanned_aerial_vehicle_uav_data_sheet_specifications_information_description_uk.html> [accessed 6 February 2017].

[71] In order to deliver its service, the Predator must then deploy multiple cameras featuring various levels of zoom from a 45° wide-angle view down to an ultra-narrow 0.2° tunnel view. On a standard 35mm camera, the equivalent lenses at these extreme ends would be a 55mm wide-angle lens and a 12,000mm telephoto. See: Hambling, p. 45.

[72] UK Royal Air Force blog, <https://www.raf.mod.uk/aircraft/mq-9a-reaper/> [accessed 6 February 2017].

[73] Source: <https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104470/mq-9-reaper/> [accessed 6 February 2017].

flexibility similarly becomes a driver to that platform's adoption and subsequent enhancement.[74] Such drivers, moreover, will be broadly similar whether AWS' tasking is primarily as a defence system (responding either automatically and/or autonomously to incoming munitions), as an anti-personnel system tackling incursions into a defined area or as a weapon tasked with offensive seek-and-destroy. The issue, again, is that the difference between 'a machine that can do these things and make its own attack decisions' is increasingly really only a matter of programming[75] and that improvement in one such capability leads to procurement expectations across other capabilities.[76]

3.2 Technology creep and dual-use technology trends

A further driver is identified by Kaspersen whereby machine autonomy is not, of course, reserved for military applications. Formerly head of Geopolitics and International Security at the World Economic Forum, Kaspersen points out that there is already significant blurring between commercial and military use of autonomous technologies and the definitions that demarcate their deployment.[77] This creates a dichotomy. While such incoherence certainly complicates arguments around banning self-directed weaponry, the role of commercial interests is pivotal in pushing technical solutions that may then be taken up in military applications. The context is that the same autonomous capabilities found in a remotely operating search-and-rescue vehicle may in due course underpin a lethal autonomous weapon.[78] Autonomy, after all, is demonstrably widespread in multiple commercial applications.[79] Civilian aviation provides a relevant example. The Airbus A320 can already take off and land itself.[80] A 2016 study by the *Humans and Autonomy Laboratory* reports that its pilots spend about three minutes per flight with their hands on the cockpit controls.[81] Similarly, Boeing 777 pilots report that they are in control for less than ten minutes in each flight.[82] System autonomy already permeates most commercial verticals. The US credit, debit and prepaid card industry is a case in point: It monitors, notes Nilson, innumerable transactions per second with autonomous tools that identify fraudulent transactions within milliseconds and can be undertaken on a customer base of more than several billion cards.[83] By way of context, IBM's *Watson for Oncology* system provides cancer doctors with recommended courses of treatment

[74] Bas Vergouw and others, '*Dr*one technology: Types, Payloads, Applications, Frequency Spectrum Issues and Future Developments*', Information Technology and Law Series*, 27, Springer, Berlin, ed. B Custers, '*The Future of Drone Use*', p. 22.

[75] T Adams*,* p. 4.

[76] For discussion on Sabin's 'revolution in expectations' see: Chapter 6 (*Wetware*), specifically: 6.1 ('*Software versus intelligence'*).

[77] Anja Kaspersen, 'Is Technology Blurring the Lines between War and Peace?*', World Economic Forum*, (12 February 2016), paras. 5 and 6 <https://www.weforum.org/agenda/2016/02/is-technology-blurring-the-lines-between-war-and-peace/> [accessed 5 February 2018].

[78] Alex Brokaw, 'Autonomous Search and Rescue drones Outperform Humans at Navigating Forest Trails'*, The Verge*, 11 February 2016 <https://www.theverge.com/2016/2/11/10965414/autonomous-drones-deep-learning-navigation-mapping> [accessed 2 January 2018].

[79] See: Zielinska Teresa, *History of Service Robots,* (USA: IGI Global Publishing, 2014) <http://www.irma-international.org/viewtitle/84885/> [accessed 12 March 2017].

[80] Source: <https://www.flightdeckfriend.com/can-a-plane-land-automatically> [accessed 3 March 2017].

[81] Drone Mag, April/May edition 2016, *Drone Mag*, p. 62 <www.drone360mag.com> [accessed 12 March 2017].

[82] Ibid., paras, 62-65.

[83] The Nilson Report, 'US general purpose cards - midyear 2015'*, NR,* Issue 1069, (August 2015), generally.

within seconds from its unstructured universe of pages of medical documents[84] and Google's machine-learning password classifiers can already authenticate users via multiple signals resulting in a ninety-nine percent reduction in Google accounts being compromised by spammers and other fraud attacks.[85] The reach of autonomous agents has clearly been disruptive within several industries[86] and, notes Worcester, it should be expected that a similarly fundamental shift will now take place in battlefield practices.[87]

More relevant to AWS deployment is that autonomy already features in several civilian uses of drone technology.[88] Non-military practices involving autonomous technology include State police, border security and parties involved in agriculture, maritime and forestry.[89] By 2014, the worldwide UAV market was already reckoned to be worth $7bn.[90] Civilian markets investing in the technology include coastguard, oil and gas, electricity grids, climate modeling and the communications industry.[91] Zenko and Kreps, furthermore, note of autonomy that 'commercial drone applications advertised by companies such as Amazon give the illusion of a technology that is ubiquitous and inevitable' adding, in the case of AWS, to the concept of threat inflation and 'a revolution in expectation' discussed in previous sections.[92] The Association for Unmanned Vehicle Systems International (AUVSI) reckons at the time of writing that some one thousand different model types are already in production.[93] Thus, development is taking place in all areas of the AUV supply chain and it is the pace of development that then encourages self-fulfilling expansion of machine capabilities from autonomy to weapons autonomy.[94] In their consideration of 'unmanned

[84] US Department of Defense, Defense Science Board, p. 8.

[85] Iain Thompson, 'AI Slurps, Learns Millions of Passwords to Work Out Which Ones You May Use Next', *The Register*, 20 September 2017 <https://www.theregister.co.uk/2017/09/20/researchers_train_ai_bots_to_crack_passwords/> [accessed 26 January 2018].

[86] EY blog, <https://www.ey.com/Publication/vwLUAssets/EY-the-upside-of-disruption/%24FILE/EY-the-upside-of-disruption.pdf>.

[87] Maxim Worcester, 'Autonomous Warfare: A Revolution in Military Affairs', *ISPSW Strategy Series: Focus on Defence and International Security*, Issue 340, April 2015, pp. 2-3 <https://www.files.ethz.ch/isn/190160/340_Worcester.pdf>.

[88] Examples abound. Drone technology is being used to provide promotional video for real estate agencies and field inspection for farmers. As above, cost is the principal driver given such UAV services can be delivered at a fraction of the cost of a helicopter. For sources, see: <http://phys.org/news/2013-12-commercial-drones-flight.html> and <http://www.globalresearch.ca/drones-from-military-use-to-civilian-use-towards-the-remote-uav-policing-of-civil-society/30876>, [both accessed 22 March 2017]. See also: Walker, pp. 65-68. Also: Tom Simonite, 'Sorry, banning 'Killer Robots' just isn't practical', *Wired*, 22 August 2018 <https://www.wired.com/story/sorry-banning-killer-robots-just-isnt-practical/> [accessed 2 January 2018].

[89] Here.com blog, 'Enabling an Autonomous World for Everyone', undated <https://www.here.com/en/vision/autonomous-world?cid=Auto-Google-MM-T2-Here-generic-BMM&utm_source=Google&utm_medium=ppc&utm_campaign=Auto_PaidSearch_Automotive_AlwaysOn> [accessed 17 July 2018].

[90] Taylor Vinters LLP, 'Qi3 Insight: Unmanned Aerial Vehicles', *Qi3 Ltd*, (February 2014), p. 5 ('*Figures*').

[91] For example: the VHALE platform as proxy satellites and MALE platform for short-term and local communications coverage.

[92] Micah Zenko and Sarah Kreps, 'Limiting Armed Drone Proliferation', p. 14. See also: Amazon blog, <https://www.amazon.com/Amazon-Prime-Air/b?ie=UTF8&node=8037720011> [accessed 24 April 2018]. The projected service is also detailed by Ed Oswald, 'Everything You Need to Know about Amazon's Drone Delivery Project, Amazon Prime Air', *Digital Trends*, 3 May 2017 <https://www.digitaltrends.com/cool-tech/amazon-prime-air-delivery-drones-history-progress/> [accessed 24 April 2018].

[93] Source: Auvsi <http://www.auvsi.org/home> [accessed 2 May 2018].

[94] Boulanin and Verbruggen, pp. 36-41. See also: Christopher Harress, 'The Rise of China's Drone Fleet and Why It May Lead to Increased Tension in Asia', *International Business Times*, 1 November 2014 <http://www.ibtimes.com/rise-

systems and the future of war', this timeline is explicitly recognised by the relevant US Representatives Committee on Oversight and Government Reform.[95] Procurement for machine autonomy is therefore already deeply intertwined and, as pointed out by the Heyns, UN's Special Rapporteur, while 'bright lines are difficult to find, lethal automated weapons have a composite nature and are combinations of underlying technologies with multiple purposes'. [96]

This posits two further features that are relevant to a discussion on AWS' drivers. Boulanin and Verbruggen highlight the phenomenon of incremental 'technology creep' (here, the rapid iterations of individual technologies but also rapid combination of developing technologies in a manner that may move their role away from their original or first intended specification). This, they note, empirically occurs before that technology 'watersheds' and when, in this case, AWS have already appeared in States' arsenals.[97] A second feature is that already adopted military hardware may exhibit tenets that would otherwise characterise autonomous systems.[98] Mine munitions exhibit similar incrementalism whereby autonomous capabilities have long enabled target engagement with no human-in-the-loop.[99] Wareham notes that, while anti-personnel landmines are prohibited by the Mine Ban Treaty, concerns remain over vehicle mines with anti-handling devices and sensitive fuses.[100] The challenge is that incrementalism may camouflage movement along that autonomy continuum; a seemingly innocuous notch up (perhaps a small mobile robot newly programmed to respond autonomously to a weapon signature to protect peacekeepers from local ensnarement) might actually represent a material discontinuity in battlecraft.

3.3 Structural and procurement drivers

While technical progress is a driver in AWS deployment, it is certainly not (to paraphrase Smith's 2009 work on the interaction of politics, resources and military procurement) the only one.[101]

---

chinas-drone-fleet-why-it-may-lead-increased-tension-asia-1535718> [accessed 20 March 2016]. See also: Michael Hoffman, 'China Reports Stealth Drone's First Test Flight', *DefenseTech*, 22 November 2013, generally <http://defensetech.org/2013/11/22/china-reports-stealth-drones-first-test-flight/>) [accessed 20 March 2016].

95 Hearing on 'The Rise of Drones; Unmanned Systems and the Future of W*ar*', *Committee on Oversight and Government Reform*, Congressional Research Service, (March 2010) <http://digitalcommons.wcl.american.edu/cgi/viewcontent.cgi?article=1002&context=pub_disc_cong> [accessed 15 June 2017] generally.

96 Christof Heyns, 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions', p. 9. See also: Markus Christen and others, 'An evaluation schema for the ethical use of autonomous robotic systems in security applications', *UZH Digital Society Initiative,* University of Zurich, White Paper 1, (2017), pp. 31-36 <https://philarchive.org/archive/CHRAES-3> [accessed 12 November 2017].

97 Boulanin and Verbruggen, p. 19. The dataset is discussed in detail in the next chapter and focusses on weapon systems rather than individual munitions.

98 Walker, *Killer Robots?*, pp. 30-31.

99 NGO Article 36 is the leading source of expertise on this area; see, for instance: Article 36 <http://www.article36.org/weapons/landmines/anti-vehicle-mines-victim-activation-and-automated-weapons/> [accessed 2 May 2016]. See also: David Larter, 'Autonomous mine-hunting boat will be delivered to British Navy this winter', *Defense News*, 13 September 2017 <https://www.defensenews.com/digital-show-dailies/dsei/2017/09/13/autonomous-mine-hunting-boat-will-be-delivered-to-the-royal-navy-this-winter/> [accessed 23 October 2017].

100 Mary Wareham, Director, Human Rights Watch Arms Division, in conversation with the author, December 2018.

101 Ron Smith, *Military Economics: the interaction of power and money,* (UK: Palgrave MacMillan, Basingstoke, 2009), p. 132 <https://books.google.co.uk/books?hl=en&lr=&id=Rd0YDAAAQBAJ&oi=fnd&pg=PP1&dq=r+smith+military+economics&ots=_uONTfBz2_&sig=_5GjeOYSLLg4T8snaahF332GAfY#v=onepage&q&f=false> [accessed 29 March 2017].

There are, then, several pathways along which to frame acceleration towards the adoption of autonomous weapons. Kostopoulos, advisor to the *AI Initiative* at Harvard's Future Society, usefully isolates certain high-level behavioural drivers. Necessary conditions precedent to AWS deployment should, she argues, include a broad 'trust' in the weapons' underlying technologies, a general 'cultural acceptance' in those technologies (by organisational, institutional and societal parties) and, finally, reasonable 'availability' of (and hence familiarity with) those underlying technologies.[102] Re-purposing her analysis to AWS deployment, 'trust' here also encompasses the weapon's fitness for purpose as well as the 'authenticity' of its machine learning environment such that operational parameters can properly be isolated and processed. Finally to this point, Kostopoulis notes that 'confidence' as a driver should incorporate a reduction in burden for both soldier and commander, reliable and value-adding knowledge transfer between colleague machines and, finally, successful training and testing of the weapon platforms in both sympathetic and non-sympathetic environments.[103]

Institutional drivers also come in different guises. To this point, it is relevant to consider military influencers in weapons autonomy. In 1958 President Eisenhower created the Advanced Research Projects Agency with a purpose to formulate and execute research and development projects that would expand the frontiers of technology and science.[104] That role had been cemented as a response to the Soviet launching of Sputnik two years before. DARPA's mission remains ensuring that US military technology retains an edge over its potential enemies and to ensure ongoing decisive military advantage. As an accelerator, DARPA's current programmes therefore provide useful insight into America's military research priorities (as well, if course, as its potential weaknesses) and, in identifying drivers to AWS, give context to how such platforms may be developed. A recent theme to the agency's work is its focus on adaptive systems, autonomous decision aids and battlefield processes that are increasingly independent of human intervention.[105] This is reflected in US weapon procurement. By 2016, US DoD was spending $3 billion on unmanned aircraft comprising, by number, forty percent of all US aircraft.[106] To this point, Boulanin notes that the underlying deployment models (involving degrees of unmanned and increasingly unsupervised UAV) have similarly proliferated.[107] This is borne out in the commercial sphere: When the American Federal Aviation Administration introduced its national framework for registering unmanned aircraft in late 2015, more than one hundred and eighty thousand drones where were registered in the first two weeks of that scheme. As evidenced in Chapter Five's discussion on proliferation, the USA is not alone in this phenomenon; as of 2018, seventy countries

[102] Lydia Kostopoulos, 'Drivers for the Deployment of Lethal Autonomous Weapons', *Medium.com*, 22 December 2017, paras. 5-9 <https://medium.com/@lkcyber/drivers-for-the-deployment-of-lethal-autonomous-weapons-systems-ae1dd6278a35> [accessed 9 March 2017].

[103] Kostopoulos, para. 10.

[104] Subsequently abbreviated to DARPA. See: <https://www.darpa.mil/about-us/darpa-history-and-timeline> [accessed 3 April 2017].

[105] DARPA website, <https://www.darpa.mil> [accessed 18 July 2018].

[106] Dan Gettinger, 'Drones in the Defense Budget', *Center for the Study of the Drone*, Bard College, October 2017, generally <http://dronecenter.bard.edu/files/2018/01/Drones-Defense-Budget-2018-Web.pdf>. Also: J Gertler, 'US Unmanned Aerial Systems', *Congressional Research Service,* CRS R42136, (3 January, 2012); and J Gertler, 'How many UAVs for DoD?', *Congressional Research Service,* CRS IN10317, (2015), generally.

[107] Boulanin and Verbruggen, p. 19. See also: Chapter 4 (*Deployment*).

operate unmanned aircraft with thirty armed UAV programmes established or in development.[108] It is against these developing set of accelerators that AWS' position in Seba's discontinuity model should be considered.

Several systemic trends comprise an institutional driver to AWS deployment. The first is run-away budgets. Miller highlights escalating *human* resource costs of having 'boots on the ground' (or, more importantly, in the air) in a general reference to manpower expenditure.[109] Between FY 2001 and FY 2012, US compensation costs per active-duty service member grew nearly sixty percent.[110] Adjusted for inflation, this equates to an annual (and unsustainable) growth of more than four percent.[111] Manned systems, moreover, are costly in *any* military environment: The human fighter must be able to breathe, eat and take care of bodily functions. His safety must be addressed with armour, redundant control apparatus and escape systems.[112] In the case of ground troops, transportation must be available to and from a contact zone as well as being available to support him while he is there. Each such subordinate system must then be supported by its own complex logistics chain. In addition, medical supplies, facilities and staff must be immediately available to evacuate and treat the injured. The human component of these chains requires extensive management, training as well as generous benefits programmes for life.[113] By comparison, the cost of robotic systems is collapsing due to increased systems commonality and production innovation.[114] A 2015 study undertaken by Duke University[115] concluded that, even by that date, the fully loaded manpower cost per hour of a UAV was ninety percent less (one hundred and fifty dollars compared to two thousand dollars) than a manned aircraft.[116] This, notes Cummings, has created an unsustainable cost differential that creates its own driver towards broad adoption of

[108] K Sayler, 'A world of proliferated drones: a technology primer', *Centre for a New American Security*, (July 2015) <www.cnas.org/world-of-proliferated-drones-technology-primer> [accessed 24 August 2016]. See also: Chapter 5 (*Obstacles*), specifically *5.7:* ('*Proliferation Constraints*').

[109] US Department of Defense, 'Cost-saving Pilot Programs to Support War-fighter Autonomy', *US DoD News feed*, <http://www.defense.gov/news/newsarticle.aspx?id=120329> [accessed 11 May 2016]. See also: Jack Miller, 'Strategic Significance of Drone Operation for Warfare', *E-International Relations Students*, (18 August 2013) <http://www.e-ir.info/2013/08/19/strategic-significance-of-drone-operations-for-warfare/> [accessed 2 October 2016].

[110] Robert Work and Shawn Brimley, '20YY; Preparing for War in the Robotic Age', *Centre for a New American Security*, (January 2014), p. 20 <https://www.cnas.org/publications/reports/20yy-preparing-for-war-in-the-robotic-age>.

[111] Over the same time period, the share of the base DoD budget for military personnel-related costs rose from 30% to 34% and is expected to consume 46% of the budget by FY 2021 even with a 2.6% historically normal real annual growth.

[112] Simon Ramo, *Let Robots Do The Dying,* (Kindle publication, 2011), 8%.

[113] The National Audit Office's analysis provides useful context here in demonstrating the current costs of manpower assets. 29% of the UK's 2017 £35.3bn budget was accounted for by service personnel, 4% by civilian contractors and 4% by administration. Pensions to service personnel comprised 2% of the UK's defence budget in 2017. See: National Audit Office, 'A Short Guide to the Ministry of Defence', *NAO*, (2017), p. 9 <https://www.nao.org.uk/wp-content/uploads/2017/09/A-short-guide-to-the-Ministry-of-Defence.pdf>.

[114] PricewaterhouseCoopers, 'The New Hire: How a new generation of robots is transforming manufacturing', *Zpryme Research survey*, (February 2014), p. 12 <https://www.pwc.fi/fi/palvelut/tiedostot/industrial-robot-trends-in-manufacturing-report.pdf>.

[115] Professor Missy Cummings, Director, Humans & Autonomy Laboratory, Duke University, in conversation with the author (Chatham House conference; Autonomous Military Technologies, February 2014).

[116] Ian Jaffe, 'Former Fighter Pilot, Duke Prof Missy Cummings talks drones', *Duke Chronicle*, (15 September 2015) <http://www.dukechronicle.com/article/2015/09/former-fighter-pilot-duke-prof-missy-cummings-talks-drones> (accessed 17 January 2017]. Note, however, that sensor multiplicity on UAV is likely to dull this advantage if further analysts are then required to process additional data.

robotics and unmanned solutions across battlefield assets.[117] Fully costed, Hambling calculates that parties will spend more than thirty thousand dollars per hour to fly their F-35 while the Reaper costs less than four thousand dollars to operate for each hour of use.[118]

The differential that is evident in the UAV's running costs also characterises its relative capital costs.[119] In this case, therefore, spiraling *hardware* costs create their own institutional driver in the face of squeezed procurement budgets.[120] Here, Wolf's work on defence inflation points to Augustine's tongue-in-cheek conclusion made in 1984 that 'in the year 2054, the entire [US] defence budget will purchase just one aircraft'.[121] Weapon autonomy, note Work and Brimley, posits an answer to escalating costs of military hardware platforms.[122] The *Centre for a New American Security* suggests that accelerating hardware costs[123] already mean that US armed forces are no longer able to replace front-line combat systems on a one-for-one basis.[124] The example of the F-35's procurement provides an appropriate case study. Thus, notes Enemark, 'when an aircraft has a pilot on board, there is a need to accommodate and protect frail human flesh in the engineering, construction and use of that aircraft'.[125] As of 2014, eight years into production, the F-35 cost some $190 million in that year's dollars.[126] This compares to the calculation in 2012 by Congress' watchdog agency that the average price for an F-35 aircraft had already doubled from $69 million since the programme's inception.[127] Hambling gauges the total cost of ownership, including maintaining and supporting the F-35 over its lifetime, to be more than $300 million per unit.[128] This masks a further driver to AWS deployment where the imperative to shrink costs is

---

[117] For Russian military cost inflation, see: *Global Security* <http://www.globalsecurity.org/military/world/russia/mo-budget.htm>. For UK cost inflation, see: RUSI, <https://www.rusi.org/downloads/assets/Comment_Defence_Inflation_Myth_or_Reality.pdf>. For Chinese military cost inflation, see: Aviation Week <http://aviationweek.com/awin/china-s-inflation-adjusted-defense-budget-75> [all accessed 18 February 2017].

[118] Hambling, p. 66.

[119] Deborah Heynes, 'Spiralling cost of weapons makes war 'too expensive', *The Times*, 26 April 2017 <https://www.thetimes.co.uk/article/spiralling-cost-of-weapons-makes-war-too-expensive-6fkzf03w6> [accessed 23 February 2017].

[120] Walker, *Killer Robots?,* pp. 28-30.

[121] Katharina Wolf, 'Putting Number on Capabilities: Defence Inflation versus Cost Escalation', *European Institute for Security Studies*, Brief Issue, 27, (July 2015), p. 1 and generally <https://www.iss.europa.eu/sites/default/files/EUISSFiles/Brief_27_Defence_inflation.pdf>.

[122] Work and Brimley, p. 22.

[123] *Centre for a New American Security*, cit. Work and Brimley, p. 22.

[124] To survive against steadily improving guided munitions all such crewed platforms will require costly stealth technology, stand-off ability and highly capable active and passive defences. In the case of the F-35 aircraft, see: <https://www.washingtonpost.com/sf/brand-connect/the-f-35-how-it-works/> [accessed 4 September 2018].

[125] Christian Enemark, 'Armed drones and the Ethics of War: Military Virtue in a post-heroic age', (UK: Routledge, Oxford, 2014), p. 98.

[126] A useful summary is provided by: Amanda Macias, 'The Pentagon is Trying to Figure Out the True Cost of its Costliest Weapon System, the F-35', *CNBC*, (28 February 2018) <https://www.cnbc.com/2018/02/28/pentagon-wants-to-know-true-cost-of-f-35-system.html> [accessed 2 September 2018]. An unmanned UAV, for instance, requires no cockpit pressurisation or temperature control and may have more space and payload capacity for fuel allowing it to stay in the air for longer (long-dwell, high altitude capabilities).

[127] Andrea Shalal-Esa, 'Insight: Expensive F-35 fighter at risk of budget 'death spiral'', *Reuters Newswire*, 15 March 2013 <http://uk.reuters.com/article/2013/03/15/us-usa-fighter-f35-insight-idUSBRE92E10R20130315> [accessed 18 April 2014].

[128] Hambling, p. 94.

reinforced by a requirement to reduce complexity across the purchasing process. The F-35's first development contract was, after all, signed in 1996; while the first unit flew in 2006, the platform did not finally arrive into service until 2015, nearly twenty years after initial consent.[129] Finally to this point, the complexity and cost of such manned platforms reduce the physical number of assets that a State can afford. As platforms become fewer and dearer, a simple numbers game of how many aircraft you can get into the sky itself becomes a driver to the deployment of cheaper and more numerous weapon systems.[130] While these drivers may not be new (Clark points to similar discussion around mechanisation and aerial warfare in the 1930s[131]), the research of Kirkpatrick and Pugh notes meaningful recent escalation in weapon unit costs as well as accelerating in-programme inflation (actual delivery costs of a weapon programme materially outstripping its initial budget).[132] Counter-point, discussed in Chapter Two, is provided by those 1,218 drones taking part in the 2018 Pyeongchang Winter Olympics Opening Ceremony, each of which likely cost less than one thousand dollars.[133]

The operation of manned weapons also acts as a driver to AWS deployment. Wheeler suggests that protecting the high value F-35 pilot through system duplication and physical shielding may account for up to sixty percent of that platform's capital cost.[134] Tellingly, this redundancy and work-round may also 'degrade [operational] performance by up to eighty percent'.[135] Unlike in its unmanned, autonomous alternative, the combat aircraft's pilot remains the most critical (and most vulnerable) component of the F-35 ecosystem.[136] In addition to creating procurement and design constraints, the pilot is a priority for adversarial targeting, becoming an increasing burden on that platform's performance in terms of additional armour, life support and limitations imposed on the manned unit by g-forces.[137] There is, notes Adams, ample incentive to exclude humans from the

[129] This has certain procurement ramifications. Only two consortia were therefore of sufficient size to compete for the contract to construct the high value aircraft. Cost and commercial risks empirically limit business competition (and any subsequent chance for market forces to reduce a programme's build cost) as well restricting the benefits of possible collaboration. See: Hambling, p. 94

[130] Matthew Alexander, 'Is the UK able to respond to the technological changes of warfare?', unpublished thesis, *University of Bath*, (15 April 2015) <http://www.bath.ac.uk/research/case-studies/centre-for-war-and-technology> [accessed 4 February 2018].

[131] Professor Lloyd Clark, in conversation with the author, January 2019.

[132] D Kirkpatrick and P Pugh, 'Towards the Starship Enterprise: Are the Current Trends in Defence Unit Costs Inexorable?', *Journal of Cost Analyses*, 2, 1 (1985), 59-60 and generally.

[133] B Barrett, 'Inside the Olympics Opening Ceremony World-Record Drone Show', *Wired.com*, 9 February 2018 <https://www.wired.com/story/olympics-opening-ceremony-drone-show/> [accessed 12 March 2018]. Also: dronesforless <http://www.dronesforless.co.uk/drones/autel-robotics-x-star-premium-quadcopter-with-3-axis-gimbal-pro-bundle/?utm_source=googleshoppingpaid&utm_medium=cpc&utm_campaign=googleshoppingpaid&utm_content=SHP2366C5&gclid=EAIaIQobChMIioqllpD72QIVkZa9Ch0gjgfxEAkYAiABEgKEHvD_BwE> [accessed 12 February 2018].

[134] For an unauthorised view of the F-35's procurement costs, see: Winslow Wheeler, 'How Much Does an F-35 Actually Cost?', *www.warisboring.com*, (27 July 2014) <https://warisboring.com/how-much-does-an-f-35-actually-cost/> [accessed 9 March 2018].

[135] Micha Zenko, *The Coming Future of Autonomous Drones,* (Council on Foreign Relations, 4 September 2012) <http://blogs.cfr.org/zenko/2012/09/04/the-coming-future-of-autonomous-drones/> [accessed 17 March 2016]. Zenko also discusses use of UAVs tasked with actions that remain short of warfare but which are nevertheless designed to secure military advantage.

[136] As much as 45% of the costs of the A-4 Skyhawk is accounted by the crew protection, system redundancy and other outlay on shielding which would be avoided if the platform was unmanned; see: US Military blog, <http://usmilitary.about.com/od/attack/> [accessed 12 January 2015].

[137] Adams, p. 7 and generally.

system.[138] Lovelace defines this driver as a process of 'unmanning the weapon platform'.[139] While this may have been a long-held aspiration in military procurement, technical innovation (and now this cumulative complexity of manned machines discussed variously by Scharre and others) that agitates for automation and autonomy across fighting systems.[140] Hambling cites Augustine's Law XV[141] to evidence that building in such layers of redundancy increases costs and compromises reliability: 'The last ten percent of performance generates one third of the cost and two-thirds of the problems'.[142] Another risk to such expensive platforms is, of course, that further technical shifts may then erode their relevance.[143] In this vein, Madrigal notes that adoption of low-cost, agile and swarm-based UAV is, at the very least, likely to make the guarantee of airspace superiority more difficult to achieve.[144]

3.4 Ethical drivers

Ethical factors may act as a driver to removing man's role in lethal engagements. As noted by Peralta, their often intangible nature (in particular the imprecision with which they relate to LOAC and their entanglement with the specific context for each such engagement) means that this section's evaluation of ethics can appear frustratingly inexact and often based on heuristics rather than worked-through evidence.[145] Nevertheless, its theoretical notion is that machine-based safeguards can be built into autonomous weapon systems to ensure compliance with international humanitarian law (IHL).[146] Here, roboticist and roboethicist Arkin, Professor of Interactive Computing at the Georgia Institute of Technology, promotes the general viability of what is termed

[138] For a general discussion on the advantages of autonomy in weapon systems, see: Adams, p. 7.

[139] Douglas Lovelace, *Autonomous and Semi-Autonomous Weapon Systems,* in Law Review, Volume 44, (Oxford: Oxford University Press, 6 October 2016), p. 70.

[140] Paul Scharre, 'Autonomous weapons and operational risk', *Centre for a New American Security*, (2016), pp. 25-34 <https://www.files.ethz.ch/isn/196288/CNAS_Autonomous-weapons-operational-risk.pdf>.

[141] David Smallwood, 'Augustine's Law Revisited', *Sound and Vibration*, (March 2012) <http://www.sandv.com/downloads/1203smal.pdf>.

[142] See: Chapter 4 (*Deployment*), specifically: 4.6 ('*Swarming models'*). Also: Hambling, p. 96. Written in 1984 by former Under-Secretary to the US Army Ralph Norman Augustine as a tongue-in-cheek set of business aphorisms, Augustine's Laws XVI states that 'in the year 2054, the entire defense budget will purchase just one aircraft. This aircraft will have to be shared by the Air Force and Navy three-and-one-half days each per week except for leap year, when it will be made available to the Marines for the extra day'.

[143] Kelsey Atherton, 'The future of the Air Force is fighter pilots leading drone swarms into battle'*, Popular Science*, 23 June 2017, paras. 3-4 <https://www.popsci.com/future-air-force-fighters-leading-drone-swarms> [accessed 6 March 2018]. See also: Ilana Freedman, 'F-22 and F-35: America's Costly Boondoggles Are the Victims of Arrogance and Appeasement'*, Gerard Direct*, 10 March 2013, paras. 7-9 <http://gerarddirect.com/2013/03/10/uss-f-35-and-f-22-americas-costly-boondoggles-the-victims-of-arrogance-and-appeasement/> [accessed 8 March 2018].

[144] Alexis Madrigal, 'Drone Swarms are Going to be Terrifying and Hard to Stop', *The Atlantic, Technology*, 7 March 2018 <https://www.theatlantic.com/technology/archive/2018/03/drone-swarms-are-going-to-be-terrifying/555005/> [accessed 19 May 2018].

[145] Eyder Peralta, 'Weighing the good and the bad of autonomous killer robots in battle'*, All Tech Considered*, 28 April 2016, paras. 21-25 of 25 <https://www.npr.org/sections/alltechconsidered/2016/04/28/476055707/weighing-the-good-and-the-bad-of-autonomous-killer-robots-in-battle> [accessed 12 February 2018]. See: Th. A, Van Baarda, *Moral Ambiguities Underlying The Laws of Armed Combat: A Perspective from Military Ethics*, (USA: The Yearbook of International Humanitarian Law, Volume 11, December 2008), p. 3.

[146] Walker, *Killer Robots?* pp. 35-36, 59-60 and 61. The role of IHL and other legal constraints in AWS deployment is discussed in detail in Chapter 5 (*Obstacles*), specifically: 5.1 ('*Geneva Convention and Laws of Armed Conflict'*) and 5.5 ('*Article 36 and LOAC-compliant weapons'*).

an 'Artificial Ethical Override'.[147] He uses, for instance, the *Mental Health Advisory Team* report from the US Surgeon General's office to assert that unmanned combat systems may obviate ethical challenges arising generally in combat conditions.[148] According to that November 2006 study, soldiers' conduct during Operations Iraqi Freedom and Enduring Freedom was often 'questionable' with some ten percent of soldiers reporting that they had mistreated non-combatants or damaged civilian property and only forty-seven percent of soldiers agreeing that non-combatants should be treated with dignity and respect.[149] Battlefield surveys, of course, risk bias and sample challenges, disputes around inappropriate structure and coding, worries around anonymity, independence and statistical rigour as well as difficulties occasioned by cross-culture and contextual issues.[150] Using such data is therefore fraught with difficulty but, for the narrow purposes of this analysis, it serves at least to highlight possible faultlines. The report, after all, found that more than thirty percent of soldiers agreed that torture be allowed in order to save the life of a fellow soldier with forty-five percent of soldiers reporting that they would not report a colleague if he had killed or injured an innocent non-combatant. Less than half of the soldiers said that they would report a team member for unethical behaviour. Empirically, this may not be unexpected given the prominent role of trust in military ethics.[151] Neuroscientists, notes Pryer, have found that human circuits responsible for conscious self-control are highly vulnerable to stress.[152] When these circuits shut down, primal impulses go unchecked. Notwithstanding the difficulty of using such primary sources, evidence nevertheless exists upon which Arkin can base his argument for an ethical driver: Properly crafted algorithms may, in theory, produce a more consistent, more compliant engagement outcome than has been achievable in a fraught combat situation where human supervision has empirically been inadequate.[153]

Accepting for the moment Arkin's use of these surveys, ethical arguments in favour of AWS deployment appear on first reading to have surprising depth. Weapons autonomy may, after all, reduce soldiers' participation on the battlefield. This has several long-recognised advantages. As far

[147] Dieter Vanderelst and Alan Winfield, 'An Architecture for Ethical Robots Inspired by the Simulation Theory of Cognition', *Cognitive Systems Research*, 48, (May 2018), pp. 56-65 <https://reader.elsevier.com/reader/sd/pii/S1389041716302005?token=C7E1EA2947EBE5A58F14432243B19081E625A4E6D56A5CC87D0ED7548B5A93DB91F0B8ADB1DA68FFA268CDC86E1319D1> [accessed 17 January 2019].

[148] Ronald Arkin, *Governing lethal behaviour: Embedding ethics in a hybrid deliberate/reactive robot architecture,* (USA: Atlanta, GA: Georgia Institute of Technology, 2007), generally. See also: Surgeon General's Office, 'Mental Health Advisory Team (MHAT) IV Operation Iraqi Freedom 05-07', *US GSO*, Final Report, (November 2006) <http://www.combatreform.org/MHAT_IV_Report_17NOV06.pdf>. HRW's Mary Wareham notes that Arkin's work dates from 2006 and that he has published little on the subject since that date. 'The attention that Arkin receives is more to do with the counterpoint his research provides to the AWS debate and the theory behind his construct'.

[149] Surgeon General's Office, 'Mental Health Advisory Team', generally.

[150] Statistical issues include sampling bias, under-coverage and social desirability in the answering of the survey's questions, non-response bias, the issue of leading questions in the absence of any control group, difficulties around apportioning causation and the dealing with dependent variables. See: Jacob Metcalf, *Ethics Codes: History, Context and Challenges*, (USA: Council for Big Data, Ethics and Society, 9 November 2014) <https://bdes.datasociety.net/council-output/ethics-codes-history-context-and-challenges/> [accessed 6 May 2017].

[151] Source: 'The Army Ethic White Paper', *Center for the Army Profession and Ethics*, 11 July 2014 <https://www.army.mil/e2/c/downloads/356486.pdf>.

[152] Lt Col Douglas Pryer, 'The rise of the machines: why increasingly 'perfect' weapons help perpetuate our wars and endanger our nation', *Military Review*, March-April 2013, p. 15.

[153] Reverend Sean Wead, 'Ethics, Combat and a Soldier's Decision to Kill', *Military Review*, March-April 2015, pp. 70-72 <https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MilitaryReview_20150430_art013.pdf>.

back as 1946 (with no substantive studies having been undertaken in the meantime), Swank put forward evidence to show that after sixty days of continuous combat ninety-eight percent of surviving soldiers suffer psychiatric trauma.[154] A further thread to this argument is the challenge, notes Hanlon, of creating soldiers that are 'fit for purpose' in the first place.[155] Again parking contextual and statistical concerns (and using contemporaneous analysis to do no more than indicate a theme), Grossman's Second World War study suggests that 'most men [sic] simply did not kill'.[156] While Marshall's US Army study has since been widely discredited (its sample size was just four hundred men and, notes Strohn, the material is unlikely to hold for 'wars of survival rather than recent wars of choice and expedition'[157]), it is nevertheless interesting to note even the questions it posed and, with caveats, its finding that only fifteen percent of those infantrymen 'interviewed' had actually fired at enemy positions on any occasion despite eighty percent of the sample having the opportunity to do so.[158] Parks and Neiss' study of the Korean War indicates that fifty percent of the F-86 pilots never fired their weapons and, ignoring area bombing, only ten percent of those had actually hit a target.[159] Finally to this point, Grossman suggests that less than one percent of Second World War pilots accounted for thirty to forty percent of all downed enemy aircraft.[160] Notwithstanding general challenges that are posed by ethical questionnaires, two inferences arise from these US sources that are then leveraged by advocates of AWS. Regardless of training, a serviceable conclusion is that it is unrealistic to expect human beings to adhere unerringly to LOAC when confronted by challenges of the battlefield. That assumption, notes Behm, is borne out by the debate that continues on ethics education by military training establishments.[161] An argument is that these same soldiers, as a general sample and when compared to a machine agent, may at best be 'a variable tool in the waging of war'.[162] Arkin then uses these findings to argue that a computational implementation of an ethical code (together, his 'Artificial Conscience'[163]), once embedded into AWS' control sequences, may in time provide enforceable

---

[154] No comparable European study exists, hence use here of another US study: R Swank, *Combat neuroses: Development of Combat Exhaustion,* Vol. 55, (USA: Archives of Neurology and Psychology, 1946), pp. 236-47.

[155] Michael Hanlon, ''Super Solders': The Quest for the Ultimate Human Killing Machine', *The Independent,* (17 November 2011), paras. 2-7 of 15 <https://www.independent.co.uk/news/science/super-soldiers-the-quest-for-the-ultimate-human-killing-machine-6263279.html> [accessed 10 January 2019].

[156] Grossman study, 1995, cit. Arkin, 'The case for Autonomy in Unmanned Systems', p. 9.

[157] Dr Matthias Strohn, in conversation with the author, January 2019.

[158] Samuel Marshall, *Men against fire: the problem of battle command in future war,* (USA: New York, William Morrow Publishing, 1947), generally.

[159] Parks and Neiss Study, 1956, cit. Arkin, 'The case for Autonomy in Unmanned Systems', p. 9. The Study is also discussed by Arkin in 'Human Failings on the Battlefield', cit. Braden Allenby, *The Applied Ethics of Emerging Military and Security Technologies*, (USA: Routledge, December 2016), chapter 12.

[160] Dave Grossman, *On Killing: The Psychological Cost of Learning to Kill in War and Society*, (Black Bay Books, 1996), generally. Despite controversy over Marshall's WW2 survey methodology, Grossman too uses Marshall's data to evidence soldiers' reluctance to kill their opponents.

[161] Beth Behn, 'The Stakes are High: Ethics Education at US War Colleges', *Air War College Publications*, Maxwell Paper Number 73, (2018), p. 2 <https://www.airuniversity.af.mil/Portals/10/AUPress/Papers/mp_0073_behn_stakes_high.pdf>. Here, Professor Lloyd Clark (in conversation with the author, January 2019) points to the 'vibrant and long-running debate at West Point, RMAS, the US War College and UK Staff College about the weighting of ethical training'.

[162] Derived from Barbara Ehrenreich, 'Do humans have a role in the robot wars of the future?', *The Guardian Newspaper*, 11 July 2011 <https://www.theguardian.com/commentisfree/2011/jul/11/human-role-robot-war-future> [accessed 2 September 2017].

[163] Arkin, *Governing lethal behaviour,* p. 61.

limits on machine actions in engagements.[164] While Arkin's construct is much weakened by its underweighting battlefield intangibles such as soldier values, leadership and other behavioural variables, the point remains that it provides an important and outwardly plausible driver advocating machine (rather than human) control in lethal engagements.

Questioned in military circles[165], Arkin's structure is based upon several assumptions, not least that lines of code will, in time and without error, be able to undertake the complex, interactive tasks required for machines to assume the capabilities currently undertaken by human soldiers.[166] Reviewing the feasibility of Arkin's construct is therefore a key research question for this thesis that occupies much of its subsequent chapters, in particular challenges around value and goal setting, the weapon's required utility function and whether these features can *dynamically* be managed. The coding complexity around ethical precepts is particularly challenging given that such programming must incorporate precepts from *all* of the human rights conventions in order to be compliant and enable the AWS to consider, in real time, the consequences of its every action. While Arkin acknowledges that it is 'too early to determine whether this software device is practicable'[167], this is not necessarily the point: The issue is whether his construct might *one day* allow better compliance with IHL than is currently capable of being exercised by a human.[168] His argument more represents a prototype and, as such, 'a preliminary version of a device from which other forms may be developed'[169]: Robots, after all, should (at least theoretically) be able to process more information faster than humans, heuristically remain uninfluenced by fear or anger[170] while at the same time monitoring the ethical behaviour of 'colleague' humans on the battlefield.[171] In this sense, Arkin's argument is frustratingly circular to those questioning its legitimacy: His framework is theoretical without having either to develop or react to contextual parameters and, suggests Clark, 'relies much upon what the construct leaves out rather than what is put in'.[172] Instead, it is

[164] The test here is that such limits should be better, or at least as good, as the limits achieved by human soldiers on the battlefield. Arkin's actionable list is comprehensive and includes, inter alia, acceptance of surrender and humane treatment of prisoners, avoiding unnecessary suffering and damage and non-use of certain weapons.

[165] Amitai Etzioni and Oren Etzioni, 'Pros and Cons of Autonomous Weapon Systems', *Military* Review, (May-June 2017), 74-77 <http://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/> [accessed 16 January 2019].

[166] Arkin 'Governing Lethal Behaviour', p. 4. Arkin justifies his framework as follows: 'It is a working assumption, perhaps naive, that the autonomous agent ultimately will be provided with an amount of battlefield information equal or greater than a human soldier is capable of managing. This seems a reasonable assumption with the advent of network-centric warfare and the emergence of the Global Information Grid'. This thesis' Chapters 6 (*Wetware*), 7 (*Firmware*) and 8 (*Software*) challenge these assumptions and the technical feasibility necessary to realise Arkin's framework.

[167] Arkin, *Governing Lethal Behaviour,* p. 211.

[168] The goal of Arkin's robotic controller design is to ensure that unethical responses are prohibited through an *Ethical Governor* and, through an *Ethical Adaptor*, 'prevent or reduce the likelihood of [an unethical action] via an after-action reflective review or an artificial affective function (guilt, remorse, grief)'. See: Arkin, 'Governing Lethal Behaviour: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture', p. 20.

[169] Ibid.

[170] Olga Khazan, 'The Best Headspace for Making Decisions', *The Atlantic, Science*, 19 September 2016 <https://www.theatlantic.com/science/archive/2016/09/the-best-headspace-for-making-decisions/500423/> [accessed 6 March 2018].

[171] Arkin, *Governing Lethal Behaviour in Autonomous Robots*, pp. 29-30.

[172] Professor Lloyd Clark, in conversation with the author, January 2019.

comprised of components which, individually sound, when aggregated to craft a wholly new battle practice, are difficult to refute.

Allen and Wallach suggest that Arkin's notion is better framed 'from the engineer's perspective, [whereby] making AWS sensitive to moral considerations will add further difficulties to the already challenging ask of building reliable, efficient and safe systems'.[173] Arkin's framework requires several building blocks that must work seamlessly and in tandem. In addition to his 'Ethical Governor' and 'Artificial Conscience', Arkin also posits a 'Responsibility Advisor' to 'make clear and explicit just where responsibility vests should an unethical action be undertaken by the autonomous robot or the robot performs an unethical act due to some representational deficiency'.[174] As explored in this thesis' later analysis, two enduring constraints remain. The first is the cumulative technical feasibility required for Arkin's mix: In discussing the challenges to action selection in machines, Martin, Secretary of the AISB[175], observes: 'I don't mean that it's too difficult like "man will never fly" or "man will never land on the moon". I'm saying it's hopelessly misguided like "man will never dig a tunnel to the moon"'.[176] A second enduring challenge remains AWS' inability to incorporate contextual analysis within the application of these technical devices.

Given human battlefield failings as chronicled by Bourke, AWS' proponents also note a moral component to these ethical drivers.[177] Slim notes, after all, that 'armies, armed groups, political and religious movements have been killing civilians since time immemorial'[178] and, notes Ohlin, shortcomings in *jus in bello* clearly abound.[179] Indeed, Arkin highlights exactly such battlefield lapses to justify AWS deployment.[180] Cummings similarly points out that current human-in-the-loop bombing practices still create unacceptably high civilian collateral damage[181], an inference being that *human-out-of-the-loop* machines might be able to perform better than (or certainly as well as) current in-loop and on-loop systems. In this way, Bourne provides support for AWS' deployment arguing that their adoption will create a set of circumstances whereby 'combatants [are] able to maintain an emotional distance from their victims largely through the application of… technology'.[182] That argument, however, is not new; used first in the nineteenth century around

---

[173] Cornelia Dean, 'A Soldier Taking Orders from its Ethical Judgement Center', *New York Times*, 24 November 2008, para. 16 <http://www.nytimes.com/2008/11/25/science/25robots.html> [accessed 7 March 2018].

[174] Arkin, *Governing Lethal Behaviour: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture,* p. 21.

[175] Source: Society for the Study of Artificial Intelligence and Simulation of Behaviour <https://www.gold.ac.uk/news/tungsten-goldsmith-ai/> [accessed 17 March 2017].

[176] Andrew Owen Martin, senior technical analyst at the Tungsten Network, cited in Ben Sullivan, 'Elite scientists have told the Pentagon that AI won't threaten humanity', *Motherboard magazine*, 19 January 2017 <https://motherboard.vice.com/en_us/> [accessed 17 March 2017].

[177] J Bourke, *An Intimate History of Killing,* (Basic Books, 1999), generally. See also: Walker, *Killer Robots,* pp. 59-60. A corollary here might be that such 'human failure' is an immutable state of the nature of war (not just the character of war). The theoretical significance of Arkin's system is therefore that it changes not just battle's character but also its nature.

[178] H Slim, *Killing Civilians: Methods, madness and morality in war,* (USA: Columbia University, New York, 2008), p. 3.

[179] Jens David Ohlin, 'Is Jus In Bello in Crisis?', *Cornell Law Faculty Publications*, (March 2013), pp. 27-29 <https://scholarship.law.cornell.edu/cgi/viewcontent.cgi?article=2475&context=facpub> [accessed 4 August 2018].

[180] Arkin, *Governing Lethal Behaviour*, p. 2.

[181] Professor Missy Cummings, Director, Humans and Autonomy Laboratory, Duke University, in conversation with the author, Chatham House Conference, February 2014.

[182] Bourke, p. xvii.

rapid-firing artillery, it has since been used for developments in tank, aircraft and naval assets. Arkin's proposals therefore create an interesting dilemma: In suggesting a prospect of more humane armed conflict (and, with it, the saving of lives), restricting battlefield autonomy by legal means 'could [in itself] amount to not properly protecting life'.[183] Notwithstanding that such constraints would presumably vanish in the event of major war, the embracing of autonomy might thus be construed as a legal imperative.

Still other constituents contribute to this ethical driver. AWS, notes Zawieska, might be able to limit their intervention to an 'appropriate' amount of force in a lethal engagement.[184] Autonomous technologies, moreover, might be capable of employing 'creative alternatives to lethality'[185] such as autonomous precision[186], the possibility of non-lethal immobilization as well as the disarming of targets that might otherwise be destroyed. In this vein, it is (again theoretically) possible that the programming of independent weapons might leave a 'digital trail' that could allow better post-facto scrutiny of their actions and thus enhance accountability.[187] From an ethical perspective, it might also be that the modus operandi of AWS will be better suited to future operations: Metz and Coker separately suggest that future conflicts, whether symmetrical or asymmetrical, will tend to be 'hide and seek' in nature rather than formal force-on-force affairs into which AWS technology 'might fit very well'.[188] In this vein, Beres suggests that any such deployment of self-directing machines might be geospatially constrained, the ethical driver being that, just as soldiers are given defined rules of engagement, AWS may be disabled once in a specified no-kill zone.[189] Finally to this point, ethical arguments are based on AWS' ability to improve battlefield practices. Economic considerations apart, unmanned systems may not 'need to protect themselves'.[190] As noted by Kirsch, self-preservation need not be an attribute in their decision-making.[191] This has two ramifications. It may influence weapons' longstanding design equation that seeks to balance an armament's protection with its firepower and mobility. They can act also conservatively. Arkin notes that AWS, appropriately programmed, might be operationally superior to human soldiers in programming out

[183] Heyns, 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions', p. 6. See also: Paul Mitchell, '"Three Laws Safe?" Autonomous Robots and Warfare', *Laurier Centre for Military Strategic and Disarmament Studies*, (15 October 2012), paras. 8-12 <http://canadianmilitaryhistory.ca/three-laws-safe-autonomous-robots-and-warfare-by-dr-paul-t-mitchell/> [accessed 1 March 2018]. Also: Damon Beres, 'The Ethical Case for Killer Robots', *Huffpost*, 3 June 2016 <http://www.huffingtonpost.co.uk/entry/lethal-autonomous-weapons-ronald-arkin_us_574ef3bbe4b0af73af95ea36> [accessed 12 March 2018].

[184] Karolina Zawieska, 'An Ethical Perspective of Autonomous Weapons', cit. *Perspectives on Lethal Autonomous Weapons,* (UN: UNODA Occasional Papers, Number 30, November 2017), pp. 49-56.

[185] Heyns, 'Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions', p. 10.

[186] See, for instance: US Patent grant US52607090A, '*Autonomous precision weapons delivery using synthetic array radar*' <https://patents.google.com/patent/US5260709A/en> [accessed 6 August 2018].

[187] Heyns, *Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions*, p. 10 and generally.

[188] Steven Metz, *Armed Conflict in the Twentieth Century: The Information Revolution and Post-Modern Warfare,* (University of California Libraries, 2000), generally. See also: Christopher Coker, *Humane Warfare,* (USA: Routledge, 23 August 2001) and London School of Economics <http://www.lse.ac.uk/researchandexpertise/experts/profile.aspx?KeyValue=c.coker%40lse.ac.uk> [accessed 2 March 2016].

[189] Damon Beres, *The Ethical Case for Killer Robots,* paras. 19-20.

[190] Andreas Kirsch, 'Autonomous Weapons will be Tireless, Efficient Killing Machines – and there is no way to stop them', *Quartz News*, (23 July 2018), generally <https://qz.com/1332214/autonomous-weapons-will-be-tireless-efficient-killing-machines-and-there-is-no-way-to-stop-them/> [accessed 12 August 2018].

[191] Ibid.

the human heuristic of 'scenario fulfillment' whereby expectation on how a set of circumstances will unfold leads directly to bias in subsequent decision making.[192] Tonkins similarly points to robots being built without faulty psychological dispositions (that might otherwise lead humans towards immoral actions).[193] AWS may also be able to disseminate information under fire to colleague machines and command structures without exaggeration, distortion or contradiction. This, notes Clark, may then materially affect the location of principal decision-makers.[194] The corollary is that such improvements in battlefield decision-making should improve outcomes in the widest sense. Just as Arkin advocates AWS deployment as a means to reduce civilian casualties[195], Freedman reworks a long-standing military aspiration of 'victimless' warfare based instead upon remote, pilotless and autonomous technology that delivers victory 'through disruption rather than destruction'.[196] The ethical facets of removing human supervision in weapons may remove certain political pitfalls: In the case of a downed Reaper, for instance, there is no high value pilot to kill or to take hostage.

3.5 **Operational drivers**

Deployment of AWS might also facilitate the regaining of 'operational initiative' given (as Worcester identifies in the case of the US) a gradual erosion in technical dominance enjoyed until the early 2000s in areas such as high-end sensors, guided weapons and stealth processes.[197] In setting out its future procurement trends, the US Department of Defense's 2013 *Roadmap* defines a broad-ranging role for autonomous weapons operating, it envisages, in seamless groups that benefit from integrated communications and shared targeting data while dynamically integrating ever more information from ever more sources.[198] The operational driver here is that refinement to weapons' accuracy will improve engagement outcomes. This, moreover, should come with much greater speed (as part, perhaps, of the US Army's current network-centric warfare concept) than could be handled by the human operator.[199] The purpose of this final section is therefore to consider operational benefits of deploying AWS including cost advantage, technical advantage and as a pathway to protecting friendly troops. It is then the role of this thesis' later chapters to consider AWS' feasibility what particularly is *lost* in removing human supervision from weapon control.

---

[192] Arkin, 'The Case for Ethical Autonomy in Unmanned Systems', *Journal of Military Ethics*, 9.4, (2010), 333 <https://smartech.gatech.edu/bitstream/handle/1853/36516/Arkin_ethical_autonomous_systems_final.pdf>.

[193] Ryan Tonkens, 'The case against Robotic Warfare', *Journal of Military Ethics*, Vol 11, No 2, (August 2012), 155.

[194] Professor Lloyd Clark, in conversation with the author, December 2018.

[195] Ronald Arkin, 'Warfighting Robots Could Reduce Civilian Casualties so Calling for a Ban is Premature', *IEEE Spectrum*, (5 August 2015) <https://spectrum.ieee.org/automaton/robotics/artificial-intelligence/autonomous-robotic-weapons-could-reduce-civilian-casualties> [accessed 8 September 2018].

[196] Lawrence Freedman, *Information warfare: Will battle ever be Joined?,* (USA: International Centre for Security Analysis, October 1996), p. 6.

[197] Worcester, p. 2. See also: A and O Etzioni, 'Pro and Cons of Autonomous Weapon Systems', *Military Review*, (May-June 2017) <http://www.armyupress.army.mil/Portals/7/military-review/Archives/English/pros-and-cons-of-autonomous-weapons-systems.pdf>.

[198] See, generally: Department of Defense, *Unmanned Systems Integrated Roadmap FY2013-2038*.

[199] DARPA (Defence Advanced Research Projects Agency) Announcement 07-52, 'Scalable Network Monitoring', cit. *International Governance of Autonomous Military Robots*, (USA: Columbia Science and Technology Law Review, Vol XII 2011), p. 280 <www.stlr.org> [accessed 2 September 2018]. See also: David Doria and others, 'Fast Computation on the Modern Battlefield', *US Army Research Laboratory*, (April 2015), p. 1 and pp. 4-5 <https://www.arl.army.mil/arlreports/2015/ARL-TR-7276.pdf>.

It is instructive first to consider the increasing *pace* in the means of combat and how this acts as a driver for removing humans from the engagement loop.[200] As satellite and surveillance technology has developed, Hampton notes that the availability of real-time battlefield information has grown exponentially.[201] While this might suggest like-for-like increases in the speed of combat operations [202], such accessibility has also increased 'both the fog and friction' of such combat.[203] By identifying that decision-making must 'still exceed this new speed of war', US Marine General Dunford acknowledges the operational problems that such developments entail.[204] The UK's Ministry of Defence's *Strategic Programme: Future Character of Conflict*[205] agrees, listing as its first theme the enduring truism that 'future conflict will not be an exact science'.[206] Adams is thus capturing a broadly accepted operational driver in his own conclusion that 'more and more aspects of war are not only leaving the realm of human senses but also crossing outside the limits of human reaction times'.[207] Notwithstanding that it is war's unchanging aspects (the 'nature of war' as discussed in the previous chapter[208]), that provides a benchmark for this exercise, there is little disagreement in either primary or secondary sources that weapon autonomy will certainly accelerate the speed of data availability.[209] Statistics illustrate the point. The human fighter pilot needs some 0.3 seconds to respond to a simple stimulus and more than twice as long to make a choice between several possible responses.[210] A robotic system faced with the same decision may, notes Singer, needs less than a millionth of one second to make that same action selection.[211] The challenge, of course , is that quickening response times (here, the speed between an event and the autonomous weapon's response) does not equate to improvements in the *quality* of that response.

---

[200] Edward Smith, *Network-centric Warfare: What is the point?,* Vol LIV #1, (USA: Naval War College, Winter 2011), p. 61. 'Pace' here refers to tempo, velocity and rate of combat and not, obviously, to the marching step of soldiers on a parade ground.

[201] Jesse Hampton, 'Space Technology Trends and Implications for National Security'*, Kennedy School Review*, (24 January 2016) <http://ksr.hkspublications.org/2016/01/24/space-technology-trends-and-implications-for-national-security/> [accessed 11 November 2017].

[202] For UK-centric narrative on future conflicts including an analysis of conflict's increasing pace, see: Ministry of Defence, 'Strategic Programmes: Future Character of Conflict'*, DCDC*, (2013) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/33685/FCOCReadactedFinalWeb.pdf>. While the document avoids definition of weapon autonomy, it details four on-going principles of battlefield deployment: i. Qualitative advantage cannot be assumed; ii. Future conflict will not be a precise science; iii. The battle of narratives will be key; iv. Maintaining public support will be key. The report's conclusion (p. 39), moreover, is titled 'The UK Must Make People its Edge' (not, by inference, its hardware or, specifically, any advantage conferred through weapon autonomy).

[203] Alex Grynkewich, 'The future of air superiority, Part IV: Autonomy, survivability, and getting to 2030'*, War on the Rocks*, (18 January 2017), para. 5 <https://warontherocks.com/2017/01/the-future-of-air-superiority-part-iv-autonomy-survivability-and-getting-to-2030/> [accessed 18 January 2018].

[204] Jim Garamone, 'Dunford: Speed of Military Decision-Making must Exceed Speed of War', *US Department of Defense*, (31 January 2017), generally <https://dod.defense.gov/News/Article/Article/1066045/dunford-speed-of-military-decision-making-must-exceed-speed-of-war/> [accessed 12 December 2018].

[205] Ministry of Defence, '*Strategic Programmes: Future Character of Conflict*', pp. 4-6.

[206] Ibid., p. 6.

[207] T Adams, p. 1.

[208] See: Introduction to Chapter 2 (*Context*).

[209] This is also acknowledged by the International Committee for Robot Arms Control. See, generally <https://www.icrac.net/statements/> and <https://www.icrac.net/research/>[both accessed 4 June 2018]

[210] Paul Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty First Century*, (USA: Penguin Publishing, 27 January 2011), p. 127.

[211] Ibid., p. 131.

Speed may only be one component of a battle plan (together, perhaps, with precision and the ability to differentiate inputs and outputs in complex environments), but it is also a component where humans have increasing difficulty to participate.[212] The development of autonomous technologies posits that machines will 'now out-speed humans'.[213] As noted, however, by Dunford, a shortened decision-space 'adds new risk with the ability to recover from earlier missteps being greatly reduced'.[214] A further example evidences this driver. The OODA Loop ('Observe, Orient, Decide and Act') concerns the techno-strategic concept first developed by fighter pilot and strategist Boyd during his time at Pentagon consultant in the 1990s.[215] Boyd's insight was that 'advantage lies with the fighter whose OODA Loop is faster and more accurate than his opponent's, and who is able to throw his opponent's OODA Loop out of sync'.[216] Just as aircraft will have to manoeuvre too quickly for a pilot to control, so too will its weapons have to be used at the same speed in order to match the 'beyond-human speed of the aircraft's own systems'.[217] The operational driver in this case is that speed becomes the imperative if friendly forces are to defeat an enemy's similarly autonomous counter-systems. Notwithstanding 'claim inflation'[218] and other biases that may influence assessment of new weaponry[219], Adams summarises this driver by noting that 'military systems now on the horizon will be too fast, too small, too numerous and will create an environment too complex for humans to direct'.[220]

Operational drivers are based largely upon technical advances and their consequences. It is the breadth of such advances that then becomes a contributory driver in feeding a 'revolution in expectation'[221] and the suggestion, notes Gibbs, that weapon autonomy is close to being feasible.[222] An inference, however, from Blake, MD Research at Microsoft, is that 'there is no scientific basis for any of this'.[223] A conclusion from this analysis is therefore that material divergence exists between

---

[212] Angam Parashar, 'How Artificial Intelligence is Outpacing Humans', *Linked In,* Oped, (11 July 2017) <https://www.linkedin.com/pulse/how-artificial-intelligence-outpacing-humans-angam-parashar> [accessed 2 June 2018].

[213] Walker, '*Killer Robots?',* pp. 23-26.

[214] Jim Garamone, 'Dunford: Speed of Military Decision-Making must Exceed Speed of War', para. 12 of 21.

[215] Frans PB Osinga, *Science, Strategy and War: The strategic theory of John Boyd,* (USA: Routledge, 2006), generally.

[216] William Marra and Sonia McNeil, 'Automation and Autonomy in Advanced Machines: Understanding and Regulating Complex Systems*', Warfare Research Paper Series,* 1-2012, (April 2012), p. 9: 'The fastest OODA Loop of the future combat plane will be an automated one – automated in both flight and weapons functions'.

[217] Anderson and Waxman, p. 5.

[218] Here defined as a tendency for interested parties to exaggerate both weapon capabilities and menace from an enemy's arsenal.

[219] Robert Szczerba, *'15 Worst Tech Predictions of All Time'*, *Forbes Magazine*, 5 January 2015 <https://www.forbes.com/sites/robertszczerba/2015/01/05/15-worst-tech-predictions-of-all-time/#65c877e91299> [accessed 5 June 2018].

[220] T Adams, pp. 57-58. See also: John Markman, 'Laser Weapons Set to Boost Military Night at the Speed of Light'*, John Markman's Pivotal Point*, (8 November 2017) <https://www.markmanspivotalpoint.com/investing/laser-weapons-set-boost-military-might-speed-light/> [accessed 18 January 2018].

[221] Professor Philip Sabin in conversation with the author, 29 June 2017.

[222] Samuel Gibbs, 'Elon Musk leads 116 Experts calling from outright ban of Killer Robots'*, The Guardian*, 20 August 2017, para. 7 <https://www.theguardian.com/technology/2017/aug/20/elon-musk-killer-robots-experts-outright-ban-lethal-autonomous-weapons-war> [accessed 3 March 2018].

[223] Professor Andrew Blake, cit. Bryan Appleyard, 'The Sheer Stupidity of Artificial Intelligence'*, Spectator Magazine*, 5 July 2014, para. 10 <https://www.spectator.co.uk/2014/07/the-sheer-stupidity-of-artificial-intelligence/> [accessed 18 August 2018].

the hype of independent weapons and the reality of AI-driven machines. In this vein, a purpose of this final section is to examine relevant technical initiatives that have been seeded by commercial parties. Examples abound and include, for instance, battery performance, a condition precedent to AWS deployment that also evidences the dual role of the private sector in accelerating advances that are relevant to AWS development. Mckinsey notes too the disruptive progress continuing pell mell in power storage and management.[224] Lithium-air batteries as at 2019 provide a viable alternative to lithium-ion and, safety issues notwithstanding, lithium-air units already generate a similar energy density to that of gasoline.[225] Other power advances provide a similar proxy for operational drivers towards AWS adoption. Researchers at George Washington University in 2016 created a molten electrolyte battery using vanadium boride enabling a battery pack to produce thirty times more energy as lithium ion[226]. Battery disruption (and, through this, a driver to unmanned weapon platforms) is also evident from lithium-sulphur chemistry, previously hamstrung by the unwelcome bi-product of lithium sulphide that had degraded cell capacity even after very few cycles.[227] Power innovation over the past five years has similarly seen disruptive development in fuel cells. In 2016, Lockheed Martin demonstrated that its Stalker XE 240 drone could stay airborne for up to eight hours using propane-driven cells.[228] The US Navy's Ion Tiger project similarly uses a Protonex hydrogen fuel cell that can already fly for forty-eight hours using a proprietary cryogenic storage system.[229] In the past three years, commercial companies have also pioneered scavenging technology whereby unmanned weapons might perch on power lines and recharge autonomously enabling missions that continue indefinitely.[230] It is the enabling of UAVs to carry out a perpetual sentry role that brings feasible AWS deployment a step closer. Commercial initiatives, moreover, continue to combine to ensure that power constraints are unlikely to be a technical bottleneck to the development of unsupervised unmanned platforms. Each such technical breakthrough acts as its own catalyst to AWS deployment given that relevant innovation is occurring across the breadth of weapon componentry.[231] Finally to this point, commercial advances

[224] McKinsey & Company, 'Disruptive Technologies: Advances that will Transform Life, Business and the Global Economy', *McKinsey Global Forum*, (May 2013), p. 5 and generally <https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Disruptive%20technologies/MGI_Disruptive_technologies_Executive_summary_May2013.ashx> [accessed 23 May 2017].

[225] Singer, *Wired for War,* p. 123.

[226] L Zyga, 'Molten Air Battery Storage Capacities' *Phys.org blog*, (19 September 2013) <https://phys.org/news/2013-09-molten-air-battery-storage-capacity-highest.html> [accessed 23 March 2018].

[227] Hambling, p. 124.

[228] Ibid., p. 125.

[229] Naval Technology, '*Ion Tiger Overview'* <http://www.naval-technology.com/projects/ion-tiger-uav/> [accessed 12 December 2016]. See also: Kelsey Atherton, 'Burning Hydrogen for Fuel, Navy Drone Flies for 48 Hours Straight', *Popular Science*, (10 May 2013) <https://www.popsci.com/technology/article/2013-05/burning-liquid-hydrogen-fuel-navy-drone-flies-48-hours> [accessed 12 December 2016].

[230] Gary Mortimer, 'Aurora Awarded AFRL Urban Beat Cop Program'*, SUAS News*, (16 December 2011), generally <http://www.suasnews.com/2011/12/aurora-awarded-afrl-urban-beat-cop-program/> [accessed 12 December 2016].

[231] Similar advances continue to be made in drone solar technology towards solving the key conundrum of 'high weight but low efficiency' power generation in unmanned vehicles, as inferred from: Mark Simon and others, 'The Relationship Between Over-Confidence and the Introduction of Risky Products*', The Academy of Management Journal*, 46.2, (April 2003). Zephyr, for instance, is part of DARPA's Vulture project and has a seventy-foot wingspan which can fly non-stop for more than three hundred hours, almost two weeks, based on solar power alone. See: Alireza Galahmar-Zavare and others, 'High Efficiency, Low Size and Low weight Vehicle Battery Chargers'*, Power, Electronics, Drive Systems and Technologies Conference*, PEDSTC, IEEE, (2015), ('*Abstract*'). Developments in solar power, furthermore, will be particularly suited for smaller swarming autonomous drones given favourable wing area ratios; an aircraft that is half the size has just one quarter of the wing area and hence only carries a quarter as many solar cells. It also has just one-eighth the weight to support. Solar cell efficiency at the time of writing has also increased to more than thirty percent

in miniturisation have also opened new possibilities across the range of AWS componentry. A laser target designator that weighed more than forty pounds in 2004 now weighs less than a golf-ball.[232] Similarly, much smaller target designators have considerably increased the range of autonomous roles for unmanned units including low altitude operation beneath cloud cover. Miniturising electronics have made much broader military options available.[233] Here, Hambling notes that laser-based radar can at the time of writing be a key component in autonomous weapons and works by bouncing a laser off thousands of points, calculating the distance to each point and then building up a three-dimensional map of the weapon's surroundings: Such units work in darkness, fog and smoke, are available with no moving parts and cost a few thousand dollars, a fraction of the price of earlier systems.[234]

In particular, it is technology arising from the ubiquitous smartphone that has acted as a driver to the military's general introduction of autonomy. In 2007, twenty million smartphones were sold. By 2015, more than two billion smartphone units were in operation.[235] Smartphone adoption conforms to the same S-curve model of technology implementation earlier posited by this chapter's discussion of disruptive AWS deployment.[236] A central context for AWS deployment is that, while it took landlines forty-five years to reach half of US households, smartphones have achieved the same penetration in just seven years.[237] Apple and Samsung each spent some $14 billion on research and development in 2016 alone.[238] As noted by Kroll and Klaus, smartphone development has accelerated improvements in algorithm construction, a central component to any AWS

using, for instance, gallium arsenide rather than traditional silicon, more than twice as efficient as earlier solar cell technology and at a fraction of the weight. See also : 'DARPA's Vulture: What Goes Up Needn't Come Down', *Defense Daily,* September 2010 <http://www.defenseindustrydaily.com/DARPAs-Vulture-What-Goes-Up-Neednt-Come-Down-04852/> [accessed 12 December 2016]. Other commercial drivers include new wing materials that allow drones actively to optimize thermals as well as software advances transforming UAV efficiency. Examples here include dynamic soaring by optimizing wind shear as well as software that allows UAVs to capitalize upon local air turbulence in urban environments. See: Altadevices <http://www.altadevices.com and http://www.altadevices.com/technology/> [accessed 5 March 2018]. See also: Fred Lambert, 'US Marines test solar powered drones at annual energy expo', *UPI*, (28 June 2015) <http://www.upi.com/Business_News/Security-Industry/2015/06/28/US-Marines-test-solar-powered-drones-at-annual-energy-expo/6911435519387/> [accessed 1 December 2016]. Finally to this point, see: Sean O'Malley, 'Research sends high-flying drones soaring', *RMIT University*, (2016) <http://www.rmit.edu.au/news/all-news/2016/april/research-sends-highflying-drones-soaring> [accessed 5 May 2016].

[232] Source: RPMclasers, <http://www.rpmclasers.com/solid-state-lasers/mil-spec-lasers/airtrac> [accessed 9 December 2016].

[233] John McHale, 'Power Electronics Design Trending Smaller and More Efficient', *Military Embedded Systems*, undated <http://mil-embedded.com/articles/power-trending-smaller-more-efficient/> [accessed 3 July 2018].

[234] The Economist, 'Cheap lasers', (29 November 2012), generally <http://www.economist.com/blogs/babbage/2012/11/cheap-sensors> [accessed 22 December 2017].

[235] Hambling, p. 161.

[236] Tony Seba, '*Clean Disruption'*. Seba identifies a 14% improvement in Lithium battery performance per $1 from 1995 to 2010 and a 20% improvement thereafter; later, at 33' 18": in 2000 one teraflop of processing power cost $46,000,000 and required housing in 150 square metres' space. In 2016, 2.3 teraflops cost just $59 and was available as a personal hard drive. By 2019, it is expected that 20 teraflops, the performance criterion for autonomous vehicles, will be available; and 32' 18": the Light detecting and Radar, LIDAR, unit that cost $150,000 in 2012 currently costs just $250.This is expected to fall to $90 in 2019.

[237] Michael DeGusta, 'Are Smartphones Spreading Faster than Any Technology in Human History*?', MIT Review*, (May 2012) <https://www.technologyreview.com/s/427787/are-smart-phones-spreading-faster-than-any-technology-in-human-history/> [accessed 6 March 2018].

[238] Bruce Upbin, 'Apple is about to become the biggest R&D spender in the world', *Tribune Interactive*, (March 2018) <https://phys.org/news/2018-03-apple-biggest-spender-world.html> [accessed 7 May 2018].

deployment.[239] An example is relevant. The replacement in 2013 of the H264 data compression standard by HEVC (High Efficiency Video Coding) means that four times as much data could by 2017 be transmitted over an identical bandwidth than was the case twenty years ago.[240] Similarly, a ten-fold improvement has taken place in the rate achieved by the 'Fourier Transform', a pivotal mathematical process that will be used by AWS in converting signals from digital to analog and back.[241]

Author and technologist Suarez highlights adjunct but quite different operational drivers to removing human oversight, noting a deluge of video output arising from military systems that outstrip humans' ability to analyse the resulting data. In 2004, just seventy-one hours of video from UAVs was produced for analysis. By 2011 this had increased to more than three hundred thousand hours of output.[242] Cummings reported in 2014 that less than ten percent of such UAV footage was then being analysed.[243] A relevant precursor to empirical AWS sensor fusion is provided by the Pentagon's Gorgon Stare and Argus programmes introduced in 2014 that provide for as many as sixty-five camera eyes on each unmanned vehicle requiring sophisticated visual intelligence software to scour the output for potentially interesting sites.[244] This exacerbates the problem sixty-five times[245] with AI already being required to aid humans what to look for in those datasets.[246] Suarez also points to the fragility of current military hardware given the threat of electro-magnetic jamming and subsequent severing of communications between unmanned weapon and operator.[247] The capture by Iran in 2011 of an American RQ 170 Sentinel drone (after adversarial compromise of its GPS) has also demonstrated the susceptibility of weapons to hostile action.[248] An overarching driver must therefore be to move decision-making further onto the weapon platform in order to remove this requirement for third party communication. The corollary, after all, is that autonomous AWS can ignore radio signals and send out few of their own. The operational driver, cites Adams, is

[239] Dennis Kroll and Klaus David, 'Measuring the Capability of Smartphones for Executing Contextual Algorithms', *Informatics LNI*, Bonn, (2017), pp. 1591-1592 <https://dl.gi.de/bitstream/handle/20.500.12116/3924/B20-1.pdf?sequence=1&isAllowed=y> [accessed 7 June 2018].

[240] Alexander Fox, 'Why is HVEC better than H.264?', *Apple Gazette*, 15 August 2018 <http://www.applegazette.com/mac/why-is-hevc-better-than-h-264/> [accessed 8 July 2018].

[241] L Hardesty, 'The Faster-than-Fourier Transfor*m', MIT News*, 18 January 2012 <http://news.mit.edu/2012/faster-fourier-transforms-0118> [accessed 17 May 2017].

[242] Ted Johnson and Charles Ward, 'The Military Should Teach AI to Watch Drone Footage', *Wired Magazine*, 26 November 2017 <https://www.wired.com/story/the-military-should-teach-ai-to-watch-drone-footage/> [accessed 12 September 2018].

[243] Professor Missy Cummings, Humans and Autonomy Laboratory, Duke University in conversation with the author, Chatham House Conference, February 2014. See also: Tim Klassen, 'The UAV Video Problem', *Military Aerospace*, (1 July 2009) <https://www.militaryaerospace.com/articles/print/volume-20/issue-7/features/viewpoint/the-uav-video-problem-using-streaming-video-with-unmanned-aerial-vehicles.html> [accessed 12 May 2017].

[244] Stephen Trimble, 'Sierra Nevada Fields ARGUS-IS Upgrade to Gorgon Stare Pod', *Flight Global*, 2 July 2014 <https://www.flightglobal.com/news/articles/sierra-nevada-fields-argus-is-upgrade-to-gorgon-stare-400978/> [accessed 18 March 2018].

[245] Open forum, Chatham House Conference on Autonomous Weapons, February 2014.

[246] Walker, '*Killer Robots?',* p. 26.

[247] JR Wilson, 'Electronic Warfare Evolves to Meet New Threats', *Military & Aerospace*, (1 August 2017) <https://www.militaryaerospace.com/articles/print/volume-28/issue-8/special-report/electronic-warfare-evolves-to-meet-new-threats.html> [accessed 12 December 2017].

[248] Iran has subsequently manufactured a clone machine; see: RT News, 'Iran Replicates CIA's RQ-170 Sentinel Drone', *RT News*, 11 May 2014 <http://rt.com/news/158272-iran-unveils-drone-copy/) [accessed 23 April 2016>.

that systems will need ever more autonomy in order just to survive[249] entailing, de facto, incrementally more decision-making being built into the UAV to facilitate mission independence but simultaneously requiring appropriate 'hand-off' routines in order to clarify the provenance of orders.[250]

A further operational driver to AWS deployment arises from the attraction of 'force multiplication'.[251] There are several facets to this driver, as noted by the UK Houses of Parliament's *Postnotes* in its discussion on the transition from manned battlespace assets to unmanned autonomous machines.[252] Uncertainties around deployment models mean, however, that predicting the manpower effect of AWS adoption is challenging. Ignoring the unfair comparison, SIPRI goes no further than noting 'combat aircraft pilots must fly in real conditions to be properly trained and to fly between ten and twenty hours a month to maintain their skill set… unmanned autonomous aircraft, on the other hand, can sit on a shelf for extended periods of time without losing their operational capability'.[253] A counterpoint (and the basis for this section's operational driver) is instead articulated by the US Military's *Future Combat System* project (FCS)[254] in its forecast that AWS will provide 'force multipliers in order to empower single soldiers on the ground to become a nexus working a cohort of divers automated weapon systems'.[255] Rather than a prescription on 'how' and 'how many', a further operational driver is that weapon autonomy allows that single soldier to do the job of what previously had taken several soldiers (justifying now such hackneyed terms such as 'expanding the battle space', 'extending the war-fighter's reach', 'autonomous casualty reduction'[256]). Lanchester's *Square Law* is a similarly long-standing but now relevant driver for AWS deployment, providing a heuristic rule of thumb for the advantage of quantity versus quality in military engagements.[257] The law states that, all things being equal, having twice as many units in a fight translates to a fourfold increase in combat power for units with aimed-fire

[249] Thomas Adams, p. 7.

[250] Patrick Beautement, 'Putting complexity to work; achieving effective human-machine teaming', *The Abaci Partnership LLP*, (2015), p. 13. See also: Chapter 8 (*Software*), specifically: 8.7 ('*Action Selection Issues'*).

[251] Tapan Bagchi, 'Force Multiplier Effects in Combat Simulation', *Proceedings of 7th Asia Pacific IEMS Conference*, Bangkok, (December 2006), pp. 1244-1245 <https://www.researchgate.net/publication/228831999_Force_Multiplier_Effects_in_Combat_Simulation> [accessed 12 September 2017].

[252] Postnotes, *Automation in Military Operations,* (UK: Houses of Parliament, Number 511, October 2015), p. 4.

[253] Boulanin and Verbruggen, p. 63.

[254] Source: 'Army Future Combat System (FCS) 'Spin-Outs' and Ground Combat Vehicle (GCV): Background and Issues for Congress', *US Congress Information Services*, RL32888, (30 November 2009) <https://opencrs.com/document/RL32888/> [accessed 13 January 2016].

[255] Taken in part from: Erin McDaniel, 'Robot Wars: legal and ethical dilemmas of using unmanned robotic systems in 21st Warfare and beyond', Unpublished thesis, *Fort Leavenworth Kansas*, (2008), p. 77. For a useful primer on Force multiplication, see: Time Magazine, 'The Reaper Revolution Revisited', 27 February 2012 <http://nation.time.com/2012/02/27/1-the-reaper-revolution-revisited/> [accessed 15 February 2016].

[256] Gary Marchant and others (including Ronald Arkin), 'International Governance of Autonomous Military Robots', *Columbia Science and Technology Law Review*, XII, (2011), p. 275 <http://stlr.org/download/volumes/volume12/marchant.pdf>.

[257] Ronald Johnson, 'Lanchester's Square Law in Theory and Practice', *School of Advanced Military Studies*, *Fort Leavenworth*, (1990) <http://www.dtic.mil/dtic/tr/fulltext/u2/a225484.pdf>.

weapons.[258] The basis for the rule is quite simple. Numerically superior forces can double up on attacking enemy units while the numerically inferior force can only attack half of the opposing force at any one time. The chief value of mass is that it can be used to impose costs on adversaries because it forces them to encounter large number of systems. Lowe identifies several examples of such force multiplication around unmanned units from very-high-altitude, ultra-endurance, 'loitering theatre' reconnaissance units to blimps to 'itty-bitty, teeny-weeny UAVs'.[259] Hambling similarly cites the role of Apache helicopter pilots using Lockheed Martin's VU-IT's collaboration with partner UAVs as a remote sensor to investigate areas too hazardous to fly the helicopter.[260] The point is again that emerging technologies are clearly creating capabilities that are required if unsupervised weapons are to be deployed.

Any analysis of drivers should also review commercial factors encouraging the adoption of weapons autonomy. A key notion is set out by Work and Brimley[261] in their contention that 'the movement toward a 'Robotic Age' is not being led by the military-industrial complex'[262] but instead by 'companies focused on producing consumer goods and business-to-business services such as advanced computing, big data, AI, miniaturization, additive manufacturing and small but high density power systems'.[263] This may be a departure from earlier procurement precedents given that AWS technologies are now coming out of a thriving commercial sector.[264] This, however, is not frictionless with commercial drivers fueling their own proliferation concerns. The threat of COTS ('commercial off the shelf') hardware is, in an age of 3d printing and Bitcoin, that 'unmanned systems [can] be assembled in disturbing anonymity'.[265] A further driver for AWS deployment is the proven lobbying powers of the US drone industry.[266] Similarly, the number of experimental military autonomy projects ('military' projects as opposed to projects explicitly around weaponisation) already in process each add towards a cumulative deployment of autonomous technologies; Singer notes, for instance, that the Pentagon's *Joint Robotics Programme* is currently developing twenty-

[258] Paul Scharre, 'Robotics on the Battlefield, Part II: The Coming Swarm', *Center for a New American Security*, p. 18, (15 October 2014) <https://www.cnas.org/publications/reports/robotics-on-the-battlefield-part-ii-the-coming-swarm> [accessed 17 January 2016].

[259] Christian Lowe, 'High Flying Secret Drone unveiled', *www.defensetech.org*, (24 July 2006), cit. *The Changing Nature of War*, eds. Hew Strachan and others, (Oxford: OUP Oxford, May 2011), p. 352.

[260] Hambling, p. 70.

[261] Robert Work and Shawn Brimley, '20YY; Preparing for War in the Robotic Age', *CNAS*, (22 January 2014), p. 6 <https://www.cnas.org/publications/reports/20yy-preparing-for-war-in-the-robotic-age> [accessed 12 August 2018].

[262] For example, missiles, guided munitions, computer networking, satellites, global positioning and stealth technologies.

[263] Work and Brimley, pp. 8-10.

[264] Walker, '*Killer Robots?',* pp. 65-68.

[265] Bill Powers, Potomac Institution for Policy Studies, in conversation with the author, Chatham House Conference, February 2014, *Autonomy: A force for Good*?

[266] See: The Association for Unmanned Vehicle Systems International, 'The Economic Impact of Unmanned Aircraft Systems Integration in the United States', (March 2013) <https://higherlogicdownload.s3.amazonaws.com/AUVSI/958c920a-7f9b-4ad2-9807-f9a4e95d1ef1/UploadedImages/New_Economic%20Report%202013%20Full.pdf>. The AUVSI report cites the economic benefit of UAS integration suggesting that more than 70,000 jobs will be created in the United States with an economic impact of more than $13.6 billion. This benefit, it claims, will grow through 2025 when the trade organisation foresees more than 100,000 jobs created and economic impact of $82 billion. Europe is considerably behind the US in their efforts and activity as evidenced by less that 15% of AUVSI's membership comes from outside the USA (Association for Unmanned Vehicle Systems International).

two different prototype 'intelligent ground vehicles'[267] ranging in size from tiny eight pound units to an autonomous 700-ton robotic dump truck that can move more than two hundred tons of earth at a time. Importantly, emerging examples of battlefield robotics will also be 'new and improved' versions of existing platforms and suddenly capable of taking on wider and more autonomous battlefield roles. Moreover, such autonomous robots will increasingly be able to carry out *multiple* roles that will likely include lethal capabilities.[268]

Finally for this section, it is the considerable *scale* of resources devoted to unmanned weapon programmes that becomes a clear driver. Over the five years to 2018, the US budget[269] for unmanned systems is expected to total US$24 billion.[270] This effort is well established; in the four years to FY 2010, flight hours for UAVs increased from 165,000 hours to more than 550,000 hours and the inventory of systems from less than 3,000 to 6,500.[271] This trajectory has been cemented with the formal embedding of UAVs in all of its Brigade Combat Teams.[272] There is also considerable room to expand further these unmanned and autonomous efforts; in 2009, US spending of two billion dollars per year on UAVs was only one tenth of the amount that the US spends on space capabilities and less than one half of one percent of the US Defence budget as a whole.[273] This has important contextual ramifications. Militaries' growing experience of UAV creates confidence in the asset class.[274] Thus, in 2010, the US Air Force trained more unmanned pilots than traditional pilots for the first time[275] while, in 2011, the University of North Dakota chartered its first four-year degree programme in UAV piloting (prompting Singer's assertion that battlefield robotics are the next major area of fundamental change in how warfare is carried out).[276]

There is, therefore, a wide set of factors which separately *and* together answer the 'why' piece in understanding motivations for AWS deployment. Analysing this deployment spur provides a pivot to this thesis; excepting this section and the previous chapter's exploration of AWS' context (here, 'how to *understand* AWS deployment'), this thesis otherwise focuses on AWS' challenges and

[267] Singer, *Wired for War,* p. 110.

[268] Foster-Miller's latest iteration of its technology is called MAARS, the Modular Advanced Armed Robotic System, capable of several hardware and software configurations. A recent partnership between Carnegie Mellon University and the Marine Corps is the Gladiator, the world's first 'multi-purpose combat robot'.

[269] Exact figures for spend appear to differ depending upon authors' definitions of particular programmes; see: US Department of Defence, 'Unmanned System Roadmap 2007-2032', p. 4.

[270] 91% of this will be allocated to aerial UAS, 8% to maritime unmanned systems with the small balance being taken up by ground systems.

[271] Evidence from Deputy Director, Unmanned Warfare (Office of the Under Secretary of Defence), 'Committee on Oversight and Governmental Reform', (USA: Congressional Research Service, March 2010), p. 4.

[272] Ibid., p. 5.

[273] Professor Philip Sabin, 'The Strategic Impact of Unmanned Aerial Vehicles', cit. Royal Air Force Directorate of Defence Studies, *Air Power; UAVs: The wider context*, (UK: MAR/UAV/113/09, 2009), p. 101.

[274] For an interesting essay on the effects of military innovation on European affairs see: David Parrott, 'The Military Revolution in Early Europe', *History Today*, Volume 42, Issue 12, (December 1992), generally <http://www.historytoday.com/david-parrott/military-revolution-early-europe> [accessed 2 February 2014].

[275] As reported to the Committee on Oversight and Government Reform, Hearing on 'The Rise of Drones; Unmanned Systems and the Future of War', *Committee on Oversight and Government Reform*, Congressional Research Service, (March 2010), generally <http://digitalcommons.wcl.american.edu/cgi/viewcontent.cgi?article=1002&context=pub_disc_cong> [accessed 15 June 2017].

[276] Singer, *Wired for War*, p. 179 and p. 203.

constraints. Chapter Three is therefore alone in chronicling the benefits posited by removing supervision from lethal engagements. The breadth of such drivers is considerable and contributes directly to expectations that these weapons have created in politicians and the wider public.[277] That same breadth mirrors, however, a wide divergence of opinion on AWS deployment from, on the hand, the *Future of Life* denouncing self-directing weapons as a danger to humanity to academic and other practitioners' broad support founded on moral, ethical and operational arguments set out above.[278] Drivers, therefore, are an important component to the subject's contextual framework against which the remainder of this thesis and its focus on deployment's problems can better be judged.

[277] Michael Horowitz, 'Public Opinion and the Politics of the Killer Robot Debate', *Sage Journals: Research and Politics Series*, (16 February 2016) <http://journals.sagepub.com/doi/pdf/10.1177/2053168015627183> [accessed 17 August 2018].

[278] Future of Life Institute, 'Autonomous Weapons: An Open Letter from AI and Robotics Researchers', *Future of Life*, (2015) <https://futureoflife.org/open-letter-autonomous-weapons/> [accessed 6 April 2016].

4. Deployment: Models for the removal of weapon supervision

Military technology has been automating for decades. This chapter considers a range of likely deployment models for AWS implementation. A starting point is States' initiatives to enhance their military capacities through pairing soldiers with technology.[1] While this may take on many forms (not all of which presage lethality), such deployment models are generally predicated on broad reduction of human control in processes and procedures.[2] Having considered drivers accelerating the move to automated and autonomous weapons in Chapter Three (and before dissecting the several challenges to such models, the aim of subsequent chapters), this chapter will consider *how* AWS might be introduced into battlefield practices. Adams highlights a dilemma. The difference between, first, a weapon that can search target areas, assist in attack decisions, select and dispense munitions and then report results and, second, 'a machine that can do these things and make its own attack decisions' is increasingly really only a matter of programming.[3]

This chapter is an important scene-setter in this thesis' overall consideration of AWS feasibility.[4] Divided into four sections, it first discusses current models around the deployment of battlefield autonomy based on review of individual capabilities that will constitute that autonomy. As part of this overview, the chapter is governed generally by a test of 'reasonableness'. It also considers possible demarcation between offensive and defensive weapon systems as well as the weight of adoption costs on the shaping of deployment models. Deployment is then considered from its proponents' perspective. Arkin, for example, is exactly highlighting the importance of deployment conditions when he posits that if AWS 'can be designed appropriately and *used in situations where they will be used appropriately*… they can reduce collateral damage significantly'.[5] In doing so, however, considerable overlap remains between the 'how' (here, deployment) and the 'under what circumstances' (here, the context of AWS deployment). Notwithstanding that future armed conflict is unlikely to have a battlefield in its traditional sense[6], much of this chapter is led by what Hickcok terms 'the *miring* effect of context over the near-future battleground'.[7] Latiff similarly points to the 1999 predictions of Chinese colonels Liang and Xiangsui that 'tomorrow's soldiers will increasingly be computer hackers, financiers, smugglers and agents of private corporations rather than members of a military'.[8] Latiff is forecasting that it will be machines that fight battles on

[1] See: RDECOM, 'Future Soldier 2030 Initiative', *US Army Soldier RD&E Centre*, (February 2009), generally <https://www.wired.com/images_blogs/dangerroom/2009/05/dplus2009_11641-1.pdf>.

[2] H Waz de Czege, 'Six compelling ideas on the road to a future army', *Army Magazine*, Vol 51, no.2, (2015), p. 3.

[3] Adams, *Future Warfare and the decline of human decision-making,* (USA: Parameters, 31.4, 2011) pp. 57-71 (p. 61).

[4] In so doing, the purpose of Chapter 4 (*Deployment*) is not to question AWS' feasibility (as analysed in subsequent chapters) but instead to assume that appropriate progress is achieved in the capabilities of computer vision, natural language processing, machine learning, search and planning, logical and symbolic reasoning, human-machine interaction, manipulation, power issues and locomotion, collaborative intelligence and general verification.

[5] Damon Beres, 'The Ethical Case for Killer Robots', *Huffington Post*, 3 June 2016, para. 8 <http://www.huffingtonpost.co.uk/entry/lethal-autonomous-weapons-ronald-arkin_us_574ef3bbe4b0af73af95ea36> [accessed 12 December 2017].

[6] Alexander Kott and others, *Visualizing the Tactical Ground Battlefield in the Year 2050*, (USA: US Army Research Laboratory, June 2015), pp. 7-14 and generally <https://www.arl.army.mil/arlreports/2015/ARL-SR-0327.pdf>.

[7] William Hickok, 'Defining War in Twenty-first Century America', *School of Advanced Military Studies, Fort Leavenworth*, (2010), p. 27.

[8] Major General Robert Latiff, *Future War*, (USA: Alfred Knopf Publishing, 2017), p. 4.

humans' behalf, machines that watch for humans and think for humans, that 'fight fast' and, moreover, that are 'not be well defined temporally'.[9] Given that such weapons will conduct operations over a 'larger, more diffuse battlefield' and that that sphere of battle will operate 'where our traditional understanding of the rules of war will be challenged', it remains through a contextual lens that deployment can be best understood.[10] The chapter's analysis also pivots around the role of supervision in battle planning and the role it plays in shaping emerging machine-machine and human-machine teaming models, the notion of *flexible* autonomy and how this might be practical in, for instance, models involving multiple AWS (here, swarming). In so doing, any review of deployment must cover the costs of machine failure as well as the challenging imperative of AWS testing and validation.

Deployment of AWS is just one element in an enduring effort to enhance a commander's frontline options. In this vein, a starting point for this chapter is to isolate autonomous functions by battlefield assets. Various weapon datasets[11] exist for this exercise that identify autonomous capabilities including, inter alia, independent data sensing, self-directing software to interpret that input data, programming to transform resulting output into plans and actions, the scope (if present) of both communication routines and human-machine interfaces and, finally, independent end-effectors that physically enable such self-directing weapons to execute derived actions.[12] As above, the Stockholm International Peace Research Institute (SIPRI)'s dataset identifies three hundred and eight-one different systems in their 2017 study on unmanned weapon systems that feature degrees of autonomy in their critical functions.[13] That number, however, is meaningless without understanding the dataset's context and methodology for recording capabilities. It is telling to note what the dataset does *not* reveal. It cannot account for task complexity. Given that autonomous systems must model battlefield tasks mathematically, a deployment challenge must be the degree of difficulty that each component task involves relative to established human standards. Nor does the dataset reflect the *precision* required by individual autonomous tasks. It ignores permissible leeway and margin of error. It overlooks any coding scale whereby the more ill-defined that task's specification, the more challenging becomes its mathematical formulation for the AWS. Datasets cannot accurately reflect task *tangibility* (can each expected outcome of the independent weapon be qualified?), *dimensionality* (can the battlefield task be carried out in a single action or does it require sequential decisions and actions?) or *interaction* (are additional assets required in order to complete the task and, in so doing, is interaction required with human or other autonomous agents?).[14] Just as the nature of the weapon's actions must be competitive, collaborative or based on instruction, each of these heuristics must in turn affect model selection. Finally, the SIPRI dataset does not capture the *dynamic* state of the AWS' operating environment, whether it is observable, cluttered, adversarial, structured or stochastic (does the AWS' action always produce the same

[9] Michael Gross, 'Ethics on the Near-Future Battlefield', *Bulletin of the Atomic Scientists*, (December 2015), generally <https://thebulletin.org/2015/12/ethics-on-the-near-future-battlefield/> [accessed 5 January 2018].

[10] Robert Latiff, *Future War*, pp. 22-25.

[11] Ariel Conn, 'The Problem of Defining Autonomous Weapons', *The Future of Life Institute*, generally (30 November 2016) <https://futureoflife.org/2016/11/30/problem-defining-autonomous-weapons/> [accessed 14 April 2018].

[12] This thesis focuses on the November 2017 work out of the Stockholm International Peace Research Institute (SIPRI): Boulanin and Verbruggen, 'Mapping the development of autonomy in weapon systems', generally.

[13] Boulanin and Verbruggen, p. 19. The dataset is focussed on weapon systems rather than individual munitions.

[14] Ibid., p. 13.

effect on it?). A key, therefore, is that each of these variables multiplies task complexity, a set of relationships that cannot be ignored by such datasets.

Within these constraints, what then is the current picture for the deployment of weapon systems without human supervision?[15] For this, it is important to understand how the SIPRI dataset is divided. Its two principal segments comprise weapons with autonomous *mobility* (forty-eight per cent of the dataset[16]) and autonomous *targeting* (twenty-three per cent of the dataset[17]). These two together groups comprise seventy-one per cent of SIPRI's November 2017 dataset, fully two hundred and seventy weapon systems. The balancing quarter of the dataset's weapon platforms also exhibit autonomous characteristics and relate to intelligence systems (ten per cent), 'interoperability' systems (ten per cent) and health management systems (six per cent). Irrespective of issues around data classification, SIPRI's breakdown provides a relevant starting point from which both to review AWS deployment as well as to identify individual weapon capabilities that must comprise that deployment. A clear heuristic emerges. Just as full autonomy is classified at one end of a capability continuum[18], there will be several intermediate and transitional deployment models that make up the rest of that continuum.[19]

This continuum is best demonstrated by identifying the autonomous capabilities that are contained within SIPRI's dataset. Mobility-related functions that are relevant to understanding deployment models comprise unsupervised homing/follow-me capabilities, autonomous navigation and, in time, takeoff/landing competences. As defined by Lekka, autonomous homing is currently associated with missile technology whereby the weapon system finds and track targets.[20] Follow-me capabilities in AWS refer to the weapon's ability to shadow a 'colleague system' or soldier. In both cases, the AWS model involves the autonomous directing of the weapon towards a targeted object that it has detected and is tracking. The capability, however, masks additional complexity. Hall notes that it will likely require automatic sense-and-avoid routines to prevent

[15] 195 (51%) of the military systems identified by SIPRI in their April 2017 dataset to be incorporating autonomy were unarmed systems. 175 (46%) systems were armed. 11 systems were classified as 'unknown'. Of this dataset of 381 weapon systems incorporating autonomous capability, 58% were air systems, 24% ground systems and 18% maritime. 225 of the 381 systems has completed their development while 131 were still undergoing development. The development status of 14 systems was unknown and 11 systems had been cancelled. See: Boulanin and Verbruggen, p. 20. By way of background, the dataset is comprised of three categories as follows: i. Unmanned weapon systems that features some autonomy in their critical functions (that is, they can autonomously search for, detect, select, track or attack targets); ii. Unmanned weapon Systems that do not have autonomy in their critical functions but feature autonomous functions in any of the other capability areas covered by the study (namely mobility, intelligence, interoperability and health management); iii. Unmanned and unarmed military systems (involved in intelligence, surveillance, reconnaissance and logistics missions).

[16] 277 systems of which 30% of that segment comprises armed weapon systems.

[17] 153 weapon systems of which a significant 85% relates to armed systems.

[18] See: National Academy of Science, Engineering & Medicine, 'Autonomies for civil aviation: Toward a new era of flight', *IAP*, (2014), pp. 12-19 ('*Autonomous capabilities and vision*') <https://www.nap.edu/read/18815/chapter/3> [accessed 12 July 2018].

[19] William Marra and Sonia McNeil, 'Understanding the loop: Regulating the next generation of war machines', *Hartford Journal of law and public policy*, volume 36, (2012), 1173-1177 <http://www.harvard-jlpp.com/wp-content/uploads/2013/05/36_3_1139_Marra_McNeil.pdf>.

[20] Anastasios Lekka, 'Guidance and path-planning systems for autonomous weapons', unpublished thesis, *NTNU*, (April 2014), pp. 6-9 <http://fossen.biz/home/PhD/thesis/Lekkas%202014.pdf>.

collision when operated in a cluttered environment.[21] Similarly, AWS' autonomous navigation will be constrained by where the weapon is operating (land, sea, air or together, adversarial or uncontested). Notwithstanding such variants, certain capabilities (for instance, navigation) are essential to AWS deployment models. Navigation must work first time, every time to ensure that the weapon can accurately determine its position in order both to plan and follow a route without supervision.[22] While the point is to chronicle these capabilities and how they combine to support particular deployment models, later chapters tackle their feasibility: The challenge of AWS' navigation (as noted by Cacca and others) is obviously resolving discrepancies occasioned by obstacles, by adversarial activities, by enduring inefficiencies generic to vision-based guidance systems[23] as well as by the complicating requirement that AWS dynamically interacts with other possibly unpredictable autonomous agents. The SIPRI dataset also masks how it is particular battlefield characteristics that influence the shape of AWS deployment models.[24] Specific tasks have solicited specific autonomous solutions in, for instance, the detection (and subsequent engagement) of perimeter intrusion, the pinpointing of delivery coordinates in gunfire as well as remote the identification and processing of objects in an ISR mission.[25] This, however, is to ignore that deployment initiatives will remain hamstrung by technical difficulties and it is this cumulative set of challenges that complicates how individual AWS 'solutions' may empirically be deployed.[26] While, for instance, biometrics and object recognition techniques have improved markedly since Karpathy's 2012 analysis, it still holds that vision technology cannot yet infer abstract meanings from images, video footage or real-life situations.[27] The corollary is that a technology which remains fundamental to the taking away of weapon supervision also remains unable to detect potential human targets based on the behaviour of those targets in a manner that is compliant with LOAC.[28]

---

[21] Brian Hall, 'Autonomous Weapon System Safe*ty', Joint Forces Quarterly*, 86, generally, (June 2017) <http://ndupress.ndu.edu/Media/News/Article/1223911/autonomous-weapons-systems-safety/> [accessed 2 February 2018].

[22] Several systems rely on 'waypoint navigation' and, as such, may not be truly autonomous. Similarly, Northrop Grumman's MQ-4C Triton UAS can autonomously plan a route but still relies on a human operator to set speed, altitude and mission objectives.

[23] Massimo Caccia and others, 'Basic navigation, guidance and control of an unmanned surface vehicle', *Autonomous Robots*, Volume 25, Issue 4, Springer US, (2008), pp. 349-365. See also: Appendix One: '*Case study*: *Automatic target recognition'*.

[24] MC Haas and SC Fischer, 'The Evolution of Targeted Killing Practices: Autonomous Weapons, Future Conflicts and the International Order*', Contemporary Policy*, 38:2, (August 2017), pp. 284-286 and 286-288 <https://www.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Haas&Fischer_2017_TargetedKillingPractices.pdf>.

[25] Examples of autonomy in intelligence roles include Israel's Counter-IED and Mine Suite (CIMS) developed by IAI, General Dynamics' unmanned ground system called Mobile Detection Assessment and Response System (MDARS, developed for the US Army), Endeavour Robotics' RedOWL artillery targeting system, Boeing's ScanEagle system that can autonomously monitor objects of interest on the sea surface and, on a research basis, the US Office foro Naval Research's collaboration that seeks to infer intentions and threats in surveillance imagery through its Automated Image Understanding Thrust (AIUT) system.

[26] See: case study on ATR, Appendix One: '*Case study: Automatic Target Recognition'*.

[27] Andrej Karpathy, 'The state of computer vision and AI: we are really, really far away*', Karpathy blog*, (22 October 2012) <http://karpathy.github.io/2012/10/22/state-of-computer-vision/> [accessed 12 December 2017]. These challenges are reviewed in detail in Chapters 7 (*Firmware*) and 8 (*Software*). See also Chapter 12 (*Appendix*), specifically: '*Case Study on Automatic Target Recognition'*.

[28] Israel's Iron Dome missile defence system can, for example, calculate where incoming missiles will detonate and suggest appropriate countermeasures based on that analysis. See: Emily Landau and Ariel Bermant, 'Iron Dome protection: Missile defence in Israel's security concept', *Lessons of Operation Protective Edge*, (2014), pp. 38-39 <http://www.inss.org.il/he/wp-content/uploads/sites/2/systemfiles/SystemFiles/Iron%20Dome%20Protection_%20Missile%20Defense%20in%20Is

Isolating autonomous capabilities in this manner allows Horowitz and Scharre to frame AWS deployment in terms of the division of human-machine tasking.[29] A 2025 vignette in the US Army's *Robotics and Autonomous Strategy* highlights broad benefits expected in urban operations from, inter alia, broad autonomous tasking such as threat avoidance, risk-reduced reconnaissance and 'contact made on our own terms'.[30] Given that military planning is often evidenced to be cumbersome, inflexible and slow[31], a deployment model that delegates tasks to autonomous machines may be an appealing first step to removing human supervision from wider battlefield tasking.[32] Without this technical lift, staff teams must continue to work with imperfect information under significant time constraints in order to arrive at what are very consequential battlefield decisions. In any such deployment model, there are clearly degrees that exist between hybrid (albeit non-lethal) weapon systems that may employ limited autonomy in certain quite specific functions and, far down the same continuum, fully autonomous weapon systems that are, by degree, self-learning and designed to operate *without* mechanisms involving human verification.[33] This continuum is the nub of the deployment challenge confronting the introduction of AWS onto the battlefield. The balance of this chapter therefore considers how such degrees of autonomy might find their way into battlefield practices.

In dissecting battlefield deployment, certain architectural tenets are important. The chapter's focus is on how autonomy *within* a weapon's individual components will *together* transform human control in an engagement sequence. This thesis therefore approaches deployment based on 'autonomy within individual weapon systems' rather than autonomy across the broad weapon type. Consideration of AWS deployment must also take into account collaborative uses of autonomy (including swarming) while at all times challenging the feasibility of what practically and technically is being proposed. It is for this reason that the chapter reviews deployment throughout with an eye to those models' reasonableness.[34] A starting point is thus to frame AWS deployment specifically by 'the type and quality of human control afforded by the different types of

rael's%20Security%20Concept.pdf>. See also: Alexander Wissner Gross, 'Datasets over Algorithms', *Edge*, June 2017, paras. 3-6 <https://www.edge.org/response-detail/26587> [accessed 2 February 2018].

[29] Michael Horowitz and Paul Scharre, 'An Introduction to Autonomy in Weapon Systems', *CNAS*, (13 February 2015), p. 3 <https://s3.amazonaws.com/files.cnas.org/documents/Ethical-Autonomy-Working-Paper_021015_v02.pdf?mtime=20160906082257> [accessed 1 June 2017]. See also: Ben Rossi, 'How industry 4.0 is changing human-technology interaction', *Information Age*, 11 November 2016 <http://www.information-age.com/industry-4-0-changing-human-technology-interaction-123463164/> [accessed 23 May 2017].

[30] US Army, 'Robotic and Autonomous System Strategy', *Army Capabilities Integration Centre*, (March 2017), p. 6 <http://www.tradoc.army.mil/FrontPageContent/Docs/RAS_Strategy.pdf>.

[31] Larry Ground and others, 'Coalition-based Planning of Military Operations: Adversarial Reasoning Algorithms in an Integrated Decision Aid', *arXiv pre-print arXiv*: 1601.06069, (2016), p. 1 <https://arxiv.org/pdf/1601.06069.pdf>.

[32] Patrick Talbot, 'Military Decision Aids – A Robust Decision-Centred Application', *TRW Systems, Technology Review Journal*, (Spring-Summer 2001), 83-84 <http://ellisinterstellar.com/DecisionAids.pdf>.

[33] Robin Geiss, *The International-Law Dimension of Autonomous Weapon Systems,* (Germany: Freidrich Ebert Stiftung, October 2015), p. 8 <http://library.fes.de/pdf-files/id/ipa/11673.pdf>. Examples include the US Navy Phalanx system which can autonomously search, detect and engage targets. Britain's 'fire and forget' Brimstone missiles can distinguish between armoured vehicles and civilian transport without human assistance and can hunt targets autonomously in pre-designated areas. Israel's Harpy missiles can detect and autonomously destroy opponents' radars.

[34] Neil MacCormick, 'Reasonableness and Objectivity', *Notre Dame Law Review*, 74, Issue 5, Article 6, (1999) <https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=1648&context=ndlr> [accessed 29 October 2017].

computerized weapon systems'.[35] In this way, classification of control can be based on each model's specific level of human supervision ranging along that continuum from humans engaging and selecting targets prior to initiating an attack (level one), algorithms suggesting targets which humans then choose which to attack (level two), algorithms selecting targets which humans must approve before an attack is initiated (level three) to algorithms selecting targets which humans then have a restricted time to veto (level four). A final level five classification of deployment is then apt where programmes select targets and initiate attacks without human involvement.[36] While this might appear compelling, any such framework is nevertheless empirically vague.[37] These boundaries are of course loose and indistinct. The more operational limitations that are built around each hypothetical deployment model, after all, the more theoretically predictable becomes that unsupervised weapon's use of force. This, however, is clearly a specious relationship.[38]

Toggling between different deployment models on that continuum also disrupts the ability to determine cause and effect. It is therefore important to understand the control triggers that will exist between, on the one hand, persistent human involvement in engagements to, on the other, complete weapon independence and the code-based management of those weapons. In models that are based upon degrees of human supervision, Knuckey summarizes this deployment continuum: While the nature of the human-machine relationship may vary model-by-model, 'the later a human leaves the loop, or the more human-limited the operational context is, the more systems may be argued to be under effective human control'.[39] For the purposes of this thesis (within its review of AWS feasibility), this is also because the weapon's task of selecting and engaging targets is comprised of a complicating variety of subtasks.[40] Deployment must be considered case-by-case across weapon control types as well as individual weapon capabilities, processes, context and task instructions. The point is that it is inappropriate to create one overarching set of deployment rules. It may, for instance, be enduringly inappropriate for self-directing weapons to be deployed autonomously in a complex urban setting given its quite different set of challenges to a remote and controlled battle space with well-understood, minimal civilian density.

In addition to empirical battlefield activities, more conceptual aspects of AWS tasking have an important part in shaping such models. Sartor defines certain abstract capabilities that must precede AWS deployment including the weapon's broad behavioural competence (its ability to

---

[35] Noel Sharkey, *Staying in the loop: human supervisory control of weapons',* cit. Nehal Bhuta and others, *Autonomous Weapons Systems: law, ethics, policy,* (Cambridge: Cambridge University Press, 2016), p. 26. Global consequences of AWS deployment (plausible denial in first use of AWS, proliferation, ethics and non-compliance) are generally ignored in any consequentialist analysis but form a key basis of argument in Chapter 5 (*Obstacles*).

[36] Noel Sharkey, *Staying in the loop,* p. 26.

[37] Ariel Conn, 'The Problem of Defining Autonomous Weapons', generally.

[37] US Department of Defense, 'Summary of the National Defense Strategies of the United States of America: Sharpening the American Militaries' Competitive Edge', US DoD, (2018) <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf> pp. 2-4.

[38] Paul Scharre, *Autonomous weapons and operational risk,* pp. 8-15.

[39] Sarah Knuckey, 'Autonomous weapon systems and transparency: towards an international dialogue', cit. N. Bhuta and others (eds.), *Autonomous Weapon Systems*, p. 164.

[40] Sensory data must first be acquired and processed. Targets must be then identified from that output and thereafter tracked, selected and prioritised before, on the basis of engagement rules, an engagement decision and resulting application of force can be undertaken. Technical ramifications are discussed in Chapters 7 (*Firmware*) and 8 (*Software).*

perform *all* required actions necessary to achieve a specific task), its epistemic competence (its ability to extract all necessary knowledge in order to act effectively and proportionately in each given context), appropriate action selection skills (an ability to design and select a plan of action that matches appropriately the weapon's goals) as well, Sartor concludes, as the AWS' ability to comply with relevant applicable norms.[41] Together, these must comprise a set of guidelines which informs a starting point for the introduction of AWS.[42] MacCormick, in particular, uses a reasonableness yardstick as a legal hurdle to deployment and points to the invariable requirement that 'impartial attention to competing values and evidences' be ensured in order to achieve compliance.[43] As later discussed, this looks in large part like an argument for maintaining meaningful human control in lethal engagements without which self-directed weapon functions come at the price of reduced predictability and reduced accountability.[44]

'Reasonable' deployment is also a central contextual matter in considering AWS adoption. It is therefore relevant to rehearse consequentialist reasons for that deployment.[45] Any such model must, after all, lever the promise of quicker, less expensive, more impactful weapon outcomes.[46] Given accurate targeting[47] and, perhaps, a more conservative decision process in advance of initiating engagement, use of autonomy should also lessen civilian harm.[48] This, however, is only one part of the deployment equation. A consequentialist framework throws up unexpected complications that must influence models. Sharkey, Professor of Robotics at Sheffield University, for example, is taking a broadly consequentialist view when he points out that fewer friendly casualties might equate to fewer State disincentives to initiate violence and start wars.[49] Similarly, deploying a swarm of aerial AWS may raise the prospect of a new arms race, frightened neighbouring nations and ensuing global destabilization.[50] AWS models must also factor in the disquieting prospect of

[41] Sartor and Omicini, cit. N Bhuta et al, *Autonomous Weapon* Systems, p. 63.

[42] MacCormick, 'Reasonableness and Objectivity', generally.

[43] Ibid., p. 1575 and p. 1581.

[44] For an analysis of the challenges to maintaining weapon predictability given AI use in weaponry, see: Chapters 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt*').

[45] Discussion of 'reasonableness' is prompted by Robert Stone, 'Puzzles of proportion and the 'Reasonable Military Commander': Reflection on the law, ethics and geopolitics of proportionality', *Harvard National Security Journal*, (2015), 332-334 <http://harvardnsj.org/wp-content/uploads/2015/06/Sloane.pdf>. A detailed review of drivers towards weapon autonomy is undertaken in the previous chapter (Chapter 3, *Drivers*).

[46] See, generally: Economist, 'Autonomous Weapons are a Game-changer', *Economist Magazine*, 25 January 2018 <https://www.economist.com/special-report/2018/01/25/autonomous-weapons-are-a-game-changer> [accessed 23 July 2018].

[47] See: Merel Ekelhof, 'Human control in the targeting process', ed. R. Geiß, *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law and Security*, (Germany: German Federal Foreign Office, Berlin, 2016), pp. 66–75. See also: explanatory slide deck, <https://www.unog.ch/80256EDD006B8954/(httpAssets)/AE38DEFD5D4D1E2CC1257F94004D4E33/$file/Presentation+CCW+M.+Ekelhof.pdf>. The model's dependence upon predefined target signatures suggests, in this instance, that it is currently more automated than autonomous. Similarly, current target prioritisation rules are based on predefined (and not dynamic) criteria.

[48] Tamburrini, p. 139.

[49] Noel Sharkey, 'Cassandra or the false prophet of doom', *International Review of the Red Cross*, Volume 94, number 886, (Summer 2012) <https://www.icrc.org/eng/assets/files/review/2012/irrc-886-sharkey.pdf>. Tamburrini highlights the destabilising potential of AWS swarms against opponents' nuclear arsenals, the impairment of an opponent's second-strike capability and the breakdown of traditional nuclear deterrence based on mutually assured destruction.

[50] Jason Le Miere, 'Russia developing a swarm of autonomous drones in the new arms race with US, China', *Newsweek*, 15 May 2017 <http://www.newsweek.com/drones-swarm-autonomous-russia-robots-609399> [accessed 2 January 2018]. Actually, evidence does not back up this assertion of a global arms race in autonomy; see: Boulanin and

asymmetrical warfare and poorer risk mitigation under these scenarios. Indeed, Kalmanovitz develops the position of Sharkey by suggesting that the most troubling aspects of AWS deployment may not be 'matters of deep ethical or legal principle but, rather, the lack of incentives for implementing effective regulations and accountability'.[51] As noted by Asaro, AWS deployment models cross a 'principled boundary' on a 'technological slippery slope'.[52] On these bases, it at once becomes difficult to posit a 'reasonable' deployment scenario where human supervision is *not* a necessary component of all battlefield processes. An analysis of targeting procedures corroborates this observation. As noted by Roff, target lists are inherently strategic and cannot be delegated to a machine without fundamental repercussions for battlefield command and organisational structure.[53] Roff terms this conundrum 'the Strategic Robot Problem'.

4.1 AWS' capabilities versus roles

Deployment models are also dependent upon the degree of technical divergence that exists between AWS *capability* and AWS *role* and the significant legal and operational ramifications that this departure has for AWS' compliant introduction. Consider, for instance, defence. What level of safeguard is required to deploy a defensive (albeit unsupervised) weapon that is narrowly programmed to engage specific enemy hardware in a specific battlefield area? Does restriction in a weapon's tasking (such as the SGR-AI discussed in Chapter Two) or the imposing of narrow defensive purpose influence the requirement for meaningful human control in lethal engagements? Johnson and Axinn conclude that an AWS securing a perimeter may be portrayed as defensive and perhaps, therefore, a 'more tolerable application of autonomy' as opposed to hunter-killer scenarios which are based on mobile weapon platforms 'in a manner that are spatially unbounded'.[54] Isolating the defensive and offensive capabilities of weapon systems might appear to be a relevant component in shaping deployment models for self-directing weapons. Levy similarly cites *mobility* and *striking power* as such deployment characteristics for an offensive weapon.[55] A third trait of *protection* is, in the case of AWS, relegated either to an economic denominator ('how many AWS can we afford?') or to a logistic denominator ('have we got enough and appropriate AWS in theatre?'). Within this argument, therefore, it might seem sensible to isolate AWS' offensive and defensive characteristics by their intrinsic features and, in so doing, ignore battlefield doctrines that would otherwise determine AWS use. Further analysis of Levy and Cole, however, refutes this notion. Here, it is not the characteristic of an individual weapon but rather 'the aggregate impact of that weapon system in a given arsenal' that should be the deployment determinant.[56] This conforms to this thesis' general emphasis on AWS' cumulative effects, a factor that merits additional weight given the likely *incremental* process of AWS implementation and the piecemeal removal of human

Verbruggen, p. 58. For a useful piece on the tactical appeal of unsupervised weapons, in particular weapons configured as a swarm, see: Emily Feng and others, 'Drone swarms versus conventional arms: China's Military Debate', *Financial Times*, 24 August 2017, generally. See also: para. 4 of 32 for an historical comparison <https://www.ft.com/content/302fc14a-66ef-11e7-8526-7b38dcaef6140> [accessed 23 February 2018].

[51] Kalmanovitz, cit. N. Bhuta and others, pp. 145-146.

[52] Peter Asaro, 'On banning autonomous systems', p. 687 and p. 707.

[53] Heather Roff, 'The Strategic Robot Problem: Lethal Autonomous Weapons in War', *Journal of Military Ethics*, Volume 13, Issue 3, (17 November 2014), 211.

[54] AM Johnson and S Axinn, 'The morality of autonomous robots', *Journal of Military Ethics*, 12, 2, (2013), 137-138.

[55] See: BHL Hart, *Aggression and the problem of Weapons,* (English Review, 55, 1932), pp. 71-73; and JFC Fuller, *Armaments and History,* (New York, Charles Schribner's Sons, 1945), generally.

[56] Jack Levy, 'The offensive/defensive balance of military technology: A theoretical and historical analysis', pp. 225-226.

oversight over a number of battlefield processes.[57] Moreover, a lack of definite milestones also complicates the definition of discrete deployment models which presuppose any such delineation between the defence and offence. Boggs cites Clausewitz to demonstrate the distinction's faultline given that every combat engagement whether 'great or small, is defensive if we leave the initiative to the enemy and wait for his appearance on our front door'.[58]

This chapter's deployment models are not, however, solely a mix of context and appropriate hardware. Challenge to the models also comes from the prerequisite that both IHL and IHLR be factors in order to establish which legal framework applies in each specific set of engagement circumstances.[59] This points to a further issue in AWS deployment. In no circumstances can the AWS make legal determination.[60] Deployment of AWS is not in itself an automated action but, rather, the deliberate decision of a group of human decision-makers.[61] In their deliberations, this group (hereafter termed the '*Delivery Cohort'*) is responsible, both morally and legally, for taking all reasonable steps to ensure that their deployment decisions comply with legal requirements. AWS deployment is, after all, an inescapably human matter. This observation requires that the Delivery Cohort fully understand in advance what is involved for IHL-compliant activation of their AWS as well as having adequate understanding of the operational boundaries required in order to meet these conditions. This is a complex set of conditions precedent[62] given that such models must be based on dependable testing and validation.[63] It is the epistemic uncertainties (principally concerning AWS technical feasibility as highlighted later in this thesis) that demonstrate that this is a challenging assumption. Indeed, it conceivably follows that military commanders and other constituents of the Delivery Cohort should refuse to deploy such weapons if material deployment criteria remain unmet. The point ignores complexity introduced by multi-national command as well as the politicisation of AWS deployment and other effects of politicians introducing themselves into this decision-making.[64] Finally to this point, cost factors also constrain deployment models.

---

[57] J Michael Cole, 'When Drones Decide to Kill on their own', *The Diplomat*, 1 October 2012, paras. 5-8 of 14 <https://thediplomat.com/2012/10/why-killing-should-remain-a-human-enterprise/> [accessed 12 October 2017].

[58] MW Boggs, *Attempts to define and limit 'aggressive' armaments in Diplomacy and Strategy,* University of Missouri 's studies*,* XVI, Number 1, (Columbia, Missouri, 1941), p. 68.

[59] Maya Brehm, 'Defending the Boundary; constraints and requirements on the use of autonomous weapon systems under international humanitarian and human rights law', *Geneva Academy of International Humanitarian Law and Human Rights*, Academy briefing No.9, (2017), p. 30. See, also: Chapter 5 (*Obstacles*), specifically: 5.1 (' *Geneva Convention and the Laws of Armed Combat'*).

[60] P Asaro, 'On Banning Autonomous Weapon Systems', pp. 2-4.

[61] The term *Delivery Cohort* is coined in Chapter 6 (Wetware, specifically *The Delivery* Cohort) to describe this decision group. It likely encompasses, inter alia, the following constituents: neurophysiologists to coordinate AWS networks, psychologists to coordinate learning and cognition, biologists for adaption strategies, engineers for control routines, logisticians, roboticists, electrical specialists, behaviorists, politicians, NGOs, sociologists, lawyers, company directors, weaponists, military tacticians, manufacturers, professionals involved in miniaturization, simulation, configuration, coding, power supply and modularity, specialists in sensors, in distributed and decentralized routines, ethicists, specialists in tooling and calibration. See also: D Floreano, 'Design, Control and Application of Autonomous Mobile Robots', *Advances in International Autonomous Systems*, (1999), pp. 159-186 <https://link.springer.com/chapter/10.1007/978-94-011-4790-3_8> [accessed 29 January 2017].

[62] See, for instance: Kelly Cass, 'Autonomous Weapons and Accountability: Seeking Solutions in the Law s of War', *Loyola of Los Angeles Law Review*, (4 January 2015), pp. 1031-1032 <https://digitalcommons.lmu.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=2941&context=llr> [accessed 15 August 2018].

[63] Tamburrini, cit. Bhuta and others, p. 127.

[64] Professor Lloyd Clark, in conversation with the author, June 2018.

Deployment costs exist, after all, in several guises from economic outlay, loss of legitimacy (should civilians suffer disproportionate harm), public opinion reversal (in cases of manifest error and the compromising of information security) and, clearly, in the scale of effort required to meet legal, technical and operational imperatives for inaugural deployment of such unsupervised weaponry. Kalmanovitz stretches this point by suggesting a further constraint on States' deployment activities, that of common interest and reciprocity; There is, he notes, 'the willingness to impose limits on one's own military means out of the expectation that one's enemies may resort to similar meanings and actions'.[65] It is against this broad context that specific deployment models can now be reviewed.

4.2 Planning tools

AWS must know *if* to decide, then *when* and *what* to decide.[66] Nearly one third of the systems identified in the SIPRI dataset already use Autonomous Target Recognition (ATR) as an autonomous 'decision aid' for human operators.[67] For the purposes of identifying possible deployment models, ATR is just one autonomous process that must be in place in order for the controlling agent (here, the commander or, eventually, the machine) to translate aims, goals and vision of an end state into a sequence of actionable tasks.[68] The programming challenge is that decision-making is both science and art. Several aspects of military operations (for instance, movement rates, fuel consumption and weapons effects) may be quantifiable and, notes O'Hanlon, reflect a *science* of war.[69] Other aspects (the impact of leadership, the general complexity of operations and uncertainty regarding enemy intentions), however, belong to the *art* of war.[70] The nub is that the starting point to such models may be the removal of supervision without compromising overall command and control in military operations. The starting premise, after all, is that operations, target selection and mission assignments will be significantly quicker using autonomy, thereby allowing commanders to respond better to changing situations.[71] What battlefield tenets then shape AWS deployment and are these canons equally common to academics and other constituents of the Delivery Cohort?[72] In this case, Boulanin and Verbruggen break up AWS deployment into a series of defined tasks centering on five decision characteristics. Their framework is based on *speed* (tasks involving cyber and air defence), *agility* (tasks that leverage a

[65] Kalmanovitz, cit. Bhuta and others, pp. 161-162.

[66] US Air University, 'The Military Decision Making Process', undated, p. 51 <http://www.au.af.mil/au/awc/awcgate/army/fm101-5_mdmp.pdf>.

[67] Boulanin and Verbruggen, p. 26. The capability is both pivotal to AWS deployment but also enduringly challenging. See Chapter 12 (*Appendix*), specifically: '*Case Study on Automatic Target Recognition*'.

[68] See also: GS Gill and JS Sohal, 'Battlefield Decision-Making: A Neural Network Approach'*, Journal of Theoretical and Applied Information Technology*, (2008), 697- 698 <http://www.jatit.org/volumes/research-papers/Vol4No8/5vol4no8.pdf>. The work's five attributes comprise Manpower Strength, Food and Ammunition States, Infantry Support, Air Support and Casualty Rates.

[69] Michael O'Hanlon, *The Science of War*, (US: Princeton University Press, 2009), pp. 1-4.

[70] Greg Simons. 'Understanding Political and Intangible Elements in Modern Wars', *Academia,* (2012), generally <http://www.academia.edu/2070261/Understanding_Political_and_Intangible_Elements_in_Modern_Wars> [accessed 5 February 2019].

[71] Amir Husain, 'AI on the battlefield: a framework for ethical autonomy'*, Forbes Technology Council*, 28 November 2016, para. 10 <https://www.forbes.com/sites/forbestechcouncil/2016/11/28/ai-on-the-battlefield-a-framework-for-ethical-autonomy/#71cdcb2c5cf2> [accessed 3 January 2018].

[72] Michael Mosser, *Puzzles versus Problems: The Alleged Disconnect between Academics and Military Practitioners*, (USA: Reflections, Volume 8, Number 4, December 2010), pp. 1077-1084.

reduced reliance on command and control) and *persistence* (tasks involving constant and consistent performance such as air defence, long ISR, enemy weapon countermeasures and tasks deep in enemy territory). Other deployment parameters revolve around general decision problems concerning *reach* (in communication denied environments) and *coordination* (force protection and the management of large groups of weapons).[73] Together, such parameters then comprise a framework within which models can properly be considered.

In considering AWS deployment, two relevant system categories exist. Weapon systems will either incorporate autonomy *at rest* (operating virtually, in software, and including expert advisory systems) or autonomy *in motion* (stand-alone machines that have a presence in the physical world including robotics, autonomous vehicles and tangible weapon systems).[74] Systems *at rest* include, for instance, monitoring and evaluation routines providing autonomous assessment to commanders, a key verification tool that can assist in swift in-loco decision-making. While this category of autonomous agent (here, an 'aid') may already be framing the actions and decisions of the commander, there exists an important distinction between commander *use* and commander *reliance* on such decision aids.[75] In *A History of the Future in 100 Objects*, Hon highlights what might happen when that line is crossed between reliance and over-reliance on technology. His illustration posits the finding by a machine learning system that attack is imminent, in this case giving the local commander just six seconds to launch an assault. In the event, human judgement is exercised, an armed response is delayed and the weapon's initial assessment is found to be plain wrong.[76]

Accepting, then, SIPRI's dataset as a plausible framework with which to consider deployment models (and assuming that the deployed weapon remains otherwise compliant), it is valuable to consider practical impediments to particular models. Given that an issue here is the degree to which applicability of the SIPRI framework can be assumed across both machine and capability type, two doctrinally defined products, the Course of Action (COA) sketch and statement, provide a relevant foundation to such analysis. Empirically, a commander develops multiple courses of action for deliberation prior to battlefield action, distinguishable from one another in force configuration and application as well as the designation of main and supporting efforts.[77] The core of the deployment model arises, of course, from the *subjective* stages of the COA that rely on the 'art of war' such that each decision can generally be deemed feasible.[78] Only then can military actions and military assets be coordinated to incorporate mission execution and, finally, an assessment of battle damage. Until this decision point, autonomous treatment of those early planning stages should plausibly improve results. AWS' deployment challenge thereafter is the selection, coding and processing of *qualitative* data (here, the 'knowledge of war'[79]) while achieving sufficiently

[73] Boulanin and Verbruggen, p. 62.

[74] US Department of Defense, Defense Science Board, p. 5.

[75] Alexander Kott and others, 'Decision Aids for Adversarial Planning in Military Operations', *arXiv preprint arXiv* 1601.06108, (2016), p. 3.

[76] Adrian Hon, *A History of the Future in 100 Objects,* (USA: Amazon/Kickstarter, 2016), generally.

[77] Alexander Kott and others, 'The decision Aides for adversarial planning', p. 4. This would also comprise consideration of, inter alia, terrain and resources, identification of objectives, target selection, intelligence support and final risk assessment.

[78] Ibid., pp. 4-6.

[79] John Saldano, 'The Coding Manual for Qualitative Researchers', *Sage Publications*, (2009), p. 2 <https://www.sagepub.com/sites/default/files/upm-binaries/24614_01_Saldana_Ch_01.pdf>.

repeatable outcomes to build trust between commanders, staff and the deployed weapon. As noted by Siniscalchi, even autonomous decision aids may be capable of state-defining actions that are not dissimilar to the deployment of full AWS while not necessarily directly initiating violence.[80]

The analogy between decision aid and AWS is revealing to AWS deployment models. The *planning* role of weapons autonomy is, after all, to ascertain the feasibility of particular courses of action, to assess their likelihood of success and to identify a range of possible executable actions and points of synchronization for participants.[81] Three deployment matters must therefore be factored into deployment models. First, the process is complex, resource consuming and dependent upon experiential (and thus subjective) inputs. This thesis' later technical review demonstrates that this is then difficult for a machine to master in isolation.[82] This is unsurprising: Regardless of rank and tasking, command and action selection is borne out of military experience built up over careers on the battlefield. While the COA is intended as a flexible but executable operational plan, the point is that it is the product of heuristics rather than strict rules-based processes.[83] The uncomfortable overlap between autonomous decision aid and wide capability AWS is best illustrated by the number of battlefield scenarios that defy being dealt with by any yes/no set of set of rules. Sharkey, moreover, notes that 'the number of such circumstances occurring simultaneously could cause chaotic robot behaviour with deadly consequences'.[84]

4.3 Machine and human teaming models

Just as force multiplication is a driver to removing human supervision from weapon routines, it also shapes the deployment models for such weapons. Its driver is founded upon the anticipated resilience, cost-efficiency and flexibility of *additional* autonomous hardware and, in order for such multiplication to be realised, Pellerin argues that it must first be harnessed to the selection, training, situational awareness and experience of the human commander.[85] Here, then, emerges the basis for the machine-human deployment model. In the case of force multiplication, the argument is that the proven effectiveness of the trained soldier can be leveraged by contiguous, accompanying hardware (designed as 'companion technology') that is designed to provide additional lethality.[86] The model's aim is to increase materially that soldier's firepower and effectiveness.[87] The technical challenge is not necessarily obvious. Mindell, Professor of Aeronautics and Astronautics at the MIT,

[80] See: Joseph Siniscalchi, 'Non-lethal Technologies: Implications for Military Strategy', *Air University, Maxwell*, Occasional Papers Number 3, (March 1998), pp. 2-3.

[81] Kott and others, p. 5.

[82] Saldano, pp. 4-8. See: Chapter 7 (*Firmware*), specifically: 7.4 ('*Attention methodologies'*). Also: Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues'*).

[83] Terry Wolff, 'The Operational Commander and Dealing with Uncertainty', *Army Command and General Staff*, Fort Leavensworth, (19 April 1999), pp. 6-13.

[84] Sharkey, 'Automated Killers and the Computing Profession', *Computer*, 40, 11, (2007), p. 122.

[85] Cheryl Pellerin, 'Work: Human-Machine Teaming represents Defense Technology Future', *Department of Defense Subscription*, (8 November 2015), paras. 1 and 3-5 <https://www.defense.gov/News/Article/Article/628154/work-human-machine-teaming-represents-defense-technology-future/> [accessed 1 September 2017].

[86] See: Susanne Biundot and others, 'Companion-technology: An Overview', *Kl-Kunstliche Intelligenz,* 30.1, (2016), pp. 11-20.

[87] Cheryl Pellerin, 'Human-Machine Teaming', paras. 5, 6-8. See also: Appendix E, *Measurement of Performance and Measurement of Effectiveness*, cit. *Making the Soldier Decisive on Future Battlefields*, (US: National Academies Press, 2012), pp. 166-170. Alternative source: <https://www.nap.edu/read/18321/chapter/11#169> [accessed 18 December 2018].

notes, however, that 'it takes more sophisticated technology to keep humans in-the-loop than it does to automate them out. History and experience show that the most difficult problem is not for autonomy but instead for the mixing of human and machine and the optimal amount of automation to offer trusted, transparent collaboration'.[88] This points to further inputs for the model, in particular around establishing an optimal human-machine ratio (here, the teaming ratio as it relates to teaming human soldiers with otherwise autonomous weapons).[89] This in turn is contingent upon the type and number of tasks to be executed, the nature and complexity of the team's operating environment, the sophistication of participating systems as well as the cognitive workload of the human operator.[90] As noted by Li, further intricacy is added to this challenge by each machine-human ratio being highly contextual.[91]

Deployment models based on teaming can also take many different forms, creating subsidiary models and adding to deployment's complexity. Notwithstanding that in *machine-machine* teaming, the most basic expression of weapon capability might be information-sharing, such collaborations will likely be armed, capable of selecting targets and initiating violence. 'Collaborative autonomy' gives rise to further complexity.[92] The challenge, notes Clark, is also reflected in commander selection at quite junior levels in order to identify individuals able to cope with the demands of teaming, a process that was formerly only undertaken at General Staff level and above.[93] As highlighted by Cuo, the model must also incorporate multiple battlefield systems in order that their actions are coordinated to achieve common goals.[94] To understand the nature of this challenge, the ratios between human operators and deployed colleague systems must be appreciated. In unmanned aerial systems this is currently *many*: 1 (that is, many human operators to just one UAV) and, for ground systems, some 2: 1. The purpose of this section is to review deployment models where those ratios now move either to 1: *many* or 0: *many*. Mindell's concept of a 'reliable human-machine ratio' points to what is a material impediment to empirical incorporation of autonomous capabilities.

A further constraint particular to AWS teaming is that configuration limits must be clearly delineated. This is not straightforward and, as noted by Chen and Barnes, requires implementation of either appropriate *collective* software architecture or implementation of an advanced 'system of systems'.[95] The implications of such complexity and arising 'technical debt' are explored in later chapters. Any such aggregation must, moreover, be sufficiently sophisticated to control a mix of

[88] David Mindell, 'Driverless cars and the myth of autonomy', *Huffington Post*, 14 October 2015, para. 12 of 14 <https://www.huffingtonpost.com/david-a-mindell/driverless-cars-and-the-myths-of-autonomy_b_8287230.html> [accessed 23 November 2017].

[89] Mustfa Demir and others, 'Team Synchrony in Human-Autonomy Teaming', *International Conference on Applied Human Factors and Ergonomics*, (January 2018), Abstract.

[90] Boulanin and Verbruggen, p. 68.

[91] Xin Li and others, 'Context Aware Middleware Architecture: Surveys and Challenges', *Sensors*, 15.8, (2015), p. 571.

[92] Kuntao Cui and others, 'The Collaborative Autonomy and Control Framework for Unmanned Surface Vehicles', *Frontier of Computer Science and Technology*, Ninth International Conference materials, IEEE publication, (2015), generally.

[93] Professor Lloyd Clark, in conversation with the author, January 2019.

[94] Boulanin and Verbruggen, p. 30.

[95] Jesse Chen and Michel Barnes, 'Supervisory control of multiple robots: Effects of imperfect automation and individual differences', *US Army Research Laboratory, Human Factors*, 54, 2, (April 2012), pp. 157-158 <https://pdfs.semanticscholar.org/4515/4ebb06b39ecdde11a820e6d7774a87af525e.pdf>.

unmanned aerial and surface systems, a swarm of low-cost systems operating as a coherent entity or the allocation of specific roles to each of that group's systems (thereby dictating an appropriate collective behaviour).[96] Complexity in this case arises from machine-machine capabilities that must comprise coordinated mobility[97], coordinated tasking and ISR (over, presumably, a large geographical area) as well as other collaborative actions, all undertaken within a dynamic anti-access and area-denial programme.[98] Challenges abound given that AWS must also be capable of *distributed* attacks whereby a higher-level UAS might act as a central authority that is able to identify targets before passing them off to lower-level but still autonomous units.

An emerging model derivative is therefore one that is based upon human-machine teaming. Current capabilities in this deployment category are quite primitive, either restricted to autonomous track-and-follow actions or the execution of straightforward pre-programmed manoeuvres as required by colleague human soldiers or other pilots.[99] SIPRI's database found no current evidence of sophisticated unmanned systems that are 'capable of acting as a loyal wingman'.[100] Indeed, it is the enduring nature of these outstanding technical challenges which supports HRW's contention that mastering such intricacy remains far-fetched.[101] Nor are such challenges restricted to broad model definitions. SIPRI notes that weapon systems incorporating human-machine autonomy include, inter alia, air defence systems, certain active protection systems, robotic sentry systems, certain guided munitions and loitering weapons. In order then to understand machine-human models, it is useful to review certain of these categories' constituents. Such platforms can be quite differentiated, specifically by the weapon system's possible range of engagement; in this way, the US' Phalanx System[102] may defend a ship by point defence while a missile defence systems such as Iron Dome[103] may protect a more substantial geographic block. How then do these precursor systems inform the introduction of broad-tasking AWS? Air defence autonomy, for instance, is currently restricted to the support of its host weapon's targeting.[104] Additional differentiation instead arises from both the type of targets that can be engaged as well as the range of autonomous countermeasures available to the weapon system. The point is that deployment models will vary by weapon type, capability and tasking, the assets involved in that collaboration but also the target set envisaged for that weapon. In the case of autonomous air defence, Turnbull notes that such model variation covers target detection and identification

[96] 'Appropriate' here refers to both the teaming model's efficacy, its operational leverage but also its ability to remain LOAC compliant.

[97] SIPRI identifies the UTAP-22 UAS developed in 2016 by Kratos that can fly autonomously in formation in several different configurations.

[98] Here, the SIPRI dataset identifies the CARACaS Project architecture (Control Architecture for Robotic Agent Command and Sensing) enabling a boat swarm fleet to conduct complex surveillance and security manoeuvres autonomously.

[99] Boulanin and Verbruggen, p. 33.

[100] Ibid., p. 34.

[101] Human Rights Watch and IHRC, 'Fully Autonomous Weapons: Questions and Answers', (October 2013), p. 2 <https://www.hrw.org/sites/default/files/supporting_resources/10.2013_killer_robots_qa.pdf>.

[102] See: Raytheon factsheet on Phalanx land-based weapon system, (2006), generally <http://www.mobileradar.org/Documents/Ray_Phalanx.pdf>.

[103] See: Rafael factsheet on Iron Dome land-based weapon system, 2007-2012 <http://www.rafael.co.il/SIP_STORAGE/FILES/6/3336.pdf, 2007-2012> [accessed 23 February 2017].

[104] J O'Halloran and others, *Jane's Land-based Air Defence,* (UK: IHS Jane's: Coulsdon, 2010), generally. The Russian S-400 Triumf (sic) can reportedly track more than 300 targets and engage with more than 36 targets simultaneously at a distance of up to 250kms.

(typically by using trajectory and velocity to ascertain target range and speed), target trajectory (tracking if the incoming weapon sticks to its predicted path) as well as autonomous target prioritisation and real-time IFF processes (Identification, Friend or Foe) in order that they be integrated into subsequent target engagement.[105]

Trying to construct a robust set of deployment models is therefore complicated[106] given SIPRI's observation that 'the one certainty is that there are many variables which need to be taken into consideration'.[107] An example of such an autonomous variable might include those exhibited by, say, Active Protection Systems (APS) that operate on a similar basis to air defence systems (combining a sensor system, a tracking, evaluation and classification system and a fire control system) in order to protect armoured vehicles against incoming hostile munitions. Similar to AWS, the deployment of APS will rely upon data from on-board sensors in order to understand incoming threats before engaging, usually autonomously.[108] Such application, however, is entirely defensive. There is no learning component to APS. These variables, however, can act to complicate AWS deployment. While not necessarily lethal, autonomous countermeasures include actions to distort the angle of an incoming missile's approach (thereby decreasing the chances of penetration) or to trigger prematurely the incoming projectile.[109] After all, it only requires a small operational stretch before the deployment, say, of robotic sentry platforms and gun turrets that can independently detect, track and engage targets.[110]

Haas and Fischer provide clarification to this deployment question by highlighting that weapon autonomy should be attached not to whole weapon systems but rather to the individual tasks.[111] In this vein, adjunct autonomous capabilities are important as they both affect *and* create deployment models. Attaching autonomy to individual tasks does not, however, reconcile the uncomfortable scenario where the weapon system is autonomous 'until a human intervenes'.[112] This is to ignore the increasing speed of machines (that empirically must leave human out-of-the-loop) as well as the challenging imperative that autonomous weapons are weapons that will learn from their

[105] Grant Turnbull, 'The realities of autonomy in unmanned aerial systems today', *Army Technology*, (9 February 2014), paras. 6, 14 and 28 of 33 <https://www.army-technology.com/features/featurethe-realities-of-autonomy-in-unmanned-air-systems-today-4175047/> [accessed 19 November 2017].

[106] George Dvorsky, 'Autonomous Killing `Machines are more dangerous that we think', *Gizmodo*, 29 February 2016, paras. 2-3 of 7 <https://gizmodo.com/autonomous-killing-machines-are-more-dangerous-than-we-1761928608> [accessed 12 January 2018].

[107] Boulanin and Verbruggen, p. 57.

[108] Ibid., p. 44.

[109] See: Battlefield Wiki, 'Active Protection Systems', <http://battlefield.wikia.com/wiki/Active_Protection> [accessed 21 July 2017]. Deployed examples include Israel's Trophy APS on its Merkava tanks. SIPRI's dataset lists only nine countries developing APS Technology and seventeen hard-kill APS products, all to degrees capable of autonomous target detection, identification, prioritisation and target engagement. As a category, APS reaction time is pivotal; just 300 milliseconds is available to intercept an anti-tank missile launched from 400 meters.

[110] SIPRI's dataset identifies just three such century weapons (Samsung's SGR-AI, Raphael's Israeli Sentry Tech and South Korea's DODAAM Super aEgis II. Guided missiles are excluded from this analysis as the sub-group is generally assigned targets in advance by human operators and only use autonomy to track, navigate or engage what is therefore already a pre-assigned (and therefore not autonomously selected) target.

[111] MC Haas and SC Fischer, 'The Evolution of Targeted Killing Practices: Autonomous Weapons, Future Conflicts and the International Order', *Contemporary Policy*, 38:2, (August 2017), pp. 283-286.

[112] Karen Petersen, 'General Concepts for Human Supervision of Autonomous Robot Teams', *Technische Universitat Darmstadt*, (23 May 2013), p. 31 and pp. 34-38 <http://tuprints.ulb.tu-darmstadt.de/3873/7/dissertation.pdf>.

surroundings. Capabilities such as a 'readjustment capacity' (allowing, for instance, flight path adjustment mid-flight) as well as autonomous target designation styles that are either *go-onto-location-in-space* (a particular geographic location) or *go-onto-target* (based on signature, heat/IR or radar) are merely a matter of instruction. They may evidence component autonomy but, for the purposes of defining deployment models, they do not comprise whole weapon autonomy.[113] Similarly, weapon tasking might permit the assignation of a *general* target area after which, in line with Haas and Fischer, the weapon autonomously find targets that match a predefined type.[114] Again, however, this ignores human's increasing inability to intervene. Moreover, all of the models thus far therefore rely on what are human-imposed criteria to manage that weapon's rules of engagement.[115]

4.4 Developing models for autonomous weapons

Casting wider for deployment models that are still akin to APS and sentry platforms leads to certain autonomous *loitering* weapons. In this case, operational utility is derived from the weapon's offensive tasking and its ability to be engaged in a geographical *area* rather than at a predefined target. While SIPRI defines loitering AWS as a discrete category which can 'conduct offensive and defensive missions that might be deemed dangerous or risky' for other weapon types[116], Sharkey notes that it is still the commander who must retain 'contextual and situational awareness of the target area at the time of initiating any specific attack'.[117] It also remains the local commander and his delegating staff (and dependent upon each scenario's context) who must 'perceive and react to any change or to unanticipated situations' that may have arisen since the planning of that attack in order dynamically to confirm that target's legitimacy.[118] This adds complexity. The suggestion is that an agent, whether the machine or the in-loco commander, must have active cognitive participation in each attack as well as sufficient time for deliberation on the nature of the targets as well as their significance in terms of the attack's overall necessity and appropriateness. In this case, Sharkey is arguing that the agent (again, whether machine or commander) must be able to abort or suspend any such action.[119] Similarly, any machine-human teaming should not prevent meaningful assessment of *incidental* effects of that attack.

---

[113] Horowitz and Scharre, *An Introduction to Autonomy in Weapon Systems,* p. 8.

[114] See: RAF factsheet, 'Dual-mode Brimstone', <https://www.raf.mod.uk/equipment/Brimstone.cfm> [accessed 23 June 2017].

[115] Kongsberg Gruppen, *Naval and Joint Strike Missile Update,* (USA: Kongsberg, 13 March 2014), pp. 18-22 <https://www.kongsberg.com/en/kds/products/missilesystems/jointstrikemissile/> [accessed 12 January 2017].

[116] Boulanin and Verbruggen, p. 54. Examples include the development (but not implementation) of the US Low Cost Autonomous Attack System (LCAAS), the US Non-Line-of-Sight Launch System (NLOS-LS) and the UK's Battlefield Loitering Artillery Direct Effect System (BLADE). By contrast, the Israeli HARPY system (and later derivatives) is in operation and performs in complete autonomy. Once launched, HARPY travels to a predetermined area and thereafter engages, on the basis of GPS and pre-programmed flight plans, potential targets.

[117] Noel Sharkey, *Staying in the loop,* p. 28.

[118] Ibid., p. 30.

[119] Noel Sharkey, *Staying in the loop,* p. 28. Paul Scharre has highlighted, however, that Sharkey's list could eliminate a number of conventional weapons already in use. The point here might be to look forward to technical developments and 'upgrade our sensibility to civilian harm as a result of [those] developments'.

A further challenge to this human-machine model is that collaboration must be based on common goals as well as appropriate 'buy-in' from the human in the loop.[120] The concept of 'centaur war-fighting' provides a relevant analogy. This arises from the game of chess where the player (here, the human commander) uses sophisticated software as an adviser but retains control of what move to make.[121] Translated to the battlefield, the partnering weapon in this human-machine team may be capable of lethal engagement but with the soldier performing three simultaneous roles.[122] He is the operator whereby the weapon system cannot effectively complete engagements without his participation. He is also the moral agent making dynamic value-based judgments about whether the use of force is appropriate and whether rules of engagement have been met. This too is not straightforward. Clark interestingly points to empirical differences (and the challenges arising) in decision-making and outcomes between male and female soldiers.[123] This evidences a critical distinction as the machine is enduringly incapable of making moral judgement. Third, technology must allow the human to act as fail-safe using his ability to intervene, alter or halt the weapon system's operation should it fail or should circumstances change such that the original engagement is no longer appropriate.[124] As noted by Scharre, however, these three role categories are not easy to maintain within the model. Removing completely the human's role as both moral agent and fail-safe in this relationship is challenging as 'humans have moral and legal judgments, responsibility and accountability, making their role as moral agents important for many tasks in war. Humans also have greater value as fail-safes, with the ability flexibly to respond to a range of unplanned scenarios'.[125] In *any* deployment model, after all, human-machine teaming is subject to the same legal, ethical and practice constraints that are in place for stand-alone soldiery. This linkage is important. As currently drafted, US Department of Defence *Directive 3000.09*[126] requires that, in the event of degraded or lost communications, weapon systems do not autonomously select or engage targets that have not previously been selected by an authorized human operator.[127] While this may appears unambiguous, system design round human-machine teaming must minimize the consequence of failure that could otherwise lead to unintended engagement or to other losses of control over that teaming system.[128]

How then can this collaboration model work in practice and what are its ramifications to weapon compliance? In the first instance, the model for machine-human teaming might instead be

---

120 Jeff Boss, 'The Army's New Decision-Making Model', *Forbes*, 8 August 2014, generally <https://www.forbes.com/sites/jeffboss/2014/08/08/the-armys-new-decision-making-model/#63585c991537> [accessed 8 December 2017].

121 Sydney Freedberg, 'Centaur Army: Bob Work, Robotics. & The Third Offset Strategy', *Breaking Defense*, 9 November 2015, paras. 4 and 9-14 <https://breakingdefense.com/2015/11/centaur-army-bob-work-robotics-the-third-offset-strategy/> [accessed 2 September 2017].

122 This chapter's analysis assumes such lethal capabilities.

123 Professor Lloyd Clark, in conversation with the author, February 2019.

124 Paul Scharre, 'Autonomous weapons and operational risk', p. 42. See also: Rami Debouck and others, 'Safety Strategy for Autonomous Systems', *Critical System Labs Inc*, Vancouver, undated, pp. 5-7 <http://www.system-safety.org/conferences/2011/papers/Safety%20Strategy%20for%20Autonomous%20Systems.pdf>.

125 Scharre, 'Autonomous Weapons and Operational Risk', p. 42.

126 See: Cryptome.org, <https://cryptome.org/dodi/dodd-3000-09.pdf>.

127 As above: Department of Defense, *DoD Directive 3000.09: Autonomy in All Weapon Systems,* (USA: Washington, 21 November 2012) <https://www.hsdl.org/?abstract&did=726163> [accessed 15 October 2017].

128 Paul Szoldra, 'An Ex-Pentagon Official thinks 'Killer Robots' need to be stopped', *Business Insider*, 9 March 2016, paras. 1-3 <http://uk.businessinsider.com/pentagon-autonomous-warfare-2016-3> [accessed 12 October 2017].

based around narrow-tasked machine-centric partnership. Pellerin then suggests that, further along this continuum, a model might comprise human-driven collaboration whereby those weapons draw on *several* sources for their operating priorities.[129] The purpose, after all, must be that autonomy enables weapons to integrate goals and actions in tandem with external information arising from the partnering human's background knowledge, his experience and situational awareness.[130] Such soldier-machine teaming, however, remains far removed from any model of AWS acting as properly 'social agents'. A relevant metric is the degree by which AWS action is *independently* shaped by wider battlefield context and deep background knowledge.[131] As noted by Chen and Barnes, the challenge is always to remove inconsistencies given that human-machine teaming is de facto based upon human intervention.[132] In the event of communications being momentarily lost between human supervisor and an otherwise autonomous system, at what point should the teaming weapon halt an engagement sequence (or, indeed, otherwise continue engaging targets as a fully autonomous machine)? Similar contradictions arise in determining protocol should the teaming weapon identify ambiguities or conflicts in its collaboration with its human colleague. The dichotomy here is that disrupted communications act both as a complication to weapon teaming as well as a key driver towards weapon autonomy.[133] Weapon reaction in the event of broken contact therefore becomes a relevant marker in deciding upon a particular deployment model. In this way, teaming weapons once isolated might be designed to stop activity (*failsafe* mode). It may also *fail dangerous* whereby the machine continues to engage targets that have been pre-authorised by human controllers. Furthermore, *fail-deadly* modes would, notes Masandu, allow the isolated weapon to engage 'emergent targets of opportunity' that have not specifically been approved by human operators.[134] Fail-deadly mode might also permit lethal force in defence. Seemingly straightforward human-machine teaming therefore masks deployment complexity. Notwithstanding outward participation of the human operator in teaming models, conflict in this case arises exactly because of complicating issues (breakdown in communication and thus a requirement to act alone) that put the machine in a position where whole-weapon autonomy may be required. As noted by Beautement, not only are these relationships neither delineated nor static but they also do not lend themselves to in-field monitoring given their mutability and the extent of overlap.[135]

4.5 **Flexible autonomy**

[129] Cheryl Pellerin, 'Human-Machine Teaming', paras. 6-8.

[130] Sean Kimmons, 'Mad Scientists' Discuss Emerging Technology and Army releases Strategy on Robots', *Defense Systems Information Analysis*, (10 April 2017), paras. 3-4 of 22 <https://www.army.mil/article/183862/mad_scientists_discuss_emerging_tech_as_army_releases_strategy_on_robots> [accessed 3 October 2017].

[131] Beautement, p. 9.

[132] Jessie Chen and Michael Barnes, 'Human-Agent Teaming for Multirobot Control', *IEEE Transactions on Human Machine Systems,* Vol.44, Issue 1, (February 2014) <http://ieeexplore.ieee.org/abstract/document/6697830/> [accessed 3 September 2017].

[133] Paul Scharre, 'Presentation at the United Nations Convention of Certain Conventional Weapons', Lecture, *Informal Meeting of Experts on Lethal Autonomous Weapons*, Geneva, (13 April 2015), p. 3 <https://www.unog.ch/80256EDD006B8954/(httpAssets)/98B8F054634E0C7EC1257E2F005759B0/$file/Scharre+presentation+text.pdf>.

[134] Nyagudi Musandu, 'Humnitarian Algorithms: A Codified Key Safety Switch Protocol for Lethal Autonomy', Nairobi, *arXiv preprint arXiv* 1402.2206 (2014), p. 6 and p. 11 <https://arxiv.org/pdf/1402.2206.pdf>.

[135] Beautement, p. 6.

Given the challenges of teaming an otherwise autonomous weapon to on-the-ground soldiers, a deployment model based instead on *flexible* autonomy might therefore be feasible whereby control of battlefield tasks, functions, and sub-systems might be passed, in theory, back and forth between soldier and weapon system as dictated by changing circumstances. The model continues to be widely investigated by parties[136] whereby a portfolio of frontline functions may be supported by varying levels of autonomy from fully-manual, decision-aiding, human-on-the-loop supervisory control to one that operates fully autonomously without any human intervention. Here, the commander will theoretically make informed choices about where and when to invoke such autonomy based on the considerations of trust in his teaming machine, the ability to verify its operation, the level of risk and risk mitigation available for operations and the degree to which partnering models are appropriate.[137] The issues, however, remain those of control (at what command level may these decisions be made and within what timescales?), feasibility, reliability and compliance. In circumstances when disaster is imminent, this same model posits that control will be removed automatically from the in-situ soldier and handed instead to the teaming weapon.

In order to assess the feasibility of flexible autonomy as a model, it is useful to review its operational characteristics.[138] Any such minimum features might, after all, be required across two, three or many teaming parties, raising issues around operator capabilities including situational awareness, informed trust, manageable workload levels and ease of interaction between colleague weapon and human party. Beautement uses the metaphor of the *machine as a team player*, 'a purposeful entity capable of contributing effectively (sic)'.[139] Two variables exist for this model. The teaming system may or may not be directly lethal. Given, moreover, that the model is predicated on flexible passing of control between man and machine, different levels of autonomy are likely to be appropriate at different times and within different teams. As such, the model contains several enduring challenges. How, for instance, can autonomy be wrested back from machine to human if that machine is incommunicado or autonomously engaged? Sharkey rightfully borrows from Kahneman to question whether the toggling of control between weapon and soldier can ever be workable given the very different methods of reasoning that these two teaming components will use in mediating such changes of control.[140] Several points arise from this uncertainty. This thesis' later analysis identifies material challenges to basing weapon reasoning upon any machine learning framework with, inter alia, ML's inappropriate suppression of doubt, disregard of ambiguity and poor inference of causes and intentions. Sharkey is particularly clear on this point around reasoning: 'An unambiguous answer pops up immediately and does not allow doubt. [It] does not search for alternative interpretations and does not examine uncertainty… If something looks like it might be a legitimate target in ambiguous circumstances, automatic reasoning will be certain that it is legitimate'.[141]

---

[136] US Air Force, 'Autonomous Horizons; system autonomy in the air force – a path to the future – human-autonomy teaming', *Office of the Chief Scientist*, AF/ST TR 15-01, (June 2015), pp. 9-15 <https://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf>.

[137] US Air Force, Office of the Chief Scientist, 'Autonomous horizons', p. v.

[138] IHLS, 'Global Powers at the Edge of Autonomous Battlefield Innovation', *IHLS*, (18 November 2017), para. 4 <https://i-hls.com/archives/79787> [accessed 30 December 2017].

[139] Beautement, pp. 3-6.

[140] D Kahneman, *Thinking, Fast and Slow,* (London: Penguin, 2011), generally.

[141] Sharkey, 'Staying in the loop: human supervisory control of weapons', cit. Bhuta and others, pp. 32-33. Also, in conversation with the author, 17 July 2017.

Research also endorses Sharkey in three important respects regarding the impact upon deployment models based on flexible autonomy. Lord, for instance, notes that decision aids founded upon automated reasoning will tend to promote automation bias (uncritical acceptance of suggested outcomes) and confirmation bias (seeking and then accommodating specific information to confirm a prior belief).[142] Parasuraman points out that the more reliable human operators judge their weapon system, the less rigourous will be their monitoring of those systems once deployed.[143] Finally to this point, research imputes that automatic reasoning will tend to return a suggested outcome *without* first considering contextual information that might be missing from inputs which are required in order for that engagement decision to be compliant.[144] In determining that automatic reasoning performs poorly where there is contradictory information on target legitimacy, Sharkey borrows Kahneman's term WYSIATI ('What You See Is All There Is') to describe this fault line. The essence of the problem becomes the behavioural issue of calibrating human control while still ensuring that the weapon remains a compliant and reliable battlefield asset. This is fundamentally a human challenge. Commenting generally on friendly-fire incidents, the US Army Research Laboratory admits 'how do you establish vigilance [after] twenty-three hours and fifty-nine minutes of boredom followed by one minute of panic?'[145] It is also a machine challenge to the extent that appropriate protocols are necessary in order to release control back to human parties.

Flexible autonomy models require that weapon control should *always* be partially manual. It is the human operator who must choreograph the task's overall performance while the weapon in turn carries out *specific* tasks. This may make theoretical sense but does not equate to weapon autonomy. Given these constraints, deployment models must also remain based upon coterminous generation of 'situation awareness support' given that sensed data is required in the partnership to drive decisions and goal states.[146] Under this refinement, further components are then required in order to allocate where control sits (between weapon and human) at every juncture. First, the weapon should autonomously provide a list of potential options and perhaps rank these options into a recommended target list or COA assessment. A second component must be robust supervisory control whereby the human operator sets goals but otherwise allows the weapon to control all aspects of function autonomously.[147] A final component for flexible autonomy might see the machine component of the teaming system having full control over all aspects of an engagement

[142] CG Lord and others, 'Human decision makers and automated decision aids: made for each other?', Raja Parasuraman, ed, *Automation and human performance: Theory and applications*, (USA: Mahwah, NJ, Laurence Erlbaum Associates, 1996), pp. 201-220.

[143] See: Raja Parasuraman and others, 'Performance consequences of automation-induced 'complacency'', *International Journal of Aviation Psychology*, 3,1, (1993). See also: Mica Endsley, 'Designing for Situational Awareness', (Boca Raton, CRC Press, 19 April 2016), 19-20 and 24 ('*Goals and Situational Awareness*').

[144] Claudio Bettini and others, 'A Survey of Context Modelling and Reasoning Techniques', *Elsevier*, (27 March 2008), pp. 3, 8 and 12 <http://ssltest.cs.umd.edu/class/spring2013/cmsc818g/files/bettinisurvey.pdf>.

[145] John Hawley, 'Automation and the Patriot air and missile defence system', *Centre for a New American Security*, Washington, (25 January 2017), p. 6 <https://www.cnas.org/press/press-release/cnas-releases-report-on-automation-and-the-patriot-air-and-missile-defense-system> [accessed 17 November 2017].

[146] US Air Force, Office of the Chief Scientist, 'Autonomous Horizons', pp. 6-7 and pp. 15-18.

[147] Ibid., p. 11.

function but with proven human ability to intervene.[148] In this manner, flexible use of autonomy becomes in theory a dynamic process dependent upon given battlefield tasks. The challenge, however, is that the level of that autonomy must change very rapidly according to battlefield developments. The model's basis should therefore be that the weapon system can be trusted to perform predictably the task at hand, where it is possible to confirm dynamically the autonomous system's performance and where probability of outsized negative risk is mitigated.

Flexible autonomy relies on the human operator being able to interact, understand but also, crucially, to *predict* his partnership with the autonomous system.[149] This prediction element is severally addressed in subsequent chapters. As noted by Hyndman, prediction routines remain enduringly difficult for machines and, for the purposes of determining deployment models, depend on important associations.[150] How well understood, for instance, are causal factors between the weapon's several data feeds and, second, how comprehensive and relevant is the sensed data collectable by the AWS for these prediction routines? A further constraint to this model then arises from such routines becoming circular whereby the weapon's output (the projections arising from its prediction routine) in turn affects the issue that the weapon is trying to forecast. Resolution here must take place notwithstanding that one or more (or all) of the weapon components (and, moreover, the soldier) might be 'operating outside its design assumptions' during certain combat operations.[151] Controlling a *group* of robotic weapons platforms creates further challenges to deployment models that are based around flexible autonomy. As inferred from Niccolini, fundamental uncertainty arises from multiple units having first to coordinate and then to process an exponentially larger (and not necessarily linear) set of weapon data points.[152] Regardless of the model, a requirement remains for complex interaction to make sense of what is an inherently dynamic environment in which teaming weapons must cooperate.[153] In this sense, it is difficult to envisage a commander having to substitute his group of traditionally trained soldiers for a troop of independent weapons. While a single robot must contend with uncertainty from sensors, effectors and representational conflicts, multiple robots deployed as a troop must also process uncertainty about partner robot states, partner actions and intentions, communications and plans. If the model is to be accretive to the Delivery Cohort, then rule sets must, after all, be sufficiently flexible to benefit from the experience and situational awareness that the human commander brings to the partnership. Teams of weapons are otherwise likely to interfere with each other's efficient battlefield operation, both in terms of physical space and, more importantly, through goal conflict and other tasking noise.[154] Predicting and managing in real-time this interference therefore become a key deployment challenge.

---

[148] The F-16's Automatic Ground Collision Avoidance System, for instance, continuously monitors for impending ground impacts and will execute recoveries at the last possible instant before returning control to the pilot. This, however, ignores counter-measures and adversarial actions.

[149] US Air Force, Office of the Chief Scientist, 'Autonomous Horizons', p. 12.

[150] Rob Hyndman, 'Why are some things easier to forecast than others?', *Hyndsight*, (18 September 2012), generally <https://robjhyndman.com/hyndsight/hardforecasts/> [accessed 23 June 2017].

[151] Daniel Serfarty and others, 'Adaption to Stress in Team Decision-making and Coordination', *Proceedings of the Human and Ergonomics Society*, 37, (1 October 1993), pp. 1228-1229.

[152] Marta Niccolini and others, 'Cooperative control for multiple autonomous vehicles using descriptor functions', *Journal of Sensor and Actuator Networks*, 3, (2014), 27.

[153] Maja Mataric, *The Robotics Primer*, TJ211.M3673, (USA: Massachusetts Institute of Technology, 2007), p. 234.

[154] Arquilla and Ronfeldt, 'Swarming and the future of conflict', pp. 45-47.

Nevertheless, flexible autonomy clearly posits operational advantages. In certain cases (operations requiring that large areas be monitored, occupied or denied), the SIPRI dataset suggests that a teaming model may out-perform in certain battlefield tasks.[155] Similarly, team operations may theoretically be better suited to multiple weapon platforms working in tandem. The model may also increase overall task robustness resulting, in part, from permitted redundancy.[156] As set out by Lachow, flexibly autonomous teaming might either be homogenous (multiple similar systems requiring less high-level coordination) or heterogeneous (with quite different and non-interchangeable weapon 'members').[157] Teams, however, whether loosely or tightly coupled, will still generate idiosyncratic complexity. Specifically, weapon platform relationships may change dynamically, either sequentially with new battlefield tasks or, more complicatedly, in-task depending on progress with that task. Teaming weapons in this model must either be deployed according to a global plan (itself a complication in a rapidly changing combat environment) or must create an ad hoc plan as team members (here, the human commander and colleague robotic weapons) coordinate their tasks and toggle autonomy accordingly. Task fluidity similarly complicates the deployment model. The human operator who is engaged elsewhere must unexpectedly take back control from an autonomous colleague system that has encountered a process or task problem. Termed the human *out*-of-the-loop control problem, Shahriari evidences that human bandwidth is particularly poor if the commander is suddenly required to change mental gears and diagnose a complex problem that was previously being addressed by the partnering machine.[158]

Two further challenges arise from the *coupled* nature of this model that questions its feasibility. As inferred from Chen, communication breakdown or incomplete data will have disproportionately adverse effect in the performance of weapons or teams of weapons operating in a flexible model.[159] This is unsurprising. It is, after all, the layering of communication (which weapon *receives* what, which weapon *received* what) that is a central deployment conundrum for planner and commander especially in an environment where an intelligent adversary is employing spoofing, decoys, cyber or electronic attack. As communication channels degrade, messages or parts of messages are likely to be lost or corrupted.[160] Given that data sharing is a fundamental precept of flexible autonomy, it follows that its deterioration must affect weapon outcomes. A second deployment problem is highlighted by Simmons and arises through the requirement for weapon

[155] Boulanin and Verbruggen, p. 36.

[156] R Arkin, 'Governing lethal behaviour: Embedding ethics in a hybrid deliberate/reactive robot architecture', pp. 115-119.

[157] Irving Lachow, 'The upside and downside of swarming drones', *Bulletin of the Atomic Scientists*, 73:2, (February 2017), p. 96 and pp. 98-99 <http://www.tandfonline.com/doi/pdf/10.1080/00963402.2017.1290879?needAccess=true> [accessed 25 June 2017].

[158] Bobak Shahriari and others, 'Taking the human out of the loop: A review of Baysian optimisation', *Proceedings of the IEEE*, 104:1, (2016), pp. 148-150 <https://www.cs.ox.ac.uk/people/nando.defreitas/publications/BayesOptLoop.pdf>.

[159] Min Chen and others, 'Machine to machine communications: Architecture, Standards and Applications', *Transactions on the Internet and Information Systems*, 6, 2, (February 2012), p. 481 <https://www.researchgate.net/profile/Jiafu_Wan2/publication/264846553_Machine-to-Machine_Communications_Architectures_Standards_and_Applications/links/550b9af60cf265693cef8967/Machine-to-Machine-Communications-Architectures-Standards-and-Applications.pdf>.

[160] Mataric, *The Robotics Primer*, p. 243.

coordination in teaming tasks. This is a longstanding dilemma in machine control and task delegation.[161] Using centralized control to enforce teaming will require precise information management whereby sensed data is collected, filtered, cleansed, checked and then prioritised before being distributed from a single source to all partnering entities. This is a complicated obligation that will add to the weapon's fragility. While this approach may allow team machines to compute an optimal solution to each problem in place, it empirically makes the centralized controller an unacceptable bottleneck in the model. More appropriate, therefore, might be the deployment of further layering that allows for decentralized collation and processing of that information. In such distributed control, each deployed weapon might then use its own controller to decide autonomously its own course of action.[162] Theoretically, teaming that involves autonomous agents will then be unaffected as a team grows or changes in size. Conversely, however, a deployment model based on distributed battlefield control requires appropriate collective behaviours to be defined, itself a bottleneck and additional source of error. Furthermore, Mataric is clear that such deployment behaviours should ideally be generated in a decentralized, non-planned fashion and directly through the interactions of individual weapon platforms that make up the team.[163] Finally to this point, this type of deployment model (involving multiple agents) will be disproportionately affected by what are local weapon behaviours that are being generated in order to optimize dynamics between the teaming units. Given the overriding importance of prediction to this model, it is precisely this extrapolation of team behaviour in, for instance, an autonomous weapon swarm (the subject of the following section) that will become ever more difficult to manage. In this case, Woods highlights the complexity of decision points, end points and, of course, the toggling of control requiring feedback and propagation mechanisms that must be shared throughout the team and other relevant control polities.[164]

A final deployment weakness for this model is identified by Mericli whereby post facto reconciliation of weapon activities and (given the model's inherently non-linear human intervention) and impact is difficult to attribute.[165] As inferred from El Deeb, the generally small number of weapon units comprising a teaming group will mean, moreover, that the data universe collected will likely be statistically irrelevant and likely prone to noise and error.[166] Just as toggling autonomous control between parties will impede team predictability, moving from local rules (one machine) to global behaviour (a swarm of machines) poses similar challenges. Egerstedt notes that

---

[161] See: Reid Simmons, 'Structured control of autonomous robots', *IEEE Transactions on Robots and Automation*, 10,1, (February 1994) <https://www.cs.cmu.edu/~reids/papers/structured.pdf>.

[162] Simmons, pp. 36-38.

[163] Mataric, '*The Robotics Primer*', p. 247.

[164] David Woods, 'Decomposing Automation: Apparent Simplicity, Real Complexity', *Automation and Human Performance Theory and Application*, Erlbaum, (1996), pp. 1-2 and pp. 3-6 <https://www.researchgate.net/profile/David_Woods11/publication/267402671_Decomposing_Automation_Apparent_Simplicity_Real_Complexity/links/546b62c60cf2f5eb18091bcd.pdf>.

[165] Cetin Mericli and others, 'Task Refinement for Autonomous Robots using Complementary Corrective Human Feedback', *International Journal of Advanced Robotic Systems*, 8.2, (2011) <http://www.cs.cmu.edu/~mmv/papers/11ijars-cetin.pdf>.

[166] Ahmed El Deeb, 'What to do with "small" data?', *Rants on Machine Learning*, 5 October 2015 <https://medium.com/rants-on-machine-learning/what-to-do-with-small-data-d253254d1a89> [Accessed 2 January 2018].

the inverse problem (when AWS' global behaviour toggles to local rules) is harder still.[167] The corollary is that it may be theoretically impossible to produce predictable group behaviour (that is required under LOAC) in a system based on ML that is both multi-party and flexible except in role-specific, narrow-task applications.[168]

4.6 Swarming model for AWS deployment

One such role-specific application is represented by the automation of small weaponised units into a swarm of self-directed armed drones.[169] Its deployment promise is that while individual elements may not themselves be threatening, they can be deployed autonomously in such numbers that, it is posited, they may be difficult to defeat. Hambling and others advocate AWS swarm characteristics of robustness, low-cost and rapid evolution notwithstanding that several of the control challenges identified in the previous section remain unchanged.[170] For the purposes of this section, swarming is best defined as a convergent attack from many directions. As identified by Arquilla and Ronfeldt, 'swarming is seemingly amorphous, but it is a deliberately structured, coordinated, strategic way to strike from all directions, by means of a sustainable pulsing of force and/or fire'.[171] How then might swarming impact upon AWS deployment models? Its theoretical advantages appear plausible. As noted by Scharre, multiple weapon systems operating as an autonomous pack could return 'mass' to the battlefield by augmenting manned combat systems with a large number of low-cost and unmanned systems that expand materially the number of sensors and shooters in the fight.[172] This in itself may be an important driver given ever-rising large-platform costs.[173] Instead, as precision-guided munitions proliferate amongst adversaries (both State and non-state)[174], Huis notes that a shrinking number of combat assets can itself become a strategic liability as adversaries concentrate increasingly accurate weapons on what is then an ever-smaller number of principal ships, bases and other high-value battlefield assets. Swarming might thus provide an appealing new paradigm.[175] As a deployment model, it hypothetically combines the highly decentralized nature of combat with mobility of manoeuvre and a high degree of organisation. In this manner, a qualitative

---

[167] Magnus Egerstedt, *Control of Autonomous Mobile Robots,* Handbook of Networked and Embedded Control Systems, (USA: Birkhauser Boston, 2005), pp. 532-534 <https://pdfs.semanticscholar.org/37cf/ab5a80cbb726799e1ebf0f4827db7585a48f.pdf>.

[168] Derived from: Mataric, *Issues and approaches in the design of collective autonomous agents,* (USA: Robotics and Autonomous Systems, 16, 1995), pp. 321-322 <http://crr.eng.auburn.edu/Bibliography/Mataric-Issues%20and%20approaches%20in%20design%20of%20collective%20autonomous%20agents.pdf>.

[169] Evan Ackerman, 'Lethal Microdrones, Dystopian Futures and the autonomous Weapon Debate', *IEEE Spectrum*, (15 November 2017) <https://spectrum.ieee.org/automaton/robotics/military-robots/lethal-microdrones-dystopian-futures-and-the-autonomous-weapons-debate> [accessed 2 February 2018].

[170] Hambling, p. 4.

[171] Arquilla and Ronfeldt, '*Swarming and the future of conflict',* generally.

[172] Scharre, 'Robotics on the Battlefield, Part II: The Coming Swarm', p. 6.

[173] For a useful primer on Force Multiplication, see Winslow Wheeler, 'Revisiting the Reaper Revolution, *Time*, (27 February 2012), generally <http://nation.time.com/2012/02/27/1-the-reaper-revolution-revisited/> [accessed 15 February 2016].

[174] Randy Huis, *Proliferation of Precision Strike: Issues for Congress,* (USA: Congress Research Services, R42539, 14 May 2012) <https://fas.org/sgp/crs/nuke/R42539.pdf>, pp. 7-8 and pp. 13-15.

[175] See, for example: Thales Aerospace Blog, 'The irresistible attraction of the drone', *Thales*, 5 April 2018, generally <http://onboard.thalesgroup.com/irresistible-attraction-drone/> [accessed 25 July 2018]. See also: Ed Yong, 'A bird-like flock of autonomous drones', *National Geographic*, 27 February 2014 <https://www.nationalgeographic.com/science/phenomena/2014/02/27/a-bird-like-flock-of-autonomous-drones/> [accessed 25 July 2018].

superiority of aggregate forces may be maintained but in a much-dispersed environment and across a greater number of platforms.[176] Swarming systems, moreover, can theoretically take considerably more battlefield risk, balancing survivability against cost. As above, greater numbers of incoming ordnance complicate an adversary's targeting priorities. This disaggregation of combat power into a larger number of less complex, less costly and less 'exquisite' systems also allows a 'family-of-systems' approach (in procurement terms) thereby increasing system diversity, reducing technology risk and, in theory, driving down costs.[177]

Swarming thus sits across those same deployment models discussed above, able to deliver lethality without supervision but also positing fundamental advantages. The swarm's ability to absorb casualties is a clear asset. A heuristic is that the morale of military units cracks on the battlefield when casualties reach a tipping point. This is a complex subject. Hambling contentiously cites a rate of some thirty percent:[178] When some third of troops are killed or incapacitated, the inference is that an attack empirically falters or defenders start retreating. With AWS, however, 'asset casualties' become less relevant. Similarly, while a single broken component can end a Reaper's mission, an autonomous swarm can continue regardless of its loss ratio. The mathematics of swarming is similarly compelling.[179] In this swarming model, combat power can be dispersed in order to force the adversary to expend more munitions on more incoming targets. Platform survivability is replaced instead with a concept of swarm resiliency.[180] As noted by Scharre, the mass of a swarming force allows the user to exhibit a 'graceful degradation of combat power as individual platforms are attrited, as opposed to a sharp loss in combat power if a single, more exquisite platform is lost'.[181] Similarly, the swarming model relies largely on overwhelming enemy defences such that 'leakers' get through, taking out the target. Notwithstanding very many variables, the model also posits saturation of enemy defences in order to exhaust both enemy logistics and replenishment. An adversary's guns can, after all, only shoot in one direction at any one time and currently require several seconds of engagement to deal with each target.

The power of the model lies in more than just the mathematics of greater numbers. Swarming's theoretical coordination is based on maintaining separation from nearest neighbours, steering with regard to the locus of neighbours (while attempting to move towards the average position of your neighbours) thereby keeping the swarm weapon flock together. Frei points to the swarm's 'self-healing networks'.[182] In this vein, Hambling notes that none of its members need be in overall charge with control therefore decentralized resulting, in theory, in no loss of cohesion during an

[176] Arquilla and Ronfeldt, 'Swarming and the future of conflict', pp. 75-79.

[177] Scharre, 'Robotics on the Battlefield', p. 6.

[178] Hambling, p. 182.

[179] Naval Postgraduate School in Monterey, 'UAV swarm attack protection system: Alternatives for destroyers', *Monterey Publishing*, (2012) <http://calhoun.nps.edu/handle/10945/28669> [accessed 4/8/16]. The modeled outcome of just ten incoming autonomous attack drones indicated that the defences of the destroyer would be overwhelmed by at least one significant hit.

[180] See: Edwin Ordoukhanian, 'Resilience Concepts for UAV Swarms', *CSSE*, USC Viterbi, (March 2016), generally <https://pdfs.semanticscholar.org/presentation/c95f/e77e25df2093fac4a5da723b43062c7e4d24.pdf>.

[181] Scharre, 'Robotics on the Battlefield', p. 14.

[182] Regina Frei and others, 'Self-healing and self-repairing technologies', *International Journal of Advanced Manufacturing Technologies*, Springer, (29 November 2012) <http://cui.unige.ch/~dimarzo/papers/JAMT.pdf>.

engagement: 'Half a swarm is still a swarm and still capable of all the same actions'.[183] There are also range benefits to swarming. It is calculated that potential distances achievable by the AWS in swarm formation will increase by the square root of the number of flyers in the formation.[184] This important model feature is evidenced by research from the US Air Force Air Vehicles Directorate which concludes that swarming units can achieve an eighty per cent increase in range over the distance they could fly alone.[185] The model's elasticity posits that rules-based operations can enable synchronized attack *or* defence as well as more efficient allocation of combat assets over a designated area. The model is already in evidence: The US Navy's low-cost UAV swarming technology LOCUST programme can manoeuvre multiple 'autonomous' units without those machines having to be individually controlled.[186] The U.S. Navy has similarly demonstrated swarm control software to manage unmanned surface vessels: Equipped with CARACaS software (Control Architecture for Robotic Agent Command and Sensing), robotic boats dutifully respond as a swarm when approached by a potentially threatening enemy ship.[187] Finally to this point, Strohn points out that Western armies do not currently have appropriate air-defence against such swarms.[188]

Such analysis, however, ignores technical and behavioural challenges to swarming deployment. As pointed out by Lachow, swarm coordination is fundamentally complicated as the required set of actions gives rise to unexpectedly complex behaviour.[189] The model also muddies attribution of responsibility; a team of weapons might be composed of individually understand units which together, as a swarm, deliver unanticipated outcomes (or simply fail). For this reason, swarm models remain in beta form.[190] Challenges to this model are also behavioural given the cognitive demands that will be placed on the Delivery Cohort (including, of course, the ceiling number of units that can be effectively be controlled *without* general autonomy). Edwards, after all, notes that even a theoretically optimal swarm model for command-and-control is neither obvious nor stable.[191] This arises from the model's requirement that it be both dynamic while still matching the level of swarm intelligence to the complexity of the underlying task. Swarming models must

[183] Hambling, p. 187.

[184] Ibid., p. 194.

[185] See: US Air Force Research Laboratory, 15 December 2014 <http://www.af.mil/AboutUs/FactSheets/Display/tabid/224/Article/104463/air-force-research-laboratory.aspx> [accessed 12 December 2016].

[186] See: popsci.com <http://www.popsci.com/navys-locust-launcher-fires-swarm-drones> [accessed 5 October 2016].

[187] David Smalley, 'The Future is Now: Navy's Autonomous Swarmboats can Overwhelm Adversaries', *Office of Naval Research, Science and Technology,* (2014) <http://www.onr.navy.mil/Media-Center/Press-Releases/2014/autonomous-swarm-boat-unmanned-caracas.aspx> [accessed 3 July 2016].

[188] Dr Matthias Strohn, in conversation with the author, January 2019. See also: UAS Vision, 'Suppressing Air Defenses by UAV Swarm Attack', *UASVision.com*, (28 June 2018), paras. 10 and 17 of 19 <https://www.uasvision.com/2018/06/28/suppressing-air-defenses-by-uav-swarm-attack/> [accessed 12 August 2018].

[189] Irving Lachow, 'The upside and downside of swarming drones', *Bulletin of the Atomic Scientists,* (4 March 2017), pp. 98-99 <https://thebulletin.org/2017/03/the-upside-and-downside-of-swarming-drones/>.

[190] For example, see: Harvard University, 'Self organizing systems research group' at <http:www.eecs.hardvard.edu/ssr> [accessed 31 August 2016], cit. Scharre, 'Robotics on the Battlefield part II: The coming swarm', p. 55. See also: Scott Maucione, 'Navy Wants to Cut Weapons Testing Time with Simulations and Modelling', *Federal News Radio*, 12 January 2018, generally <https://federalnewsradio.com/defense/2018/01/navy-wants-to-cut-weapons-testing-time-with-simulations-and-modeling/> [accessed 6 June 2018].

[191] Sean Edwards, 'Swarming and the Future of Warfare', *Pandee Rand Graduate School*, (2005), pp. 104-105 <http://www.dtic.mil/dtic/tr/fulltext/u2/a434577.pdf>.

depend, task by task, on appropriate information (arising, in this case, from each battlefield brief as well as the consequences of that brief). To retain compliance under LOAC, this must be enacted upon *in advance* of each mission. It must also account for any state changes encountered during the mission, the speed of reaction required by the swarm to adapt as well as the extent to which cooperation among swarm components is required in order to complete the task. Any deployment model must thus account for the degree of risk (both in terms of probability and consequences) of task failure.[192]

Other issues combine to undermine the swarm model. A trade-off exists between a swarm's attribute of mass and coordination on the one hand and the speed of decision-making for local commanders and policymakers on the other. Enduring behavioural obstacles remain notwithstanding, notes Sadler, that swarming methods continue to undergo material development and overhaul.[193] The challenge here to removing human supervision is that traditional tasks must then be defined and allocated in new ways.[194] The US military, moreover, is heavily invested, both financially and bureaucratically, in its *current* in-the-loop equipment and methods of fighting. Ferrell notes that only one dollar out of every twenty dollars spent by the US Department of Defense on R&D and procurement currently goes into developing unmanned systems.[195] As above, by 2018 global spending on military robotics is estimated to reach $7.5bn per year but this compares to the $43bn that is spent worldwide on commercial and industrial robotics. Sadler points instead to the flux that exists around swarming's command and integration issues.[196] Such mechanisms also require reliable high bandwidth to move data and instructions between host and swarm. As Scharre notes, 'the problem [here] is that the enemy gets a vote'.[197]

4.7 Operations and causes of failure in AWS models

Deployment models must also be reviewed through what Ingersoll terms 'the empirical lens of battlefield practicality'.[198] This requires taking measure of relative fragility within these models as well as the likely drivers to their operational failure. It requires review of ramifications arising from testing complexities and the challenges of validating and maintaining these models' integrity. For the purposes of this section, failure may occur at multiple points in a deployment model. There is

[192] Scharre, 'Robotics on the Battlefield', p. 40. This is discussed in this chapter's subsequent section: 4.7 ('*Operations and causes of failure*').

[193] Brent Sadler, 'Fast Followers, Learning Machines, and the Third Offset Strategy', *National Defense University Press*, Joint Force Quarterly 83, 1 October 2016, generally <http://ndupress.ndu.edu/Media/News/Article/969644/fast-followers-learning-machines-and-the-third-offset-strategy/> [accessed 23 August 2018].

[194] Although published in 1993, for discussion on task allocation and definition challenges, see: Myron Hura and others, 'Intelligence Support and Mission Planning for Autonomous Precision-guided Weapons', *RAND*, *United States Airforce, Library of Congress Publishing*, (1993), pp. 5-7, 14-16 and 42 <https://apps.dtic.mil/dtic/tr/fulltext/u2/a282344.pdf>.

[195] David Klein, 'US the Department of Defense 2015 budget analysis', *www.auvsi.org*, (May 2014) <www.auvsi.org/Mississippi/blogs/david-klein/2014/05/02/us-department-of-defense-2015-budget-analysis> [accessed 31 August 2016].

[196] Brent Sadler, generally. Also, Lloyd Clark in conversation with the author, 17 July 2018. See also: Michael Pilling, 'Issues regarding the Future Application of Autonomous systems to Command and Control', *Australian Government Department of Defense*, Joint Operations Division, DSTO-TR-3112, pp. 11-12 (June 2015) <https://pdfs.semanticscholar.org/f4cd/63e3db777cc2f43d98aa82875802a9079494.pdf>.

[197] Scharre, 'Robotics on the Battlefield', p. 11.

[198] Geoffrey Ingersoll and Robert Johnson, 'The 25 Most Effective Weapons in the US Arsenal', *Business Insider*, 14 December 2012 <https://www.businessinsider.com/most-effective-weapons-in-the-us-arsenal-2012-12?IR=T> [accessed 30 July 2018].

the risk of *practical* failure (not accomplishing a mission and therefore being a waste of time and resources), *technical* failure (the AWS fails to behave as envisaged) and *legal* failure (whereby LOAC and civilians are excessively compromised).[199] Understanding, also, the risks of low-probability, high-consequence events is an unlikely key to analysing AWS deployment models, the more so as militaries' track records in managing risks of this type are generally poor.[200] Commenting on NASA's 1986 Challenger accident, physicist Feynman highlights the wide set of views in articulating the probability of such accidents where estimates ranged from one-in-one hundred (engineering departments) to one-in-one hundred thousand (management departments).[201] The challenge that emerges is that it is difficult to quantify such risk. For AWS deployment, calculating risk of accident actually depends on that risk in any given instance multiplied by the number of *exposures* to that risk over a given period. Thus, a 1-in-10,000 chance of fratricide might convert into an impressive 99.99% safety rate which, if verified by testing, might lead policymakers to conclude that such a weapon system without supervision is appropriately safe. If, however, the number of potential weapon *interactions* with friendly forces in a combat environment is sufficiently large (the case with the US Patriot system and certainly the case with roving AWS), this would translate into an actual number of fratricides in the hundreds, enough to have operational impact and to distort popular contextualization of weapon autonomy. This is both a deployment and operational point. Even a very low probability of failure can result in an unacceptably high number of fratricides if the number of possible interactions with friendly systems is high.[202] Akin to the sale of lottery tickets, there exists a deployment paradox whereby even very low probability events can become effectively inevitable given enough exposure that then makes unlikely accidents 'normal' in complex (and, here, autonomous) systems. As Perrow explains of AWS, 'these systems are currently too complex and tightly coupled to prevent accidents that have catastrophic potentials. We must live and die with their risks, shut them down or radically redesign them'.[203]

This coupling phenomenon is pivotal to understanding deployment risk within AWS models. Ever more complex weapons are potentially vulnerable to system failure simply due to their components interacting in unexpected ways whether within the system itself, within human operators or within the weapon's deployment environment. Collins and Thompson point here what they term as 'risk non-linearity' to be a key challenge.[204] The phenomenon must influence the practical shape of deployment models, the subject of this section. The tightly coupled nature of AWS technology removes, for instance, any time slack in its routines, impinging on the exercise of external judgement or the 'bending' of rules that might otherwise alter AWS' behaviour. The challenge is that system failures within the AWS may rapidly cascade from one silo to the next with

[199] Cynthia Ferrell, 'Failure recognition and fault tolerance of an autonomous robot', *MIT, Adaptive Behaviour*, 24, (1994), p. 3 and pp. 4-5 <http://web.media.mit.edu/~cynthiab/Papers/Breazeal-AB94.pdf>.

[200] Scharre, 'Autonomous weapons and operational risk', p. 49.

[201] Richard Feynman, 'Volume 2: Appendix F – personal observations on reliability of the shuttle', p. 49 <http://history.nasa.gov/rogersrep/v2appf.htm> [accessed 12 August 2018] cit. Scharre, 'Autonomous weapons and operational risk'.

[202] Although written in 2004, see, generally: Bennie Sanchez, *Fratricide, Technology and Joint Doctrine*, (USA: US Naval War College, February 2004), http://www.dtic.mil/dtic/tr/fulltext/u2/a422756.pdf.

[203] Charles Perrow, *Normal accidents; living with high risk technologies,* (USA: Princeton NJ, Princeton University Press, 1999), p. 354.

[204] RJ Collins and R Thompson, 'Systemic Failure Modes: A model for Perrow's Normal Accidents in Complex, Safety Critical Systems', *Advances in Safety and Reliability*, (1997), pp. 361-363 <https://pdfs.semanticscholar.org/18d7/8946bc8bb1f58f0df6e57a7cce8fcf65f0aa.pdf>.

no chance for external agents either to react, reboot or otherwise intervene. In this environment, accidents will be inevitable and, suggests Scharre, become the new deployment 'normal'.[205] Gu's analysis using 'normal accident' theory suggests that in tight coupled complex systems, accidents may not only be likely but will be inevitable over a sufficient time horizon and, given unpredictable interactions within the AWS subsystems, the rate of such accidents is likely to be understated in advance.[206] Perrow's point is also that hidden failure modes may lurk undetected.[207] Underlining the issue's significance, UNIDIR has undertaken specific work to identify examples of potential failure paths in AWS, looking at interactions at machine, user and operator levels.[208] Their finding is that any of these factors might interact with any other (or others) to compound an initial failure. It is not, notes Borrie, solely the machine or machine code that is the relevant determinant in predictable weapon performance.[209]

Systemic weaknesses across deployment models are not simply conjectural. Even by 2006, seventy-seven robot-related incidents were reported in the UK in that year.[210] A worked example is helpful. The USS Vincennes is a Ticonderoga Class Cruiser equipped with the Aegis Special Weapon System intended to counter multiple air, surface and sub-surface targets with variable autonomy in its prioritising of these targets.[211] In July 1988, its systems wrongly identified an inbound civilian aircraft carrying nearly three hundred passengers as an Iranian F-14 fighter plane.[212] The systems then engaged the aircraft and there were no survivors. Subsequent investigations found that procedural errors had failed to reset the Aegis computer which was still displaying data relating to an earlier grounded F-14. Nor did this event represent 'peak complexity'.[213] As detailed above, AWS will operate in exponentially more complex environments that will increase materially the ways in which failures can occur, whether through incidents of incomplete information, the accelerated pace of interaction, unanticipated connections between component systems or through adversarial meddling.[214] The Vincennes tragedy thus provides a further marker for what might go wrong with weaponry without supervision. It, as well as the 2010 Flash Crash in the US equity markets, act as reminders that interactions between individually simple components can produce complicated and

[205] Scharre, 'Autonomous weapons and operational risk', p. 25.

[206] Weiquing Gu and others, 'Towards modelling the behaviour of autonomous systems and humans for trusted operations', *Naval Research Laboratory*, (2014), p.2 <https://pdfs.semanticscholar.org/89e1/a67f72d6feb4dc4b8e30d35dea626eda6c42.pdf>.

[207] Perrow, generally.

[208] UNIDIR here considers the overall *context* of the autonomous weapon including its environment, the behaviour of adversaries, the actions of friendly forces as well as the socio-technical background in which any weapon system is used.

[209] John Borrie, Chief of Research, UNIDIR, '*Security, unintentional risk and system accidents', UNIDIR briefing note*, Geneva, (15 April 2016), p. 2 and in conversation with the author, April 2016.

[210] Singer, *Wired for War*, p. 195.

[211] See: Walker, *Killer Robots?,* p. 85.

[212] Marten Zwanenburg, 'Human agents and international humanitarian law: dilemmas in target discrimination', (2008), <http://www.estig.ipbeja.pt/~ac_direito/ZwanenburgBoddensWijngaards_LEA05_CR.pdf>.

[213] Tom Goodwin, 'We're at Peak Complexity. And It Sucks', *TechCrunch*, (18 October 2016) <https://techcrunch.com/2016/09/03/were-at-peak-complexity-and-it-sucks/?guccounter=1> [accessed 23 July 2018].

[214] Scharre, *Autonomous weapons and operational risk,* p. 34.

unexpected effects.[215] Systemic risk can, moreover, build up in system configurations as *new* elements are introduced comprising risks that are not obvious until after deployment and after something goes wrong (and, notes Bostrom, sometimes not even then).[216] Context suggests that such coupling is inevitable. While personal computer software used to be crammed onto just 1 MB of RAM memory, the same processing and spreadsheet tasks now require a minimum of 256 MB of RAM memory.[217] Windows 95 required just 4 MB of RAM at launch but this increased to 32 MB on the introduction of Windows 2000. Windows 8 currently requires 1,000 MB of memory simply to turn on the computer.[218] Just as the general issue of 'technical debt' informs much of this thesis' later review on feasibility, it is also an important factor in deciding deployment models.

Risk of malfunction is therefore a key deployment consideration, framing the purpose of this thesis' technical review into AWS' intrinsic faultlines. Malfunction in all its forms is, however, only one further component of the Cohort's deployment equation. An adjunct complication is provided by enemy actions designed to foil such deployment. As inferred from Gherman, AWS communications, navigation, logistics and engagement processes all depend on reliable access to the battlefield's electromagnetic spectrum.[219] Koerner's conclusion is that it is only command of this spectrum that allows combatants to operate hardware from a remote location.[220] The models' converse is that preventing enemies from using this asset remains a critical goal. As highlighted by Paul, the point here for AWS deployment is that adversaries have strong incentives to hack into combat systems, either directly through malware or via behavioural interference to turn them on friendly forces.[221] While humans can exhibit inherent resilience against such hacking (after all, they can ignore orders and use common sense to adapt to the situation at hand), autonomous systems will likely lack flexibility to consider any such broader context.[222] In this case, effects might include compromised system integrity and unintentional system availability, occasioning errors in judgement, missed opportunities, inappropriate targeting, return-to-base sequences and other measures of poor performance. Additionally, an adversary may target AWS' behaviours whereby engineering a misleading set of circumstances might, notes Garfinker, '*mis-train'* that weapon's systems. Hacking might also compromise system confidentiality (matters of intelligence value such as command intent, mission orders, doctrine, rules of engagement and target parameters).[223] This

[215] See: Jill Treanor, 'The 'Flash Crash' of 2010: How it unfolded', *Guardian*, 22 April 2015 <https://www.theguardian.com/business/2015/apr/22/2010-flash-crash-new-york-stock-exchange-unfolded> [accessed 12 December 2017].

[216] Nick Bostrom, *Superintelligence; Paths, Dangers, Strategies,* (Oxford: Oxford University Press, 2014), p. 17.

[217] See: Computer History, 'Timeline of Computer History', <http://www.computerhistory.org/timeline/memory-storage/>.

[218] Hambling, p. 162.

[219] Laurian Gherman, 'Electromagnetic Spectrum Domination', *Review of the Air Force, Air Force Academy*, Romania, No. 1, 28, (2015), pp. 23-24 <http://www.afahc.ro/ro/revista/2015_1/23.pdf>.

[220] Brendan Koerner, 'Inside the new arms race to control bandwidth on the battlefield', *Wired Magazine*, 18 February 2014 <https://www.wired.com/2014/02/spectrum-warfare> [accessed 11 July 2016].

[221] Kari Paul, 'When Killer Robots Arrive, They'll Get Hacked', *Motherboard*, 24 February 2015, paras. 1-4 and 11 <https://motherboard.vice.com/en_us/article/ezvknz/when-the-killer-robots-arrive-theyll-get-hacked> [accessed 12 January 2018].

[222] Scharre, *Autonomous weapons and operational risk*, p. 39.

[223] Simon Garfinker, 'Hackers are the real obstacle for self-driving vehicles', *MIT Technology Review*, (22 August 2017) <https://www.technologyreview.com/s/608618/hackers-are-the-real-obstacle-for-self-driving-vehicles/> [accessed 18 December 2017].

may be exacerbated by AWS' modular nature and the broad manufacturing provenance of their components that have often been procured from commercial suppliers whose rewards are determined by designing *in* capabilities rather than designing *out* vulnerabilities.

How might deployment models be impacted by such adversarial actions? The technique of spoofing involves, inter alia, the creation of false GPS signal in order to trick the UAV on both location and time.[224] Moreover, the price of waging spectrum warfare has become cheaper in step with technical sophistication within retail electronics.[225] The battlefield consequences of quite simple electronic assault on AWS might therefore result in silent radios, inoperable drones and smart bombs that fail to find targets.[226] Weaknesses are similarly highlighted in the US National Security Agency's leaked report, *Threats to Unmanned Aerial Vehicles*, that sets out the risks to UAV deployment of commercially available 'lasers and dazzlers' in disabling drones by blinding cameras and sensors.[227] A second deployment complication arises from jamming, broadly defined as rendering electronically either a circuit or network unusable through disruption.[228] Such attack can be directed against any segment of the AWS' communication system and be of variable duration. It is often difficult for the weapon platform to distinguish between adversarial jamming and other information flow disruptions caused by malfunction, cryptographic reset, system management changes or other quite natural phenomena. This is a material deployment consideration. More than eighty per cent of current military transmission still travels on vulnerable commercial satellite communications channels and, as evidenced by Forest, only one per cent of defence communication is protected against even modest jamming.[229] Indeed, given exogenous influences, 'failure' instances are rarely easy to diagnose. Signal fratricide, after all, occurs when friendly antennas unintentionally overpower other friendly communications. Spectrum management is complex, contextual and requires laborious intelligence before mapping of an enemy's electromagnetic activity can be undertaken.[230] Whitlock similarly notes that lost link incidents, triggered when a

[224] Cockrell School of Engineering, 'School researchers demonstrate first successful spoofing of UAVs', 27 July 2012 <http://www.engr.utexas.edu/features/humphreysspoofing> [accessed 10 July 2016].

[225] Communications jammers can be readily assembled using power amplifiers and other off-the-shelf components. Similarly, GPS spoofers are universally available. For a useful discussion on spectrum warfare, see: Brendan Koerner, 'Inside the New Arms Race to Control Bandwidth on the Battlefield', *Wired Magazine*, 18 February 2014, generally <https://www.wired.com/2014/02/spectrum-warfare/> [accessed 12 April 2018].

[226] Koerner, 'Inside the new arms race to control bandwidth on the battlefield'. By 2009, Iraqi insurgents were using a commercially available program called SkyGrabber to intercept video feeds from Predator UAVs requiring American forces to encrypt its files on-going and introducing further complexity into battlefield drone operations.

[227] Barton Gellman, 'US documents detail Al Qaeda efforts to fight back against drones', *Washington Post*, 3 September 2013 <https://www.washingtonpost.com/world/national-security/us-documents-detail-al-qaedas-efforts-to-fight-back-against-drones/2013/09/03/b83e7654-11c0-11e3-b630-36617ca6640f_story.html?utm_term=.83448110bc4c> [accessed 12 December 2017]. The report also predicts that both State and non-state players might use rudimentary acoustic receivers to detect drones before interfering with their navigation and communication through simple jammer techniques.

[228] Ronald Wilgenbusch and Alan Heisig, 'Command and control of vulnerabilities to communications jamming', *JFQ*, ndupress.ndu.edu, Issue 69, (June 2013) <http://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-69/JFQ-69_56-63_Wilgenbusch-Heisig.pdf>.

[229] Benjamin Forest, *An analysis of the military use of commercial satellite communications,* (USA: Naval Postgraduate School, Monterey, California, September 2008), pp. 14-19 <https://calhoun.nps.edu/bitstream/handle/10945/3991/08Sep_Forest.pdf?sequence=1>.

[230] See: Walker, *Killer Robots?,* pp. 82-85. In addition to tracking emissions of every piece of friendly military hardware, spectrum managers 'must compile dynamic frequency lists that account for a galaxy of cheap civilian devices also used by friendly forces'.

satellite moves out of range or a machine otherwise drops such signal, are already common.[231] AWS deployment models must recognise that spectrum supremacy can never be more than *fragile* and is difficult to ascertain when it has even been attained; as noted by Koerner, 'there is no quick formula for evaluating when an enemy has been entirely ejected from an immense, invisible battle space'.[232] Equally, even a reeling opponent can rebound quickly given that an electromagnetic counter-attack requires technical (and invariably human) capital rather than high-value hardware. The importance of the problem to compliant deployment is illustrated by DARPA's current call for software that can recognise when radio waves are encountering interference.[233]

Lastly, all such deployment models must factor in *operational* in-field challenges such as appropriate AWS re-supply, ammunition replenishment, inventory control and, as above, in-situ servicing.[234] Given that AWS will be sensor and software-intensive, deployment considerations must include the requirement for frequent subsystem upgrades in order to maintain military advantage while also beta-testing (with appropriate safeguards) new capabilities. The conclusion is therefore that deployment models must themselves be challengingly dynamic. Likewise, they must be suitably flexible to reflect AWS reliance on rapidly developing technologies that include cyber warfare[235], protected communications[236], advanced computing and big data. Deployment models must thus ensure that such integration occurs without any performance variability and upon the invariable basis that autonomous weapons will be fault-free from their first deployment on the battlefield.[237] Operational conditioning, otherwise termed trial and error, is clearly an inappropriate basis for AWS deployment. Several pre-cursor systems appear, moreover, to work well in some but *not all* aspects of their operation. An example is the US Department of Defence's request for more than $8bn to fund its semi-autonomous Reaper procurement programme between 2010 and 2015.[238] Several fault lines would appear to disqualify the current system from fitting the deployment models set out in this chapter. To this point, the Reaper remains incapable of detecting other aircraft while in flight. Its operation must be curtailed in high winds, snow and rain. As a deployment model, it is unknown how the system will perform in a hostile airspace and the issue of its integration with manned aircraft also remains outstanding. The example highlights the importance of this chapter's findings that AWS deployment depends upon layers of contextual and technical challenge. It is upon this foundation that this thesis can now turn to consider other levels

---

[231] Craig Whitlock, 'US military drone surveillance is expanding to hot spots beyond declared combat zones', *The Washington Post*, 20 July 2013 <https://www.washingtonpost.com/world/national-security/us-military-drone-surveillance-is-expanding-to-hot-spots-beyond-declared-combat-zones/2013/07/20/0a57fbda-ef1c-11e2-8163-2c7021381a75_story.html?tid=a_inl> [accessed 10 June 2016].

[232] Koerner, 'Inside the new arms race to control bandwidth on the battlefield', generally.

[233] Jamie Lantino, 'DARPA to reinvent GPS navigation without the use of satellites'*, DARPA Information Bulletins,* undated <http://www.darpa.mil/program/adaptable-navigation-systems> [accessed 16 July 2016].

[234] Walker, '*Killer Robots?',* pp. 100-101.

[235] For an excellent overview, see: PW Singer and Allan Friedman, *Cybersecurity and Cyberwar: What everyone needs to know,* (New York: Oxford University Press, 2013), generally.

[236] See: Todd Harrison, 'The Future of MILSATCOM', (Washington: Centre for Strategic and Budgetary Assessments, 2013), generally. See also: Kenneth Neil Cukier, *Big Data: A revolution that will transform how we live, work and think,* (New York: H.M. Harcourt, 2013), generally.

[237] Inferred from: Jeff Guo, 'Google's new artificial intelligence can't understand these sentences. Can you*?', Independent*, *Indy100*, (May 2016) <https://www.indy100.com/article/googles-new-artificial-intelligence-cant-understand-these-sentences-can-you--Zy9gs38g7Z> [accessed 20 May 2017].

[238] US Department of Defence, 'Unmanned System Roadmap 2007-2032', p. 16.

of constraint that challenge AWS deployment. Chapter Five (*Obstacles*) reviews legal, behavioural, moral and ethical difficulties to AWS adoption before the thesis then turns in its following four chapters to consider technical obstacles frustrating compliant removal of human supervision from machines initiating violence.

5. Obstacles: General challenges to the removal of weapons supervision

There are reasons to assume that progress towards weapons autonomy will actually be quite muted.[1] This thesis' next five chapters seek to dissect fault-lines likely to impact on AWS deployment. The aim of this first chapter is to consider deployment impedimenta that are *not* rooted in technology.[2] Its first two sections assess the existing legal framework facing States considering AWS deployment. In subsequent sections it is then able to review political, behavioural and contextual factors that will impact the removal of human supervision from battlefield weapons while still remaining compliant to that legal basis. The chapter concludes by considering challenges to AWS deployment that are created by proliferation, by ethical constraints and, finally, by the overarching requirement for accountability if autonomous weapons are to be used. In order to isolate these behavioural and often intangible faultlines, the current chapter is making an assumption, later to be undone, that AWS deployment is technically feasible in order to question whether the politician or battlefield commander can reasonably devolve responsibilities to independent weapons. Sartor and Omicini capture the two opposing ends of this argument: 'On the one hand, it has been claimed that autonomous weapons may be as good as humans in implementing the laws of armed combat… On the other hand, it has been observed that few machines can today proficiently replicate human cognitive skills that are needed in many contexts to use force consistently'.[3]

An underlying premise for this thesis is that of an ongoing transformation in the character of warfare with evermore control of weapons being delegated to computer systems.[4] The magnitude of this assumption and what it may entail comprises much of the discussion below. Chapter Four, for instance, discusses several halfway houses from task-specific autonomy taking place in specific weapon sub-components to wholly unsupervised weapons operating independently whereby several choices of action are available. AWS function is neither linear nor obvious along that continuum. This chapter therefore aims to demonstrate that different factors come in and out of play as an individual weapon's tasking changes, often in direct response to legal, technical and behavioural constraints upon its deployment model. This, moreover, can happen moment by

[1] Walker, *Killer Robots?,* pp. 75-78 and pp. 103-107.

[2] These subsequent five chapters review AWS feasibility and comprise Chapters 5 (*Obstacles*), 6 (*Wetware*), 7 (*Firmware*), 8 (*Software*) and 9 (*Hardware*).

[3] Sartor and Omicini, cit. Bhuta et al (eds), p. 66. Sartor notes that machines already excel in a number of LOAC-relevant cognitive skills including the calculation of positions and trajectories, recognizing relevant patterns in large data sets as well as applying complex rules to a given situation. Conversely, the author also notes machine weakness in reading peoples' intentions and attitudes, anticipating behaviours, reacting creatively to unexpected circumstances, applying latitude and rule-deviation in exceptional circumstances and then assessing significance in battlefield gains, losses and other attributes.

[4] Models for this control shift are discussed throughout chapter 5 (*Deployment*), specifically: 4.3 ('*Machine and human teaming model*') and 4.5 ('*Flexible autonomy'*). Here, the mechanics of human-machine autonomies, hybrid systems, joint-cognitive systems, co-agency and other interaction models severally complicate any such deployment. See: E Hollnagel and DD Woods, *Joint Cognitive Systems,* (New York: Basic Books, 2005) <https://epdf.tips/joint-cognitive-systems-patterns-in-cognitive-systems-engineering.html> [accessed 1 July 2018]. Hollnagel and Woods consider in detail the matter of agency and the importance of weapon configurations in determining relationships between operator and technology. For consideration of the nature and developing character of war, see: Chapter 2 (*Context*), specifically: chapter introduction.

moment.[5] It is such discussion of non-technical challenges that is complicated throughout this continuum by difficulties around definition of terms and processes. This is unavoidable given the pace of innovation in AWS componentry as well as the maze of possible deployment models for independent weapon systems which, notes Sharkey, 'no one yet knows how to create'.[6]

At least one overarching factor acts as an anchor for this chapter's analysis: What is the legal bar that an unsupervised weapon must achieve in order to ensure that a target is properly legitimate? Weizmann from Geneva Academy provides appropriate narrative with which to frame the issue of these legal obligations whereby the AWS must be able 'to evaluate a person's membership in the State's armed forces (as distinct from a police officer), his or her membership in an armed group (with or without a continuous combat function), whether or not he or she is directly participating in hostilities, and whether or not he or she is *hors de combat*… [The weapon] would also need to be able to, first, recognise situations of doubt that would cause a human to hesitate before attacking and, second, refrain from attacking objects and persons in those circumstances'.[7] Given these difficulties around clear demarcation, discussion on parameters and their place in this analysis is relevant to understanding the complexity of AWS' compliance.[8] In the first place (and as evidenced in the previous chapter), analysis of weapons autonomy should not distinguish between AWS in a defensive mode and those intended for hostile application. Targeting, furthermore, should be defined as the weapon's determination between several presented objects as to which, when and where to engage chosen targets with chosen force. A requirement for AWS analysis requires, after all, that lethality take place without human involvement. This reading excludes remotely control weapons systems where a human operator has undertaken prior determination as part of a whole engagement process.[9] This is not always clear and, for the purposes of this chapter, is taken to mean where the human operator has (or, in the case of pushing a button, has not) meaningful participation in the process.[10] Similarly, it is useful to restrict parameters around the engagement process. Instances where autonomous technologies do not compromise LOAC (such as the autonomous refueling and autonomous navigation) are not relevant to this analysis. This treatment generally accords with broad commentary.[11] In order now to review legal challenges, deployment models are also ignored where a weapon is activated without supervision but in a very limited environment without civilian

[5] Paul Scharre, 'Making Sense of Rapid Technological Change', *Center for a New American Security*, (19 July 2017), generally <https://www.cnas.org/publications/commentary/making-sense-of-rapid-technological-change> [accessed 3 July 2018].

[6] Noel Sharkey, 'Staying in the loop: human supervisory control of weapons', cit. Bhuta and others, p. 23.

[7] Nathalie Weizmann, 'Autonomous weapon systems under international law, Academy briefing 8', *Geneva Academy*, (2014), p. 14 <https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf>.

[8] For a useful overview, see: Christopher Ford, 'Autonomous Weapons and International Law', *University of South Caroline Law Review*, 69. 5. Car. 413, (11 April 2017), pp. 413-424.

[9] This accords with definitions of weapon autonomy in Chapter 1 (*Introduction*). See, also: Robert Sparrow, 'Robots and respect: Assessing the case against Autonomous Weapon Systems', *Ethics and International Affairs*, 30, 1, (2016), p. 97.

[10] Mark Horowitz and Paul Scharre, 'Meaningful Human Control and Weapon Systems: A Primer', *Centre for a New American Security*, Working Paper, (March 2015), p. 2 and pp. 4-5.

[11] Robert Sparrow, 'Robots and Respect', generally.

exposure or where the weapon can be deactivated should targeting parameters change mid-engagement.[12]

5.1 The Geneva Convention and Laws of Armed Conflict

The St. Petersburg Declaration of 1868 was the first formal treaty prohibiting the use of certain weapons in war, in this case banning exploding bullets weighing less than four hundred grammes.[13] Since that date, the international community has attempted two structures to regulate new technologies in warfare.[14] International Humanitarian Law (IHL) broadly consists of a body of rules[15] that apply during armed conflict with the aim of protecting persons who do not, or no longer, participate in the hostilities.[16] Such rules regulate the conduct of hostilities.[17] IHL sets limits on armed violence in order at least to reduce suffering and, notes the ICRC, 'is based on long-standing norms that are rooted in the tradition of all societies'.[18] The rules of IHL have been developed and codified over the past century in international treaties, notably the 1949 *Geneva Conventions* and their *Additional Protocols* of 1977.[19] Droege notes widespread acceptance that human rights protection does not cease in times of armed conflict and that IHL and International Human Rights legislation (IHRL) apply concurrently.[20] Certain of these rights (including, inter alia, the right to life) are, moreover, *not* subject to derogation, whatever the circumstances. The crux in this case is that the extent to which AWS can legally be deployed therefore depends on this interplay between IHL and IHRL.[21] The distinction is relevant. IHL, for instance, tends to provide

[12] Christopher Ford, 'Autonomous Weapons and International Law', p. 419.

[13] For a full text of the agreement, see: IHL <https://ihl-databases.icrc.org/ihl/INTRO/130?OpenDocument> [accessed 23 July 2018]. The origin of the restriction was in Russia's 1863 invention of a bullet exploding on contact with soft substances. Given that such ordnance would have been 'an inhuman instrument of war', the Russian Government suggested that its use be banned by international statute.

[14] Jakob Kellenberger, 'International Humanitarian Law and New Weapon Technology', *ICRC*, 34th Round Table, (10 September 2011), pp. 1-3 <http://www.icrc.org/eng/resources/documents/statement/new-weapon-technologies-statement-2011-09-08.htm> [accessed 2 January 2017].

[15] As set out in its introduction, this dissertation focuses exclusively on State players (that is, nations) and their LOAC-compliant involvement of weapon autonomy. This is appropriate given current absence of evidence for non-State involvement. It is also beyond the scope of this paper to consider in detail 'hybrid warfare' (see: Lt Gen James Mattis and Lt Col Frank Hoffman, 'Future Warfare: The rise of Hybrid Wars', *USNI, Proceedings Magazine*, Volume 132/11/1233, (November 2005) <http://milnewstbay.pbworks.com/f/MattisFourBlockWarUSNINov2005.pdf>. It is noteworthy that the role for weapon autonomy within these doctrines has yet to be set.

[16] For example, civilians and wounded, sick or captured combatants.

[17] The published work of the International Review of the Red Cross is a key reference for this study; in particular, see: International Committee of the Red Cross, 'A guide to the legal review of new weapons, means and methods of warfare: Measures to implement Article 36 of Additional Protocol I of 1977', *ICRC, Geneva*, 88, 864, (December 2006), p. 932.

[18] International Committee of the Red Cross, '*The Basics of International Humanitarian Law – December 2017', ICRC*, (27 January 2018) <https://shop.icrc.org/l-039-essentiel-du-droit-international-humanitaire-2507.html> [accessed 1 May 2017].

[19] These are complimented by a number of other treaties dealing with specific matters such as international criminal justice and certain weapons.

[20] Cordula Droege, 'The interplay between International Humanitarian Law and Human Rights Law in situations of armed conflict', *Israel Law Review*, 40 (02), (2007), p. 311.

[21] For a discussion on how IHL and IHRL can be applied jointly in a complementary fashion, see: G Gaggioli and R Kolb, 'A right to life in armed conflict? The contribution of the European Court of Human Rights', *Israel Yearbook on Human Rights*, (2007), pp. 115-161.

stronger protection than IRHL against lethal force and the destruction of civilian property.[22]

The purpose of this section is to provide sufficient foundation to enable legal challenges to AWS deployment to be understood. It is not to do more than this, as the subject is complex and well documented.[23] Instead, this section focuses on the two frameworks as they relate to AWS deployment. It looks at obligations conferred by the Geneva Conventions.[24] It also considers the nexus that exists between, on the one hand, specific legal requirements arising from IHL and IHRL and, on the other, relevant technical bottlenecks that exist to the fulfillment (in AWS operation) of those parallel requirements. In this vein, the *Hague Conventions* of 1899 and 1907 are a series of international declarations, the first formal statement of the laws of war and war crimes in the body of secular international law.[25] They define the qualifications of belligerents including what must comprise AWS characteristics such as acceptably proportionate methods of engaging the enemy and, tangentially, the prohibition of pillage within seized territory as a result of war. The *Geneva Conventions* then comprise four treaties and three additional protocols that establish standards within the international transaction of war. Signed in 1949 by 195 countries, the documents define the basic, wartime rights of prisoners (*Convention* I), protections for the wounded (*Convention* II) as well as for civilians in a combat zone (*Convention* IV). The combined corpus is relevant to AWS deployment as a set of rules maintaining human dignity and protecting the vulnerable and defenceless during conflict. The point is that it is these restrictions which directly comprise the legal framework to which AWS must comply. Sitting underneath these two conventions is the *Law of War*, a legal term of art that refers to that aspect of international law that concerns acceptable justifications to engage in war (*jus ad bellum*[26]) and the limits to acceptable conduct once war is being fought (*jus in bello* and International Humanitarian law, the IHL detailed above).[27] An

[22] International Committee of the Red Cross, 'The Basics of International Humanitarian Law – December 2017', *ICRC*, (27 January 2010) <https://reliefweb.int/sites/reliefweb.int/files/resources/0850_002-IHL_web.pdf>.

[23] For a useful overview, see: Jarna Petman, 'Autonomous Weapon Systems and International Humanitarian Law: 'Out of the Loop'?', *Helsinki*, pp. 24-52 <https://um.fi/documents/35732/48132/autonomous_weapon_systems_an_international_humanitarian_law__out_of_the> [accessed 12 May 2017]. For a compendium of reports since 2012, see also: Human Rights Watch and Harvard Law School International Human Rights Clinic, 'The Need for New Law to Ban Fully Autonomous Weapons', *Memorandum to Convention for Conventional Weapons Delegates*, (November 2013), pp. 2-3 <https://www.hrw.org/sites/default/files/supporting_resources/11.2013_memo_to_ccw_delegates_fully_autonomous_weapons.pdf>.

[24] For a primer of the *Conventions*, see: International Committee of the Red Cross, 'Geneva Conventions and Commentaries', *ICRC,* undated <https://www.icrc.org/en/war-and-law/treaties-customary-law/geneva-conventions> [accessed 12 May 2017].

[25] By way of context, *Convention* II of the Hague's 1899 document specifies the treatment of prisoners of war, forbids the use of poisons and killing combatants who have surrendered as well as the attack of undefended towns. Convention IV codifies 'the prohibition of the discharge of projectiles and explosives from balloons or by other new analogous methods', 'the prohibition of the use of projectiles spreading asphyxiating poisonous gases' and 'the prohibition of bullets which can easily expand or change their form inside the human body'.

[26] Comprising 'proper authority and public declaration', 'just cause', 'proportionality', 'last resort', 'reasonable probability of success' and 'right intention'; See also: Keigh Abney, 'Autonomous Robots and the Future of Just War Theory', cit. *Routledge Handbook of Ethics and War*, eds. Fritz Allhoff et al, (UK: Routledge, Oxford, 2013), p. 340. The definition of *jus in bello* is best provided by the International Committee of the Red Cross; see: Jasmine Moussa, 'Can "Jus ad Bellum" override "Jus in Bello"? Reaffirming the Separation of the two Bodies of Law', *ICRC*, (31 December 2008), generally <https://www.icrc.org/en/international-review/article/can-jus-ad-bellum-override-jus-bello-reaffirming-separation-two-bodies> [accessed 4 May 2017]. *Jus in Bello's* position here is set out in the peer-reviewed Encyclopaedia of Philosophy: <http://www.iep.utm.edu/justwar/> [accessed 13 May 2017].

[27] For a useful discussion of Just War Theory, see: Keith Abney, 'Autonomous Robots and the Future of Just War Theory', cit. *Routledge Handbook of Ethics and War*, p. 339.

understanding of these concepts is therefore important to evaluate and then weight challenges to the compliant deployment of unsupervised weapons. Assessing then the technical feasibility of lethal machines conforming to LOAC then comprises the basis of subsequent chapters.

Conforming to such principles provides a broad obstacle to compliant AWS deployment.[28] In the first place, action selection for self-directing weapons must be capable of distinguishing between combatants and civilians.[29] Similarly, combatants who surrender must be spared from harm. The *Conventions* forbid methods (here, AWS action sequences) that inflict unnecessary human suffering or physical destruction. Even if an affirmative obligation is retained by colleague humans, AWS routines must still adhere to such obligations facilitating, for instance, the provision for wounded combatants and the sick to medical attention.[30] Similarly, AWS procedures must acknowledge that captured combatants and civilians are protected against acts of violence. As noted by Petman, the empirical implementation of these measures is complicated by lack of definition.[31] The LOAC framework is rarely unambiguous given its lengthy evolution and, as recognised in the US Army Field Manual, the 'customary' nature of its treaty law.[32] An example test from the US Army's *Field Manual* would be that an unsupervised weapon must 'refrain from employing any kind or degree of violence which is not actually necessary for military purposes and that they conduct hostilities with regards to the principles of humanity and chivalry'.[33] While LOAC may not specify that decision-making must be carried out by a human, a series of quite definite decisions is nevertheless required to verify targets before they are attacked, to establish how each target is to be attacked, to minimise civilian collateral damage, to confirm that each attack is proportionate, to establish that the prioritisation of targeting is in line with military advantage and necessity[34] and, finally, to incorporate assessment whether any attack should be suspended in light of the latest intelligence available.[35]

The challenge, however, is that ambiguity characterises much of the decision process which precedes an engagement sequence. Interpretation of legal frameworks, of operational norms and the constituents of the targeting sequence[36] are complex and, as discussed in later chapters, not at all conducive to coding. This uncertainty is itself an obstacle. Either those tenets have not recently been tested (whether in a combat environment or in a court of law) or the existing legal frameworks just do not cover emerging practices and technology. Dimitovski notes in this instance

[28] International Committee of the Red Cross, 'History of Humanitarian Law: The Essential Rules', *ICRC,* (2004), generally <https://www.icrc.org/eng/resources/documents/misc/5zmeem.htm> [accessed 23 July 2017].

[29] International Committee of the Red Cross, 'Basic Rules of the Geneva Convention and Additional Protocols', *ICRC*, (December 1988), generally <https://www.icrc.org/eng/assets/files/other/icrc_002_0365.pdf>.

[30] Combatants (in this case, the AWS) must also be able to distinguish the universal Red Cross with combat engagement on facilities and vehicles displaying this universal symbol being forbidden.

[31] Petman, pp. 15-24.

[32] US Army, 'Field Manual, FM 27-10 as amended', *Department of the Army Field Manual*, (6 April 1976) <https://www.loc.gov/rr/frd/Military_Law/pdf/law_warfare-1956.pdf>, paragraph 1.

[33] Ibid., paragraph 3.

[34] For ICRC's discussion on 'military necessity' and its sanction, see; International Committee of the Red Cross. 'Military necessity', *ICRC Casebook,* undated <https://casebook.icrc.org/glossary/military-necessity> [accessed 18 July 2017].

[35] Royal Air Force Directorate of Defence Studies, *Air Power – UAVs: The wider context*, p. 72.

[36] For the purposes of this dissertation, the targeting sequence is comprised of three components that run in succession: the *identification* of a target, the *selection* of a target to engage and the *deployment of lethal force* to kill that selected target.

that several parameters (operating at different levels of the engagement sequence) comprise each 'kill chain' and each decision tree.[37] Furthermore, these engagement parameters tend to multiply quicker than the decision waterfalls running these parameters can manage.[38] As evidenced in the preceding chapter, the degrees of human input across these chains will be both variable and ill-defined as well as being dependent upon the technologies involved in each such engagement. It is for this reason that Schulzke points to the impact that such instability (here, AWS inconsistency where behaviours and outcomes are the variable consequence of the weapon's machine learning) can have upon operator confidence given this explosion of parameters that must be incorporated in the engagement sequence.[39] This constraint is well captured by the Department of Defense's 2012 *Directive on Autonomy in Weapon Systems* whereby 'persons who authorise the use of, direct the use of, or operate autonomous and semi-autonomous weapon systems must do so with appropriate care and in accordance with the law of war, applicable treaties… and applicable rules of engagement'.[40] The conundrum is also evidenced by the counter-factual (put forward by Moyes and others) that the level of casualties still inflicted on non-combatants in armed conflict testifies to the on-going ambiguity, limitation and constraint of both IHL and its current framework.[41]

In weighting legal constraints on AWS deployment, Moyes also notes that previous control regimes have had a poor reputation.[42] In previously deployed weapon types, subsequent legal restrictions have tended to be weak, ineffective or illusory.[43] The ICJ's inconclusive 1996 ruling on nuclear weapons is an example of an institutional posture 'not to decide'.[44] This section therefore focuses on the *Convention's* existing provisions that are relevant to compliant deployment of AWS, in particular the strictures it creates that are based on proportionality and distinction. This gives rise to a further conundrum. While the rules of proportionality can be written down on a piece of

[37] Romy Dimitovski, 'Removing Humans from the Kill Chain: the Legality of (Semi-) Autonomous Weapon Systems under International Law', *University of Tilburg, Law Faculty*, (June 2017), pp. 19-20 <http://arno.uvt.nl/show.cgi?fid=142798>. For additional discussion on this term and its ramifications, see: Directorate of Defence Studies, 'Unmanned Aerial Vehicles – the legal perspectives', Wg Cdr Allison Mardell, cit. *Air Power – AUV: the wider Context,* p. 68.

[38] Walker, *Killer Robots?,* pp. 47-49.

[39] Marcus Schulzke, 'Autonomous Weapons and Distributed Responsibility', *Philosophy and Technology,* SpringerLink, 26, 2, (June 2013), pp. 203-219.

[40] Department of Defense Directive, 'Autonomy in Weapon Systems', p. 3.

[41] Brian Rappart, Richard Moyes and Thomas Nash, 'The roles of civil society in the development of standards around new weapons and other technologies of war', *International Review of the Red Cross*, 94, 886, (Summer 2012), p. 767.

[42] Tommi Koivula and Katariina Simonen, 'Arms Control in Europe: Regimes, Trends and Threats', *National Defence University, Helsinki*, Series 1, Research Publication 16, (2017), pp. 116-117 <http://www.doria.fi/bitstream/handle/10024/144087/Arms%20control%20in%20Europe_netti.pdf?sequence=1> [accessed 15 July 2017].

[43] For example, nuclear weapons are not subject to an explicit legal prohibition. The International Court of Justice in July 1996 reported as follows on the issue: 'The threat or use of nuclear weapons would generally be contrary to the rules of international law applicable in armed conflict, in particular the principles and rules of humanitarian law. However, the Court cannot conclude definitively whether the threat or use of nuclear weapons would be lawful or unlawful in an extreme circumstance of self-defence… so while the threat or use of nuclear weapons was generally held to be against international law, the judges could not determine that it would always be'. See: ICJ, 'Legality of the threat or use of Nuclear Weapons', ICJ *Reports*, (1996), p. 266.

[44] Louis G Maresca, '20 Years since the ICJ advisory opinion and still difficult to reconcile with international humanitarian law', *Humanitarian Law and Policy, blog*, 8 July 2016 <http://blogs.icrc.org/law-and-policy/2016/07/08/nuclear-weapons-20-years-icj-opinion/> [accessed 12 April 2018].

paper[45], the empirical calculation that must precede authorisation of an engagement (and under which, for instance, an autonomous weapon must assess in advance civilian collateral damage) is intrinsically complicated. In practice, the real-time analysis required to determine whether an attack is proportionate is profoundly contextual and cannot reliably be captured by current coding models.[46] The regime requires, inter alia, the attachment (as well as dynamic amendment) of variable values to targets, objects and even categories of human beings before calculating probabilistic assessments that must account for almost limitless contextual factors.[47] Similar to military necessity, Brown highlights that proportionality also pivots on subjective assessment of tactical advantage.[48] The whole process of analysis that attempts to define this advantage evidences the subject's complexity.[49] As noted by Driggs-Campbell, such processes revolve around probabilities and smart prediction.[50] This then posits a question of policy on how, where and how often such confidence thresholds should be reset in AWS routines. The complexity also encompasses the embedding of control routines that refine these inputs in real-time. Lede points out, furthermore, that such recalibration will be particularly challenging in an environment where communications have likely been denied.[51]

5.2 Proportionality and distinction in AWS deployment

Proportionality is a fundamentally *dynamic* construct. Priorities and available assets change moment-by-moment. Kalmanovitz notes the complexity that such flux introduces into judgement of proportionality: battlefield situations, he posits, are 'context specific because both civilian risk and military advantage are highly situational, uncertain, complex and dynamic, and they are holistic because the military advantage of tactical actions has to be assessed relative to broader strategic

[45] Weapons Law Encyclopaedia, 'Proportionality in attacks (under International Humanitarian Law)', <http://www.weaponslaw.org/glossary/proportionality-in-attacks-ihl> [accessed 12 April 2018]: 'The international humanitarian law of proportionality in attack holds that in the conduct of hostilities during and armed conflict parties to the conflict must not launch an attack against lawful military objectives if the iPAQ and may be expected to result in excessive civilian harm (death, injuries or damage to civilian objects, or a combination thereof) compared to the concrete and direct military advantage anticipated. If conducted intentionally, a disproportionate attack may constitute a war crime'.

[46] See: Chapter 8 (*Software*), specifically 8.1 (*Coding methodologies*). PAX uses the equation of 'one low-level terrorist versus three children?' to illustrate this conundrum; moreover, a slight change in circumstances (is the soldier's hands up in surrender?) might fundamentally change the legally-compliant response. See also: Walker, *Killer Robots?,* pp. 47-59: 'How, for instance, will a *collection* of autonomous machines, each self-learning based on individually set confidence limits and working without external interference, independently and identically assess the likely harm that may be caused to civilians and then decide on an engagement's proportionality in relation to any anticipated military advantage before undertaking that lethal action?'

[47] Based, for instance, on uniform, posture, actions, geographical position, trajectory and associations.

[48] Bernard Brown, 'The Proportionality Principle in the Humanitarian Law of Warfare: Recent Efforts at Codification', *Cornell International Law Journal*, 10, 1, (December 1976), Article 5, 140-142.

[49] See: UK Army, 'Land Operations', *Land Warfare Development Centre*, Army doctrine publication AC 71940, <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/605298/Army_Field_Manual__AFM__A5_Master_ADP_Interactive_Gov_Web.pdf> Chapters 5 and 8. See also: Nobuo Hayashi, 'Contextualizing Military Necessity', *Emory International Law Review*, 27, (2013), pp. 192-195.

[50] Katherine Driggs Campbell, 'Tools for Trustworthy Autonomy: Robust Prediction, Intuitive Control and Optimized Interaction', *Electronic Engineering and Computer Science, UCAL Berkeley*, (9 May 2017), pp. 5-9 <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-41.pdf>.

[51] See: Jean-Charles Lede, 'Collaborative Operations in Denied Environment', *DARPA*, undated <https://www.darpa.mil/program/collaborative-operations-in-denied-environment> [accessed 10 September 2018].

considerations'.[52] The importance of an opposition leader identified by an AWS in a crowded market square is not a constant and is likely to change on several levels as a battle unfolds.[53] The process requires flexibility, without which it will be compromised. Indeed, proportionality is already a difficult matter for the human chain of command to decide, usually requiring real-time inputs from several agencies.[54] It is for this reason that the US Air Force Judge Advocate General focuses on the matter of *context* when assessing the authorization of violence: As set out in *Air Force Operations and the Law,* proportionality is 'an inherently subjective determination'.[55] This is unsurprising given that tactical advantage is connected first to strategy and thence directly to political (and therefore human) goals. As notes in this instance by Clark, tactics generally serve operational ends and can often appear dislocated from strategic ambitions.[56] Later chapters will also argue that is likely infeasible for an algorithm to make such system adjustments that are nevertheless appropriately anchored to such a changing landscape.[57] It is for this reason that Kalmanovitz concludes that 'AWS estimates [must] necessarily belong to narrow tactical actions'.[58]

The issue is one of empirics and reasonableness. Anderson and Waxman point out that the AWS' ability to assess proportionality is both a *technical* issue (the design of software capable of measuring predicted civilian harm) as well as an *ethical* issue whereby weightings must be attached to relevant variables.[59] How many individuals should make up that grouping in that market square before the autonomous weapon considers it sufficiently significant to amend its action selection?[60] The setting of these threshold values would indicate that human judgement should remain an imperative in the process. It should also factor in potential effects from AWS malfunction. It is, after all, not the weapon that is making such judgement but human-made algorithms and human choices on threshold values that will frame weapon outcomes.[61] It will be a matter of indifference to victims whether the threat they are exposed to comes from manned, unmanned, supervised or

[52] Kalmanovitz, pp. 150-151. This is corroborated by the ICRC's conclusion that 'an attack carried out in a concerted manner in numerous places can only be judged in its entirety'.

[53] Here, for instance, on local, tactical and strategic levels as well as the prioritising of assets given limited resources and risk-tolerance levels for own force engagements. See also: Adam Rawnsley, 'CIA Drone Targeting Techniques', *Wired*, 7 August 2009 <https://www.wired.com/2009/07/infrared-beacons-guiding-cia-drone-strikes-qaeda-claims/> [accessed 1 April 2018].

[54] International Committee of the Red Cross, 'Decision making in military combat operations', *ICRC Publications*, (October 2013), pp. 13-14 ('*Commander's direction and review')* and p. *15 ('Evaluation of* factors') <https://www.icrc.org/eng/assets/files/publications/icrc-002-4120.pdf>. See, also, plot summary of Gavin Hood's *Eye in the Sky* movie, 2016, at <https://www.imdb.com/title/tt2057392/plotsummary> [accessed 12 September 2018].

[55] Tonya Hagmaier, *Air Force Operations and the Law; a Guide for Air and Space Forces*, (USA: Air Force Advocate General's School Press, First Ed., 2002), generally.

[56] Professor Lloyd Clark, in conversation with the author, January 2019.

[57] The issue here of anchoring relates to the degree of change that is appropriate given system experience and subsequent amendment to system settings.

[58] Kalmanovitz, p. 151.

[59] Anderson and Waxman, *Law and Ethics for Autonomous Weapon Systems,* p. 23.

[60] See, for example: Human Rights Watch, 'A wedding that became a funeral: US drone attack of marriage procession in Yemen', (2013), generally <https://www.hrw.org/report/2014/02/19/wedding-became-funeral/us-drone-attack-marriage-procession-yemen> [accessed 12 May 2018].

[61] This of course holds for machine learning in AWS routines. See: Chapter 7 (*Firmware*), specifically: 7.2 (*Firmware ramifications of machine learning'*) and 7.3 ('*Reasoning and cognition methodologies'*).

unsupervised weapons.[62] Establishing then whether such machine-generated predictions are reasonable will be intractably complex[63] and minimally demands that the weapon can navigate with acceptable care between contradictory values and interests on what is a fast-moving and contested battlefield.[64] It is also legally challenging to define reasonableness *ex-ante* an event. Nor can it be sufficient that the rules of engagement for an AWS rely solely on the test of a 'reasonable attacker' as proposed by Schmitt and Thurnher.[65] This is difficult ground as it assumes that the unsupervised weapon has processed *all* appropriate information for such a decision process. The relevant question for any such reasonableness test is therefore 'not what the machine should do but rather what the human beings who plan or decide upon an attack should do it before deploying AWS'.[66] Given the probabilistic nature of distinction and proportionality assessments, Schutzke is correct in noting that AWS routines must be governed by confidence levels that are anchored to relevant contexts, relevant eventualities and missions.[67] It is the nature of such weightings that they be dynamic, changing minute-by-minute, hour-by-hour. As inferred from McNeal, the critical decision for AWS' Design Cohort (here, the engineers, commanders and policy-makers) is once again how and how often these levels should re-based.[68] Ignoring for a moment the technical complexity of this requirement, this is not straightforward on any practical level. Will, moreover, thresholds be set differently for democratic States than non-democratic states? To what degree should popular opinions be factored into these thresholds? Will UK AWS have exactly similar threshold levels to French AWS?[69] How are differences to be understood and communicated?

The empirics of AWS' legal compliance also give rise to contextual complications. Game theory might posit a consequence whereby a State, in time, judges itself legally compelled to deploy its AWS assets if such weapons are deemed to confer military superiority to the side that first uses them. Krishnan notes in this the possibility of 'a new international dynamic whereby war becomes both increasingly automated and increasingly recurrent simply by the principle of necessity'.[70] A second constraint arises from human rights protection of life and physical integrity requiring not only that parties must refrain from arbitrary deprivation of life but also that *positive steps* must be

[62] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[63] Kalmanovitz, p. 152.

[64] For a useful primer of such adversarial combat environments, see: Alexander Kott, 'Challenges and Characteristics of Intelligent Autonomy for Internet of Battle Things in Highly Adversarial Environment', *US Research Laboratory, Adelphi MD, arXiv preprint arXiv*: 1803.11256, (2018), unnumbered, generally <https://arxiv.org/pdf/1803.11256.pdf>.

[65] Michael Schmitt and Jeffrey Thurnher,' Out of the loop: autonomous weapon systems and the law of armed combat', *Harvard National Security Journal*, 4, (2012), 279-281 <http://harvardnsj.org/wp-content/uploads/2013/01/Vol-4-Schmitt-Thurnher.pdf>.

[66] Kalmanovitz, p. 152.

[67] Marcus Schulzke, 'Autonomous Weapons and Distributed Responsibility', *Philosophy and Technology,* 26, 2, (2013), pp. 203-219.

[68] Gregory McNeal, 'Are Targeted Killings Unlawful? A case study in Empirical Claims without Empirical Evidence', cit. Claire Finkelstein, Jens David Ohlin, eds., *Targeted Killings: Law and Morality in an Asymmetrical World,* (Oxford: Oxford University Press, 2012), pp. 327-346.

[69] Clearly the application of force has much to do with the politics, national morale, historical tradition and cultures of warring states. See, for instance, the different stances adopted towards Germany by France and Great Britain between 1919 and 1939 (Professor Lloyd Clark, in conversation with the author, 13 September 2014).

[70] Armin Krishnan, *Killer Robots,* UG479. K75 (UK: Ashgate Publishing, 2009), p. 91.

taken to secure the right to life within States' jurisdiction.[71] The upshot is that AWS deployment must adhere to States' duty to control the use of lethal force including 'careful assessment of surrounding circumstances'.[72] This then becomes an important dynamic in States' deployment decision. An obligation of States, after all, might be to control their security actions so as to minimise recourse to legal force and incidental loss of life.[73] Case law, moreover, extends this obligation to conducting investigation upon individuals' deaths in order to secure accountability.[74] The obstacle for AWS deployment is that failure to fulfill these positive obligations constitutes a human rights violation. Legal and ethical assessment cannot simply be binary (is the system acceptable or unacceptable?) but instead must be contextual and specific to each environment.

In order to incorporate the concept of distinction into AWS routines, parties to a conflict 'must at all times distinguish between civilians and combatants. Attacks may only be directed at combatants. Attacks must not be directed against civilians'.[75] While this appears a straightforward obligation, compliance is complicated. As noted by Mattis and Hoffman, moves from State-on-State warfare to conflict types that are characterised instead by urban battles between civilian populations (but still, of course, bound by those legal constraints under LOAC) have made the distinction between legitimate targets and non-combatants very difficult.[76] Combatants in unconventional armed conflict may not be wearing uniforms or insignia, making it testing for weapon routines to judge whether an individual (or, more difficult still, an individual within a body of persons) should be categorised as a relevant combatant. Krishnan is succinct: 'Distinguishing between a harmless civilian and an armed insurgent would be beyond anything machine perception could (sic) possibly do. In any case, it would be easy for terrorists or insurgents to trick these robots by concealing weapons or by exploiting their sensual and behavioural limitations'.[77] Unambiguous definition of this 'civilian' within the legal principle of distinction does not exist and, again, frustrates any compliant paring of recognition algorithm and recognizing sensor.[78] In considering civilian populations, the 1949 Geneva Convention instead requires the use of 'common

[71] Jean-Francios Akandji-Kombe, 'Positive Obligations Under European Conventions on Human Rights', *Directorate General of the Human Rights Council of Europe*, Strasbourg, (January 2007), pp. 7-10 <https://www.echr.coe.int/LibraryDocs/dg2/hrhand/dg2-en-hrhand-07(2007).pdf>. For further analysis of positive obligations, see: Sandra Krahenmann, 'Positive obligations in human rights treaties', *Graduate Institute of International studies, Geneva,* PhD Thesis no. 949, (2012), generally.

[72] ECtHR, 'Natchova et al vs Bulgaria', *European Court of Human Rights,* App. nos 43577/98 and 53579/98, Judgement, (6 July 2005).

[73] Theresa Reinold, 'Sovereignty and the responsibility to protect', *Routledge Advances in International Relations and Global Politics*, (2013), p. 50 ('*The responsibility to protect*') and p. 119 ('*The duty to prevent*').

[74] ECtHR, 'McKerr vs United Kingdom', *European Court of Human Rights,* App. no 28883/95, (4 May 1995) and ECtHR, 'Al-Skeini vs United Kingdom', *European Court of Human Rights,* App. no 55721/07, Grand Chamber, Judgement, (7 July 2011), para. 164. See also: this chapter, specifically: 5.8 ('*Ethical and accountability constraints*').

[75] Rule 1, *St Petersburg Declaration,* <www.icrc.org> [accessed 30 December 2013]. Now codified in Articles 48, 51 and 52 of *Additional Protocol I of the Geneva Regulations* (against which, interestingly, no objects or reservations have been made).

[76] HRW, *Losing Humanity, the Case against Killer Robots,* p. 30. As above, it is beyond the scope of this paper to consider in detail hybrid warfare. See instead: Lt Gen James Mattis and Lt Col Frank Hoffman, 'Future Warfare: The rise of Hybrid Wars', *USNI, Proceedings Magazine*, (November 2005), Vol 132/11/1,233.

[77] Armin Krishnan, *Killer Robots,* p. 99.

[78] Suchman, 'Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy', p. 4

sense' while the 1977 Protocol 1 defines a civilian in the *negative* sense as someone who is not a combatant.[79]

A further confounding variant arises from whether a combatant has become hors de combat, a key tenet of battlefield law.[80] Under Article 41 of *Protocol* I, an individual is considered to be *hors de combat* if he is in the power of an adverse party, has clearly expressed an intention to surrender or has been incapacitated by wounds and is therefore incapable of defending himself (provided in all cases that that person abstains for hostile acts and refrains from escaping).[81] The test, in order to be LOAC compliant, is that the unsupervised weapon's targeting routines must perform these tasks at least as well as a human soldier; it cannot be assumed that everyone present in a war zone is a combatant. What then are the empirical challenges for AWS deployment? The self-directing weapon must at once be able to attribute *intention* to those in its immediate arena. Sharkey notes in *Killing Made Easy* that 'humans understand one another in a way that machines cannot; cues can be very subtle and there are an infinite number of circumstances where lethal force is inappropriate'.[82] Obligations for distinction, moreover, contain a further complication to AWS deployment. Adopting Goodman's position (whereby '[t]he purpose of military hostilities in warfare is per se not to kill combatants'[83] but to defeat the enemy, even if this requires the killing of combatants'), this posits AWS routines that can comprehend tactics, wider strategy and planning. It would obviate deployment models based on reflexive, cursory instructions that occasion, for instance, that there be 'no survivors' or for weapons to conduct hostilities on that basis.[84]

Sharkey points out that the principle of distinction is a particularly subjective process.[85] HRW's *Killer Robots* lists several scenarios to illustrate such engagement dilemmas, epitomized by the frightened mother running after her two children who are playing with toy guns near a soldier.[86] Adjuncts to this challenge arise. Unable to 'identify' with humans, Adams notes that AWS will be incapable of showing compassion, a likely check (suggests Adams) on the willingness to kill.[87] In this manner, while an AWS might self-authorise engagement of a child holding what is identified as a gun, a human soldier in the same circumstances might remember his own children and hold fire,

---

[79] Article 50 (1) of the *Protocol additional to the Geneva conventions* of 12 August 1949 and relating to the protection of the victims of international armed conflicts, (8 June 1977).

[80] ICRC, 'Customary IHL: Rule 47, Attacks against Persons Hors de Combat', *IHD Database*, undated <https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule47> [accessed 2 August 2018].

[81] *Protocol I,* Article 41(2); the ICJ in its case 'Legality of the Threat or Use of Nuclear Weapons' affirmed the importance of the Martens Clause 'whose continuing existence and applicability is not to be doubted' (*Advisory opinion*, 8 July 1996, para 87). The judges also found that the Martens Clause represents customary international law: Ibid., para 84.

[82] Noel Sharkey, 'Killing made easy', Lin, Abbey and Bekey, eds., *Robot Ethics*, (USA: MIT Press, 3 January 2012), generally.

118. A useful summary of the topic can also be found at: David J. Gunkel, Joanna J. Bryson, and Steve Torrance, 'The Machine Question: AI, Ethics and Moral responsibility', *Society for the Study of Artificial Intelligence and Simulation of Behaviour*, 978-1-908187-21-5, (2013) <http://www.cs.bath.ac.uk/~jjb/ftp/MQ2012-frontmatter.pdf>.

[83] Ryan Goodman, 'The Power to kill or capture Enemy combatants', *European Journal of Law*, 24, 2, (2013), 819–853.

[84] Director General for External Policies, European Parliament, 'Human Rights issues of the usage of drones and unmanned robots in warfare', *DROI* (2013), para 3.1.4, p. 26.

[85] Walker, *Killer Robots?,* p. 51.

[86] Human Rights Watch, 'Losing Humanity – the Case against Killer Robots', p. 46.

[87] Thomas Adams, *Future warfare and the Decline of Human Decision Making,* (USA: Parameters, 2001), pp. 4-6 and p. 8 <http://ssi.armywarcollege.edu/pubs/parameters/articles/2011winter/adams.pdf>.

seek the child's capture or avoid that child. While the issue's complexity may be well understood, technical solution to the hurdle is enduringly challenging. In this vein, the International Committee of the Red Cross took until 2009 to provide what is still convoluted guidance on how best to classify civilians engaging in hostilities[88]:

> In order to avoid erroneous or arbitrary targeting of civilians, parties to a conflict must take all feasible precautions in determining whether a person is a civilian and, if that is the case, whether he or she is directly participating in hostilities. In case of doubt, the person in question must be presumed to be protected against direct attack.[89]

This is clearly a difficult judgement for an experienced human soldier to undertake.[90] It becomes a question of technical feasibility, explored in later chapters, for machine systems to perform the same.

Other obligations arise from the *Conventions* that cumulatively create enduring obstacles to AWS deployment. As above, the matter of proportionality in violent engagement is covered by *Article 51* which states that 'an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated is forbidden'.[91] Its legal basis is corroborated by subsequent precedent: Under the *Statute of the International Criminal Court*, 'intentionally launching an attack in the knowledge that such an attack will cause incidental loss of life to civilians or damage to civilian objects… which would be clearly excessive in relation to concrete and direct overall military advantage anticipated… constitutes a war crime'.[92] State practices, furthermore, establish this ruling as a norm of customary international law, applicable in both international and non-international armed conflict.[93] The deployment issue for AWS is therefore unambiguous. How can the constituents of proportionality be coded such that an independent machine can comply with these legal requisites? How can these constituents then be appropriately governed within the situational awareness that is mandatory for proportionality's satisfactory practice? The deployment challenge is best summarized by the ICRC's own definition that proportionality is fundamentally subjective and 'above all must be a question of common sense

[88] Chris Cole, 'Towards the Next Defence and Security Review on the use of armed Unmanned Aerial Vehicles', *Submission from Drone Wars UK to the Defence Select Committee Inquiry*, 7th Report of the Session, 2013-14, 2, (18 December 2013) <https://publications.parliament.uk/pa/cm201314/cmselect/cmdfence/197/197vw.pdf>, p. 6.

[89] ICRC, 'Direct participation in hostilities: questions and answers', (2 June 2009), cit. Chris Cole, 'Towards the Next Defence and Security Review on the use of armed Unmanned Aerial Vehicles', *Submission from Drone Wars UK to the Defence Select Committee Inquiry*, p. 6.

[90] Frans Osinga, *Targeting: The challenges of Modern Warfare*, ed. Paul Ducheine and Michael Schmitt, (USA: Asser Press/Springer, 3 November 2015), pp. 193-195 ('*9.5.2 Subjectivity in targeting*').

[91] The principle of proportionality in attack is also contained in *Protocol II* and *Amended Protocol II* to the *Convention of Certain Conventional Weapons*.

[92] ICC Statute, Article 8(2)(b)(iv), UNTAET Regulation 2000/15.

[93] ICRC, *Rule 14, Proportionality in Attack* <www.icrc.org> [accessed 30 December 2016].

and good faith for military commanders'[94] and, as noted by HRW in *Losing Humanity*, any such test requires '*more* than a balancing of quantitative data'.[95]

5.3 Accountability in AWS deployment

Legal obligations together comprise what Beard terms AWS' 'framework of responsibility'.[96] In this vein, Sartor identifies clear accountability as the essential component in responsible weapon operation, linking functional failure with causality: 'Harm would not have resulted had the responsible component correctly exercised the function attributed to it'.[97] A second notion within this responsibility framework relates to 'blameworthiness' when harm occurs through a system fault ('a substandard behaviour in a moral agent', here defined as the AWS designer).[98] Other footings exist to this framework. Tort law, for instance, widely extends liability into areas that clash with removal of human supervision. These include causality of harm (strict liability), ownership or custody of the weapon that caused the harm, being 'principal to the agent who caused the harm' (vicarious liability) as well as to general product, design, negligence and organisational liabilities. Davey also points out that AWS' intent will be very difficult to determine given the human's increasing remoteness and the consequent diffusion of agency that blurs any attribution of purpose.[99] *Animus belligerendi* (the intention of a party to fight against a chosen opponent, in this case an armed AWS that has been specifically tasked) cannot be proved without impracticable real time analysis. Establishing direct connection between action and intention will likewise be complicated when these means of warfare are either 'spatially distributed' or 'temporally deferred'.[100] As noted by Marauhn, AWS' independence means that it will also be impossible to rely on robust legal links between AWS actions and State intent, itself a prerequisite in establishing violation of international law.[101] With action sequences depending upon precise deployment models, Marauhn notes that an erroneous lethal engagement by an AWS in a neighbouring territory would not necessarily constitute a legal trigger for conflict where that mistake was due solely upon insufficient care being taken on that weapon's targeting parameters.[102] Moreover, AWS programming must ensure throughout that those action sequences are governed by IHRL standards on the use of force and not just by the law of hostilities. For the law of hostilities to govern that use of force, sufficient *control* over the weapon's actions must be demonstrable in order to have that

[94] ICRC, *Rule 14, Proportionality in Attack* and attendant ICRC's authoritative commentary on the 1977 *Additional Protocol* (See: http://www.loc.gov/rr/frd/Military_Law/pdf/Commentary_GC_Protocols.pdf) generally but specifically pp. 231-237 ('*Identification'*) and 237-241 ('*Neutrals'*).

[95] Ibid. This author's *italics* for emphasis.

[96] Jack Beard, 'Autonomous Weapons and Human Responsibility', *University of Nebraska, Georgetown Journal of International Law*, 617, (2014), 642.

[97] Sartor and Omicini, p. 62.

[98] Ibid., p. 62. The definition here of 'fault' is broad; it may relate to design, the machine's inability to exercise with the abstract function attributed to it, faulty integration or inappropriate specification.

[99] Tucker Davey, 'Who is Responsible for Autonomous Weapons?*', Future of Life Institute*, (21 November 2016), para. 7 of 23 <https://futureoflife.org/2016/11/21/peter-asaro-autonomous-weapons/> [accessed 12 January 2018].

[100] The link is also difficult to establish when the weapons systems are 'victim-activated' such in the case of proximity or contact-mines.

[101] See: Thilo Marauhn, 'An analysis of the potential impact of lethal autonomous weapons systems on responsibility and accountability for violations of international law'*, Presentation*, *CCW Meeting of Experts on Lethal Autonomous Weapons Systems*, Geneva, (May 2014), pp. 1-2 <http://bit.ly/2dGUOzc> [accessed 12 March 2017].

[102] Marauhn, p. 5.

weapon classified as a targeted 'means of warfare' against another party. Brehm's point is that empirical deployment of unsupervised weapons makes it difficult to determine when AWS (or parts thereof) are being used as that relevant 'means of warfare' (used by Brehm to denote the method of lethal engagement).[103] This challenge is also exacerbated by the increasingly blurred distinction between war-fighting, policing and other State uses of violence.[104]

In considering the laws of hostilities as a constraint to AWS deployment, two questions arise. When does AWS deployment constitute a legal use of a means of warfare? Second, in the use of an AWS, what constitutes an attack to which legal rules on targeting then apply? Establishing in real time the appropriate framework (within which the laws on hostilities must apply) therefore becomes a further constraint on lawful deployment; the difficulty is that AWS' actions will be both context-dependent and, crucially, not under direct human control. The framework around 'means of warfare' is more tied to the IHL notion of 'attack' to which, mutatis mutandis, apply the *full* rules of targeting including obligations around proportionality, distinction and the need to demonstrate precautions in attack.[105] It is for this reason that Moyes questions how compliance can be ensured when neither scope nor context (nor, indeed, the spatial or temporal boundaries of specific acts of violence) can be understood from the outset of AWS' deployment.[106] Notwithstanding that several legal precepts in IHL and IHRL may remain ill-defined or in conflict with one another[107], States remain bound by obligations regardless of circumstance, a further intractable challenge to AWS deployment models.[108] It is therefore important to this analysis to acknowledge the overlapping components that constitute this legal framework.

At a supra-national level, it is the UN Charter's 'collective measures'[109] and arms control treaties that address how Nation States either initiate or forbid armed conflict, 'harmonizing the actions of nations in the attainment of common ends'.[110] These have clear consequences on the use of AWS and complicate their deployment. Would, for instance, cross-border use of autonomous weapons amount to *control* of that area such that the agent who is responsible for that weapon has established a jurisdictional link for extra-territorial application of IHRL treaties? By way of context, Brehm frames the issue as whether precursor use of unsupervised drones constitutes sufficient

---

[103] Brehm, p. 24.

[104] Keith Krause, 'War, Violence and the State', *Securing Peace in a Globalised World*, (2009), pp. 183-184 and generally <http://graduateinstitute.ch/files/live/sites/iheid/files/sites/admininst/shared/doc-professors/forthcoming%20war,violence%20and%20state%20HW70_ch9_krause_corrected%5B1%5D.pdf>.

[105] Such conditionality also includes *feasible* measures to cancel or suspend that attack should a target's legality or proportionality become questionable.

[106] Richard Moyes, 'Key Elements of meaningful human control', *Article 36 Briefing Paper, CCW Meeting of Experts on Lethal Autonomous Weapon Systems*, (April 2016), p. 3 <http://article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf>.

[107] The onus is on the State to show that exceptional circumstances exist that limits its responsibilities. See: ECtHR, 'Sargsyan vs Azerbaijan', *European Court of Human Rights,* App. no 40167/06, Grand Chamber, Judgement, (16 June 2016), paras. 126-131.

[108] D.Hart, 'War remains inside the courtroom: Jurisdiction under ECHR', *UK Human Rights Blog*, 11 September 2016 <https://ukhumanrightsblog.com/2016/09/11/war-remains-inside-the-court-room-jurisdiction-under-echr/> [accessed 18 June 2017]. The US, however, has always rejected that ICCPR applies outside its own borders.

[109] In particular, Sections 1 and 7, UN <http://www.un.org/en/documents/charter/index.shtml> for full text [accessed 12 June 2017].

[110] See: *United Nations Charter 2006*, Section 3 <http://www.un.org/en/sc/repertoire/2012-2013/Part%20III/2012-2013_Part%20III.pdf>.

control over either area or individuals that would otherwise be akin to effecting custody.[111] Use of AWS thus creates new legal complexity. This is evidenced by Melzer who notes that 'a State exercising sufficient control or power to carry out a targeted killing will also exercise sufficient factual control (sic) to assume legal responsibility for its failure to respect with the right to life of the targeted person'.[112] A premise for compliant deployment of AWS may be that 'seeing and knowing triggers obligations' on the basis that sufficient control over territories and persons evidently exists. As such, extra-territorial application of human rights law would be triggered by AWS deployment, greatly complicating the placement of independent weapons without human supervision and leading Haas and Fischer to argue that the *entire* process of targeting represents a form of control by very definition.[113] In arguing the violation of LOAC regardless of deployment model, Brehm points to broad legal consensus that neither the level of lethality nor the duration of hostilities affect the immediate application of IHL to armed conflict.[114]

5.4 Martens Clause

In considering legal impediments, two further legal nuances require review. The Martens Clause[115] has formed a part of the laws of armed conflict since its appearance in the preamble to the 1899 Hague Convention.[116] *Protocol* I Article 1(2) states that means of warfare (and, by extension, AWS deployment) must be evaluated according to the 'principles of humanity' and the 'dictates of public conscience'. While there is no certain interpretation of the Clause[117], at its most restricted it serves as evidence that customary international law is over-arching and continues to apply in all battlefield circumstances.[118] A wider interpretation of the Clause is that conduct in armed conflict should not only be judged according to treaties and custom but also to the principles of international law as referred to by the Clause.[119] This interpretation is borne out by the International Law Commission's 1999 statement that the Martens Clause 'provides that even in cases not covered by specific international agreements, civilians and combatants remain under the protection and authority of the principles of international law derived from established customs, from the principles of humanity and from the dictates of public conscience'.[120] Building upon this clarification, Queguiner then highlights other requirements[121] whereby commanders must follow legal precautions in order to remain LOAC-compliant.[122] It is Article 57 that deals widely with the

111 Brehm, p. 20.

112 Nils Melzer, *Targeted Killing in International Law*, (Oxford: Oxford University Press, 2008), p. 20.

113 Haas and Fischer, generally.

114 Brehm, p. 22.

115 Walker, *Killer Robots?,* p. 57.

116 See: International Committee of the Red Cross Resource Centre, <www.icrc.org> [accessed 30 November 2015].

117 Human Rights Watch, 'Heed the Call: A moral and Legal Imperative to Ban Killer Robots', *HRW Publications*, pp. 9-11 (September 2018) <https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots> [accessed 12 October 2018].

118 Human Rights Watch, 'Heed the Call', pp. 13-16 ('*Applicability and Significance of the Martens Clause'*).

119 HRW, 'Losing humanity; the case against killer robots', p. 35.

120 UN Report of the International Law Commission, 'Work of its 46th Session'*, UN Publications*, GAOR A/49/10, (May-July 1994), p. 317. See, also: Human Rights Watch, 'Heed the Call', p. 3.

121 Jean-Francois Queguiner, 'Precautions under the law governing the conduct of hostilities'*, International Review of the Red Cross*, 88, 864, (December 2006) <http://www.icrc.org/eng/assets/files/other/irrc_864_queguiner.pdf>.

122 Source: Article 36 <http://www.article36.org/weapons-review/autonomous-weapons-meaningful-human-control-and-the-ccw/> [accessed 12 March 2017].

protection of civilians in a conflict situation. How might this affect AWS deployment? Its first section requires that 'constant care shall be taken to spare the civilian population, civilians (sic) and civilian objects'. In this case, AWS' Delivery Cohort must ensure compliance with a portfolio of legal prerequisites that are, in programming terms (as well as in semantics), fundamentally abstruse:[123] To this point, *Article 57(2)* requires that 'all feasible precautions in the choice of means and methods of attack' be taken in order either to avoid or to minimise loss of civilian life, injury and damage to civilian objects'.[124] *Article 57(2) b* then stipulates that an attack shall be cancelled if it becomes apparent that the object is not a military one, or that the attack may be expected (following due analysis[125]) to cause incidental loss of civilian life, injury to civilians and damage to civilian objects that would be excessive in relation to the direct military advantage anticipated. While this thesis' later chapters address the technical challenges that such requirements entail[126], these ancillary constraints must govern AWS' operational deployment. *Article 57(3)*, for instance, stipulates that when a choice is possible between several military objectives, then the objective to be selected shall be the attack that is expected to cause the least danger to civilian life and civilian objects. Certain conditions, moreover, appear to represent a direct proscription to AWS deployment. *Article 54(b)(4)* of *Protocol 1* bans inherently indiscriminate weaponry while *Article 35(2)* of *Protocol* 1 rules against weapons that cause unnecessary suffering or superfluous injury. Notwithstanding arguments on the constitution of a target[127], a weapon, after all, may be deemed indiscriminate simply if it cannot be aimed *specifically* at that target.[128]

5.5 Article 36 and LOAC-compliant weaponry

Other legal hurdles exist that complicate States' attempts to develop compliant AWS. Article 36 of *Additional Protocol* I to the *Geneva Conventions* confers the obligation on States signatories to evaluate new or modified weapons in order to ensure their compliance with the provisions of humanitarian law.[129] States' deployment of AWS must be managed within the prescriptions of such weapon reviews. These should also be a continuous procedure taking place throughout States' procurement processes. Given that Article 36 forms part of the agreed *Conventions* (signed by 196 States signatories), much of HRW's *Case against Killer Robots*[130] rests on this assertion of best

[123] For a discussion here of programming challenges, see: Chapter 8 (*Software*), specifically: 8.1 ('*Coding methodologies'*), 8.6 ('*Value setting issues'*) and 8.8 ('*Behaviour setting and coordination'*).

[124] ICRC IHL database, 'Rule 14 – Proportionality in Attack', <https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_cha_chapter4_rule14> [accessed 23 September 2018].

[125] Eva Svoboda and Emanuela-Chiara Gillard, (eds.), 'Protection of Civilians in Armed Conflict: Bridging the Gap between Law and Reality', *Humanitarian Policy Group, Overseas Development Institute*, (2015), p. 3 and generally <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9876.pdf>

[126] See: chapters 6 (*Wetware*), 7 (*Firmware*) and 8 (*Software*).

[127] For a detailed discussion on targeting's ramifications to compliant AWS deployment, see: Chapter 10 (*Oversight*), introduction and 10.1 ('*Meaningful human control*').

[128] Autonomous weaponry must also comply with this rule but the mere feature of autonomy does not per se determine either compliance or non-compliance if sufficiently robust targeting algorithms are in place and work all of the time.

[129] Protocol 1 Additional to the *Geneva Convention*s of 12 August 1949, and Relating to the Protection of Victims on International Armed Conflicts (Protocol 1) adopted June 8, 1977, 1125 UNTS 3 (entered into force, 7 December 1978).

[130] For a useful distillation of the arguments see: Denise Garcia, 'The Case Against Killer Robots', *Foreign Affairs magazine*, 10 May 2014 <http://www.foreignaffairs.com/articles/141407/denise-garcia/the-case-against-killer-robots> [accessed 2 March 2014]. For the original report, see: Bonnie Docherty, 'Losing Humanity – the Case against Killer Robots', *Human Rights* Watch, (2012) <http://www.hrw.org/reports/2012/11/19/losing-humanity-0> [accessed 2 June 2015].

practice. Removing weapon supervision will thus require that States undertake formal impact assessments[131] regardless of whether specific weapon systems are yet available.[132]

Boulanin notes that few States have appropriate review mechanisms in place.[133] Were *Protocol 1*'s Article 36 to be properly enforced, it would impose material conditionality upon AWS' deployment and, as such, merits further analysis in order to evaluate the significance of that challenge.[134] Under the Article, the 'study, development, acquisition or adoption of a new weapon, means or method of warfare' places any States party 'under an obligation to determine whether its employment would, in some or all circumstances, be prohibited', either by *Protocol* 1 or by 'any other rule of law applicable' to such party.[135] The scope of an Article 36 legal review is broad[136] covering weapons of all types[137] and the way in which these weapons are to be used.[138] It also captures *any* weapon in a State's arsenal.[139] Review under Article 36 requires legal analysis identify where and under what legal limitations the use of that weapon is lawful and, in so doing, Anderson and Waxman note that States must consider whether the use of these new technologies would be contrary to international law *in some or all* circumstances.[140] Sandoz highlights that failure to do this renders a State internationally responsible for a breach of its obligations vis-à-vis the other parties to that Additional *Protocol* 1.[141] For those States that are not signatories to Additional *Protocol* 1, Rappart and Moyes argue that such review should still be undertaken 'as a corollary to other international obligations'[142] and as a matter of best practice.[143] Article 36, moreover, refers to 'any other rule of international law applicable' to the State including weaponry already subject to a

[131] The effect of both the weapon and the ways that the weapon can be used are subject to this mandated review.

[132] HRW, 'Killer Robots, the Case against Killer Robots', p. 3.

[133] Boulanin and Verbruggen, 'Article 36 Reviews', *SIPRI*, (December 2017), p. 15.

[134] According to the ICRC's Commentary on the *Additional Protocols*, Article 36 'implies the obligation to establish internal procedures for the purpose of elucidating the issues of legality and the Other Contracting Parties can ask to be informed on this point'.

[135] US Defense Department lawyers have rejected various proposed new weapons on this basis including blinding laser weapons in the 1990s and, reportedly, various cyber-technologies for use in cyber-conflict. See: 'Legal reviews of weapons and cyber capabilities'*, Air Force Instructions,* number 51-402, (27 July 2011) <http://www.fas.org/irp/doddir/usaf/afi51-4w02.pdf>.

[136] International Review of the Red Cross, 'A guide to the legal review of new weapons, means and methods of warfare: Measures to implement Article 36 of Additional Protocol I of 1977'*, ICRC Geneva*, 88, 864, (December 2006), p. 938.

[137] 'lethal', 'non-lethal' or 'less lethal'.

[138] Pursuant to military doctrine, tactics, rules of engagement, operating procedures and counter-measures.

[139] It is useful to provide evidence for Article 36's scope: 'All weapons to be acquired', 'further to research and development' or 'off-the-shelf', 'acquired for the first time without necessarily being new in a technical state', 'and existing weapon that is modified in a way that alters its function', 'and existing weapon where a State has joined a new treaty'; when in doubt, moreover, Article 36 stipulates that legal advice should be sought.

[140] Anderson and Waxman, 'Law and Ethics for Autonomous Weapon Systems: Why a ban won't work and How the Laws of War Can', *National Security and Law Essay, Hoover Institute*, Stanford University, (2013), p. 11.

[141] Yves Sandoz, ed., *Commentary on the Additional Protocols of 8th June 1977 to the Geneva Conventions of 12 August 1949,* (Geneva: Martinus Nijhoff Publishers, 1987), p. 423.

[142] Brian Rappart, Richard Moyes and Thomas Nash, 'The roles of civil society in the development of standards around new weapons and other technologies of war'*, International Review of the Red Cross*, 94, 886, (Summer 2012), p. 779.

[143] International Committee of the Red Cross, 'A Guide to the Legal Review of New Weapons, Means and Methods of Warfare: Measures to Implement Article 36 of Additional Protocol 1 of 1977'*,* p. 940.

ban.[144] Article 36 is therefore deliberately catch-all.[145] As well as examining the legality of the weapon's design and characteristics[146], the reviewing authority must also look at how the weapon under review *might* be used.[147] The reviewing body must thus take into consideration a wide range of military, technical, health and environmental factors[148] in their commission.[149] There are, indeed, several ways for a weapon system to fail its Article 36 review.[150] Testing, however, is time-consuming. As noted by Cummings, it is 'thankless, expensive and empirically offers little benefit to the budgetary constraints, tactical interests or sense of urgency of deploying States'.[151] It has therefore been tempting for States to opt for expediency and short cuts.[152] The process contains several practical difficulties. Its conditions are problematic to enforce as they exactly counter States' efforts to gain technical advantage over adversaries. They have also been flouted since their inception with neither specific policing nor sanctions included in their definition.[153] Empirically,

[144] In chronological order, the following specific weapons have been considered under international instruments: 1868 – explosive projectiles under 400g weight, 1899 – Asphyxiating gases, 1899 – expanding or dum-dum ammunition, 1907 – poisoned weapons, 1907 – automatic submarine contact mines, 1907 – Martens Clause: weapons according to 'principles of humanity and the dictates of public conscience', 1925 and 1972 – biological weapons, 1976 – Environmental modification techniques, 1977 – means of warfare unable to be directed precisely at a specific military objective and might thus include civilians and civilian objectives without distinction, 1977 – bombardment which treats as a single military objective a number of clearly separated military objectives located in a town, city, village or other area containing a similar concentration of civilians or civilian objects, 1977 attacks expected to cause excessive damage to civilians and civilian objects (excessive in relation to concrete and direct military advantage anticipated ('rule of proportionality'), 1980 – non-detectable fragments, 1980 – mines, booby-traps, 1980 – incendiary weapons, 1993 – chemical weapons, 1995 – Blinding lasers, 1997 – anti-personnel mines, 1998 – International Criminal Court definitions on poisons, weapons causing superfluous injury or unnecessary suffering, 2003 – Explosive remnants of war. 2008 – Treaty to ban cluster munitions. Mary Wareham notes that such weapons in this list have not been prohibited. For instance, non-detectable fragments (CCW Protocol I) are not regarded as a weapon. Nor are such statutes watertight. CCW Protocol II and Amended Protocol II do not prohibit antipersonnel mines or anti-vehicle mines. CCW Protocol III, for instance, prohibits the use of air-delivered incendiary weapons in the civilian areas but not gound-incendiary weapons.

[145] Vincent Boulanin, 'Implementing Article 36 Weapon Reviews in the Light of Increased Autonomy in Weapon Systems', *SIPRI*, 2015/1, (November 2015), generally <https://www.sipri.org/sites/default/files/files/insight/SIPRIInsight1501.pdf>.

[146] The 'means' of warfare.

[147] The 'method' of warfare, bearing in mind that any weapon's effects will result from a combination of its design and the manner in which it is used.

[148] Action 2.5.2 of Agenda for Humanitarian Action adopted by the 28th International Conference of the Red Cross and Red Crescent notes the importance of ensuring a multi-disciplinary approach to the review of weapons

[149] It is telling that the US Department of Defense's 2012 *Directive on autonomous weapons* contains a separate enclosure defining such a set of review guidelines for future AWS development. See: Department of Defense Directive, 'Autonomy in Weapon Systems', *US DoD,* Number 3000.09, Enclosure 3, (21 November 2012) p. 7. These specify policies on system capabilities, doctrines, training, appropriate levels of human judgement and care in deploying these systems as well as security and testing of the systems.

[150] Examples include predicted reliability of targeting mechanisms, evidence that the foreseeable effects of an unsupervised weapon can be limited to the target and can be controlled in time or space. Similarly, the Article's prescriptions on precise injury and damage levels are complicating, especially in their treatment of mortality rates and 'whether the weapon would cause anatomical injury or anatomical disability or disfigurement'.

[151] Professor Missy Cummings, Director, Humans and Autonomy Laboratory, Duke University, in conversation with the author (Chatham House Conference, 'Autonomous Military Technologies: Policy and Governance for Next Generation Defence Systems', February 2014).

[152] Harold Hutchinson, 'Russia says it will ignore any ban of killer robots', *Business Insider Tech*, 30 November 2017 <http://uk.businessinsider.com/russia-will-ignore-un-killer-robot-ban-2017-11?r=US&IR=T> [accessed 3 March 2018].

[153] As an example, see: Charles Clover, 'Chinese ships accused of breaking sanctions on North Korea', *Financial Times*, 27 November 2017 <https://www.ft.com/content/21a0407e-eadd-11e7-bd17-521324c81e23> [accessed 23 August 2018].

they are weakened too by their basis that 'all relevant scientific evidence pertaining to the foreseeable effects on humans has been gathered' including empirical (rather than desktop) evidence on 'what is the expected field mortality' as well as consequences that account for 'alteration to the victims' psychology or physiology'.[154]

Deployment processes are also challenged by current legal frameworks becoming, notes HRW, increasingly unfit for purpose.[155] There are, for instance, practical problems of definition. The iRobot Warrior is an example of an unmanned weapon that falls outside the *UN Register's* technical specifications[156] of either 'Battle Tanks' or 'Armoured Combat Vehicles'. Revision and the contentious updating of these legislative listings is clearly required.[157] Identifying circumvention of Article 36 is also complicated by the pace of weapon development as well as the modular nature of their procurement. Nor is there a set manner (that can then be regulated by statute) in which new weapons are developed and, given multiple contractor relationships, no common build system or indexing procedure, no repository for source code nor any recognised (and continuous) testing infrastructure. Procurement of the US Raven UAV[158] demonstrates this compliance challenge: As noted by Hambling (and characteristic of a *spiral* development process), Raven's deployment has followed a portfolio of incremental developments that have been shaped by rapidly changing end-user requirements and technology rather than any traditional Version 1 release being followed by a Version 2.[159]

Legal ambiguity may also lead to policy ambiguity. The US Department of Defense's *Directive Number 3000.09* was published in November 2012 and is the first significant policy announcement by a developed State on the development of semi-autonomous and autonomous weaponry.[160] The document instigates a temporary ten-year moratorium[161] under which the DoD should only develop autonomous systems that deliver non-lethal force.[162] That suspension, however, is not clear-cut. There is a carve-out, for instance, that allows a waiver 'in cases of urgent military need'.[163] The Directive also defines wide testing requirements. Given the challenges in testing unsupervised

[154] States generally agree that suffering which has no military purpose is generally a legal violation.

[155] Human Rights Watch and Harvard Law School International Human Rights Clinic, 'The Need for New Law to Ban Fully Autonomous Weapons', *Memorandum to Convention for Conventional Weapons Delegates*, (November 2013), pp. 2-3 <https://www.hrw.org/sites/default/files/supporting_resources/11.2013_memo_to_ccw_delegates_fully_autonomous_weapons.pdf>.

[156] It does not have the 'high cross-country mobility' or 75mm bore canon of the tank nor is it able to 'transport a squad of four or more infantrymen'. See: (UNODA (n.d.), 'Categories of equipment and their definitions' <http://www.un.org/dept/ddar/Register/Categories.html> [accessed 9 March 2018].

[157] For instance, the Switchblade and Fire Shadow weapon systems currently fall into a grey area between 'munitions' (under the less tightly controlled *Article 3* of the treaty) and a 'combat aircraft'.

[158] See: Global Security, 'RQ-11 Raven Procurement', undated, <https://www.globalsecurity.org/intell/systems/raven.htm [accessed 6 March 2018].

[159] Hambling, p. 70.

[160] Department of Defense Directive, '*Autonomy in Weapon Systems',* Number 3000.09, 21 November 2012, generally.

[161] Section 7 of the Directive states that it must be 'reissued, cancelled or certified current within 5 years of its publication' or will expire on 21 November, 2022 (10 years after it took effect).

[162] Department of Defense Directive, 'Autonomy in Weapon Systems', generally.

[163] This term remains undefined in the Directive's text.

weapons[164] (and the requirement that this be undertaken within a 'realistic operational environment'), HRW rightly points out that significant practical difficulties arise in the directive.[165] Given also that the point of AWS trialing 'is precisely to determine the probabilistic range of action without which no appropriate limits can be defined *ex-ante*', Kalmanovitz highlights the implausibility of any 'test-as-you-go' regime.[166] The conflict in this case is that any elasticity of standards only underpins the requirement for human judgement in what is the deployment of independent weapons *without* human judgement. Finally, the Directive only relates to the US DoD and notably excludes other procuring parties such as the Central Intelligence Agency.

5.6 Behavioural constraints to AWS deployment

As noted by Mickeviciute in *Lessons from the Past for Weapons of the Future*, legal challenges to AWS deployment tend to coalesce around behavioural constraints against their adoption.[167] History suggests that adoption of new methods of war is rarely straightforward, not least the raising of 'unprecedented issues that make the legality of an attack more complex'.[168] Lorber might therefore concur with this thesis' Chapter Two that AWS deployment will also be challenged by the experiences of past procurement programmes where technical developments have empirically failed to deliver on promised performance.[169] Even where military technologies may have made a decisive contribution to defeating an enemy, measuring the effects of individual weapon innovation is tricky. This missing correlation between weapon system and outcome is important.[170] A further aim of this section is therefore to identify behavioural forces, exogenous and often socio-political, that should constrain removal of human supervision in lethal engagements.[171] Chapter Two, after all, argues that the role of context in creating impedimenta to that removal (in what Black terms 'new military history', the social context arising from the 'position, experience and relationships of the rank and file'[172]) are similarly important to technical hurdles considered in this thesis'

[164] See: Chapter 10 (*Deployment*), specifically: 10.2 ('*Validation and testing*').

[165] Human Rights Watch and IHRC, Harvard, '*Review of the 2012 US Policy on Autonomy in Weapon Systems',* April 2013, p. 5.

[166] Kalmanovitz, cit. Bhuta and others (eds.), p. 160.

[167] Neringa Mickeviciute, *Lessons from the Past for Weapons of the Future*, (USA: International Comparative Jurisprudence, 2, Elsevier, 2016), pp. 99-100.

[168] For discussion on 'new methods of warfare' see: ICRC, 'International Review of the Red Cross', *ICRC Publications,* 886, 94, (2012) pp. 457-467 <https://www.icrc.org/en/international-review/new-technologies-and-warfare> [accessed 24 September 2016]. See also: Chapter 2 (*Context*), specifically 2.2 ('*The role of context in AWS' argument'*).

[169] Azriel Lorber, *Misguided Weapons: Technological failure and surprise on the battlefield,* (USA: Potomac Books, January 2003), generally. An interesting narrative to this point is provided at: Judy Dutton, 'Nine Bizarre Weapons that Failed Spectacularly', *Mental Floss Blog*, (29 April 2014) <http://mentalfloss.com/article/30669/9-weapons-failed-spectacularly-and-1-possibly-didn%E2%80%99t> [accessed 24 February 2015].

[170] See, for example: Defence Synergia, 'UK Air Defence: A forgotten capability gap', *National and Defence Strategies Research Group,* (4 February 2014) <http://www.defencesynergia.co.uk/uk-air-defence-a-forgotten-capability-gap/> [accessed 3 March 2018].

[171] Terry Moe, 'Vested Interests and Political Institutions', *Department of Political Science, Stanford University*, (May 2014), pp. 1-2 <https://politicalscience.stanford.edu/sites/default/files/images/vested%20interests_psq%20final%20June%202014.pdf.

[172] Jeremy Black, 'Military Organisations and Military Change in Historical Perspective', *The Journal of Military History*, Lexington, 62, 4, (1 October 1998), 871.

subsequent chapters.[173] This behavioural analysis is underpinned by Tilford's conclusion that the battlefield's changing character is 'about people, not systems. Armies, Air Forces, and Navies function with people who use and employ machines and weapons'.[174]

States' militaries have often failed to exploit technical innovation. In relative terms, Boot notes that the Mongols missed the Gunpowder revolution, the Chinese, Turks and Indians let pass the Industrial Revolution while the Soviets missed the Information Revolution.[175] The behavioural point for AWS adoption is that military procurement is rarely a straightforward exercise and that inertia is to be expected not least because past performance is no guarantee to future returns in defence planning. As noted by Penna, priorities in a military budget are not obvious in a landscape of fast-moving technologies.[176] To this point, Brannen highlights that while the US may enjoy a current lead in unmanned technology[177], there is no certainty that this will continue. It may be that a rival State (or even non-state grouping) introduces new variants or counter-measures that fundamentally and very quickly alters the current balance of power around such technologies. It may also be that a particular State devotes resources to besting a new weapon class; in this vein, Weinberger in *Foreign Policy* contends that the US has already lost its edge in drone technology.[178] Strohn notes, moreover, that 'it is not only about procurement but the philosophy behind this: The French had tanks in 1940 but used them incorrectly'.[179] It is, therefore, as much about the conceptual component of procurement as it is about the physical. A challenge to the adoption of AWS is thus the basic difficulty of predicting the trajectories in battlefield technologies and, in the case of AWS as noted by Kovic in *The Strategic Paradox of Autonomous Weapons,* their certain place in battlefield structures.[180] A feature of AWS deployment is therefore that behavioural imperatives will challenge its procurement. Rudichhauser from Germany's *Federal Academy for Security Policy* highlights here that certain applications of weapon autonomy will be 'close to impossible to control because they often represent off-the-shelf civilian technology with nearly infinite purposes'.[181] Similarly, the deployment of other new and easy-to-disseminate weapon technologies (germs,

---

[173] Scharre, 'Autonomous weapons and operational risk', pp. 6-8. For detailed review of technical issues facing compliant removal of human-weapon supervision, see: chapters 6 (*Wetware*), 7 (*Firmware*) and 8 (*Software*).

[174] Earl Tilford, *Reviewing the Future,* (USA: Parameters. 2002), p. 151. The point is made in a different manner by Toby Walsh: 'There isn't this separate part of the world called the battlefield that's signposted "Battles over here please". Battles are fought in cities right where we live'. TedX Berlin, 'How can you stop Killer Robots?', 12.40 minutes, 8 October 2015, <https://www.youtube.com/watch?v=c277ynyRPgs> [accessed 12 October 2017].

[175] Max Boot, *War Made New; Weapons, and the Making of the Modern World,* (USA: Gotham, Penguin, New York, 2006), generally.

[176] Charles Penna, 'A Reality Check on Military Spending', *Issues in Science and Technology*, XXI, 4, (Summer 2005), generally <http://issues.org/21-4/pena/> [accessed 19 June 2018].

[177] Samuel Brannen and others, 'Sustaining the US Lead in Unmanned Systems'*, Center for Strategic and International Studies*, (27 February 2014), pp. 1-4 <https://csis-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/publication/140227_Brannen_UnmannedSystems_Web.pdf>.

[178] Sharon Weinberger, 'China Has Already Won the Drone Wars'*, Foreign Policy*, 10 May 2018 <https://foreignpolicy.com/2018/05/10/china-trump-middle-east-drone-wars/>.

[179] Dr Matthias Strohn, in conversation with the author, January 2019.

[180] Marko Kovic, 'The Strategic Paradox of Autonomous Weapons'*, Zurich Institute of Public Affairs Research*, (22 February 2018), p. 11 <https://zipar.org/policy-brief/strategic-paradox-autonomous-weapons/> [accessed 24 June 2018].

[181] Wolfgang Rudwichhauser, 'Autonomous or Semi-Autonomous Weapon Systems: A Potential New Threat of Terrorism'*, Federal Academy for Security Policy*, Security Working Papers 23, (2017), p. 2 <https://www.baks.bund.de/sites/baks010/files/working_paper_2017_23.pdf>.

chemicals, cyber-viruses?) will likely obfuscate what passes for the status quo and what should therefore inform States' procurement priorities.[182] As noted by Boot, 'the end can come with shocking suddenness even after a long streak of good fortune'.[183]

A second diminishing heuristic for AWS deployment is that the introduction of new battlefield technology rarely gives lasting advantage.[184] Parties facing autonomous weaponry without technical or budgetary wherewithal to field similar weapons must instead pursue Brodie's logic, developing in its place battlefield behaviour 'that works well enough, be it ever so inelegant and probably decidedly irregular' in order to defeat that new technology.[185] This, after all, is a repeated theme characterising the deployment of disruptive weaponry.[186] In the case of the air power theorists of the 1930s, Douhet grossly under-estimated the neutralising effect of a defending air force and anti-aircraft fire on attacking bombers.[187] MacIsaac similarly notes that planners persistently over-estimated the effect of bombing.[188] Defeat has also often been a spur to innovation.[189] The Israelis almost lost to Egyptian and Syrian anti-tank and anti-aircraft missiles in the 1973 Yom Kippur War having beaten them just six years previously. Technologies and scientific concepts rarely remain the property of one power for a protracted length of time. In this vein, Boot points out that France matched the needle gun less than four years after Konniggratz, Germany matched Britain's Dreadnoughts some three years after their first launch while the USSR had its own atomic bomb four years after Hiroshima.[190] Just as Douhet's advocacy of pre-emptive aerial attack was compromised by the arrival of radar, an argument is that AWS may be compromised by the introduction of some other quite exogenous technology.[191] In the same way that trench systems were to offset the impact of artillery and quick-firing machine guns (and improvements in radar negate the edge enjoyed by first-generation stealth aircraft[192]), combatants have a portfolio of alternative tactics to blunt a State's investment in autonomous technologies. Margaritoff highlights that this may take on several guises, from neutrino beams, electromagnetic pulsing, flux compression technologies, microwave and other high frequency interference, energy lasers and

[182] Ibid., pp. 2-3.

[183] Max Boot, 'What the Past Teaches About the Future', *JFQ*, 44, (2007), p. 109 <http://indianstrategicknowledgeonline.com/web/MIL%20HIS%20JFQ44%20boot.pdf>.

[184] 'Lasting' is defined here as enduring over the course of a campaign (such timescale continues to be compressed by increases in the 'pace' of warfare discussed earlier in this chapter. For definition of this thesis' timelines, see: Chapter 1 (*Introduction*) and thereafter). See also: Walker, *Killer Robots?,* pp. 71-72.

[185] Gray, 'Another Bloody Century', p. 52.

[186] See: Introduction to Chapter 2 (*Context*).

[187] David MacIsaac, *Voices from the Central Blue: The Air Power Theorists*, p. 634.

[188] Ibid. See also: Tammi David Biddle, *Rhetoric and Reality,* (USA: Princeton University Press, 10 January 2009), pp. 69-128.

[189] See, for instance: Kenneth Miller, 'Is China Winning the Innovation Race?', *LeapsMag*, 19 June 2018 <https://leapsmag.com/is-china-winning-the-innovation-race/> [accessed 6 September 2018].

[190] Greg Satell, '4 Innovation Lessons from the History of Warfare', *Forbes*, 14 March 2015, generally <https://www.forbes.com/sites/gregsatell/2015/03/14/4-innovation-lessons-from-the-history-of-warfare/#508dec4e73f3> [accessed 6 March 2018].

[191] MacIsaac, p. 630.

[192] David Szondry, 'Quantum Radar to Render Stealth Technologies Obsolete', *New Atlas*, (26 April 2018) <https://newatlas.com/quantum-radar-detect-steath-aircraft/54356/> [accessed 6 September 2018].

signal jamming.[193] Parties that find themselves challenged by new technologies will hunt for asymmetrical and equalizing tactics, operations and strategies. In this vein, Gray notes that parties who create slender advantage in a narrow (albeit relevant) combat field are likely to be able to stall others' more general technical superiority[194] and, on this basis, it may be wrong to 'conclude that high tech can only be defeated by similar high tech'.[195] This argument is supported by Clark (the example, for instance, of orders in December 1944 being disseminated by hand-written note as opposed to Ultra[196]) and also by what Boot terms 'psychological nullification' in which he notes 'the first time an army faces a major new weapon – the needle gun at Konniggratz, the machine gun at Omdurman, the tank, the smart bomb in the Gulf War – it is likely to be caught off guard… The next time the other side is likely to be less impressed'.[197] The behavioural constraint for AWS deployment is that a series of quickly-learned tactical innovations by an adversary will likely dull the effect of weapons autonomy.[198] Indeed, it is AWS' psychological facets (better characterised by Pollick as 'behavioural aspects'[199]) that complicate the role of the Delivery Cohort in considering AWS' several fragilities.

In identifying behavioural challenges to AWS feasibility, further cause-and-effect examples are relevant.[200] While Handel argues that high-tech confrontation (taken here as a proxy for AWS deployment) will typically be most effective in major, conventional warfare against an enemy who takes a similar approach[201], Hoffman instead points to the growing toolkit of distinctly low-tech adversaries. In this case, 'conventional, irregular, catastrophic, disruptive… or [custom-build] hybrid strategies' might, by inference, reduce materially the effect of unsupervised weaponry and its realistic deployment.[202] Lonsdale highlights the Serbian use of UN hostages as human shields in

[193] For a general discussion, see: Marco Margaritoff, 'The Seven Most Significant Anti-Drone Weapons', *The Drive*, (21 June 2017) <http://www.thedrive.com/aerial/11505/the-7-most-significant-anti-drone-weapons> [accessed 12 April 2018].

[194] Gray, *Another Bloody Century,* p. 52.

[195] Dan Clavin, 'These Were The Most Unfair, Technologically-Lopsided Battles in History', *Warped Speed*, 20 September 2016 <http://www.warpedspeed.com/these-were-the-most-unfair-technologically-lopsided-battles-in-history/> [accessed 8 September 2018].

[196] Lloyd Clark, in conversation with the author, September 2018. See: Richard Hayes and others, 'The State of the Art and the State of Practice, Battle of the Bulge: The Impact of Information Age Command and Control on Conflict – Lessons Learned', *CCRTS*, (2006) <http://www.dodccrp.org/events/2006_CCRTS/html/papers/206.pdf>.

[197] Caroline Houck, 'New Report Notes Erosion of Pentagon's Technological Advantage', *Defense One*, 22 February 2018 <http://www.defenseone.com/threats/2018/02/new-report-quantifies-erosion-pentagons-technological-advantage/146162/> [accessed 24 April 2018].

[198] Walker, *Killer Robots?,* p. 71. See also: Rosa Brooks, 'Why Sticks and Stones Will Beat Our Drones', *Foreign Policy*, 4 April 2013 <http://foreignpolicy.com/2013/04/04/why-sticks-and-stones-will-beat-our-drones/) [accessed 5 April 2018>. Measures here might include electromagnetic pulsing, simple spray paint on drone sensors, improved frequency jamming, innovative camouflage, ghosting, low-cost deception and obstacle building.

[199] Amy Pollick, 'Behavioural Science and National Security', *Association for Psychological Science*, (1 June 2008) <https://www.psychologicalscience.org/observer/behavioral-science-and-national-security> [accessed 1 April 2018].

[200] See: Economist Magazine, 'The Last Manned Fighter', *Economist*, 14 July 2011 <http://www.economist.com/node/18958487>.

[201] Michael Handel, *Masters of War,* p. xxii.

[202] Frank Hoffman, 'Complex Irregular warfare: The Next Revolution in Military Affairs', *Foreign Policy Research Institute*, Orbis, (Summer 2006), p. 398 <http://indianstrategicknowledgeonline.com/web/hoffman.complexirregularwarfare.pdf>.

Bosnia[203] to illustrate how a simple although asymmetrical act can quickly negate the advantages conferred by investment in technologically advanced equipment.[204] The difficulties in assessing cause-and-effect are demonstrated by Brooks' citing Einstein's advice to President Truman: 'I know not what weapons World War Three will be fought but World War Four will be fought with sticks and stones'.[205]

Roff identifies a similar challenge to AWS feasibility in what might be a reluctance by parties to set up and engage in decisive battle. 'The rub', she concludes, 'is that if two countries have roughly equal military capabilities, any subsequent fight would be long, drawn out and may not bend to either side's favour – that is, it would end up being a war of attrition'.[206] Quintana cites innovations such as cyber and electronic counter-measures to suggest that effects of AWS technology may be similarly blunted;[207] the evidence, after all, is that 'game-changing' military technology is disseminated quickly and widely.[208] Boot's notion of nullification[209] therefore echoes Gray's corollary that 'the history of war is not primarily the history of weaponry but instead of the history of the person who wields that weapon'.[210] Deployment of AWS will clearly require material adjustments in the way States organise their military efforts in order properly to integrate the technologies' capability. While the US *Third Offset* may demonstrate considerable flux taking place in strategic direction[211], little 'near-future' defence work exists in the public domain on these grass root organisational and battlecraft changes that AWS deployment will require.[212] This is perhaps unsurprising given current lack of definition around the weapons' classes but, as noted by Conn, this also extends to the political sphere where complications await politicians trying to frame the deployment of AWS.[213]

AWS deployment is complicated by how States organise their strategic imperatives. The US has chosen to be strong in every area of combat – land, sea, air and cyber-space – in order to defeat a

[203] Joel Br, 'Bosnian Serbs seize more UN troops', *Washington Post*, 29 May 1995 <https://www.washingtonpost.com/archive/politics/1995/05/29/bosnian-serbs-seize-more-un-troops/991628ef-8469-436d-8759-470fe4ab11d4/?utm_term=.d6bb7243d9e3> [accessed 4 April 2017].

[204] Lonsdale, cit. Strachan and Herberg-Rothe, p. 242.

[205] Rosa Brooks 'Why Sticks and Stones Will Beat Our Drones', *Foreign Policy*, (4 April 2013) <https://foreignpolicy.com/2013/04/04/why-sticks-and-stones-will-beat-our-drones/> [accessed 6 April 2017].

[206] Heather Roff, 'Killer Robots on the Battlefield*', Slate, New America*, (7 April 2016) <http://www.slate.com/articles/technology/future_tense/2016/04/the_danger_of_using_an_attrition_strategy_with_autonomous_weapons.html> [accessed 27 February 2018].

[207] Elizabeth Quintana, Director, Military Sciences, Royal United Services Institute, in conference with author (Chatham House Conference, 'Autonomous Military Technologies: Policy and Governance for Next Generation Defence Systems', February 2014).

[208] Boot, *War Made New*, generally.

[209] Boot here uses the term *nullification* to refer to new methods, hardware and work-round solutions annulling existing technical advances.

[210] Gray, *Another Bloody Century,* p. 61.

[211] Paul McCleary, 'The Third Offset May Be Dead But No One Know What Comes Next', *Foreign Policy*, 18 December 2017 <https://foreignpolicy.com/2017/12/18/the-pentagons-third-offset-may-be-dead-but-no-one-knows-what-comes-next/> [accessed 2 September 2018].

[212] Logistics (distributed replenishment and service facilities), control and monitoring, tactical incorporation.

[213] Ariel Conn, 'The Problem of Defining Autonomous Weapons'*, The Future of Life Institute*, (30 November 2016) <https://futureoflife.org/2016/11/30/problem-defining-autonomous-weapons/> [accessed 17 April 2018].

portfolio of different potential foes (from a rising superpower in China, medium-sized powers such as North Korea and Iran, to non-State actors such as al Qaeda).[214] In so choosing, it must be ready for every type of warfare from peace-keeping to high-intensive conflict. The behavioural constraint is that this leads to funding, prioritisation and allocation challenges that then might dilute efforts to integrate AWS into States' battlecraft.[215] Such dilution has consequences, not least a likely compromise in the setting of common standards and ethical-legal criteria that are a prerequisite for compliant removal of weapon supervision. Black thus points to challenges arising from the integration of wholly new procedures, from vested interests within particular military services and from new command structures that must be integrated in order to maximize weapon effectiveness by being able to synchronize attacks in time and space.[216]

A quite different set of behavioural challenges arises from the tight correlation that will exist between intended weapon function and that weapon's transparency and accountability. In this vein it is impossible to punish a lethal robot for unlawful acts that it carries out. HRW highlights that AWS would, by definition, be free from human supervision and, lacking emotion, they can feel no remorse that might change subsequent behaviour.[217] As noted by Cummings, unless the weapon can 'understand' (not merely respond to certain impulses as a blind trigger for subsequent action) that it will be admonished for breaking IHL then its decisions will not be influenced by the threat of that accountability.[218] Suarez' work is based on the principle that transparency and responsibility are the 'long-held cornerstones of representative government and that both of these pillars are undermined by autonomous weaponry'.[219] His point (and a further challenge to AWS deployment) is that removal of human supervision paradoxically invests too much power concentrated in the hands of too few unseen hands: Binary deployment decisions become the preserve of a small cohort. For Suarez, transparency is the key behavioural challenge to AWS deployment: 'No robot should have the expectancy of privacy in a public place'.[220] This points to a further paradox arising from machine autonomy. As Epstein notes, 'the irony is that the military want the robot to be able to learn and react in order to do its mission well but won't want it to be too creative. But once you reach a space where it is really capable, how do you limit them? To be honest, I don't think that we

[214] US Department of Defense, 'Summary of the National Defense Strategies of the United States of America: Sharpening the American Militaries' Competitive Edge', *US DoD*, undated, pp. 2-4 <https://www.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>.

[215] See: Joe Anselmo, 'Defense Contractors need looser R&D purse strings', *Aviation Week and Space Technology*, 23 April 2014 <http://aviationweek.com/defense/opinion-defense-contractors-need-looser-rd-purse-strings> [accessed 20 April 2018].

[216] Jeremy Black, 'Military Organisations and Military Change in Historical Perspective', *Journal of Military* History, 62, 4, (1998), 871-892.

[217] HRW, 'Losing humanity', p. 44. HRW are careful to phrase this as a conditional for the reason that fully autonomous weapons are not yet deployed.

[218] Missy Cummings, 'The Human Role in Autonomous Weapon Design and Deployment', *University of Pennsylvania*, undated, pp. 4-5 <https://www.law.upenn.edu/live/files/3884-cummings-the-human-role-in-autonomous-weapons> [accessed 23 April 2016].

[219] Daniel Suarez, 'The kill decision shouldn't belong to a robot', *Ted Talk,* (June 2013) <https://www.ted.com/talks/daniel_suarez_the_kill_decision_shouldn_t_belong_to_a_robot?language=en> [accessed 15 April 2016].

[220] Ibid., generally.

can'.[221] Current use of unmanned technology has been restricted to benign environments largely free from electronic counter measure or sophisticated electronic attack; questioned about this correlation between AWS function and accountability, one such drone pilot told *Der Spiegel* in 2010, 'we operate in a war that highlights the strengths of remotely-piloted aircraft; their weaknesses are not much of a problem right now'.[222]

There is also overlap between behavioural and technical challenges and how these intersect with AWS' likely operational routines. To this point, Leboucher notes the 'extraordinary complexity' that arises from the simple dynamic that targets move, especially if that movement is oblique to the AWS' visual sensors.[223] Moving targets and their treatment create a further unexpected and behavioural hurdle which Durlach and others term 'change blindness', categorized here as the weapon's systemic failure to react to what should otherwise be obvious visual changes in how that autonomous weapon sees its immediate environment.[224] The point is behavioural as the phenomenon is likely to arise from how AWS focus is managed. Durlach notes that it may occur because of some visual transient which has occurred at the same time that masks the relevant visual change. Such circumstances are already difficult for human operators to master.[225] Quite different from the human brain, AWS' visual scenery will not be stored as sequences of actual imagery but instead only as general ideas about 'what is where'.[226] The point here is that change blindness is particularly prevalent in code-based processes characterised by complex, layered tasks.[227] While its associated technical constraints form the basis of later chapters[228], AWS' tendency towards change blindness arise in large part from quite behavioural functions that are intrinsic to AWS deployment including the need for human programmers to manage 'meaningfulness' in the autonomous weapon's sensed data.[229]

---

[221] Peter Singer, 'In the Loop: Armed Robots and the Future of War', *Brookings Institute*, (28 January 2009), para. 13 of 26 <https://www.brookings.edu/articles/in-the-loop-armed-robots-and-the-future-of-war/>. A detailed discussion on creativity is undertaken in Chapter 6 (*Wetware*), specifically: 6.5 ('*Missing pieces'*).

[222] Christian Enemark, *Armed drones and the Ethics of War: Military Virtue in a Post-Heroic Age,* (Routledge, Oxford, 2014), p. 99.

[223] Cedric Leboucher and others, 'A Two-Step Optimisation Method for Dynamic Weapon Target Assignment Problem', *Recent Advances in Meta-Heuristics, Intech*, (2013), pp. 109-110 <http://cdn.intechopen.com/pdfs/42284.pdf>. See also: Chapter 9 (*Hardware*), specifically: 9.4 ('*Operational hardware issues'*). See Appendix One: *Case study; automatic target recognition'*.

[224] Paula Durlach, 'Change Blindness and its Implications for Complex Monitoring and Control Systems Design and Operator Training', *US Army Research Institute, Human-Computer Interaction*, 19, (2004), p. 423 <https://www.researchgate.net/profile/Paula_Durlach/publication/234791524_Change_Blindness_and_Its_Implications_for_Complex_Monitoring_and_Control_Systems_Design_and_Operator_Training/links/0c960528e1742cbafa000000/Change-Blindness-and-Its-Implications-for-Complex-Monitoring-and-Control-Systems-Design-and-Operator-Training.pdf>.

[225] Ibid., generally.

[226] Pentti Haikonnen, *The cognitive approach to conscious machines,* (UK: Imprint academic, 2003), p. 54.

[227] Durlach, p. 424.

[228] See: Chapter 8 (*Software*), specifically 8.4 ('*Software processing functions'*) and 8.7 ('*Action selection issues'*).

[229] Durlach, pp. 433-439. See also: Chapter 7 (Wetware), specifically: 7.4 ('*Attention Methodologies'*) and 8.7 ('*Action Selection Issues'*). This crossover between behavioural and technical constraints will likely result in biases in each weapon's management of priorities, a phenomenon that Durlach terms 'exogenous attentional capture'.

A further behavioural tenet is that predicting the human *workload* around AWS deployment models will be challenging.[230] Kania notes that AWS' independence does not, of course, 'equate on the ground to absence of human engagement'.[231] Calculating such capacity is a variable over different parts of the same task and, notes Stimpson, will vacillate widely depending on that task's contingencies.[232] Indeed, if AWS' deployment models do not deliver materially more efficient task-loading, then why should a commander allocate resource or permissions to such assets? Forecasting workloads in AWS deployment is complicated by subtle shifts from observable manual tasks to not so observable cognitive tasks and, in the case of flexible autonomy, by erratic changes in the supervisory demands of human handlers.[233] Endsley and Jones question such human involvement in otherwise autonomous engagements citing handlers' likely biased understanding of given battlefield situations.[234] This is all unhelpful to AWS' Delivery Cohort and, by challenging the notion of workload/autonomy correlation, it rightly complicates decisions to deploy unsupervised weapons. Deci and Flaste instead point to a material *increase* in the time required by human parties to make decisions in scenarios where autonomy is deployed given that those operators must now factor additional unknowns and sources of information into their processes.[235] An associated behavioural twist arises from what Harford terms the *automation conundrum* whereby the better a system's autonomy, the more out-of-practice a human operator quickly becomes and the more extreme the situations he or she will have to face.[236] As noted by the psychologist Reason, author of *Human Error*, mixing human and autonomous agents clearly creates behavioural complexity:[237]

> Manual control is a highly skilled activity, and skills need to be practiced continuously in order to maintain them… yet an automatic control system that fails only rarely denies operators the opportunity for practicing these basic control skills. When manual takeover is necessary

[230] Jack Beard, 'Autonomous Weapons and Human Responsibility', *University of Nebraska-Lincoln, College of Law Faculty Publications*, 196, (2014), pp. 614-625 <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=http://scholar.google.co.uk/&httpsredir=1&article=1196&context=lawfacpub>.

[231] Elsa Kania, 'The Critical Human Element in the Machine Age of Warfare', *Bulletin of the Atomic Scientists*, (15 November 2017), generally <https://thebulletin.org/2017/11/the-critical-human-element-in-the-machine-age-of-warfare/> [accessed 4 September 2018].

[232] Alexander Stimpson and others, 'Assessing the Workload in Single Pilot Operations with Advanced Autonomy', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, (15 September 2016), 675 <http://journals.sagepub.com/doi/pdf/10.1177/1541931213601155> [accessed 3 May 2018].

[233] US Air Force, 'Autonomous horizons; system autonomy in the air force – a path to the future – human-autonomy teaming', *Office of the Chief Scientist,* AF/ST TR 15-01, (June 2015), p. 7.

[234] MR Endsley and DG Jones, *Designing for situation awareness: An approach to human-centered design,* (London: Taylor and Francis, Second Edition, 2012). See also: Nicolas Cote and others, 'Integrating the human recommendations in the decision process of autonomous agents', *Robot Teamwork in Dynamic Adverse Environment*, Papers from the 2011 AAAI Fall Symposium, (2011), pp. 2-3 <https://pdfs.semanticscholar.org/f346/759a7fc08dbdc9fca96689b678d5baf33096.pdf>.

[235] Edward Deci and Richard Flaste, *Why We Do What We Do: The Dynamics of Personal Autonomy,* (USA: GP Putnam's Sons, NY, 1995), generally. See also: M Hurston, 'Even Artificial Intelligence can acquire Biases against Race and Gender', *Science Magazine*, Science AAAS, (13 April 2017).

[236] Tim Harford, 'Crash: how computers are setting us up for disaster', *The Guardian*, 12 October 2016, p. 4 <www.theguardian.com/technology/2016/oct/11/crash-how-computers-are-setting-us-up-for-disaster> [accessed 12 October 2016].

[237] Chapter 4 (*Deployment*), specifically: 4.7 ('*Operations and causes of failure'*).

> something has usually gone wrong. Operators need to be more rather than less skilled in order to cope with atypical conditions.[238]

Merely automatic systems may 'accommodate' incompetence by being easy to operate and, as highlighted by Kinni, by remotely correcting mistakes.[239] A behavioural consequence is that inept operators and erroneous practices can function for some time before systemic incompetence becomes apparent. A further challenge is that automatic and then autonomous systems erode skills by removing the need for practice even if operators were once expert.[240] Riegler instead points to the challenge of goal divergence arising from imbedded conflict between machine (a focus on the control of input states) and human engineer (a corresponding focus on what Riegler terms 'output of artifacts').[241] Autonomous systems, after all, will tend to fail either in unusual situations or in ways that produce unusual situations.[242] In either case, an unrealistically skillful response is required from the human intervening to correct that failure in what will be a bewildering evidence set. While the technical foundation for these behavioural constraints forms the basis of later chapters, Weiner's Laws of Aviation provides a useful generalisation noting that a digital device 'tunes out small errors while creating opportunities for large errors'.[243] Further context to this point is provided by Harford's observation that a computer (which is broadly one hundred times more accurate than a human and one million times faster) will statistically still make ten thousand times as many mistakes.[244]

Still other behavioural constraints create challenges to AWS deployment. The adoption of autonomy may be influenced by the phenomenon of *threat inflation* whereby 'national commentary represents threats as much larger than they truly are'.[245] Milliman's *Psychological Rationales for Threat Inflation* explains this behavioural challenge as a consequence of preeminent States having no locally relevant enemies but nevertheless being 'suffused by an inherent insecurity' whereby the mere existence of a threat, in this case weapon autonomy, cannot ever be eliminated.[246] Any sudden embrace of compensating technologies must also have regard for economic factors. Even in the relatively benign environments that remote weapons have thus far been operating, semi-

---

[238] James Reason, *Human Error,* (Cambridge: Cambridge University Press, 1990), generally and: James Reason, *Human Error: Models and Management*, 320/7237, (UK: British Medical Journal, 2000), 768-770.

[239] Theodore Kinni, 'Beware the Paradox of Automation'*, MIT Sloan Management Review*, (20 October 2016) <https://sloanreview.mit.edu/article/beware-the-paradox-of-automation/> [accessed 20 April 2018].

[240] Harford, p. 4.

[241] Alexander Riegler, 'The Paradox of Autonomy: The Interaction between Humans and Autonomous Cognitive Artifacts', *Center Leo Apostel for Interdicsciplinary* Research, CEPA-Info E-Print 292, (2008) <http://www.univie.ac.at/constructivism/archive//292> [accessed 12 January 2018].

[242] See: Chapter 7 (*Firmware*), specifically: 7.1, ('*Sources of Technical Debt*').

[243] Madeleine Eish and Tim Hwang, 'Praise the Machine! Punish the Human!' *Comparative Studies in International Systems*, Working Paper Number 1, Society Research Institute, (24 February 2015), generally <https://www.datasociety.net/pubs/ia/Elish-Hwang_AccountabilityAutomatedAviation.pdf>.

[244] Harford, p. 7. See also: Chapter 6 (Wetware), specifically: 6.6 ('*AWS Control Methodologies'*).

[245] Noah Milliman, 'Psychological Rationales for Threat Inflation'*, The American Conservative*, (5 May 2014) <http://www.theamericanconservative.com/millman/psychological-rationales-for-threat-inflation/> [accessed 29 December 2017]. See also: Alan Stephens and Nicola Baker, *Making Sense of War: Strategies for the 21st Century,* (Cambridge: Cambridge University Press, 13 November 2006), p. 190.

[246] Ibid., p. 190 and generally.

autonomous UAVs are reported to be suffering a 'disproportionately significant' accident rate.[247] Assuming for this purposes AUVSI's calculations that the capital cost of a UAV is broadly one third less than its manned equivalent[248], Carpenter notes that losses at several times the rate of the manned equivalent might quickly erode any cost benefit.[249] Crash rates vary by aircraft type but magazine *Drone360* estimates that the US military's larger UAV 'crash three times more often than the remainder of the fleet'[250]; to this point, Hanson reports that more than four hundred large military UAVs crashed between 2001 and 2016.[251]

Deployment constrains are also created by political considerations. In considering the use of State force, 'maximum operational efficiency' is rarely achieved in battlefield assets.[252] In this vein, van Riper and Scales argue that removing human supervision from battlefield weapons may similarly provoke political sensitivities 'routinely precluding the unconstrained employment of military means'.[253] A behavioural upshot for AWS deployment is therefore captured by Lonsdale whereby 'the mere possession of advanced technology is no guarantee of its *practical utility'*.[254] Concern over casualties in Kosovo obliged ground attack bombers to fly above fifteen thousand feet thereby reducing greatly their effectiveness.[255] Other political restrictions in that conflict, reports Lonsdale, arose from the coalition-nature of that war whereby clear decision-making was complicated by NATO's unanimity principle and similar constraints are likely to compromise deployment of AWS.[256] Lonsdale observes the irony in politicians' desire for stand-off, autonomous weapons while political restrictions on their operation may diminish their chances of success. The first high profile mistake involving unsupervised weapons is likely to be front-page news precipitating hiatus in their deployment.[257] This political dimension has further effects. Given that

[247] Royal Air Force Directorate of Defence Studies, 'Air Power; UAVs: The Wider Context', Report on AUV Progress and Challenges, Professor RA Mason, p. 119.

[248] AUVSI News blog, 'Are UAS More Cost Effective Than Manned Flights?', *AUVSI,* (24 October 2013) <http://www.auvsi.org/are-uas-more-cost-effective-manned-flights> [accessed 15 April 2018]. See also: Conversation blog, 'Drones and Cheap: Soldiers are Not: A Cost-Benefit Analysis of War', *The Conversation blog*, 26 January 2014 <http://theconversation.com/drones-are-cheap-soldiers-are-not-a-cost-benefit-analysis-of-war-27924> [accessed 15 April 2018].

[249] Charli Carpenter, 'Don't Confuse me with the Facts: Costs of Lethal Autonomous Weapons', *Duck of Minerva*, (11 June 2014), generally <http://duckofminerva.com/2014/06/dont-confuse-me-with-the-facts-costs-of-lethal-autonomous-weapons.html> [accessed 6 March 3018].

[250] Dronemag, March/April 2016, p. 56 <www.drone360mag.com. The publication's media study from 2010 showed that thirty eight Predator and Reaper units crashed in combat in Iraq and Afghanistan with nine further craft crashing in training on US soil. Each Predator costs between $3.7m and $5m. The Washington Post also cites mechanical and pilot error being responsible for thirty large Air Force UAV crashes in 2015.

[251] George Hansen and others, 'Reliability of UAVs and Drones', *DSIAC*, 4, 2, (Spring 2017) <https://www.dsiac.org/resources/journals/dsiac/spring-2017-volume-4-number-2/reliability-uavs-and-drones> [accessed 12 February 2018].

[252] See: 'Political Constraints on Urban Operations', *International Law and the Politics of Urban Air Operations*, undated, pp. 25-26 <https://www.rand.org/content/dam/rand/pubs/monograph_reports/MR1175/MR1175.chap3.pdf>.

[253] Paul van Riper and Robert Scales, p. 9.

[254] Any deployment risk also reduces the appeal of any new weapon system to the procurement chain.

[255] Lonsdale, *Clausewitz and Information Warfare,* cit. Strachan and Herberg-Rothe, p. 241.

[256] Benjamin Lambeth, *NATO's Air War for Kosovo: A Strategic and Operational Assessment,* (USA: Santa Monica, CA, 2001), p. 185.

[257] Rebecca Crootof, 'War Torts: Accountability for Autonomous Weapons', *University of Pennsylvania Law Review*, 164, 6, (May 2016), pp. 1350-1351 <https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=9528&context=penn_law_review> [accessed 17 June 2018]. See also: Neal Boudette, 'Fatal Tesla Crash Raises New

autonomous machines can, presumably, continue to work indefinitely, it follows that combat involving AWS (at this level, without human involvement) 'cannot practically be won or lost'.[258] On a battlefield dominated by AWS, once machines have defeated opposing machines, humans will still need to move onto that battlefield to negotiate and settle the dispute themselves.[259] It is for this reason that Beard notes that, while legal and ethical issues of AWS deployment are widely covered[260], there is little secondary source material yet available on empirically *how* autonomous weapons may be used on the battlefield.[261] Two quite separate points arise. Gray notes the 'inescapable reality of geography and the ubiquitous nature of the elements'.[262] Clausewitz, after all, highlights that 'geography and the character of the ground bear a close and ever-present relation to war'.[263] A second exogenous challenge comes from the psychological cost of AWS deployment. UAV pilots are showing signs of equal or greater stress from combat compared to traditional pilots; although an over-simplification, the stress of fighting a war several thousand miles away then joining one's family at the dinner table is already presenting significant mental health challenges.[264] In this vein, Sharkey similarly that operating a drone is already akin to a video game in its detachment from the act of killing.[265]

5.7 Proliferation constraints

An adjunct constraint arises from the challenges of UAV and AWS proliferation.[266] Just as Krupp, Winchester and Armstrong were happy to sell advanced munitions to competing nations, the marketplace for drone manufacturers is already global, an inference here being that States will be able to level any technical playing field by purchasing comparable AWS technology.[267] Boyhan, after all, notes that States who fail to adopt such new technologies may be disadvantaged as autonomy spreads in weaponry.[268] Moreover, horizontal proliferation (that is, to other countries) of UAV

Questions About Autopilot System', *New York Times*, 31 May 2018 <https://www.nytimes.com/2018/03/31/business/tesla-crash-autopilot-musk.html> [accessed 5 July 2018].

[258] See: Tom Simonite, 'Sorry, banning 'Killer Robots' just isn't practical', *Wired Magazine*, 22 August 2018 <https://www.wired.com/story/sorry-banning-killer-robots-just-isnt-practical/> [accessed 16 March 2018].

[259] Tonkens, p. 163.

[260] A high-level search on 'ethics', 'legal', autonomous weapons' in the British Library's in-house indexing system returns more than 1,400 citations and references [accessed 7 December 2016].

[261] Beard, pp. 622-625.

[262] Colin Gray, 'Inescapable Geography', *Journal of Strategic Studies*, 22, 2-3, (1999), 161-177.

[263] Karl Clausewitz, *On War*, V, 17, p. 348.

[264] Hearing on 'The Rise of Drones; Unmanned Systems and the Future of War', *Committee on Oversight and Government Reform*, Congressional Research Service, (March 2010) <http://digitalcommons.wcl.american.edu/cgi/viewcontent.cgi?article=1002&context=pub_disc_cong> [accessed 15 June 2017] generally.

[265] Noel Sharkey, 'Saying No to Lethal Autonomous Targeting', *Journal of Military Ethics*, 9, 4, (2010), 372.

[266] Lynn Davis and others, *Armed and Dangerous? UAV and US Security*, (USA: Rand Corporation, 2014), pp. 7-10 <http://www.dtic.mil/dtic/tr/fulltext/u2/a599239.pdf>. The document also provides a useful general primer into AWS development and proliferation.

[267] Micah Zenko and Sarah Kreps, 'Limiting Armed Drone Proliferation', C*ouncil on Foreign Relations, Center for Preventative Action*, Council Special Report 69, (June 2014), pp. 6-8 <https://cfrd8-files.cfr.org/sites/default/files/pdf/2014/06/Limiting_Armed_Drone_Proliferation_CSR69.pdf>. The report cites drone exports rising from $5.2bn in 2013 to $8.4bn in 2018. See also: Jason Koebler, 'Oregon Company To Sell Drone Defence Technology to Public', *US News*, 15 March 2013 <http://www.usnews.com/news/articles/2013/03/15/oregon-company-to-sell-drone-defense-technology-to-public> [accessed 9 June 2017].

[268] David Boyhan, 'Autonomous Weapon Proliferation', *NY University*, undated <https://itp.nyu.edu/classes/foti-fall2011/files/2011/11/autonomous-weapons-dboyhan.pdf>. Proliferation is also discussed by: Noel Sharkey, 'The

technologies is already widespread; research from Dortmund University suggests that some twenty countries as at 2019 export such systems.[269] Horizontal proliferation already takes the form of qualitative improvements to UAV technology through collaborative development (such as between Israel and India). Technology development and collapse in cost curves, both characteristics underpinning Seba's S-Curve adoption model, are also evident. [270] The proliferation issue here is that an air-frame can already be built on a 3D printer[271] and a cheap version of Hobby King Software can provide users with a useable software skeleton to create an autonomous flight path.[272] Unregulated proliferation of AWS' underlying technologies[273] is thus clearly widespread. A further proliferation challenge arises from the ease with which adversaries might duplicate and improve on weapon capabilities through espionage, graft, State-sponsored malfeasance or an AWS being captured.[274] Examples abound.[275] The Chukar, a target drone system used by the US, provided the basis for the Israeli unmanned aerial system having crashed in the waters off Tel Aviv in 1968 and immediately reverse-engineered to provide the basis for future Israeli efforts in the technology.[276] Sharkey similarly points to developments in AWS technologies that have been helped by easy availability of physical parts as well as the increasing cross-border ease of sourcing necessary components.[277]

An obvious proliferation concern is that unmanned and, eventually, autonomous technology might be exploited by terrorist and non-state players.[278] Notwithstanding that 'threat inflation'[279] plays its part in keeping that possibility in the public eye, assassination, dissemination of chemical and biological agents and unmanned access into secure areas are, notes Masters, likely to become

'Evitability' of Autonomous Robotic Warfare', *International Review of the Red Cross*, 94, 886, (Summer 2012), p. 798 <https://www.icrc.org/eng/assets/files/review/2012/irrc-886-sharkey.pdf>.

[269] Jurgen Altmann, '*Preventive Arms Control for Uninhabited Military Vehicles',* Experimentelle Physik III, Technische Universitat Dortmund, 2004, p. 78.

[270] Tony Seba, 'Clean Disruption' <https://www.youtube.com/watch?v=2b3ttqYDwF0>.

[271] Gizmag blog, <http://www.gizmag.com/3d-printed-uav-airframe/31473/> [accessed 24 May 2015].

[272] Paul Scharre, Director, 20YY Warfare Initiative, *Centre for a New American Security*, in conversation with the author, February 2014 Chatham House conference on Autonomous Weaponry.

[273] See: Economist Magazine, 'From here to Autonomy', *Economist*, 1 March 2018 <https://www.economist.com/news/special-report/21737420-making-vehicles-drive-themselves-hard-getting-easier-autonomous-vehicle-technology> [accessed 1 April 2018].

[274] Sharon Weinberger, 'China Has Already Won the Drone Wars', *Foreign Policy*, 10 May 2018, para. 4 of 11 <https://foreignpolicy.com/2018/05/10/china-trump-middle-east-drone-wars/>.

[275] For an early discussion on relevant structural changes in the arms industry, see: Richard Bitzinger, '*The Globalization of the Arms Industry: The Next Proliferation Challenge',* International Security, 19, 2, (Fall 1994), 170-198.

[276] Source:, 'MQM/BQM-74', *Directory of US Military Rockets and Missiles* <http://www.designation-systems.net/dusrm/m-74.html> [accessed 2 January 2014].

[277] Sharkey, 'Automated killers and the computing profession', generally.

[278] In December 2009 the Wall Street Journal reported that Iraqi insurgents had used a $26 kit to tap into live video feeds from a circling Predator drone; see: Siobhan Gorman and others, 'Insurgents Hack US Drones', *Wall Street Journal*, 12 December 2009 <https://www.wsj.com/articles/SB126102247889095011> [accessed 27 October 2017].

[279] The role of 'threat inflation' in accelerating AWS development is widely discussed whereby constituencies exaggerate the strength of possible threats and capabilities and the ramification of this upon national security in order to create a self-fulfilling policy escalation; see: Sean Lawson, 'Domestic 'Drones' Are the Latest Object of Threat Inflation', *Forbes Magazine*, 18 April 2014 <http://www.forbes.com/sites/seanlawson/2014/04/18/domestic-drones-are-the-latest-object-of-threat-inflation/> [accessed 12 February 2018].

ever easier for such non-State parties to undertake.[280] Michelson also points out that such capabilities are unlikely now to disappear given that unmanned systems are increasingly based on universally available platforms.[281] The hazard is acknowledged by the US Department of Defense in its 2012 *Directive on Autonomous Weaponry* that requires clear 'safeties' and 'anti-tamper mechanisms'[282] in the event of a loss of system control to 'unauthorised parties'. In this vein, Altmann argues that proliferation of AWS might induce States to start wars with greater ease, particularly in industrialised democracies where casualties may quickly erode support for such action.[283] This trend reinforces what is an age-old security dilemma: Each State must build up its own AWS forces and, by doing so, increases the threat to its neighbours who, in turn, adopt this new technology. The behavioural challenge is that security is decreased for all.[284]

5.8 Ethical constraints to AWS deployment

Ethical arguments around the removal of human decision-making in lethal engagements are similarly polarized but, as evidenced broadly by Lin, comprise a key challenge to AWS deployment.[285] Tonkens and Arkin occupy the opposite reaches of this debate. While Arkin's advocacy of an *Ethical Governor* and *Artificial Conscience* is explored in this thesis' Chapters Three and Eight[286], Tonkens' rebuttal posits that lethal robots on the battlefield will instead create new levels of uncertainty.[287] Robots, moreover, will not be capable of morally praiseworthy behaviour such as courage (both moral and physical) or the ability to go, as Tonkens explains, 'above and beyond duty'.[288] At its core (and validated in later chapters[289]), his argument is that critical ethical notions are incapable of capture by machine code. These tenets include heuristics fundamental to battlefield function (and, argue Sturchler and Slergrist, pivotal to the compliant use of force[290]) such as guilt, empathy, responsibility, duty and restraint. The list includes precepts such as concern, unease, disquiet and compassion. This analysis is not to argue that all human soldiers unerringly display such characteristics but rather that AWS will fail to meet appropriate benchmarks from the outset.

---

[280] Jonathon Masters, 'Targeted killings', *Council on Foreign Relations*, Backgrounders, (May 2013) <http://www.cfr.org/counterterrorism/targeted-killings/p9627>.

[281] Robert C. Michelson, 'Micro Flyers and Aerial Robots, Missions and Design Criteria', *Georgia Tech Research Institute*, (2009) <http://www.dtic.mil/dtic/tr/fulltext/u2/p010759.pdf>.

[282] Department of Defense Directive, 'Autonomy in Weapon Systems', para 4 (*Policy*), point 2.

[283] Jurgen Altmann, 'Preventive Arms Control for Uninhabited Military Vehicles', AKA Verlag Heidelberg, (2009), 70 <https://e3.physik.tu-dortmund.de/p&d/pubs/0909_ethics_and_robotics_altmann.pdf>.

[284] Kate Allen, 'Now For An Arms Trade Treaty', *The Guardian*, 24 May 2012 <https://www.theguardian.com/law/2012/may/24/global-arms-trade-treaty> [accessed 19 March 2018].

[285] For a useful primer, see: Patrick Lin, 'Killer Robots: New Reasons to Worry about Ethics', *Forbes Tech*, 4 January 2016 <https://www.forbes.com/sites/patricklin/2016/01/04/killer-robots-new-reasons-to-worry-about-ethics/#276edb9727e5> [accessed 23 April 2018].

[286] See: Chapter 3 (*Drivers*), specifically: 3.4 (*Ethical drivers*) and 8 (*Software*), specifically: 8.1 (*Coding methodologies*).

[287] Tonkens, p. 149.

[288] Ibid., p. 151.

[289] See: Chapter 8 (*Software*), generally but specifically: 8.1 ('*Coding methodologies')* and *8.4 ('Software processing functions'*).

[290] Nikolas Sturchler and Michael Slergrist, 'A 'Compliance-base' Approach to Autonomous Weapon Systems', *Blog of the European Journal of International Law*, (1 December 2017) <https://www.ejiltalk.org/a-compliance-based-approach-to-autonomous-weapon-systems/> [accessed 7 May 2018]. Also: Major-General Patrick Cordingley, in conversation with the author, May 2017.

In considering deployment challenges, it is therefore instructive to unpick certain of the ethical drivers discussed in Chapter Three.[291] First, Atkin's arguments around AWS' inevitability' are unfounded given that humans always retain the choice in practice to retain human supervision. The entire human chain around AWS deployment (those comprising the Delivery Cohort, here the battlefield commander, the maintenance, control and logistics staff[292]) must, after all, be governed by its own set of values, standards and experiences, each affecting how and to what ends autonomy is deployed. Nor does AWS' monitoring of the battlefield (and, implicitly, Arkin's notion of reprimand in cases of *unethical* behaviour) actually require the autonomous machine to have its own independent lethal capability.[293] This is a specious connection. Indeed, the shortcomings of Arkin's framework are recognised with its author acknowledging that finding ways for AWS to adhere to LOAC remains an 'outstanding issue'[294] and that the challenge for AWS to distinguish a soldier from a civilian is one of several 'daunting problems'.[295] Controversy also exists within the issue of what comprises a 'correct' ethical framework for AWS protocols given the multiple candidates in available philosophies.[296] Different States, different politicians and different programmers are unlikely to have a consistently similar ethical framework, an acute deployment issue given the importance of dynamically tuning the weapon's dynamic thresholds and confidence levels discussed in later chapters.[297]

It is also relevant to review ethical challenges in their empirical context. Arkin's model does not suggest AWS might reliably prevent humans committing atrocities.[298] Indeed, it can be assumed that most instances of human unethical activity will occur 'out of sight' of the autonomous machine. Improving the moral calibre of battlefield combat is, moreover, already irrelevant if the use of AWS takes place in a war that is unjust in the first place. This portents clear difficulty for AWS deployment that is highlighted by Volker, former US Permanent Representative to NATO, in his 2012 conclusion that 'drone strikes allow our opponents to cast our country as a distant, high-tech, amoral purveyor of death. It builds resentment and alienates those we should seek to inspire'.[299] As noted by Dworkin, ethical challenges to removing weapon supervision are thus properly fundamental.[300] There are, of course, ways other than autonomy of improving ethical behaviour on

[291] Chapter 3 (*Drivers*), specifically: 3.6 ('*Ethical drivers'*).

[292] Together, the 'Delivery Cohort'; see: Chapter 6 (*Software*), specifically: 6.3 ('*Delivery Cohort'*).

[293] See: Arkin, *Governing Lethal Behaviour,* p. 39.

[294] Ibid., p. 126 and p. 211.

[295] Arkin, 'The Case for Ethical Autonomy in Unmanned Systems', pp. 11-12 <https://www.cc.gatech.edu/ai/robot-lab/online-publications/Arkin_ethical_autonomous_systems_final.pdf>.

[296] Unitarian, Kantian, Social Contract, Virtue Ethics, Cultural Relativism et al. Arkin's answer to this is to start from a 'first, do no harm' strategy that sets the default for AWS.

[297] See, inter alia: Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues*') and 8.6 ('*Value setting issues'*).

[298] Indeed, Arkin identifies specific problems to the creation of ethically sensitive machines including the abstract nature of laws, codes and principles of military law, the variety of interpretation of these laws depending on context and the frequent conflict that exists between these abstract rules. See: Tonkens, p. 158.

[299] Kurt Voller, 'What the US risks by relying on drones', *Washington Post*, 26 October 2012 <http://articles.washingtonpost.com/2012-10-26/opinions/35500650_1_drone-strikes-drone-attacks-guantanamo-bay> [accessed 12 April 2014].

[300] Gerald Dworkin, *The Theory and Practices of Autonomy,* (Cambridge: Cambridge University Press, 20 August 1988), pp. 62-65.

the battlefield[301] such as shorter tours, improving therapeutic resources, tightening up on soldiers' psychological screening (indeed, even under the US Surgeon General's research[302], more than ninety per cent of soldiers have displayed conduct that *is* acceptable) by improved training and, as well, by revising rules of engagement. More importantly, it is unfounded to state *without doubt* that AWS will be 'ethically more sound' than human soldiers, the more so, notes Widdows, given the number and scope of morally indeterminate situations that arise in battle.[303] It is for this reason that Schulzke highlights an enduring fit between human participants (as opposed to machines) and LOAC where combat is based upon an ethical 'do-the-best-you-can' framework.[304]

A further ethical constraint then arises from humans' investigative role should violations of IHL take place. Under the *Geneva Conventions*, the notion of accountability is central to deterring future harm to civilians while providing victims with a means of retribution:[305] Critics can thus argue that autonomous weapons should be banned because they 'inherently preclude the fair attribution of responsibility. Fair criminal liability presupposes that commanders can foresee the outcome of [AWS'] actions'.[306] In this vein, Kalmonovitz concludes that deploying lethal AWS under 'complete uncertainty would clearly be wrong, comparable to releasing a toxic substance in inhabited environment'.[307] Primarily, facility must exist to identify responsible parties when AWS processes go awry. The corollary (taken together, the transparent and collective responsibility of AWS' Delivery Cohort[308]) is captured by former UN Special Rapporteur Heyns when he notes that '[t]here is clearly no point in putting a robot in jail'.[309] This matter of responsibility gives rise to important ancillary challenges. Even when constituents of that Cohort are judged to exercise material control over AWS, uncertainty will exist 'whether each or any individual should be held to account in a specific case'.[310] After all, if everyone (in this case, the Delivery Cohort) is deemed to be responsible, then no one is responsible. Heyns also notes the ethical constraint that it is unclear *who* will investigate civilian deaths following erroneous engagement by an AWS. Uncertainty around accountability otherwise creates an unacceptable 'responsibility gap'.[311] Either the AWS has

[301] See: Peter Olsthoorn, *Military Ethics and Virtues: An interdisciplinary approach for the twenty-first century,* (USA: Cass Military Studies, Routledge, 2011), p. 81 and generally <https://philpapers.org/archive/OLSMEA.pdf>.

[302] Surgeon General's Office, Mental Health Advisory Team, 'IV Operation Iraqi Freedom 05-07, Final Report'*, MHAT*, (November 2006), generally. See: Chapter 3 (*Drivers*), specifically: 3.4 ('*Ethical drivers'*).

[303] Heather Widdows, 'The ethics of warfare: Is it ever morally right to kill on a massive scale?'*, University of Birmingham, Perspectives*, undated <https://www.birmingham.ac.uk/research/perspective/ethics-of-warfare-heather-widdows.aspx> [accessed 18 April 2018].

[304] Marcus Schulzke, 'Ethically Insoluble Dilemmas in War'*, Journal of Military Ethics*, 12, 2, (2013), generally. See also: Eric Jackson, 'Sun Tsu's Art of War'*, Forbes*, (23 May 2014) <https://www.forbes.com/sites/ericjackson/2014/05/23/sun-tzus-33-best-pieces-of-leadership-advice/#19c3ac7d5e5e> [accessed 2 August 2017].

[305] HRW, 'Losing humanity; the case against killer robots', p. 42.

[306] Kalmanovitz, cit. Bhuta and others, p. 154.

[307] Ibid.

[308] See: Chapter 6 (*Software*), specifically: 6.3 ('*Delivery Cohort'*).

[309] Christof Heyns, 'Autonomous Weapon Systems: living in a dignified life and dying a dignified death', cit. Nehal Bhuta and others, p. 12.

[310] Ibid.

[311] Heyns and other also use the term *accountability vacuum*.

been deployed unlawfully or its programming is unlawfully careless or it has been illegally designed or neglectfully maintained.[312]

The current process of determining accountability in examples of AWS malfunction is also ambiguous. *Protocol* I, Article 85(3) of the *Geneva Conventions*, specifies that individuals can only be held liable if any such carelessness is 'willful' and intentional. Manufacturers have empirically gone unpunished for how their weapons have been used, particularly having first disclosed during their procurement the risks associated with those weapons. It is the unpredictable bases of AWS that will ensure manufacturers set out the widest possible specification caveats. Prosecution of product liability also entails a civilian lawsuit that puts the onus to act upon the victims, often wretched and displaced by conflict. Boulanin and Verbruggen note, furthermore, that it is AWS' technical conformation which creates accountability challenges;[313] while it may be theoretically possible to separate out obligations attached a particular software engineer (or, post facto, to identify those responsible for discrete elements of the system's componentry), Sparrow points out that these individuals cannot be expected to predict the weapon's subsequent learning or its decision outcomes in battle.[314] After all, 'the possibility, however far-fetched, that an autonomous system will make choices other than those predicted and encouraged by its programmers is inherent in the claim that it is autonomous'.[315] In highlighting the large number of agencies comprising AWS' Delivery Cohort[316], Epstein is making the ethical and behavioural point that the absolute numbers involved may collectively make an accident more likely as well as blurring the lines of responsibility when that accident occurs.[317] It is also unworkable that any one element of this Delivery Cohort can conceptualize all of the complexities that are involved either in AWS deployment or in the battlespace in which that machine must properly operate.[318] McDaniel thus notes that 'steady advances in technology will reveal legal and ethical issues that are currently unimaginable'.[319] An ethical adjunct to the matter of accountability is the notion of 'plausible denial', a circumstance that will occur when no single State then admits to a lethal engagement using AWS or when it is difficult to determine *whose* weapon system is responsible in an attack, an unwelcome consequence, notes Suarez, of States' moves to independent weapons.[320] Such an anonymous act of war has the potential to disrupt the geopolitical balances and turn, perhaps, 'States' long-standing onus from defence into one of attack'.[321]

---

[312] Walker, *Killer Robots?,* pp. 77-78.

[313] Boulanin and Verbruggen, pp. 7-12.

[314] Robert Sparrow, 'Killer Robots', *Journal of Applied Philosophy*, 24, 1, pp. 69-70 <http://staffwww.dcs.shef.ac.uk/people/A.Sharkey/Sparrow.pdf>.

[315] Ibid., p. 70.

[316] See also: Chapter 6 (*Wetware*), specifically: 6.3 ('*Delivery Cohort'*).

[317] Richard Epstein, 'The Case of the Killer Robot', (West Chester, PA: University of Pennsylvania, January 1997) <https://www.researchgate.net/publication/242362942_The_case_of_the_killer_robot>.

[318] Scharre, 'Autonomous weapons and operational risk', *Centre for a New American Security*, (2016), 25-34 <https://www.files.ethz.ch/isn/196288/CNAS_Autonomous-weapons-operational-risk.pdf>.

[319] Erin McDaniel, 'Robot Wars: legal and ethical dilemmas of using unmanned robotic systems in 21st Warfare and beyond', unpublished MA, Military Art and Science, (*Fort Leavensworth, 2008*), p. 70.

[320] Suarez, 'The kill decision shouldn't belong to a robot', generally.

[321] Ibid.

A final argument for this section is provided by the notion of 'human dignity' acting as an umbrella concept which, according to Birnbacher,

> … bridges seemingly insurmountable ideological gulfs [and provides] a basis for consensus and compromise as a foundational principle that overarches, as it were, all constitutional and other political principles, a common reference point that is beyond the controversy and conflict and plays the role of an *a priori* to which all other political ideas are subject.[322]

An ethical issue is also whether autonomous weapons and, specifically, the removal of human supervision from lethal violence, infringes human dignity. In judging AWS to be *mala in se* under Just War theory ('evil in themselves'), Horowitz cites Asaro's overarching conclusion that 'justice cannot be delegated to automated processes'.[323] The construct is given weight by UNIDIR's discussion of 'an instinctual (sic) revulsion against the idea of machines "deciding" to kill humans'[324], prompting Birnbacher to posit human dignity as a basic challenge to AWS deployment given the 'openness of its content' and human dignity's 'independence of any particular metaphysical background theory'.[325] This 'openness' is relevant precisely because human dignity cannot be pigeonholed by definitions around human rights.[326] Furthermore, unlike the concept of morality (that comprises both rights and duties), human dignity implies rights against others but no duties against others. In this sense, there is obvious disconnect between human dignity and lethal engagement undertaken by unsupervised weaponry. The context of human dignity, moreover, is unlimited as it contains 'an intrinsic evaluative component' which, notes Birnbacher again, gives to human beings an exclusive value 'on which the exceptional normative status of human beings is assumed to depend'.[327] The ethical challenge in this case is that such value is incapable of appropriate capture in machine code.[328]

The human rights that are implied by human dignity (and which will likely be compromised by AWS function) include the right to avoid humiliation, the right to a minimum freedom of action and decision, to receive support in situations of severe need, the right to a minimum quality of life and the relief of suffering and, crucially, the right not to be treated merely as a means to other people's

---

[322] Dieter Birnbacher, 'Are Autonomous Weapons a threat to human dignity*?', cit.* Bhuta et al, p. 105. The concept of human dignity was first introduced into Article 1 of the United Nations Universal Declaration of Human Rights in December 1948 (UN GA Res 217 III A). Similarly, the Vienna Declaration of the 1993 World Conference on Human Rights affirmed that 'all human rights derive from the dignity and worth inherent in the human person'.

[323] Michael Horowitz, 'The Ethics and Morality of Robotic Warfare: Assessing the Debate over Autonomous Weapons', *American Institute of Arts and Sciences*, (February 2016), pp. 12-14 <http://www.michaelchorowitz.com/Documents/HorowitzLAWSEthicsDraftFeb2016.pdf>. See also: Asaro, '*On Banning Autonomous Weapon Systems'*, p. 710.

[324] UNIDIR, 'Safety, Unintentional Risk and Accidents in the Weaponization of Increasingly Autonomous Technologies*', UNIDIR*, 5, (2016), p. 9 <http://www.unidir.org/files/publications/pdfs/safety-unintentional-risk-and-accidents-en-668.pdf>

[325] Birnbacher, pp. 105-107.

[326] Human Rights Watch, 'Heed the Call', pp. 19-22 ('*Humane* treatment') and 23-27 ('*Respect for Human Life and Dignity'*).

[327] Birnbacher, pp. 105-107.

[328] Birnbacher, p. 106. See also: Chapter 8 (*Software*), specifically: 8.1 ('*Coding Methodologies'*) and 8.5 ('*Anchoring and Goal Setting Issues'*).

ends.[329] Human dignity incorporates a civilian's right to privacy as well as rights to live without severe harm or risk of harm.[330] The deployment challenge in this instance is that removing weapon supervision instead risks making civilians 'the mere means of aims that are in no way their own aims, with the risks of serious harm to life and physical and mental integrity'.[331] This manifests itself in three ways. As noted by Heyns, AWS are 'invulnerable by being free from fear', with their deployers carrying 'no cost except the economic'.[332] Second, AWS' intrinsic unpredictability as well weapons' likely randomness of attack exacerbates the threat to civilians. Finally to this point, Sharkey notes that AWS' 'illusion of accuracy' will foster over-confidence and distort the Delivery Cohort's judgement.[333] A corollary, therefore, to this chapter's analysis of AWS' legal and behavioural challenges is articulated by Brehm. It is the degree of *control* that military commanders exercise over their weapons, and the challenges arising from that relationship, which will 'affect their ability and, by extension, that of the State on whose behalf they act, to perform their legal duties and to be accountable for the consequences'.[334] It is this portfolio of legal, political, ethical and behavioural constraints (as well as their intrinsic intractability) that underpins civil society's opposition to AWS deployment[335] and which now provides relevant narrative to the remainder of this thesis' analysis, the identification of *technical* bottlenecks that challenge the compliant adoption of AWS.

[329] See, generally: 'Human Rights and Armed Conflict', *Icelandic Human Rights Centre*, <http://www.humanrights.is/en/human-rights-education-project/human-rights-concepts-ideas-and-fora/human-rights-in-relation-to-other-topics/human-rights-and-armed-conflict> [accessed 3 May 2018].

[330] Articles 8 and 21 of the Rome Statute of the International Criminal Court prohibit and make punishable 'committing outrages upon personal dignity, in particular humiliating and degrading treatment'.

[331] Birnbacher, p. 116. Birnbacher focuses in particular on the asymmetry of facing forces if autonomous weapons are employed only on one side. 'Morally, symmetry and asymmetry are crucial moral variables', he claims, 'and they are deeply influenced by the use of robots'.

[332] Christof Heyns, 'Report of the Special Rapporteur', p. 12.

[333] Noel Sharkey, 'Automating warfare: lessons learned from the drone', *Journal of Law*, Information and Science, (2012), 12 <www.austlii.edu.au/au/journals/JlLawInfoSci/2012/8.html> [accessed 12 January 2018]. See also Chapter 6 (*Software*), specifically: 6.3 ('*The Delivery Cohort*').

[334] Maya Brehm, 'Defending the Boundary; constraints and requirements on the use of autonomous weapon systems under international humanitarian and human rights law', *Geneva Academy of International Humanitarian Law and Human Rights*, Academy briefing No.9, (2017), p. 19.

[335] Peter Asaro, 'Why the World Needs to Regulate Autonomous Weapons, and Soon', *Bulletin of the Atomic Scientists*, 27 April 2017, <https://thebulletin.org/2018/04/why-the-world-needs-to-regulate-autonomous-weapons-and-soon/> [accessed 15 May 2018].

6. Wetware: Design challenges to AWS function

Having assessed 'why' and 'how' a weapon without human oversight might be deployed, this thesis now considers *technical* obstacles to compliant deployment of weapons-directing artificial intelligence. Only once such 'build' challenges have been identified can a review of operational considerations to AWS deployment reasonably be carried out.[1] This technical analysis is therefore undertaken in the next four chapters titled *Wetware*, *Firmware, Software* and *Hardware*. Three assumptions underlie the review. Sartor and Ominici define a teleological system (here, the AWS) by its cognitive states having a definable representation such as goals (objectives to be achieved by the weapon), beliefs (the tracking of its environment) and plans (paths specifying how to reach the goals, given the beliefs, through actions of the weapon within a belief-desire-intention architecture).[2] Any faultlines so identified are deemed to be agnostic to the precise *degree* of autonomy in the underlying weapon platform. For the purposes of this thesis, the test remains the 'absence of meaningful human control'[3] in an engagement sequence described by Air Force Colonel Riza as 'the string of events [that] we technological warriors facetiously call the "consecutive miracles" that comprise the effective functioning of technologically advanced weapon systems'.[4] The purpose of this chapter is thus to review AWS' most fundamental technical basis, the transforming of data from its battlefield environment into purposeful plans and compliant actions based on appropriate integration of capabilities that sense, decide and then act in an independent and lethal manner.[5] The outwardly similar models used by Russell and Norvig to classify such 'intelligent' agents (here, reactive rule-based systems or deliberative goal-based systems) highlight the complexity masked by this standpoint whereby very different decision outcomes may arise from exactly similar input data.[6] This is a key assumption for this thesis as AWS deployment may take many forms[7]

from individual independent componentry within an otherwise supervised platform to, Oberhaus posits, an off-the-shelf drone kit where targets are selected according to broad search instructions (ethnicity, location, gender, gait, age, even 'target reaction').[8] That these represent just two points of

[1] This comprises the subject of Chapter 4 (Deployment), generally.

[2] Sartor and Omicini, p. 51. Sartor and Omicini usefully state that 'in order to realise its desires (goals), the system constructs plans of action on the basis of its model of the relevant facts (beliefs) and commits itself to act according to the chosen plans (intentions). Note that by using the terminology of beliefs, desires and intentions to denote cognitive structures of artificial systems, we are not assuming that such structures are similar to those existing in the human mind'. This is a pivotal distinction.

[3] For analysis of this notion, see: Chapter 10 (*Oversight*), specifically 10.1 ('*Meaningful human control*').

[4] Shane Riza, *Killing without Heart,* (USA: University of Nebraska Press, Potomac Books, 2013), p. 4.

[5] Boulanin and Verbruggen, p. 7. Capabilities around sensing (to complete a task autonomously, the weapon must be able to perceive the battlefield environment in which it operates), around deciding (data from the weapon's surroundings, once processed by the machine's sensing software, then serves as input for the AWS' decision-making processes which in turn are overlooked and 'assured' by its control systems) and around acting (the decisions made by the AWS' control systems are then executed through computational or physical means) may provide the skills basis for autonomous AWS but this masks considerable technical uncertainty and likely differences to be overcome before a working set of machine routines can be achieved.

[6] Stuart Russell and Peter Norvig, *Artificial Intelligence: A modern Approach,* (UK: Pearson Education, Harlow, 2014), p. 35 and p. 49.

[7] See: Chapter 4 (*Deployment*), generally.

[8] Daniel Oberhaus, 'Watch 'Slaughterbot': A Warning about the Future of Killer Robots', *Motherboard*, (13 November 2017) <https://motherboard.vice.com/en_us/article/9kqmy5/slaughterbots-autonomous-weapons-future-of-life>

what is an *evolving* continuum is evidenced by debate in the UN's CCW that after five years of discussions cannot yet agree on a working definition on what constitutes a lethal autonomous weapon.[9] More likely, it looks like a colleague machine with autonomous and lethal capabilities but, as part of a human-machine 'team', operating nevertheless with, in theory, a human overlay whereby control is then toggled between human and weapon 'according to circumstance'.[10]

How then are these four chapters organised? First, in *Wetware*, this chapter isolates basic complexities arising from the plausible fundamental architecture in autonomous weapons. Its aim is to establish likely architectural bases for AWS in order to underpin the analysis of subsequent chapters. What might an appropriate artificial intelligence look like for an independent weapon system? Given that this weapon is being tasked with independent lethal operation, how might its high-level architecture learn, reason and undertake cognition in order to fulfill such assignment? *Firmware* (Chapter Seven) then develops this analysis to review the architectural bases that may comprise AWS function. For the purposes of this thesis, firmware here relates to the permanent routines (weapon learning, reasoning and direction) underpinning AWS' architecture. In *Software* (Chapter Eight), the thesis then identifies fault lines arising from the weapon's volatile operational routines that are likely to comprise its function.[11] Finally, in *Hardware* (Chapter Nine), challenges arising from AWS' physical properties are discussed. Only then is assessment possible on whether unsupervised weapons can be fit for purpose given the nexus between compliant operation, technical challenge and the role expected of AWS. Taken together, the review seeks to pinpoint discrepancies that exist between *capabilities* that may be feasible and *tasks* that will be required of unsupervised weapons.

Wetware[12] is vernacular to describe the *human* element of information technology architecture. The human brain is composed of some seventy-five per cent water.[13] The term is useful in describing programmers' efforts to replicate the essentials of human intelligence and, crucially, to effect this through code.[14] Indeed, the narrative to the next four chapters is informed by

---

[accessed 17 November 2017]. For video presentation, see: <https://www.youtube.com/watch?v=2CYvjjOwcWQ> [accessed 17 November 2017]. 'Slaughterbots' is presented by Professor Stuart Russell, Director of Computing at Berkeley University, who concludes 'this is not speculation. It is the result of integrating and miniaturizing technologies that we already have'. The theme also frames much of the introduction to Chapter 2 ('Context').

[9] Campaign to Stop Killer Robots, (17 November 2017), paras. 5 and 7 of 10 <https://www.stopkillerrobots.org/2017/11/gge/, 17 November 2017> [accessed 18 November 2017]. HRW's Mary Wareham points out that the CCW has met six times since 2014. She also notes that the stated goal of the CCW deliberations has not been to produce a working definition. The consensus-agreed 2018 CCW GGE Report finds that: 'For some delegations, a working definition of lethal autonomous weapons systems is essential to fully address the potential risks posed. For others, absence of an agreement on a definition should not hamper discussions or progress within the CCW. Characterisation, or working definitions, should neither predetermine nor prejudge policy choices; they should be universally understood by the stakeholders' (Mary Wareham, Director, HRW Arms Division, in conversation with the author, December 2018).

[10] See: Chapter 4 (*Deployment*), specifically: 4.3 ('*Machine and human teaming models*') and 4.5 ('*Flexible autonomy'*). For definitions of in-the-loop, on-the-loop and out-of-the-loop weapon controls, see: Chapter 1 (*Introduction*) p. 16. See also: 'Superbots' at <https://www.stopkillerrobots.org/2017/11/gge/> [accessed 5 October 2017].

[11] Volatility in this case relates to the weapon's temporary generation of outcomes. Unlike the AWS' *firmware* (and arising from the weapon's set of battlefield inputs), this is obviously liable to rapid change.

[12] See: Urban Dictionary <http://www.urbandictionary.com/define.php?term=wetware> [accessed 12 August 2017].

[13] USGS, para. 2 of 8 <https://water.usgs.gov/edu/propertyyou.html> [accessed 18 August 2017].

[14] One component of AWS' 'Delivery Cohort'. For the purposes of this thesis, the broad Delivery Cohort describes the several parties responsible for the design, implementation and deployment of AWS. The term is deliberately undefined but will include, inter alia, the following tasks: neurophysiologists to coordinate AWS networks, psychologists to

AWS' independent and learning capabilities and, it is posited, an element of sentient initiative. Given its basis around cognitive traits of the human brain, wetware is thus a central component in weapon feasibility. The aim, however, of the chapter is really to evaluate Krishnan's suggestion that this basic construct is simply 'a holy grail in AI research; highly desirable but unattainable'.[15] Sharkey concurs in reckoning that developing the required independent intelligence will 'remain science fiction – at least for the next one hundred years and maybe always'.[16] This thesis' technical review is generally framed by citing Moravec's paradox[17] and the challenging links that exists between a weapon's low-level sensor-motor skills, the requirement for disproportionate computational resources and the ensuing fragility that this entails.[18] The relevant context is that '[i]t is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible' concludes Moravec 'to give them the skills of a one-year-old when it comes to perception and mobility'.[19] The same context is noted by Pinker*: '*The mental abilities of a four-year-old that we take for granted – recognizing a face, lifting a pencil, walking across a room, answering a question – are in fact solve some of the hardest engineering problems ever conceived'.[20] It recently took Berkeley University's towel-folding robot more than ten hours to replicate the human folding of just twenty-five towels.[21] This wetware issue is common to AWS and arises from '*complexity layering'.*[22] Independent weapons must be reliably capable of actioning known tasks and known unknown tasks but must also have routines immediately available to process and derive actions from unknown unknown tasks. Various factors require emphasis in this introduction. Technical delineation between subject matters in chapters Six through Nine (*Wetware, Firmware, Software* and *Hardware*) is not clear-cut; after all (and as inferred from Nwana), the role of software components throughout AWS operation is pervasive.[23] Finally to this point, the chapters are intended to plot a likely but not definite set of weapon architectures and do so from a deliberately behavioural rather than technical perspective: the

coordinate learning and cognition, biologists for adaption strategies, engineers for control routines, logisticians, roboticists, electrical specialists, behaviorists, politicians, NGOs, sociologists, lawyers, company directors, weaponists, military tacticians, politicians, civil servants and diplomats, manufacturers, professionals involved in miniaturization, simulation, configuration, coding, power supply and modularity, specialists in sensors, in distributed and decentralized routines, ethicists, specialists in tooling and calibration. See also: 6.3 ('*The Delivery Cohort'*).

[15] Krishnan, *Killer Robots,* UG479.K75, (UK: Ashgate Publishing, 2009), p. 48.

[16] Human Rights Watch, 'Losing Humanity*'*, p. 29.

[17] Kelly Clancy, 'A Computer to Rival the Brain'*, New Yorker Magazine*, 15 February 2017, para. 2 of 9 <http://www.newyorker.com/tech/elements/a-computer-to-rival-the-brain> [accessed 16 August 2017]. For a discussion of the *Paradox*, see: introduction to Chapter 9 (*Hardware*).

[18] Sally Doherty, 'Narrow versus General AI – Is Moravec's Paradox still relevant?*', Graphcore Magazine*, January 2017, para. 4 of 7 <https://www.graphcore.ai/posts/is-moravecs-paradox-still-relevant-for-ai-today> [accessed 16 August 2017].

[19] Hans Moravec, *Mind Children,* (USA: Harvard University Press, 1988), generally.

[20] Steven Pinker, *The Language Instinct,* (USA: William Morrow, 1994), p. 191 <http://www.unc.edu/~moeng/teaching/Pinker%20-%20Language%20Instinct.pdf>.

[21] Elizabeth Kolbert, 'Our Automated Future'*, The New Yorker Magazine*, 19 December 2016, p. 7.

[22] Aiciel blog, 'Does Using Machine Learning Mean More Layers of Complexity in Scientific Study?', undated, paras. 3-6 of 12 <https://aicial.com/blog/does-using-machine-learning-mean-dealing-with-more-layers-of-complexity-in-scientific-study> [accessed 3 March 2017].

[23] Hyacinth Nwana, 'Software Agents: An Overview*', Knowledge Engineering Review*, II, 3, (October 1996), pp. 205-208.

sections are about concepts, abstracts and structures rather than coding lines and detailed composition.[24]

Further narrative is useful to frame this chapter. As noted by Parloff, developments in multi-layer neural networks has transformed the field of AI.[25] These advances are relatively new (Parloff cites 2011 as the relevant turning point) and have been occasioned by, first, progress in fast hardware graphics processor units (GPUs)[26] allowing the training of larger and much deeper networks[27] and, second, by very large labelled datasets available at the time of writing as training test beds. It is their combination, notes Wang and colleagues, that has permitted recent progress in Deep Learning (DL) on deep neural networks (DNN).[28] Their relevance to AWS deployment is that deep-learning processes attempt to mimic human brain activity, specifically the neuron layers of the human neo-cortex, the 'crinkly eighty per cent of the brain', cites Hof, where human thinking takes place.[29] The issue for this chapter is that AWS' various deployment models (the use of software learning to leverage patterns in digital representations of sounds, images and other data) is easily stated but very complex to effect. Given also that DL methods already outdo humans in certain types of image recognition, language recognition and game playing, it is this thesis' assumption that DL represents an inflexion point in AI.[30] As inferred from Ackerman's work on autonomous driving, it is advances in reinforcement learning, in graphical and probabilistic modeling (as well as in file manipulation enabling machine training to be undertaken with materially *smaller* data sets) that support a premise that DL has reached a pivot-point whereby the technology might now be a relevant foundation for weapon autonomy.[31] It is the jump, however, between DL's place within AI and that same facility being a determinant for removing weapon supervision that occupies much of the following technical analysis. In challenging this link, O'Neil, author of *Weapons of Math Destruction*, notes that 'some call this form of capability "artificial narrow intelligence" but here the word "intelligent" is being used much as Facebook uses "Friend"

[24] See: Richard Taylor, 'Software architecture: Foundations, theory and practice', *School of Information and Computer Science*, UCal at Irvine, (October 1999) <https://www.ics.uci.edu/~taylor/Architecture.pdf>.

[25] Roger Parloff, 'Why Deep Learning is suddenly changing your life: Decades-old discoveries are now electrifying the computing industry', *Fortune*, 28 September 2016, paras. 1-5 <http://fortune.com/ai-artificial-intelligence-deep-machine-learning/> [accessed 12 August 2017].

[26] H Mujtaba, 'NVIDIA Pascal shatters 3 GHz GPU frequency record –highest clock speed ever recorded on a graphics chip', *WCCFTech*, (18 December 2016), paras. 4-7 <http://wccftech.com/nvidia-pascal-gpu-frequency-world-record-3-ghz/>) [accessed 18 August 2017].

[27] Ang Li, 'GPU performance Models and Optimization', *Technische Universiteit Eindoven*, (18 October 2016), pp. vii-viii ('Abstract') <https://pure.tue.nl/ws/files/39759895/20161018_Li.pdf>.

[28] Linnan Wang and others, 'SuperNeurons: Dynamic GPU Memory Management for Training Deep Neural Networks', *Proceedings of 23rd ACM Symposium on Parallel Programming*, (2018), pp. 1-3 <https://arxiv.org/pdf/1801.04380.pdf>. For a useful discussion on developments in military AI, see: JASON program, 'Perspectives on research in artificial intelligence and artificial general intelligence relevant to DoD', *US Department of Defense,* JSR-16-Task-003, (January 2017), pp. 1-5.

[29] Robert Hof, 'Deep Learning: with the massive amounts of computational power, Machines can now recognise objects and translate speech in real time. Artificial intelligence is finally getting smart', *TechnologyReview.com*, (June 2016), paras. 3-4 <https://www.technologyreview.com/s/513696/deep-learning/> [accessed 16 August 2017].

[30] Brenden Lake and others, 'Building Machines that Learn and Think Like People', *Behavioural and Brain Sciences*, (2016), pp. 6-8, ('*Cognitive and Neural Inspiration in Artificial Intelligence'*) <https://arxiv.org/pdf/1604.00289.pdf>. See also: Robert Hof, 'Deep Learning', para. 5.

[31] E Ackerman, 'How Drive AI is mastering autonomous driving with deep learning', *IEEE Spectrum*, (11 March 2017), paras. 1-3 <http://spectrum.ieee.org/cars-that-think/transportation/self-driving/how-driveai-is-mastering-autonomous-driving-with-deep-learning> [accessed 24 August 2017].

to imply something safe and better understood than it is. Why? Because the machine has no context for what it's doing and can't do anything else… We might as well call an oil derrick or an aphid intelligent'.[32] It is this dichotomy that forms the basis for this thesis' next three chapters.

Given such divergence, additional context is relevant. The role of DL in AWS deployment might, after all, be 'just an aggressive system of statistics' and, suggests Asaro, merely similar to other uses of arithmetical logic in battlefield engagements.[33] In considering the link between DL and compliant deployment of AWS, the US Department of Defense's recent JASON programme also notes caution[34]: 'DNN are function *approximators* that perform in a very high dimensional space. The manifolds whose shape and extent they are attempting to approximate are almost unknowably intricate, leading to failure modes for which currently there is very little human intuition and even less established engineering practice'.[35] As a basis to this thesis' technical analysis, this is a key observation from which two vectors arise. First, it is necessary to explain why an approximate answer is not good enough for the purposes envisaged for unsupervised weapons.[36] The previous chapter's analysis of deployment challenges identifies predictability and robustness as key weapon requirements if AWS deployment is to be compliant under LOAC.[37] Second, while breakout technologies may promise disruption through a 'golden age of AI'[38], it is necessary to underline that progress in AI is fundamentally different from advances in Artificial General Intelligence (AGI). As discussed in this chapter's following sections, it is AGI and the advent of genuinely general *cognitive* abilities that comprise those capabilities required for human supervision to be removed from lethal engagements.[39] In this vein, the overarching context is provided by Stanford University's Department of Computer Science that that 'there are no present signs of any corresponding revolution in AGI'.[40] Its 2015 report, *Artificial Intelligence and Life in 2030,* concludes that 'no machines with self-sustaining long-term goals and intent have been developed, nor are they likely to be developed in the near future'.[41]

---

[32] Cathy O'Neil, cit. Andrew Smith, 'Franken-algorithms: the deadly consequences of unpredictable code', *The Guardian*, 30 August 2018 <https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger> [accessed 12 September 2018].

[33] Peter Asaro, see: <http://www.peterasaro.org> in conversation with the author, UN CCW GGE Meeting, 16 November 2017.

[34] Federation of American Scientists, Fas.org <https://fas.org/irp/agency/dod/jason/> [accessed 3 May 2017].

[35] JASON program, p. 2.

[36] Ibid. The JASON Study was sponsored by the Assistant Secretary of Defence for Research and Engineering within the DoD.

[37] See: Chapter 5 (*Obstacles*), specifically: 5.1 ('*The Geneva Convention and Laws of Armed Combat'*).

[38] Dom Gohd, 'Amazon's CEO Says We're Living in the Golden Age of AI', *Futurism*, (9 May 2017) <https://futurism.com/amazons-ceo-says-were-living-in-the-golden-age-of-ai/> [accessed 12 October 2017].

[39] See: Chapter 8 (*Software*), specifically: ('*Value Setting* and *Anchoring'*). See also: Cassio Pennachin and Ben Goertzel, *Contemporary Approaches to Artificial General Intelligence,* (USA: AGIRI, Springer Publishing, XVI, 2007), p. 509.

[40] Source: Stanford University <https://ai100.stanford.edu> and, for an executive summary of the Stanford University report on the future of AI, see: Stanford University <https://ai100.stanford.edu/2016-report/executive-summary> [accessed 3 March 2017].

[41] Stanford University, 'Artificial Intelligence and Life in 2030'*, 2015 study panel*, (June 2016) <https://ai100.stanford.edu> [accessed 4 March 2017]. For a useful discussion on long-term planning considerations, see: D Stojkovic and B Dahn, 'Methodology for long-term Defence Policy'*, Norwegian Defence Research Establishment*, (28 February 2007) <http://www.ffi.no/no/Rapporter/07-00600.pdf>.

6.1 **Software 'versus' intelligence**

While this thesis refers throughout to the term AI, this distinction between AI and AGI in AWS deployment is therefore key. The International Panel on the Regulation of Autonomous Weapons (IPRAW) even suggests that the umbrella term of AI as applied to AWS should 'be used with prudence and parsimony'.[42] The terms AI and ML imply a deeper meaning behind what are, after all, statistical methods, leading, notes Angelov and Sperduti, to the false impression of intention and purpose.[43] For this reason, Burgess therefore prefers the umbrella term 'computational methods'.[44] As inferred from Kruchtren, even the most sophisticated range expansion of battlefield tasks[45] (as undertaken by AI routines) does not approach even the most basic representation of artificial *general* intelligence.[46] Indeed, Yampolskiy and Fox note that popular demarcation between AI, AGI and 'enhanced' software is inappropriately imprecise.[47] For the purposes of this thesis, AI refers just to computational methods that have more generally applicable problem-solving capacities than conventional software. This falls considerably short of Muehlhauser's definition of AGI which, adopted by this thesis, is 'the ability to achieve complex goals in complex environments with limited computational resources'.[48] The *assumed* trajectory, after all, for weapon processes is one of seamless progression (mirroring, for instance, the same timeline that Deloitte conjectures in business applications) from rules-specific conventional software to machine-learning AI to independent and then sentient AGI.[49] It is this promise of software (that it can learn on its own to do new tasks) which then provides the catalyst for weapon models where human can be moved out-of-the-loop.[50] Its theoretical attractions are that such systems may learn and plan and, importantly, do so in a powerful and robustly cross-domain manner.

In line with Braga and Logan, a purpose of this section is to evidence that AI capability does *not* presage machine cognition, sentience or reasoning.[51] As noted by Yampolskiy, while AI is orientated

---

[42] International Panel on the Regulation of Autonomous Weapons (IPRAW), 'Executive Summary Number 2', *Computational Systems in the Context of Autonomous Weapon Systems*, November 2017, generally <https://www.swp-berlin.org/en/projects/international-panel-on-the-regulation-of-autonomous-weapons-ipraw/> [accessed 17 August 2017].

[43] Plamen Angelov and Alessandro Sperduti, 'Challenges in Machine Learning', *European Symposium on Artificial Neural Networks ESANN*, (2016), pp. 490-491 <https://pdfs.semanticscholar.org/4e9a/7debe3df64e0bd8e861f7680dc6aa42ab954.pdf>.

[44] Matt Burgess, 'Killer Autonomous Weapons are coming, but they're not here yet', *Wired magazine*, Technical opinion, 12 August 2017, paras. 4 and 6 of 10 <http://www.wired.co.uk/article/killer-robots-elon-musk-autonomous-weapon-systems-uk> [accessed 2 November 2017].

[45] Philippe Kruchtren and others, 'Technical Debt: From Metaphor to Theory and Practice', *IEEE Software*, University of British Columbia, (2012) <https://www.computer.org/csdl/mags/so/2012/06/mso2012060018.pdf>.

[46] Nick Bostrom, *Superintelligence; Paths, dangers, strategies*, p. 151.

[47] For a general discussion, see: Ibid., pp. 16-23. Also: Roman Yampolskiy and Joshua Fox, 'AI versus AGI', (Singularity Hypotheses, Springer, Berlin, 2012), pp. 130-131 <https://intelligence.org/files/AGI-HMM.pdf>.

[48] Luke Muehlhauser, 'What is AGI?', *MIRI Machine Intelligence Research Institute*, (11 August 2013) <https://intelligence.org/2013/08/11/what-is-agi/> [accessed 2 February 2017].

[49] Deloitte, 'Artificial Intelligence Innovation Report', *Deloitte Innovation*, (2016), p. 2 <https://www2.deloitte.com/content/dam/Deloitte/at/Documents/human-capital/artificial-intelligence-innovation-report.pdf>.

[50] See, generally: Chapter 4 (*Deployment*).

[51] Adriana Braga and Robert Logan, 'The Emperor of Strong AI Has No Clothes: Limits to Artificial Intelligence', *Information*, 8, 156, (2017), pp. 1-3 and generally. The intelligent decision-making of such machines, notes Dr Hongbo Du, fundamentally consists of a function sequence comprised largely of pre-processing of input data, extraction of useful

towards specific tasks, this is considerably removed from the general cognitive abilities that are assumed for AGI.[52] Some context is useful on this point.[53] The US Military's JASON programme concludes that 'on account of this ambitious goal, AGI has a high visibility, disproportionate to its size and present level of success, among futurists, science fiction writers, and the public'.[54] In this vein, Proctor coins the term 'agnotology' to describe the study of the 'cultural production of ignorance and its effect on both individual and collective decision-making processes.'[55] This is analogous to Sabin's 'Revolution in Expectation'.[56] Public embrace of AI and AGI is also promoted by what Tamburrini terms a 'temporal framing mechanism' where there is common misunderstanding on timescales around necessary technological development.[57] This arises in part from a failure to convey clearly the distinction between long-term and visionary research goals versus expected short-term outcomes from such research. In considering AGI, the implausibility of this is evidenced by the work of Grace and colleagues and their clear demonstration of the dispersion in expectations about AGI's arrival; in collating evidence on timelines from experts in this field, Grace finds, for instance, that Asian communities forecast high-level machine intelligence (HLMI) a telling forty-four years earlier than their US counterparts.[58]

Four characteristics should be assumed for independent weapon platforms that are to operate without human oversight.[59] The autonomous weapon should be independent. It should be capable of intrinsic and robust analysis based on comprehensive data and able to operate without supervision based on this data. The weapon must then have a capacity to deduce and decide courses of action driven by this data. The machine must also be 'aware' of its surroundings and, based on data analysis, will be able to carry out tasks that prioritise mission performance rather

features from that input data stream and then a decision-making or classification routine concerning the input stimuli: 'Such sequential structure of these systems means that any small error in its early stages will propagate into larger errors as the stages unfold. The more complex is the processing function, the more probable of such error occurrence'. Here, Du notes that self-correction of such errors is challenging given that each discrete stage is handling a different type of 'sub-problem'. Initiatives to combine neural networks in order to merge such stages into 'an end-to-end solution' confound attempts then to explain how and where these errors are corrected. It is, notes Du, 'premature to even conclude that such an architecture will overcome this problem at all'. Source: Dr Hongbo Du, School of Computer Science, Buckingham University, in conversation with the author, January 2019.

[52] Roman Yampolskiy and Joshua Fox, 'Artificial General Intelligence and the human mental model', *Machine International Research Institute*, (2012), p. 7 <https://intelligence.org/files/AGI-HMM.pdf>.

[53] See: Dave Michels, 'AI Heading Back to the Trough: Expectations Over Artificial Intelligence are Becoming Too Inflated', *Networld World*, (11 July 2017) <https://www.networkworld.com/article/3206313/internet-of-things/ai-heading-back-to-the-trough.html> [accessed 12 March 2018].

[54] JASON program, p. 1. See also: Richard Potember, 'Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD', *US Department of Defense Publications*, JRR-16-Task-003, p. 2 and pp. 28-32 ('*Why the "Ilities" may be intrinsically hard for Machine Learning'*) <https://fas.org/irp/agency/dod/jason/ai-dod.pdf>. The 'ilities' here refer to 'reliability, maintainability, accountability, verifiability, evolvability and attackability'. Potember in this case concludes that 'ML is weak on the "ilities"'.

[55] Robert Proctor, 'A missing term to describe the cultural production of ignorance', in Proctor and Schiebinger, eds., *Agnotology: The making and unmaking of Ignorance*, (USA: Stanford University Press, 2008), p. 1.

[56] Professor Philip Sabin, Professor of Strategic Studies at KCL, in conversation with the author, 29 June 2017.

[57] Guglielmo Tamburrini, 'On banning autonomous weapon systems: From deontological to wide consequentialist reasons', cit. Bhuta and others, p. 134.

[58] Katja Grace and others, 'When Will AI Exceed Human Performance? Evidence from AI Experts', *arXiv preprint arXiv*: 1705:08807, (2017), pp. 3-4 and generally <https://arxiv.org/pdf/1705.08807.pdf>.

[59] The analysis of AWS system fundamentals is then relevant to this thesis' review of *firmware* (chapter 7), *software* (Chapter 8) and *hardware* (Chapter 9).

than simpler mission execution.[60] Macdonald suggests that it will be the combination of these two characteristics that provides AWS with elements of sentience and self-awareness based on system sensing, system perceiving and self-learning.[61] A third tenet that requires restatement is that the AWS must still obey the same physical laws that govern all the physical polities: It cannot change shape and size arbitrarily, it must use the effectors to move itself around based on reason-based rules, it requires an energy source to think, sense and move, and when moving, it will take time to speed up and slow down.[62] A final characteristic relates to its *internal models of the world.* In order for such unsupervised weapons to be battlefield-ready, units must rely upon and be able to manipulate and search through broad possible solutions that are based on these statistical and logic models.[63] AWS' performance must, after all, be based upon its ability to solve problems by referring first to internal current models and then to the dynamic updating of these models. A constraint to emerge is thus the balancing between model and sensed data, a set of processes that, in large measure, will determine AWS feasibility once deployed.[64]

The issue for this review is therefore the *extent* of the platform's tasking and the correlation between the weapon's empirical competences and that tasking. In order to understand whether a weapon might act outside the boundaries of that initial tasking, it is necessary first to understand the basis of such autonomous capabilities.[65] In this vein, Chalmers defines machine-autonomy as the faculty 'to discriminate, characterise and react to environmental stimuli; the integration of information by a cognitive system; the reportability of mental states; the focus of attention and deliberate control of behaviour'.[66] Extending this to the battlefield, a deduction (also inferred from Kaspersen) must be that compliant removal of weapon supervision requires more than rules-based routines.[67] While Chalmers' characterization might depict a high-level model of a weapons platform, the parallel still ignores important deficiencies facing weapons-directing architecture, principally the absence of 'qualia' (the *feel* of precepts), the concept of learned experience and, literally, the requirement in independent weapons of 'phenomenal consciousness'.[68] It is these tenets that define the scope of subsequent chapters. They also inform this thesis' conjecture that

[60] For the avoidance of doubt, the terms 'weapon', 'machine' and 'platform' often refer throughout this analysis to a deployed autonomous weapon system (AWS).

[61] Fiona MacDonald, 'A robot has just passed a classic self-awareness test for the first time', *Science Alert*, (17 July 2015), paras. 2-4 of 13 <https://www.sciencealert.com/a-robot-has-just-passed-a-classic-self-awareness-test-for-the-first-time> [accessed 3 February 2016].

[62] Vincent Boulanin and Maaike Verbruggen, 'Mapping the development of autonomy in weapon systems', *SIPRI*, (November 2017), pp. 7-11 ('*Unravelling the machinery*') <https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_0.pdf>.

[63] Bernard Marr, 'What is the Difference between Artificial Intelligence and Machine Learning?', *Forbes Magazine,* Technology, 6 December 2016 <https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#1a4e99432742> [accessed 14 October 2017].

[64] See Chapter 7 (*Software*), specifically: 8.5 ('*Anchoring and gaol-setting issues*').

[65] It is noteworthy that States comprising the Group of Government Experts (GGE) at the UN's CCW have not yet agreed on a definition of what constitutes an AWS after meeting annually since 2014.

[66] See, generally: Nick Bostrom, *Superintelligence,* pp. 200-219.

[67] Anja Kaspersen, 'We're on the brink of an artificial intelligence arms race. But we can curb it', *World Economic Forum*, (15 June 2016), paras. 7-9 of 41 <https://www.weforum.org/agenda/2016/06/should-we-embrace-the-rise-of-killer-robots/> [accessed 18 August 2017].

[68] Pentti Haikonnen, *The cognitive approach to conscious machines,* (UK: Imprint academic, 2003), p. 145.

the prerequisite of workable AGI is a key constraint to removing human supervision.[69] Here, Lu and Gamez also challenge the elided claim that machine intelligence equates to machine consciousness.[70] Rossi, member of IBM's AI Ethics Committee, dismisses here the notion that machines will 'one day wake up and change their minds about what they will do'.[71] This argument has been used by the British Foreign Office to underpin its opposition in the UN's CCW to statutory controls over AWS deployment.[72] In this case, *Empty hangar syndrome* suggests that certain scenarios are too far-fetched to warrant current consideration or statutory documentation and unrealistic, in this case, for a commander to wander into the weapons hanger to find that his AWS has decided under its own volition to depart unexpectedly on an unsupervised mission.[73] The inference relevant is that human engagement must *always* remain fundamental in weapon use.

6.2 Architectural approaches to AWS deployment

A pervasive challenge for compliant deployment of autonomous weapons, identified by Melli, is the quantity of components, routines and techniques that must reliably be in place to realise this state.[74] This evidences a broad architectural conundrum in AWS deployment.[75] Failures anywhere in this sequence will likely impact weapon performance in a material (perhaps catastrophic) manner and, central to this thesis, prejudice that weapon's legal compliance.[76] This section's purpose is to demonstrate that none of the architectures envisaged to remove human oversight is straightforward.[77] A common starting point is the human mind as it is this that must be substituted by machine processes in AWS. A recent estimate for the human brain is that the cerebral cortex contains some thirty billion neurons in the part of the brain associated with consciousness and intelligence. Those neurons in turn contain some one thousand trillion synapses, the connections

[69] Economist Magazine Special Report, 'Artificial intelligence: From not working to neural networking', *Economist*, 25 June-1 July 2016, p. 13.

[70] See, for instance: Clara Lu, 'Why We are Still Light Years Away from Full Artificial Intelligence', *TechCrunch*, (2016), generally <https://techcrunch.com/2016/12/14/why-we-are-still-light-years-away-from-full-artificial-intelligence/> [accessed 5 March 2018]. Although written in 2008, see also: David Gamez, 'Progress in Machine Consciousness', *Consciousness and Cognition*, 17.2, (2008), pp. 12-16 ('*Criticism of Machine Consciousness: Hard problems of consciousness*') <http://davidgamez.eu/papers/Gamez07_ProgressMachineConsciousness.pdf>.

[71] Source: Francesca Rossi, IBM's AI Ethics committee, cit. Economist, '*Artificial intelligence*', pp. 13-15.

[72] Until publication in August 2018 of 'Human Machine Touchpoints: The United Kingdom's perspective on human control over weapon development and targeting cycles', the Foreign Office had adopted a broadly negative negotiating position in discussions on banning weapons autonomy at the UN's Convention for Conventional Weapons, 2014-2017. See: United Nations Office at Geneva, (8 August 2018) <https://www.unog.ch/80256EDD006B8954/(httpAssets)/050CF806D90934F5C12582E5002EB800/$file/2018_GGE+LAWS_August_WP_UK.pdf>.

[73] The concept is discussed in Chapter 11 (*Conclusion*).

[74] A useful primer is provided in Roberto Melli, 'Artificial Intelligence in Component Design', *Exergy* (sic); energy system analysis and optimization, III <http://www.eolss.net/sample-chapters/c08/E3-19-04-04.pdf>.

[75] Vera Lucia Menezes de Oliveira and others, 'The Complex Nature of Autonomy', *Delta Online*, (2008), pp. 442-445 and pp. 445-449 <http://www.scielo.br/pdf/delta/v24nspe/04.pdf>.

[76] See: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt*').

[77] US Air Force, Office of the Chief Scientist, 'Autonomous horizons; System autonomy in the Air Force – a path to the future – human-autonomy teaming', *AF/ST TR*, 15-01 (June 2015), p. 4. See also: Isabelle Guyon and others, *Active Learning Challenge: Challenges in Machine Learning, Volume 6*, (USA: Microtome Publishing, 2012), generally <http://www.mtome.com/Publications/CiML/CiML-v6-book.pdf>.

between neurons.[78] While it might appear feasible for a machine eventually to 'copy' that mind, Epsten disagrees and notes that the infant brain neither arrives with nor subsequently develops 'lexicons, representations, algorithms, programmes, processors, subroutines, encoders or buffers'.[79] Unlike idealized connections that are typical of an artificial neural network, human synapses are particularly variable in nature based upon different and quite undifferentiated neurotransmitters with different cycle times.[80] Information data in the human brain, more than a megabyte for each connection that is being generated by each synapse, is being processed in real time within each synapse cycle, itself more than one thousand bursts per second.[81] How can this magnitude best be understood in the context of machines copying such a model? If each such synapse were handled by the equivalent of only a single line of code, the programme to simulate a human cerebral cortex would be some twenty-five million times larger than reputedly the largest software product written to date, Microsoft Windows, which is estimated to be some fifty million lines of code.[82] Given Kassan's conclusion that the probability of successfully completing such emulation is 'effectively zero', this then becomes the appropriate marker for subsequent discussion on creating weapons-directing intelligence.[83]

As posited by Bringsford and Arkoudas, there are different methods envisaged for AWS machine intelligence which justify architectural analysis.[84] Engineers can look to replicate entirely the brain, copy certain characteristics of the brain or apply theoretical neural processing in order to ape brain processes. No such methodology is trivial and, argues Bindi, none is currently remotely available.[85] Less conventional approaches might instead tackle machine intelligence using massively parallel computers in an effort to force biological cognition through superfast iterations that run, perhaps, genetic algorithms.[86] Reddy, however, notes the considerable gulf that exists

[78] Gerald Edelman and Giulio Tononi, 'A Universe of Consciousness: How Matter Becomes Imagination' (cit. Kassan, 'AI gone awry: futile quest for artificial intelligence', *The Skeptics Society and Skeptic Magazine*, undated, pp. 3-9 <http://www.skeptic.com> [accessed 14 September 2016).

[79] Robert Epstein, 'The Empty Brain', *Aeon*, (18 May 2016), generally <https://aeon.co/essays/your-brain-does-not-process-information-and-it-is-not-a-computer> [accessed 6 July 2018]. Epstein's article provides a useful primer on challenges to machine emulation of brain function.

[80] Doo Seok Jeong and others, *Towards Artificial Neurons and Synapses: A materials point of view,* (RSC Publishing, DOI 10.1039/c2ra22507g, undated), Abstract and generally.

[81] Helen Philips, 'Introduction: The Human Brain', *New Scientist*, (4 September 2006), generally <https://www.newscientist.com/article/dn9969-introduction-the-human-brain/> [accessed 7 July 2018].

[82] Charles Choi, 'Too Hard for Science: Simulating the Human Brain', *Scientific American*, (9 May 2011), generally <https://blogs.scientificamerican.com/guest-blog/too-hard-for-science-simulating-the-human-brain/> [accessed 9 July 2018]. See also: Code.org, para. 1 of 6 <https://code.org/loc> (accessed 2 August 2017]. Also; Cade Metz, 'Google is 2 Billion Lines of Code – and It's All in One Place', *Wired*, 16 June 2015, generally, https://www.wired.com/2015/09/google-2-billion-lines-codeand-one-place/> [accessed 12 July 2018].

[83] Kassan, p. 2.

[84] Selmer Bringsford and Konstantine Arkoudas, 'The Philosophical Foundations of Artificial Intelligence', *Department of Cognitive Science, RRI*, Troy NY, (October 2007), pp. 2-6 <http://kryten.mm.rpi.edu/sb_ka_fai_aihand.pdf>.

[85] Tas Bindi, 'True Artificial Intelligence cannot be developed until the 'brain code' has been cracked', *zdnet.com*, (9 August 2017), paras. 6-8 of 35 <http://www.zdnet.com/article/true-ai-cannot-be-developed-until-the-someone-cracks-the-brain-code-starmind/> [accessed 24 August 2017].

[86] Source: Mathworks <https://www.mathworks.com/discovery/genetic-algorithm.html> [accessed 2 February 2017]. A genetic algorithm is a way to solve optimization problems based on a natural selection process mimicking biological evolution whereby an algorithm repeatedly modifies a population of individual solutions.

between desktop and in-field applications of these approaches.[87] Similarly, Bindi focuses on the absent qualitative relationships that must underpin such models in order to demonstrate the far-fetchedness of, for example, expecting computers artificially to evolve machine intelligence through a process of raw iteration.[88] Schulman and Bostrom concur and note that the computational resources needed to copy the relevant evolutionary processes that produced our own human-level intelligence 'are severely out of reach'.[89] A simple honeybee brain has some ten-to-the-power-of-six neurons.[90] Regardless of the level of detail included in that simulation, the current computational cost of simulating just a *single* neuron suggests that any approach based on superfast iteration is unfeasible even as a premise upon which to remove supervision from machine actions.[91] It is therefore necessary to look elsewhere for AWS' feasible architecture.[92]

How else then might appropriate weapon AI be constructed? Most methodologies still start with that human brain as a template for machine intelligence. This might comprise 'whole brain emulation' (WBE) through meticulous scanning and close modeling of the computational structure of a biological brain in order then to create a serviceable facsimile in the independent weapon.[93] As above, the route is unlikely to be practicable.[94] Pennachin and Goertzel note that the process will require considerably advanced enabling technologies, not yet available even in the laboratory, including high-throughput scanning detection, automated image translation and deep simulation processes.[95] Nor does the existence today of very simple prototypes constitute proof of eventual success in the endeavour. Sharkey highlights a clear difference between creating a single laboratory-based machine with autonomous traits versus the deployment of multiple AWS based on that same intricate technology.[96] Instead, the process will be dogged by what Sandberg and Bostrom term 'chaotic dynamics', already a feature in test systems with just a handful of neurons.[97] Mitchum also evidences that simple mapping of neuron connections in the brain does little to advance weapons-ready machine intelligence.[98] The challenge, after all, is to achieve *structural*

[87] Raj Reddy, 'Foundations and Grand Challenges of Artificial Intelligence', *AI Magazine*, 9, 4, (Winter 1988), p. 8 and pp. 17-18.

[88] Bindi, paras. 19-20 of 35.

[89] Carl Shulman and Nick Bostrom, 'How hard is Artificial Intelligence? Evolutionary arguments and selection effects', *Journal of Consciousness Studies*, 19, 7-8, (2012), 17-18.

[90] Nick Bostrom, *Superintelligence*, p. 24.

[91] Kurzweil Accelerating Intelligence, 'IBM simulates 530 billion neurons, 100 trillion synapses on supercomputer' (19 November 2012) <http://www.kurzweilai.net/ibm-simulates-530-billon-neurons-100-trillion-synapses-on-worlds-fastest-supercomputer> [accessed 2 March 2017].

[92] ZH Zhou, 'Machine Learning Challenges and Impact: An interview with Thomas Dietterich', *National Science Review*, 5, 1, (January 2018). Dietterich is Professor of Computer Science at Oregon State and former President of the AAAI.

[93] Nick Bostrom, *Superintelligence*, p. 29.

[94] Anders Sandberg and Nick Bostrom, 'Feasibility of Whole Brain Emulation', *Future of Humanity Institute*, Theory and Philosophy of Artificial Intelligence, SAPERE, Berlin, Springer, (2013), p. 3

[95] Cassio Pennachin and Ben Goertzel, p. 17 and p. 19.

[96] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[97] Sandberg and Bostrom, 'Feasibility', p. 19.

[98] Rob Mitchum, 'Can the Connections between the 100 Billion Neurons in the Brain be Mapped?', *Forefront*, University of Chicago Medicine, (1 June 2018), https://www.uchicagomedicine.org/neurosciences-articles/can-100-billion-neurons-be-mapped> [accessed 15 July 2018]. See also: The Human Connectome Project, <http://www.humanconnectomeproject.org> [accessed 5 March 2017].

validity as opposed to just *replicative* validity. Weapon designers using this methodology would need, for example, to understand which synapses may be excitatory and which are inhibitory. They would then need to model exactly the strength of these connections as well as understand the dynamical properties of underlying brain subsystems.[99] As inferred from Frankel, it should not be assumed that WBE models can scale to battlefield applications.[100] Cattell and Parker, moreover, note that WBE approaches more resemble 'extended' software rather than any expression of machine intelligence.[101] This is because code-based artificial neurons act more as generalized Boolean logic gates (characterised by an underlying rules-driven basis)[102] than actual neurons.[103] More fundamental to these models, it is insufficient to ascribe single weightings to each such artificial neuron in order to manage its threshold or to each synapse in order to reflect a new signal strength: As noted by Seaok Jeong, these linear relationships just do not exist.[104] While a human synapse has an estimated minimum of ten thousand connections (just to be able to undertake this capability), Kassan highlights that this is 'at least six hundred billion times more complicated than any artificial neural network yet devised' in the case of the whole human cortex.[105] Finally to this point, Sandberg notes the extensive list of technologies still outstanding if WBE is to be effected including general data 'interpolation', geometric 'adjustment', parameter 'estimation' as well as resolution 'controls'.[106] Again, therefore, it is necessary to look elsewhere for a feasible architecture.

There is no obvious path to search as an inescapable challenge to all learning models is that of scaling. Wolpert's *No-Free-Lunch* theorem, still observed after two decades[107], states that *general-purpose* learning algorithms cannot exist 'in the sense that for every learning model there is a data distribution on which it will fare poorly on both training and test'.[108] Marsh, in *Can Man Ever Build a Mind,* points to two systemic logjams to this approach: The 'Binding Problem' questions 'how does all of this disparate neuronal activity, spread out in both time and space, produces coherent experience'.[109] 'Von Neuman's Bottleneck' then arises from the extraordinary electrical

[99] Nick Bostrom, *Superintelligence*, p. 35.

[100] Stuart Frankel, 'Data Scientists don't scale'*, Harvard Business Review*, (22 May 2015), paras. 2 and 5-6 of 15 <https://hbr.org/2015/05/data-scientists-dont-scale> [accessed 12 December 2016].

[101] Rick Cattell and Alice Parker, 'Challenges for Brain Emulation: Why is it so difficult?'*, Natural Intelligence*, INNS, 1, 3, (2012), pp. 18-19 <https://pdfs.semanticscholar.org/a0d4/71388ae0db850419c7e854e603b8198f8930.pdf>.

[102] Brighthub Engineering blog, 'Logic gates – the gateways to intelligence in machines'*,* (8 July 2008) <http://www.brighthubengineering.com/diy-electronics-devices/3349-logic-gates-the-gateways-to-intelligence-in-machines/> [accessed 25 October 2017].

[103] For a useful overview on gates and gate functions see: The University of Surrey, Department of Electrical and Electronic Engineering <http://www.ee.surrey.ac.uk/Projects/CAL/digital-logic/gatesfunc/> [accessed 2 January 2017].

[104] Doo Seok Jeong and others, *Towards Artificial Neurons and Synapses: A materials point of view,* DOI 10.1039/c2ra22507g, (USA: RSC Publishing, undated), Chapter 5 and generally.

[105] Kassan, p. 2.

[106] Sandberg and Bostrom, 'Feasibility', p. 18.

[107] Leon Fedden, 'The No Free Lunch Theorem (or why you can't have your cake and eat it)', *Medium.com blog*, (6 October 2017) <https://medium.com/@LeonFedden/the-no-free-lunch-theorem-62ae2c3ed10c>[accessed 8 January 2018].

[108] David Wolpert, 'The lack of distinction between learning algorithms'*, Neural Computations*, 8, 7, (1996), 1341-1390 <http://www.mitpressjournals.org/doi/abs/10.1162/neco.1996.8.7.1341> [accessed 23 November 2017].

[109] Henry Marsh, 'Can Man Ever Build a Mind?', *Financial Times*, (10 January 2019), para 9 of 24 <https://www.ft.com/content/2e75c04a-0f43-11e9-acdc-4d9976f1533b> [accessed 2 February 2019].

requirements of the approach whereby 'an exascale computer, capable of a quintillion calculations per second, scaled up to the size of a human brain, would consume hundreds of megawatts [of power]'.[110] A ramification for AWS development is also that every one of its learning models must contain restrictions on the class of functions that it can learn and it cannot therefore be assumed that basic algorithms can be scaled in AWS deployment without material performance cost. Bengio and LeCun demonstrate in this case that current scaling approaches to AI remain very limited.[111] In particular, 'kernel methods' that avoid assumptions about frequency distribution must all be founded, observes Copeland, on inappropriately 'shallow' architecture.[112] While learning models are considered below[113], a relevant deduction is that the basis of trainable coefficients supporting layers of template matchers is fundamentally inefficient and incapable of scale.[114] To this point, Bengio and Lecun highlight their 'depth-breadth trade-off' and 'curse of dimensionality' and, as demonstrated by Simonite, architectures become exponentially more far-fetched as laboratory-based prototype is transitioned to scaled-up emulation.[115] Engineers, moreover, are unlikely to understand in advance just how many flaws persist in late beta versions of their models.[116] While Moore's Law might be invoked, doubling speed and capacity does not technically solve problems posed by system complexity. A second Law, attributed this time to Wirth, is perhaps more relevant.[117] Wirth instead suggests that 'software gets slower faster than hardware gets faster'.[118] While Moore's Law suggests that the personal computer should be some hundred thousand times more powerful than it was twenty-five years ago, the computer's word processor certainly is not. The inference is that Moore's law does not apply to weapon *software*.[119]

[110] Henry Marsh, 'Can Man Ever Build a Mind?', para 20 of 24.

[111] Yoshua Bengio and Yann LeCun, 'Scaling Learning Algorithms towards AI', *Large Scale Kernal Machines, MIT Press*, (2007), pp. 16-18 <http://yann.lecun.com/exdb/publis/pdf/bengio-lecun-07.pdf>.

[112] Michael Copeland, 'What's the difference between artificial intelligence, machine learning and deep learning?', *nvidia blog,* (29 July 2016) <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/> [accessed 2 October 2017]. Shallow networks have fewer hidden network layers (and can function with a single layer) but require unwieldy multiplicity of data points.

[113] See: Chapter 7 (Firmware), specifically: 7.2 ('*Firmware ramifications of learning methodologies'*).

[114] Yoshua Bengio and Yann LeCun, 'Scaling Learning Algorithms towards AI', *MIT Press*, (2007), p. 1 ('*Abstract'*), pp. 12-14 ('*Depth-Breadth tradeoff'*), p. 16 ff ('*Fundamental limitations of local learning'*) and p. 21 ff ('*Curse of dimensionality'*) <http://yann.lecun.com/exdb/publis/pdf/bengio-lecun-07.pdf>.

[115] Tom Simonite, 'Thinking in Silicon'*, MIT Technology Review*, (December 2013) <https://www.technologyreview.com/s/522476/thinking-in-silicon/> [accessed 15 March 2017].

[116] Source: Wikipedia, 'List of unsolved problems in neuro-science' <https//:en.wikipedia.org/wiki/List_of_unsolved_problems_in_neuroscience> [accessed 28 December 2016].

[117] *Source: Public Archive of the IEEE Computer Society History Committee* <http://history.computer.org/pioneers/wirth.html> [accessed 21 August 2017]

[118] Kassan, p. 2.

[119] J Vincent, 'These are three of the biggest problems facing today's artificial intelligence' *The Verge*, (10 October 2016), paras. 1-4 of 5 <https://www.theverge.com/2016/10/10/13224930/ai-deep-learning-limitations-drawbacks> [accessed 7 July 2017].

6.3 **The AWS' Delivery Cohort**

AI frameworks based on probabilistic modelling have emerged as the lead theoretical approach for designing machines that learn[120] and, by extension, a principal pipework upon which to build independent weapons. A faultline is noted by Kalmonovitz whereby AWS will be 'non-deterministic and non-scripted to various degrees, in all cases their range of action should be both bounded and probabilistically estimated. Deploying them under second-levels of uncertainty (uncertainty about the probabilistic range of action) would create in inestimable levels of risk on a civilian population and would consequently be illegal'.[121] Put otherwise, all parts of all weapon chains (unsupervised and supervised) must (legally) have reasonable epistemic confidence about the precise range of weapon action. From this notion arises the concept of 'role responsibility'[122], the subject of this section, and the implication that it is a broad cohort who are involved in the deployment of new weapons technologies. Those fielding AWS[123] might assume that their superiors and legal experts have appropriately overseen those involved in the design, programming and testing of the weapon. It is a class effort. The notion of an AWS design and implementation 'team' is referred throughout this thesis as the *Delivery Cohort*.

This important piece of shorthand refers to the extended group of experts and other parties that will be needed to implement independent weaponry.[124] It masks certain procurement shortcomings including the diffusion of responsibility between a raft of participating agencies to the point that any meaningful attribution has been obscured. As evidenced by section sub-headings[125] to Chapters Six to Nine of this thesis, the list of required competencies and the control mechanisms necessary to deliver these capabilities is very broad.[126] The role of this Delivery Cohort in the deployment of compliant AWS is correspondingly wide. It must manage both those high-level software and hardware challenges identified in these chapters in order to field a weapon that is appropriate to local commander *and* wider community. The Cohort's responsibility is, after all, highly complex.[127] It must ensure its unsupervised weapon can synthesise all key tenets that characterise human supervision including, inter alia, reliable scaling of weighting factors and the

---

[120] For a useful primer, see: Z Ghahramani, 'Probabilistic Machine-learning and Artificial Intelligence*',* *Nature*, 521 University of Cambridge, 28 May 2015), pp. 452-459 <https://www.repository.cam.ac.uk/bitstream/handle/1810/248538/Ghahramani%202015%20Nature.pdf>.

[121] Kalmanovitz, cit. Bhuta and others, p. 156.

[122] H Hart and J Gardner, *Punishment and Responsibility: Essays in the Philosophy of Law',* 2nd Edition, (Oxford: Oxford University Press, 2008), p. 212.

[123] For the purposes of this thesis, the notion of the *Delivery Cohort* includes, inter alia, the following tasks: neurophysiologists to coordinate AWS networks, psychologists to coordinate learning and cognition, biologists for adaption strategies, engineers for control routines, logisticians, roboticists, electrical specialists, behaviorists, politicians, NGOs, sociologists, lawyers, company directors, weaponists, military tacticians, manufacturers, professionals involved in miniaturization, simulation, configuration, coding, power supply and modularity, specialists in sensors, in distributed and decentralized routines, ethicists, specialists in tooling and calibration.

[124] For an introduction to these required tasks, see: D Floreano, 'Design, Control and Application of Autonomous Mobile Robots', *Swiss Federal Institute of Technology, Lausanne*, undated <https://infoscience.epfl.ch/record/63893/files/aias> [accessed 7 July 2017].

[125] See: *Contents*, pp. 2-4 of this thesis.

[126] See: Note 123 above. See also: Floreano, pp. 159-186.

[127] For a useful primer of such responsibilities, see: Ezio Nucci and Filippo Santoni (eds), 'Drones and Responsibility: Legal, Philosophical and Socio-technical Perspectives on Remote Controlled Weapons', *Routledge*, (2016), Part III (*'Design and socio-technical perspectives') and Part IV ('Autonomous Killer Drones').*

calculation of prior probabilities to each possible battlefield outcome and at any given time.[128] How might this work in practice? Those more likely 'favoured' worlds (the multiple set of actions that most closely align weapon outcome to the Delivery Cohort's intended purposes) will be given higher probabilities by the Cohort.[129] Later chapters, however, demonstrate that this process is not obvious[130] and demands compromise, negotiation and management by the Cohort.[131] It will, moreover, be similarly non-trivial to achieve technical consensus between the Cohort's competing parties as well as for them to define intended outcomes in appropriate detail. The role of the Delivery Cohort at both design and deployment stages is therefore fundamental.

It is also the *dynamic* nature of the Cohort's task that creates challenge. Datasets, the central element to machine learning, are dynamic and, Shah notes, prone to rapid but ambiguous obsolescence.[132] The complex systems of an AWS will be nested, that is they 'unfold from and are enfolded in other another'[133] and their network of dynamic processes must be both integrated into a dynamic whole and rely on inbuilt capacity to adapt to changes in those datasets. Given that battlefield events will be chaotic, such changes are unlikely to be linear and effects rarely proportionate to outcomes.[134] In this case, fractional changes in a weapon's set 'initial conditions' may drastically alter the long-term behaviour of that weapon.[135] The dynamic nature of the Cohort's task requires it layer the AWS with feedback mechanisms either to identity performance-variation or confirm adherence to its intended goal state. It is such feedback routines, notes Sculley, that complicate its operation.[136] They also lead to 'system patching' in order for the Cohort to deal with new and evolving modes of AWS operation.[137] As noted by Smith and colleagues, patching is empirically complicated.[138] At best, it should involve unambiguous identification of coding problems, creation of an unambiguous fix to that issue, delivery of the patch and then application of that patch once it has been dispatched to *all* host weapons. Patch management then becomes a key Cohort activity to coordinate weapon stability, version control, feedback facilitation and the

---

[128] Dinesh Nirmal, 'How to Decide: Machine Learning and the Science of Choosing', *Medium.com blog*, (20 March 2017) <https://medium.com/inside-machine-learning/how-to-decide-machine-learning-and-the-science-of-choosing-7a0d70059079> [accessed 4 September 2018]: 'If you choose not to decide, you've still made a choice'.

[129] Ibid. Inferred from Nirmal's analysis.

[130] MP Huerta, 'Assessing Difficulties of Conditional Probability Problems', *University of Valencia*, EDU/2008-03140/Edu Project, (2012), p. 4 <http://www.cerme7.univ.rzeszow.pl/WG/5/CERME_Huerta-Cerdan-Lonjedo-Edo.pdf>.

[131] See: Chapter 7 (Firmware), specifically: 7.1 ('*Sources of technical debt*).

[132] Tarang Shah, 'About Train, Validation and Test Sets in Machine Learning', *Towards Data Science*, (6 December 2017) <https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7> [accessed 6 May 2018].

[133] Lucia Menezes de Oliveira and others, pp. 447-448.

[134] Inferred from: Soham Chatterjee, 'Good Data and Machine Learning', Towards Data Science, 24 August 2017 <https://towardsdatascience.com/data-correlation-can-make-or-break-your-machine-learning-project-82ee11039cc9> [accessed 4 September 2018].

[135] See: Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues*').

[136] D Sculley and others, 'Machine learning: The high-interest credit card of technical debt', *Google Inc*, SE4ML: Software Engineering for Machine Learning, NIPS, (2014), p. 3 <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/43146.pdf>.

[137] Kemal Altinkemer and others, 'Vulnerabilities and Patches of Open Source Software: An empirical study', *Journal of Information System Security*, 4.2, (2008), 5-7 <https://www.krannert.purdue.edu/academics/mis/workshop/papers/ars_092305.pdf>.

[138] Edward Smith and others, 'Is the Cure Worse Than the Disease? Over-fitting in Automated Program Repair', *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*, ACM, (2015), Abstract <https://www.cs.cmu.edu/~clegoues/docs/smith15fse.pdf>.

synchronising of subsequent intervention. Patching, points Sculley, is a recognised source of machine variability that downgrades overall cohesion of a system.[139] The issue for the Cohort is that it also reduces system *predictability* given the universe of new tasks (and an increased vulnerability to the system) that this entails.[140]

6.4 AWS learning architecture

What architecture might therefore be appropriate for weapons where human supervision has been to be removed?[141] The assumption for this review is inferred from McCarthy that AWS must at least be able to effect deduction and interpretation whereby whatever the weapon has experienced in prior cases should inform but not decide what the weapon should expect now.[142] Booch, chief scientist on IBM's Watson programme, concludes that such 'reasoning and learning are the litmus test to defining an AI'.[143] Autonomous weapons' reasoning capabilities must thus encompass not less than an understanding of a *known case* whose relationships can then be carried over to the *present case*; In considering the issue, Haikonnen argues that two-plus-two bananas should be analogous to two-plus-two apples.[144] The nub of the next four chapters is to evidence why this will be particularly complicated for AWS routines. Currently, system routines in machines are handcrafted whereby human programmers are responsible for defining tasking and the way that solutions are to executed.[145] The issue for AWS deployment (as noted by Boulanin and Verbruggen) is that autonomy's inherent limitations are only revealed as an environment become too complex to be captured in such models' programming.[146] It is for this reason that the Delivery Cohort must then rely upon ML to underpin that deployment. Humans, after all, are evolutionarily capable of reasoning when available information is imperfect, formulating deductions that are based on knowledge that is 'generally true'. In the case of AWS, working with incomplete information will require cascading through multiple routines in parallel (each with their own filters, error bias screens, weightings and confidence predictions).[147] Sombe identifies that machine outcomes will also be 'inappropriately transient' as weapon sensors contribute new data to refine previously available information (without necessarily contradicting it).[148]

---

[139] Sculley and others, p. 5. See also: Chapter 10 (*Oversight*), specifically 10.2 (*Validation and testing'*).

[140] US Air Force, Office of the Chief Scientist, 'Autonomous horizons', p. 6.

[141] As discussed in Chapter One's Introduction, this thesis' technical analysis assumes the deployment of broad-task unsupervised machines with wide weapon autonomy in order to identify the widest record of possible procedural faultlines.

[142] John McCarthy, 'What is Artificial Intelligence?', Stanford University Department of Computer Science, (12 November 2007) <http://www-formal.stanford.edu/jmc/whatisai/> [accessed 20 January 2017].

[143] Grady Booch, Chief Scientist, IBM Watson/M, Department of Embodied Cognition, RUSI/Institute for Life Conference Collaboration, in conversation with the author, 8 November 2017.

[144] Haikonnen, p. 93.

[145] For a useful primer on the ramifications of pervasive software, see: James Somes, 'The Coming Software Apocalypse', *The Atlantic*, 26 September 2017 <https://www.theatlantic.com/technology/archive/2017/09/saving-the-world-from-code/540393/> [accessed 7 September 2018].

[146] Boulanin and Verbruggen, p. 16.

[147] These routines are generally discussed in the following chapter. See: Chapter 8 (*Software*).

[148] Lea Sombe, 'Reasoning under incomplete information in Artificial Intelligence', *International Journal of Intelligent Systems*, 5, (September 1990), 423.

Several issues clearly arise from this review of AWS architecture. First, it is not obvious which system architecture can provide those capabilities that are required to remove machine supervision.[149] Second, architectures must facilitate other counterintuitive capabilities such as detection of contradictions, evaluation of significance and, complicatedly, the efficient *rejection* of those alternatives that leave the weapon with foreseen unsatisfactory outcomes. Given this combinatorial complexity, a material leap from current software processes is required for AWS deployment to be realised. Since such computation takes up time and memory, such architecture must also include (as inferred from Brom and Bryson) appropriate bias in order to *constrain* weapon's processes, either in a highly distributed manner (whereby that bias is programmed across the weapon's principal routines, a complicated exercise of balance) or by the addition to AWS architecture of special purpose modules.[150] Gosavi notes that each such routine introduces additional 'noise'[151] and complexity, either from the weapon's reconciliation of disparate data, from integrating subsequent learning output in its succeeding sequences or, lastly, from smoothing interventions that must take place in order to achieve system stability.[152] Subsequent chapters demonstrate, after all, that 'machine learning is messy'[153], especially when it involves uncertain, dynamically changing environments that are characterized either by hidden or partially observable states.[154] Bosquet notes that the battlefield demonstrates exactly this lack of order.[155]

An inquiry into AWS architecture must therefore review the likely programming spine of AWS, the artificial neural network (ANN).[156] Whether this spine then constitutes machine learning or, notes Warner and Misra, other similarly 'statistical tools' is less relevant to an analysis on machine feasibility than identification of enduring technical constraints.[157] The broad construct for AWS deployment must nevertheless be for a general information-processing model that is based on the

[149] While the purpose here is to identify high-level system challenges in AWS, it is useful to list current research directions in order to highlight the complexities involved in achieving degrees of machine reasoning. These include 'default logic, nonmonotonic modal logic, auto-epistemic logic, circumscription and circumscription-like approaches, supposition-based logic, conditional logic, logics of uncertainty as well as belief functions, numerical quantifier logic and fuzzy logic'.

[150] Cyril Brom and Joanna Bryson, 'Action Selection from Intelligence Systems', *European Network for the Advancement of Cognitive Systems*, (2006) <https://pdfs.semanticscholar.org/3419/955ab2c59b029dcda41904d46b763018de79.pdf>.

[151] Abhijit Gosavi, 'The Effect of Noise on Artificial Intelligence and Meta-heuristic Techniques', *Proceedings of the Artificial Neural Network in Engineering Conference*, 12, (2002), p. 1 and p. 7.

[152] Gavin Taylor and Ronald Parr, 'Value Function Approximation in Noisy Environments', *Cornell University Press*, arXiv preprint arViv 1210.4898, (2012) <https://arxiv.org/abs/1210.4898> [accessed 5 September 2017].

[153] Sculley and others, p. 2. See: Chapter 7 (Firmware), specifically: 7.2 ('*Firmware ramifications of machine learning*') and 7.3 ('*Reasoning and cognition methodologies*').

[154] Carla Brodley and others, 'Challenges and Opportunities in Applied Machine Learning', *AI magazine*, Association for the Advancement of Artificial Intelligence, 33, (2012), pp. 11-12.

[155] Antoine Bousquet, 'The Scientific Way of Warfare: Order and Chaos on the Battlefield of Modernity', *LSE*, PhD thesis, (2014), p. 3 and pp. 189-195 <http://etheses.lse.ac.uk/2703/1/U615652.pdf>.

[156] A Google search using 'autonomy' and 'artificial neural network' returns 741,000 references [accessed 25 March 2017]. For discussion on ANNs' role in AWS, see also: Ben Farmer, 'Prepare for the rise of 'Killer Robots' says former Defence Chief', *Telegraph Newspaper*, 17 August 2017, paras. 4-7 <http://www.telegraph.co.uk/news/2017/08/27/prepare-rise-killer-robots-says-former-defence-chief/> [accessed 29 August 2017].

[157] Brad Warner and Manavendra Misra, 'Understanding Neural Networks as Statistical Tools', *American Statistician*, 50, 4, (November 1996), p. 284 <ftp://gis.msl.mt.gov/Maxell/Models/Predictive_Modeling_for_DSS_Lincoln_NE_121510/Modeling_Literature/Warner%20and%20Misra_neural%20networks.pdf>.

way biological nervous systems process information. An essential element to this paradigm would be the novel structure of the weapon's processing system which should be composed of very many highly interconnected processing elements (neurones) working in unison to solve specific problems.[158] The architectural intention is that ANNs, like people, learn by example.[159] The unsupervised weapon's network will then be configured for *specific* individual applications, such as pattern recognition or data classification, through a learning process that subsequently knits together outputs in order to generate an action path. While learning in biological systems involves adjustments to the synaptic connections that exist between the neurones, this process is then to be copied in machines in their ANN where, in the case of AWS, the machine's neurons can either be in 'training' or 'using' mode. The Cohort's goal for its weapon networks is that it recreates a learning system in machine code by developing specific framework programmes to tackle specific problems.[160] Translating the work of Stergio and Siganos to the battlefield, AWS' models will rely upon receiving system feedback on how its programme routines are performing.[161] The weapon's neural network can theoretically then optimize its response by doing the same problem thousands of times and adjusting its response according to this feedback. A computer, in this case the routines within an AWS, can then theoretically be given a different problem which it can approach in the same way as it learned from the previous one.[162] By varying the problems and the number of approaches to solving them that the computer has learned, the theory is that the Delivery Cohort can teach the AWS to be necessarily generalist and, again in theory, more broadly and appropriately 'intelligent' to its battlefield surroundings and task lists.[163] Whether this is achieved through machine learning or other statistical and logic tools does not alter the architectural, behavioural and technical issues that must all be integrated prior to the weapon's compliant deployment.[164]

In evaluating AWS deployment, there has been considerable recent progress in such models.[165] Since 2009, deep-learning algorithms based on ANNs have run on graphical processing units (GPUs), specialized chips used in PCs and video games consoles.[166] This recent marriage has

---

[158] RT Networks, 'Kalashnikov develops fully automated neural-network-based combat module', (5 July 2017) <https://www.rt.com/news/395375-kalashnikov-automated-neural-network-gun/> [accessed 4 March 2018].

[159] For a primer on ANN, see: Christos Stergio and Dmitrios Siganos, 'Neural Networks', undated <https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html> [accessed 12 March 2018].

[160] Anish Talwar and Yogesh Kumar, 'Machine Learning', pp. 3400-3402. See also: Anish Talwar and Yogesh Kumar, 'An Artificial Intelligence Methodology', *International Journal of Engineering and Computer Science*, 2, 12, (2012), Abstract.

[161] Stergio and Siganos, 'Neural Networks', generally. See also: Vidushi Sharma and others, 'A Comprehensive Study of Artificial Neural Networks', *International Journal of Research in Computer Science and Software Engineering*, 278-279 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.9353&rep=rep1&type=pdf>.

[162] See, generally: Simon Haykin, *Neural Networks and Learning Machines*, 3rd Edition, (USA: Person Prentice Hall, 1999), pp. 1-6.

[163] Techopedia Staff, 'What Is the Difference Between Artificial Intelligence and Neural Networks?', (6 December 2017) <https://www.techopedia.com/2/27888/programming/what-is-the-difference-between-artificial-intelligence-and-neural-networks> [accessed 20 January 2017].

[164] Bernard Marr, 'What is the Difference between Artificial Intelligence and Machine Learning?', *Forbes Magazine*, Technology, 6 December 2016 <https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#5c79a3082742> [accessed 6 March 2018].

[165] See, for instance: Martin Wielomski, 'The GPU: Powering the Future of Machine Learning of AI', *Phoenix NAP Publications,* (21 September 2018), generally <https://phoenixnap.com/blog/future-gpu-machine-learning-ai> [accessed 12 February 2019].

[166] M Janakiram, 'In the Era of Artificial Intelligence, GPUs are the New CPUs', *Forbes Magazine*, 7 August 2017 <https://www.forbes.com/sites/janakirammsv/2017/08/07/in-the-era-of-artificial-intelligence-gpus-are-the-new-

increased deep-learning system processing nearly a hundredfold and allowed, notes the *Economist*, the training of a multi-layer neural network to take less than a day, a procedure which had previously taken several weeks.[167] Such deep-learning systems have also become exponentially more powerful with networks of double digit layers currently being worked upon by researchers.[168] Using deep networks, the JASON Study[169] reckons that the error rate of image capture has fallen from twenty-five per cent to some three per cent. This is better than the accepted figure for human performance of five per cent. Difficulties, however, abound. While such processes may be surprisingly accurate on an individually equational basis, *systemic* problems arise from the manner in which such networks misclassify datasets.[170] Nguyen highlights that such ranking errors are particularly likely to occur in ways that are unpredictable and unfamiliar to humans.[171] In this case, the architectural issue is that the trait will likely complicate how the weapon's neural network and other statistical tools might classify battlefield objects since outcomes arising from any such cognitive approaches will, as inferred from Berreby, likely be materially different from those expected by the Delivery Cohort.[172]

Weapon architecture that is based upon neural networks will involve intractable complexity. The AWS' neural network must have a number of inputs, each of which Haikonnen notes must have three characteristics. First, each input (here, the weapon's original representation or, more likely, subsequently sensed data derived from its battlefield surroundings) must have its own weight (or synaptic) value. Second, it must have a summing function and, third, it will have a threshold-based output function.[173] The model's premise is that all such inputs are largely free from noise and sufficiently full in detail.[174] A further challenge is that these weapon inputs are the *right* inputs (indeed, *all* of the right inputs) necessary (as well as being sufficiently comprehensive) to allow that weapon to derive applicable intelligence from its battlefield surroundings and, crucially, to divine meaning from those inputs. But the process thus far is incomplete. Based on continuous signals (that each have variable intensities), the value of each input signal must then be multiplied by its related weight value with the results then summed together.[175] In AWS deployment, this is complex given the intractable imprecision of these inputs upon which, for instance, engagement calculations

---

cpus/#705bb4955d16> [accessed 3 September 2018]. See also: Simon Haykin, '*Neural Networks and Learning Machines'*, pp. 21-24 ('*Network architectures'*) and 94-100 ('*Unconstrained optimization: A review'*).

[167] Economist Magazine Special Report, 'Artificial intelligence'*, Economist,* p. 4. There is, however, no evidence that this learning process can be undertaken in real-time.

[168] Carole Lundgren, 'Recent Development in Neural Networks', Appen, (23 March 2018), generally <https://appen.com/recent-developments-neural-networks/> [accessed 5 September 2018].

[169] JASON program, p. 5. The report cites networks with thirty such layers.

[170] Anh Nguyen, 'Deep networks are easily fooled: high confidence predictions for unrecognizable images'*, Computer vision and pattern recognition*, IEEE, (2015) <http://arxiv.org/pdf/1412.1897v4.pdf>.

[171] Ibid.

[172] David Berreby, 'Artificial intelligence is already weirdly inhuman: what kind of world is our code creating?'*, Nautilus*, (6 August 2016) <http://nautil.us/issue/27/dark-matter/artificial-intelligence-is-already-wierdly-inhuman> [accessed 12 May 2017]

[173] Haikonnen, p. 31.

[174] Eugenio Culurciello, 'Neural Network Architectures', *Towards Data Science*, (23 March 2017) <https://towardsdatascience.com/neural-network-architectures-156e5bad51ba> [accessed 3 September 2017].

[175] Avinash Sharma, 'Understanding Activation Functions in Neural Networks', *Medium.com blog*, (30 March 2017) <https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0> [accessed 3 September 2017].

must be based. The process, notes Dietterich, is neither rapid nor real-time.[176] To enable this model, this sum value is then integrated with the weapon's threshold circuit with an artificial numeral providing appropriate output if specific thresholds are exceeded. A further challenge then arises where *no* combination of weight values meets the thresholds that have been set for the AWS by its Delivery Cohort (AI's *Exclusive-Or Problem*).[177]

It is also useful to consider how these statistical tools might apply in a battlefield context.[178] Within the weapon's network, each neural unit must be connected with innumerable others with such statistical links either having an enforcing or an inhibitory effect on the activation state of the weapon's neural units. As above, each individual neural unit will have a broad summation function, a threshold function or a limiting function on each connection *and* on the unit itself. In this way, a battlefield signal must exceed a limit that, in theory, has be defined by the Cohort before being able to propagate on to other neurons. It is thus envisaged that the unsupervised weapon system will become 'trained' rather than being explicitly programmed. A constraint inferred from Han and others, however, is that the unsupervised weapon must minimally have its architecture fixed *before* training starts. In other words, training cannot subsequently improve the weapon's architecture.[179] Nevertheless, the intended benefit in this case is that machine-learning will enable the AWS to 'excel in areas where the solution or feature detection is difficult to express in a traditional computer program'.[180] How then might this work? Neural networks typically consist of multiple layers (or a cube design) with signal paths traversing from front to back. This is necessary given that the structure will be predicated on training the AWS by running and re-running very large sets of 'experienced data' in an iterative process.[181] This repeating allows those layers of neurons to adopt and refine 'prioritizing weights' so that the weapon system might then make sense of *new* data sets on the same basis that it has encountered, honed and weighted previous training sets of data. Back propagation is therefore a vital operation[182] for AWS function whereby stimulation is used to reset weights within the framework's neural units.[183] The challenge is to regiment the weapon's network connections in order to prevent interaction in a chaotic and complex fashion, the

[176] Thomas Dietterich, 'Machine Learning for Sequential Data: A Review', *Structural, Syntactic and Statistical Pattern Recognition*, (2002), pp. 5-7 <http://web.engr.oregonstate.edu/~tgd/publications/mlsd-ssspr.pdf>.

[177] Haikonnen, p. 33.

[178] The weapon's neural network will, after all, be a connectionist system based on a very large assemblage of artificial neural units that are loosely modelled on the way a human brain solves problems. See: Jerry Fodor and Zenon Pylyshyn, 'Connectionism and Cognitive Architecture: A critical analysis', *Rutgers University*, undated, pp. 2-4 <http://www-cogsci.ucsd.edu/~sereno/170/readings/02-FodorPylyshyn.pdf>.

[179] Song Han and others, 'Learning both Weighting and Connections for Efficient Neural Networks', *Advances in Neural Information Processing Systems,* (2015), Abstract <https://papers.nips.cc/paper/5784-learning-both-weights-and-connections-for-efficient-neural-network.pdf>.

[180] See: Royal Society, 'The Power and Promise of Computers that Learn by Example', *Royal Foundation*, (April 2017) <https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf>. For an assessment of ML deficiencies see: p. 30.

[181] Roumen Trifonov and others, 'Artificial Neural Network Intelligent Method for Prediction', *AIP Conference Proceedings*, (6 July 2017), pp. 1-2 and generally <https://aip.scitation.org/doi/pdf/10.1063/1.4996678?class=pdf>.

[182] V Preetham, 'Back Propagation – How Neural Networks learn Complex Behaviours', *Autonomous Agents #AI,* (9 August 2016) <https://medium.com/autonomous-agents/backpropagation-how-neural-networks-learn-complex-behaviors-9572ac161670#.qwz64wcu6> [accessed 2 February 2017].

[183] Yoshua Bengio, 'Challenges of Training Deep Neural Networks', *Montreal, Course notes IFT6266*, (Winter 2012) <https://www.iro.umontreal.ca/~bengioy/ift6266/H12/html.old/deepchallenge_en.html>) [accessed 3 August 2017]..

forming of new connections and even new neural units while disabling others.[184] Current network projects, furthermore, work with up to a few million neural units and connections which equates to the computing power of a worm.[185] For this reason, Vincent notes the challenge presented by the massive scaling-up in capability that is required if human supervision is to be amended.[186]

In order properly to identify architectural constraints, a further purpose of this section is to impose on these processes an empirical lens that is based on battlefield *practices*. Given that AWS learning will be fundamentally based on mathematical techniques involving statistics and probability, several challenges can be inferred to any widespread use of this model on the battlefield.[187] For a weapon network that will operate on sensed images, *each* visualization input represents a single pixel. For reference, a standard single TrueColour digital image currently requires thirty megabytes of platform memory. An HDTV clip from just one sensor and at just 1920 x 1080 pixels (capturing its environment at sixty frames per second) requires more than twenty gigabytes memory for every minute of video input.[188] Operating continually in real time, recording and processing input from multiple visualization camera positions[189], this equates to at least three hundred gigabytes of relevant information per minute for the deployed AWS. Input neurons, moreover, have the broadest application and may also represent audio sample for speech recognition, character sample for natural language understanding or chemical sample for hazard or some other olfactory measure. The model for AWS is that discrete data (here, characters or numbers) must be represented in a dynamic series of one-shot representations where a separate neuron is used for each possible symbol in each position.[190] Given the number of individual data points that will comprise a single engagement sequence, in can be inferred from Sigh and Wood that the architectural convolution of processing this information (while accounting for contextual sensitivities, clutter, bottlenecks and significance) is in itself an intractable proposition.[191] A recurring theme of this thesis is that maintaining human control over such systems is not simply a matter of ensuring compliance but is a matter of feasibility.

Central to weapon learning capabilities (again, whether by neural networks or other logic framework) will be its training processes in which the millions of weights connecting its neurons

---

[184] Stergio and Siganos, 'Neural Networks', generally.

[185] Artificial Brains blog, 'OpenWorm', (14 August 2012) <http://www.artificialbrains.com/openworm> [accessed 12 February 2017].

[186] See: James Vincent, 'These are Three of the Biggest Problems Facing artificial Intelligence', *The Verge*, 10 October 2016, paras. 4-5 of 13 <https://www.theverge.com/2016/10/10/13224930/ai-deep-learning-limitations-drawbacks> [accessed 1 April 2017].

[187] Bengio, 'Challenges of Training Deep Neural Networks', paras. 2-3. For a general discussion of weapon network architecture, network training and back-propagation, see: JASON program, pp. 6-19.

[188] The TrueColour photo assumes a 2736 x 3648 pixel format. See: Ken's Image Gallery <http://kias.dyndns.org/comath/44.html> [accessed 3 November 2017].

[189] The current Tesla S Class car requires 8 cameras, 1 radar unit and 6 ultrasonic units just to auto-drive; *Electrek,* <https://electrek.co/2016/10/20/tesla-new-autopilot-hardware-suite-camera-nvidia-tesla-vision/> [accessed 12 October 2017].

[190] LR Medsker and LC Jain, 'Recurrent Neural Network Design and Application', *CRC Press International Series on Computational Intelligence*, (20 December 1999), p. 227.

[191] Inferred from: David Wood and others, 'Can We Ever Escape From Data Overload? A Cognitive systems Diagnosis', *Cognition, Technology and Work*, 4.1, (2002), p. 3 and pp. 7-10. See also: Shailesh Singh and others, 'Training of Artificial Neural Networks using Information-Rich Data', *Hydrology*, 1(1), (2014), p. 42 <https://pdfs.semanticscholar.org/cbb2/38ded06c5d17ca8f749a917828ba57ab8c03.pdf>.

are assigned values. Indeed, the *only* adjustable parameters in this process are these weights. Discovering optimal values for these weights will occur during the process of training very large data sets of input/output pairs.[192] For example, in order to train a weapon network to recognise images, the x-input function might be intensity values of the image pixels while the y-input function the picture's description. The model is to match these two inputs. The architectural challenge to AWS is that the performance of individual neurons must be quantified using an error function with the goal of training being to *minimise* this function. In AWS, this will be achieved using techniques such as a *gradient descent* whereby the whole network's weighting is updated again and again using the gradient of the error function.[193] The architectural model in this case is for such iteration to continue repeatedly until the weapons network's weights find the global minimum of the error function averaged over all of that training data. Nasiriany and others, however, note that additional routines are required: As part of this process, a sub-set of the weapon's original raw data might be hived off from the main training set to be used as validation data in order to measure how well the training has progressed and to determine whether it is necessary to adjust parameters such as learning rate and architecture.[194] The efficiency of this process depends on two further characteristics. There must be an ability to propagate *forwards* through the weapon's network to calculate the current values of the weapon system. Repetton then notes that there must be a subsequent ability to propagate *backwards* the error function through the network in order to update the weapon's learning weights.[195] While each iteration through the entire set of training data is termed an *epoch*, current networks are targeted at a single pre-determined task which, once learned, require that network's connections be frozen on deployment. As again noted by Repetton, this creates an inappropriately 'one trick pony' with no facility for additional learning.[196] Notwithstanding the static nature of current networks, the dynamic nature of battlefield conditions will require the infeasible collection of epochs in order to improve the weapon's error function.

What then might the operational ramifications be of such data training? A systemic issue can be inferred from Angelov and Sperduti that such networks will be inappropriately unpredictable for battlefield deployment.[197] After training, some neurons become demonstrable problem solvers while others empirically do not perform as well. As above, several thousand cycles of interaction must typically occur. A ramification here is then a plateauing in performance due to ever reducing learning gradients and ever-smaller changes to the model's weights.[198] Each new layer to a

[192] For a detailed explanation of back-propagation and DNN training, see: JASON program, pp. 61-68.

[193] Soroush Nasiriany and others, 'A Comprehensive Guide to Machine Learning', *University of California at Berkeley*, (13 August 2016, Section 4.3 (*'Gradient Descent'),* pp. 82-88 <http://snasiriany.me/files/ml-book.pdf>.

[194] Several online resources are available on the division of labelled datasets into training sets, validation sets and test sets. See, for instance, Stanford University videos on Machine learning, *Coursera* <https://www.coursera.org/learn/machine-learning/lecture/QGKbr/model-selection-and-train-validation-test-sets> [accessed 24 May 2017].

[195] Anthony Repetton, 'The Problem with Back-Propagation', *Towards Data Science*, (18 August 2017), generally <https://towardsdatascience.com/the-problem-with-back-propagation-13aa84aabd71> [accessed 29 July 2018].

[196] Ibid., para. 3 of 23.

[197] Plamen Angelov and Alessandro Sperduti, 'Challenges in Deep Learning', *Proceedings of the European Symposium on Artificial Neural Networks*, Bruges, (April 2016), pp. 489-491 <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2016-23.pdf>. The section highlights challenges arising from the need for regularization schemes, system complexity and a requirement to use off-line methods to ensure computational efficiency (which would not be appropriate for AWS).

[198] Xavier Girot and Yoshua Bengio, 'Understanding the Difficulties of Training Deep Feed-forward Neural Networks', *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, (2010), p. 249 and generally

weapon's neural network also means an extra layer of non-linearity which, inferred from Bengio, increases the difficulty of optimizing the weapon's learning process. Such models, moreover, readily restrict their learning to just the network's top layer 'while lower layers remain random transformations that do not capture much input'.[199] Gradients are also prone to inherent dilution in these lower layers and may provide unpredictable and weak guidance to the overall learning process.[200] If AWS are to be accretive to the Delivery Cohort, it is machine learning's intrinsic instability that questions its attraction as an appropriate technical spine for AWS.

The efficacy of the architecture is dependent upon the fit between weapon *tasking* (image classification or sound classification or written word manipulation) and weapon *training*.[201] As noted by Dohler (and evidenced in later sections[202]), a marginally different set-up or a marginally different training dataset will likely lead to very substantial output discrepancies.[203] To this point, Murphey confirms that identifying scenario variation in this primary data is beyond human intervention given the specificity of the raw dataset.[204] Training methodology is, after all, critical to weapon feasibility. If incorrectly commissioned, battlefield features with only a small number of examples in that training set (but possibly of critical importance) may likely be ignored by the process. The architecture therefore relies on very clear data *definition* which, suggests Sharkey, is easily confounded by Sun Tzu's maxim for the successful commander who must be 'extremely subtle, even to the point of formlessness'.[205] This problem of discounting evidence is, moreover, multifaceted.[206] Singh notes that it may be caused by an overwhelming number of learning examples in one set of the weapon's sensor data undoing the training effect on the learning examples in a different data set. It may also be caused by incorrectly set model sensitivity ('detection rate') and incorrectly set model specificity ('false alarm rate'). Longer time series data does not, furthermore, equate to more useful information; if there is considerable repetition in the weapon's sensed data then the AWS' ANN may not become any 'wiser' from additional training.[207]

<http://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf?hc_location=ufi> [accessed 23 July 2017]. See also: Narayanan Manikandan, 'Software design challenges in time series prediction using parallel implementation of artificial neural networks', *The Scientific World Journal*, 2016, Article ID 670 9352 <http://do.doi.org/10.1155/2016/6709352> [accessed 26 July 2018].

[199] Guillaume Alain and Yoshua Bengio, 'Understanding Intermediate Layers', *ICLR Paper*, (2017), pp. 7-8 (see also '*Abstract*') <https://pdfs.semanticscholar.org/2706/77b5c44ea0c93313f41db2f885fef305bbcc.pdf>.

[200] Bengio, 'Challenges of Training Deep Neural Networks', generally.

[201] Although relating to the medical field, see: G Auda and others, 'Improving the accuracy of an artificial neural network using multiple differently trained networks', *Neural Network Proceedings*, IEEE World Congress on Computational Intelligence, (1998), 2, p. 1356 <http://ieeexplore.ieee.org/document/685972/> [accessed 4 September 2017].

[202] See: Chapter 7 (*Firmware*), specifically: 7.2 ('*Firmware ramifications of machine learning*') and 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues*').

[203] Mischa Dohler, Professor in Wireless Communications, King's college, London, Fellow IEEE and RSA, in conversation with the author, 20 January 2017.

[204] Yi Murphey and others, 'Neural Learning from Unbalanced Data', *Applied Intelligence,* 21, (2004), pp. 117-118 <http://sci2s.ugr.es/keel/pdf/specific/articulo/NL-Unbalanced-data.pdf>.

[205] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017. See also: Eric Jackson, 'Sun Tsu's Art of War', *Forbes Magazine*, 23 May 2014 <https://www.forbes.com/sites/ericjackson/2014/05/23/sun-tzus-33-best-pieces-of-leadership-advice/#19c3ac7d5e5e> [accessed 2 August 2018].

[206] Greg Williams, 'Wise up, Deep Learning May Never Create a General Purpose AI', *Wired Magazine*, 28 January 2018, generally <https://www.wired.co.uk/article/deep-learning-automl-cloud-gary-marcus> [accessed 5 September 2018].

[207] Singh, p. 42.

The architectural ramification of learning instability is also compounded by data noise. Murphey notes that minute distortion in AWS classification of its sensed data will likely lead to different data classes being inseparable in the space where such variables are processed.[208] In other words, if a weapon's dataset is noisy, the class boundary that separates different class examples is almost impossible for the weapon to define and separate for ongoing statistical analysis. If, as suggested by Beradi and Zhang, such misclassification is 'inevitable'[209], then further supervision sequences are required in order to train the AWS to make 'favourable classification decisions towards a particular class'.[210] Termed the weapon's *matching* challenge, Forgy highlights that it is also a source of inappropriate third-party biases.[211] An instance might be the over-fitting of training data that will render the weapon's training model brittle. The sequences might, for instance, require unknown and incremental processing steps in what should otherwise be a seamless, real-time action series. Other processing pitfalls then arise. Given that as much as ninety per cent of the weapon's run-time might be taken up with pattern matching[212], Kirkpatrick notes that part-processed data might also be hived off into interim (possibly off-line) storage between matching cycles. In this case, complicating 'weight decay' routines[213] may then be required to regularize training sequences and to manage what Kirkpatrick terms 'catastrophic forgetting'.[214] Such ancillary routines usually involve compromise: Adjusting training weights may make the weapon less sensitive to noise from data inputs but correspondingly less likely to learn from that noise.

Such weapon architecture will be disproportionately affected by biases. As highlighted by Bullinaria, they may push the weapon's neurons into saturation which then desensitizes those neurons to *all* inputs.[215] A further architectural complication then arises over 'dropout' whereby learning routines regularly omit randomly selected neurons from the weapon's training process in order to reduce over-fitting and false correlation.[216] This additional step is likely to impact machine outcomes and so introduce inappropriate added variability into the weapon's engagement routines. It is also unknowable from the outset if the weapon's training data is both sufficiently relevant to its y-function (the task that the Delivery Cohort has for each network) or of sufficient size appropriately to train the network. Mu and others note that the architecture's descent gradient (here, determining weapon system data fit) is *systemically* unstable 'and therefore incapable of

[208] Murphey, p. 118.

[209] Victor Beradi and Peter Zhang, 'The Effect of Misclassification Costs on Neural Network Classifiers', *Researchgate.net*, (June 1999) <https://www.researchgate.net/publication/227807698_The_Effect_of_Misclassification_Costs_on_Neural_Network_Classifiers> [accessed 5 May 2017].

[210] Murphey, pp. 117-118.

[211] Charles Forgy, 'A fast algorithm for the fast pattern/many object pattern match problem', *Artificial intelligence*, 19, (1982), pp. 17-18 <https://pdfs.semanticscholar.org/464e/b245ff9822defa8db82c385ff1fd0b0b6ffe.pdf>.

[212] Forgy, p. 22.

[213] James Kirkpatrick and others, 'Overcoming Catastrophic Forgetting in Neural Networks', *PNAS*, *Proceedings in the Natural Academy of Sciences*, 14, 13, (March 2017), p. 3521 <http://www.pnas.org/content/114/13/3521.full.pdf>.

[214] Kirkpatrick and others, pp. 3524-3525.

[215] For a general primer, see: John Bullinaria, 'Biases and Variances, Under-fitting and over-fitting', *Birmingham University*, Lecture 9, (2015) <http://www.cs.bham.ac.uk/~jxb/NN/l9.pdf>.

[216] See: Amar Bughiaja, 'Dropout in (Deep) Machine Learning', *Medium.com*, (15 December 2016), generally <https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5> [accessed 8 September 2018].

remediation'.[217] This is unsurprising given that such gradients are derived directly from the product of terms in all subsequent network layers and, as above, the product of many of these terms will themselves be similarly volatile. In considering the architecture of AWS, the JASON report concludes that 'different layers will learn at different rates, and the learning will be unbalanced. This problem becomes worse as the number of layers becomes large and is thus a particular challenge for deep neural networks'.[218] In settling that weapon learning derives 'an approximate answer [that] is usually good enough; when it works, it is not necessary to understand why or how', JASON is unintentionally confirming the unsuitability of machine learning as the AWS' technical spine.[219]

There are also *operational* difficulties with this architectural learning model that are relevant to AWS deployment. This is unsurprising given JASON's conclusion that current learning technology 'has not systematically addressed the engineering priorities of reliability, maintainability, debug-ability, evolvability, fragility and attackability'.[220] That report also questions whether existing AI models are systematically amenable to validation or verification.[221] It is the *number* of parameters (including attendant weights and biases) that is operationally problematic. Outputs will depend on the training data used, the order in which this data is processed, the training algorithm employed and, as noted by Matthews, the frequency with which machine templates are updated.[222] In highlighting that the function is non-convex (while the model's optimization has been designed for convex problems such as the gradient descent, for instance, as set out above), Van den Berg cites this as a key architectural shortcoming.[223] Moreover, the operational significance of such variability is unlikely to be spotted and remediated in any timely manner. Regularization routines will also act as an aggressive and inappropriate edit on the AWS' primary sensed data while Wei Pan notes that the practice of dropping neurons ignores the non-zero weighting of *all* such connections.[224] While Sharkey's conclusion is that the fragility of learning networks can only increase as that model is shoehorned into weapon management[225], Bartlett notes that sample complexity requires that 'the number of training examples should grow *at least linearly* with the number of adjustable

[217] Yadong Mu and others, 'Stochastic Gradient Made Stable: A Manifold Propagation Approach for Large Scale Optimisation', *arXiv*:1506.08350v2, (12 January 2016), pp. 1-2 <https://arxiv.org/pdf/1506.08350.pdf>.

[218] JASON program, p. 66.

[219] Ibid., p. 25.

[220] Ibid., p. 27.

[221] See: Chapters 10 (*Oversight*), specifically: 10.2 ('*Validation and testing'*) and 9 (*Hardware*), specifically: 9.4 ('*Operational hardware issues'*).

[222] Iain Matthews and others, 'The Template Update Problem', *Robotics Institute, Carnegie Mellon University*, (2004), pp. 13-14 <http://www.ri.cmu.edu/pub_files/pub4/matthews_iain_2004_1/matthews_iain_2004_1.pdf>.

[223] Ewout van den Berg, 'Training variance and performance evaluation in Neural Networks in speech', *IBM Watson Group*, ICLR, (2016), p. 1 and p. 5 <https://arxiv.org/pdf/1606.04521.pdf>. Linear functions are convex. Linear programming problems are therefore convex problems where there can only be one optimal solution. Non-convex problems are, reports Berg, more complex and often intractable.

[224] Wei Pan and others, 'DropNeuron: Simplifying the Structure', *arXiv*, (23 June 2016), pp. 3-4 <https://arxiv.org/abs/1606.07326> [accessed 15 July 2017]. Other strategies include pruning datasets based on a test of 'small weighting' and amending the training run's cost function based on 'sparse regression'.

[225] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

parameters in the network'.[226] Efficiencies, therefore, appear unlikely from expanding data inputs in these weapon processes. Finally to this point, current precedents remain laboratory-bound[227] and, notes Wray, largely restricted to limited tasks in a controlled environment that is amenable to human intervention and supervision. Evidence is scarce from its practical deployment in settings that are subject to arbitrary occurrences.[228] The conclusion for this section is therefore provided by Asaro whereby 'the decision to kill a human can only be legitimate if it is a non-arbitrary and there is no way to guarantee and that the use of force is not arbitrary without human control, supervision and responsibility'.[229]

6.5 Missing pieces

Another way to look at the feasibility of AWS' likely architectural basis is to consider instead those critical technologies that currently remain outstanding. Selam notes in this case that several such capabilities appear out of reach including, inter alia, reliable processing of abstract imagery, robust summarization skills as well as proven system tools that permit scene and episode understanding.[230] As noted in SIPRI's November 2017 report (and in line with Sabin's 'Revolution in Expectation'[231]), 'delivery of such technology trails behind expectations'.[232] In particular, Srinivasan highlights that computers struggle to interpret wider context.[233] Vision software may identify a soldier walking but is unable to determine *why* the soldier is walking.[234] Given the difficulty of representing abstract relationships between objects and people in models of the real world, the same happens in computer speech recognition where the computer may understand what is said but not 'what is being discussed'.[235] This likely renders autonomous systems particularly vulnerable

[226] Peter Bartlett, 'The sample complexity of the pattern classification with networks; the size of the weights is more important than the size of the network', *IEEE Transactions on information*, 44, 2, (March 1998), p. 1. See, also, Chapter 1 (*Introduction*), specifically 1.7 ('*Statement of Methods*').

[227] Barry Wray and others, 'An artificial neural network approach to learning from factory performance from a Kanban based system', *Journal of International Information Management*, 12, 2, article 7, (2003) generally.

[228] Jason Brownlee, 'How to Improve Deep Learning Performance', *Deep Learning*, (21 September 2016) <https://machinelearningmastery.com/improve-deep-learning-performance/> [accessed 2 August 2017].

[229] Peter Asaro, 'On banning Autonomous Weapon Systems: Human rights, automation and the dehumanisation of lethal decision making', *International review of the Red Cross*, 94, 886, (Summer 2012), p. 693 <https://www.icrc.org/eng/assets/files/review/2012/irrc-886-asaro.pdf>.

[230] B Selam (moderator) and others, 'Challenge problems for Artificial Intelligence', *13th National Conference on AI*, AAAI-96, paras. 7-17 of 40 <http://erichorvitz.com/selman.htm> [accessed 16 June 2017].

[231] See: Chapter 1 (*Introduction*), specifically: 1.7 ('*Statement of* methods'). See also introduction to Chapter 11 (*Conclusion*).

[232] Boulanin and Verbruggen, p. 65.

[233] See: Venkat Srinivasan, 'Context, Language and Reasoning in AI: Three Key Challenges', *MIT Technology Review*, (14 October 2016) <https://www.technologyreview.com/s/602658/context-language-and-reasoning-in-ai-three-key-challenges/> [accessed 12 September 20127]. Although written in 1999, see: Michael Schrage, 'The Real Problem with Computers', *Harvard Business Review*, (October 1997) <https://hbr.org/1997/09/the-real-problem-with-computers> [accessed 8 September 2017].

[234] Ibid., p. 15.

[235] Venkat Srinivasan, 'Context, Language and Reasoning in AI', generally. Srinivasan notes that, in the absence of Natural Language Understanding (NLU), learning systems must convert text into data and, in that conversion process, lose all context within that text.

to trickery.[236] Commenting on such missing pieces, Haikonen, Principal Scientist at Nokia Research's cognitive technology lab, observes of AI that 'performance has varied from barely acceptable to outright ridiculous. As for general intelligence and true creativity, there has been definitely none'[237] and concluding that 'AI and the networks have produced some remarkable results but these approaches do not excel in applications where a true understanding is needed'.[238] AWS, after all, must be able to *sense, think and decide*, *act* and *team*.[239] Without a facility to generalize, Cummings concludes that AWS deployment must be limited to known situations and environments.[240] It is also inferred from Knight that AWS' features are themselves in flux with capabilities under development including 'culturally informed' and values-based reasoning as well as an ability to integrate social with behavioural models.[241] There are, moreover, seeming roadblocks to several of these required capabilities where, notes Bodel, progress has been nonexistent.[242] An example is the enabling of statistical and logic agents to display creativity. While Vego argues that creativity in battlefield routines is a pivotal advantage[243], very little relevant progress is evident in the coding and assimilation of *combinatory* creativity (the novel and likely improbable amalgamation by the weapon of otherwise familiar routines), *exploratory* creativity (the generation of novel strategies through adhoc exploration of what might be conceptual spaces) or, more importantly, *transformational* creativity (the modification of arguments allowing new structures to be generated that would otherwise not have been available to the weapon).[244]

Any such absent capabilities must also *all* be available at the moment of AWS deployment in order for that weapon to be feasible, relevant and compliant. While navigation and routing are already established framework capabilities, Ackerman demonstrates that obstacle avoidance, agility and dexterity remain work in progress.[245] Navigational intent and the ability to exhibit independent actuator control also remain outstanding notwithstanding their importance to AWS

[236] US Army, 'Robotic and Autonomous System Strategy', *Army Capabilities Integration Centre*, p. 4, (March 2017) <http://www.tradoc.army.mil/FrontPageContent/Docs/RAS_Strategy.pdf>. See also: Chapter 8 (*Software*), specifically: 8.2 ('*Coding errors*').

[237] Haikonnen, p. 3.

[238] Ibid., pp. 3-5.

[239] US Department of Defense, 'Summer Study on Autonomy', *Defense Science Board*, Office for Acquisition, Technology and Logistics, Washington, (June 2016), p. 11.

[240] Missy Cummings, 'Artificial intelligence and the future of warfare', *Chatham House*, Research Paper draft, (January 2017), generally.

[241] Will Knight, '5 Big Predictions for Artificial Intelligence in 2017', *MIT Technology Review*, (4 January 2017), paras. 2 and 6 of 16 <https://www.technologyreview.com/s/603216/5-big-predictions-for-artificial-intelligence-in-2017/> [accessed 30 August 2017]. See: Chapter 7 (*Firmware*, specifically: '*Statistical frameworks*'). For specific analysis of technical challenges to AWS routines, see: Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues*') and 8.6 ('*Value setting issues*').

[242] Margaret Bodel, 'Creativity and Artificial Intelligence', *School of Cognitive and Computer Science*, Brighton, Artificial Intelligence, 103, (1998), p. 347.

[243] Milan Vego, 'On Military Creativity', *ndupress*, 70, (Third Quarter 2013), pp. 83-86 <http://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-70/JFQ-70_83-90_Vego.pdf>.

[244] Bodel, p. 348.

[245] Evan Ackerman, 'Algorithms allow Micro Air Vehicles to avoid obstacles with single camera and neuro-morphic hardware', *IEEE Spectrum*, (6 November 2012), paras. 3 and 5 of 9 <http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/algorithms-allow-mavs-to-avoid-obstacles-with-single-camera> [accessed 4 September 2017].

deployment.[246] An additional challenge is that the weapon's sequence of observations and actions are only revealed incrementally.[247] Similarly outstanding is an architecture that can reliably capture *qualia* within weapon systems. For the purposes of this thesis, qualia can best be defined as the subjective or qualitative properties of a lethal engagement.[248] They are the introspectively accessible, phenomenal aspects of what might be the human operator's mental sequences. Zakaria notes the current absence of either an agreed definition for their functional characterization or, more importantly, 'a set of engineering principles for [such] synthetic phenomenology'.[249] The verso is what Suchman and Weber note to be 'the simplistic and shallow [nature] of world models' whereby current machine representations all operate at a 'metric level that precludes reasoning, or at a cognitive level without a physical grounding'.[250]

6.6 AWS control methodologies

In understanding AWS architecture, it is helpful in this chapter's final section to consider from a technical perspective how (and where) the Delivery Cohort can *control* its independent AWS which is being deployed, after all, in order that it adhere to and then further the Cohort's broad plans. AWS actions must be intentional and in accordance with those plans.[251] The contradiction is that such systems are by definition independent. Yampolskiy and Fox note here that there are several variants to such intervention.[252] The Cohort might seek to exert broad *capability* control whereby undesirable engagement outcomes might be avoided by limiting what a weapons-directing artificial intelligence system might accomplish. Physical containment does not equate to locking the weapon system in a box. Instead, it relates to suppression of certain capabilities in preventing its interaction with the external world other than by specific and restricted input and output channels.[253] This clearly contradicts the fundamentals of machine learning as well as AWS' practical relevance. Other restraints must therefore be investigated.

It might instead be theoretically possible to incorporate strongly specific reasoning into the weapon's sequences not to engage in harmful behaviour (*incentive* method). Again, this second model fails on a practical level. Human values regularly contain contradictions that may not be mutually reinforcing; many people enjoy eating meat but cannot imagine killing animals from whence it comes. Implementing any such control structure across a weapon's AI will be intrinsically

[246] Mor Vered and Gal Kahminka, 'Online recognition of navigational goals through goal mirroring', *Proceedings of 16th Conference of Autonomous Agents and Multi-agent Systems*, Extended Abstract, AA-MAS, (2017), p. 1748 <http://www.ifaamas.org/Proceedings/aamas2017/pdfs/p1748.pdf>.

[247] Fangzhen Lin, 'On Moving Objects in Dynamic Domains', *Association for the Advancement of Artificial Intelligence,* (2012), p. 1 <http://commonsensereasoning.org/2011/papers/Lin.pdf>.

[248] Source: Stanford Encyclopaedia of Philosophy <https://plato.stanford.edu/entries/qualia/> [accessed 23 August 2017].

[249] Norodin Zakaria, 'Thoughts of Qualia in Machines', *rxiv.org*, (2005), pp. 2-4 <http://vixra.org/pdf/1505.0146v1.pdf>.

[250] Suchman and Weber, 'Human-machine autonomies', cit. Bhuta et al (eds), p. 93.

[251] Lise Verdiesen, 'How do we ensure that we remain in control of our autonomous weapons?', *AI Matters*, 3, 3, p. 47 and generally <https://sigai.acm.org/static/aimatters/3-3/AIMatters-3-3-11-Verdiesen.pdf>.

[252] Roman Yampolskiy and Joshua Fox, 'Safety Engineering for Artificial General Intelligence', *Machine International Research Institute*, (2012), p. 1 and pp. 9-10 (*Grand Challenges*') <https://intelligence.org/files/SafetyEngineering.pdf>.<https://intelligence.org/files/SafetyEngineering.pdf>

[253] Yampolskiy and Fox, pp. 7-8.

delicate and prone to error and unforeseen outcomes.[254] The challenge to the 'specification route' comes from determining which rules and values are appropriate and then expressing them in code.[255] Borrowing from Russell, 'everything is vague to a degree you do not realise until you have tried to make it precise'.[256] An adjunct architectural method might be to limit the internal capacities of the weapon platform (*stunting*). In this case, however, Bostrom notes that too much stunting produces a weapon platform that is simply 'another piece of software'.[257] Furthermore, it will not be obvious which information should be rationed, either in AWS' data-gathering phase or subsequent engagement sequence. Indeed, Etzioni and Etzioni correctly point out that expansive situational awareness requires that weapon to have all possible information at its immediate disposal.[258] Given this, action by the Delivery Cohort to reduce either sensor or processing bandwidth must weaken LOAC compliance and oblige review elsewhere for an appropriate model of control.

Capability control methods require architecture that automatically detects and reacts to attempted transgression (termed *tripwires* by Bostrom).[259] An overly sensitive tripwire, however, will interfere with intended operations and reduce weapon certainty while too much latitude might enable poor (and therefore illegal) behaviour. Furthermore, the Cohort's monitoring must be properly powerful in order to scan, for instance, all of the weapon's cognitive processes for deception or other coding vulnerabilities. Weapon restraint is therefore an unexpected architectural challenge in AWS deployment.[260] As pointed out by Boddens Hosang, containment strategies involving incentives, curbs or tripwires must still adhere to adopted rules of engagement.[261] A conundrum then arises from striking balance between weapon *control* and weapon *functionality*, a recurrent thread to the challenges facing the removal of supervision across autonomous machines.[262] Containment strategies, moreover, may generally encourage what proves to be a false sense of security in the unsupervised weapon, especially in times of battlefield stress.[263] An adjunct method, therefore, might be *informational* containment which would aim to

[254] Verdiesen, 'How do we ensure that we remain in control of our autonomous weapons?', pp. 49-50.

[255] See: Chapter 8 (*Software*), specifically: 8.1 ('*Coding methodologies*').

[256] Bertrand Russell, *The Philosophy of Logical Atomism,* The Collected Papers of Bertrand Russell, (Boston: Allen & Urwin, 1986), p. 161

[257] Bostrom, *Superintelligence,* p. 136.

[258] A and O Etzioni, 'Pro and Cons of Autonomous Weapon Systems'*, Military Review*, (May-June 2017), p. 78 <http://www.armyupress.army.mil/Portals/7/military-review/Archives/English/pros-and-cons-of-autonomous-weapons-systems.pdf>.

[259] Bostrom, *SuperIntelligence*, p. 129.

[260] Boulanin and Verbruggen, pp. 72-73.

[261] JFR Boddens Hosang, 'Rules of Engagement; rules on the use of force as linchpin for the international law of military operations'*, UvA-DARE*, (8 February, 2017), '*Military Operational Context*'.

[262] Paulina Hensman and David Masto, 'Impact of Imbalanced Training Data'*,* Degree project, Royal Institute of Technology, Stockholm, (May 2015), p. 5. See also: Norman Cook, 'Correlations between Input and Output Units in Neural Networks'*, Cognitive Science*, 19, (1995), pp. 563-564 and pp. 573-574 <http://onlinelibrary.wiley.com/store/10.1207/s15516709cog1904_4/asset/s15516709cog1904_4.pdf?v=1&t=j92drf8l&s=e020a71c328cea489ebd9adda64fe63338ff9177> [accessed 10 June 2017]. See also: Chapter 10 (*Oversight*), generally.

[263] Ryan O'Hare, 'Armed drones and military robots have 'limitless potential for disaster': Experts fear that we are being lulled into a false sense of security by autonomous machines', *Daily Mail newspaper,* 4 March 2016 <https://www.dailymail.co.uk/sciencetech/article-3476870/Armed-drones-military-robots-limitless-potential-disaster-experts-warn.html> [accessed 24 July 2018].

pre-filter and otherwise control what information is allowed to *exit* from the weapon. Still other capability control tools exist on paper. *Reward* mechanisms might appear unlikely but rely, for instance, upon diffuse 'social' services[264] to reward (as well as mechanisms that penalize) the weapon's AI. These models currently remain untested as do associated mechanisms that require machine validation from its Delivery Cohort which would then incentivise, notes Grossman and Hart, the weapon to act in the interests of that principal.[265] The boundaries of weapon interaction and the level of priority given to such intervention (in order not to adulterate the weapon's wider independent operation) will also require definition, implementation and monitoring. The law of unintended consequences lurks: An unexpected outcome here might be the AWS taking on disproportionate risk in exchange for a small chance of increasing its sphere of influence. It can, for instance, be inferred from Bostrom that it would be 'expensive' to offer a weapons-directing AI any higher-than-expected utility as a reward for cooperation than the weapon could itself hope to achieve by pursuing a nefarious end.[266] Restraint thus becomes a trial-and-error matter of juggling confidence levels and weightings and, as such, is inappropriate for a model that must work first time and every time in order for the Cohort to remain compliant with LOAC.[267] A final architectural alternative exists in order to ensure weapon control. Specifying a process for *arriving* at an appropriate battlefield standard rather than specifying the standard itself might in theory achieve appropriate control over the independent weapon. Under such indirect normativity, the AWS might then be motivated to carry out this process and adopt whatever standard the process has imputed.[268] It does, however, nothing to dilute the complexity and variability of specifying and then overseeing such systems. Russell's concern is how corruption can be prevented from spreading throughout AWS action routines as that weapon's overall cognitive capabilities theoretically mature with battlefield experience?[269] Any enhancement of a system's overall cognition (whether through learning routines, from external updates or from other manipulation) is likely to affect the AWS' motivation in ways that are impossible either to predict or detect.

Architectural flux is therefore a key challenge to AWS deployment and one that is compounded by the CACE phenomenon whereby 'changing anything changes everything'.[270] CACE renders prior training datasets immediately obsolete, it may require abrupt reset of the AWS and is a source of machine instability and operational randomness. Consequences arising from such flux are exacerbated by deployment experience being unique to each unsupervised weapon. One AWS will

[264] See: Chapter 8, specifically: 8.3 ('*Utility Function'*).

[265] Sanford Grossman and Oliver Hart, 'An Analysis of the Principal-Agent Problem', *Econometrica*, 51, 1, (January 1983), p. 10 <http://brousseau.info/pdf/cours/grossman_hart_83.pdf>.

[266] Bostrom, *Superintelligence*, p. 132.

[267] Given possible incapacity in its principal (and this need only be implied), subsequent disagreement about the AWS' performance or any *change* in its 'agreed' measurement regime might lead the weapon's learning mechanism no longer to *trust* that principal to deliver its promised rewards. Finally to this point, neither the outcomes produced by the AI nor the end-state of those outcomes may be obvious to the battlefield commander or anyone in the Delivery Cohort.

[268] Nick Bostrom, 'Hail Mary, Value Porosity, and Utility Diversification', *nickbostrom.com*, (19 December 2014), p. 4 <https://nickbostrom.com/papers/porosity.pdf>.

[269] Stuart Russell and others, 'Research Priorities for Robust and Beneficial Artificial Intelligence', *Association for the Advancement of Artificial Intelligence*, (Winter 2015), pp. 105-16 <https://ocs.aaai.org/ojs/index.php/aimagazine/article/download/2577/2521> [accessed 8 June 2017].

[270] See Chapter 7 (Firmware), specifically: 7.1 ('*Sources of technical* debt') and 7.2 ('*Firmware ramifications of machine* learning'). Also: Slav Ivanov, '37 Reasons your Neural Network is not Working', *M Slav blog*, (25 July 2017) <https://blog.slavv.com/37-reasons-why-your-neural-network-is-not-working-4020854bd607>) [accessed 2 October 2018].

be materially different from a second 'colleague' weapon within moments of first deployment. This creates immediate heterogeneity (both within and without autonomous weapon categories) reducing perhaps the attraction AWS to the Delivery Cohort unless those weapons' tasking is so tight to make any such machine sovereignty almost meaningless. It is therefore unclear whether an architectural basis is empirically available for the Cohort that is both feasible *and* appropriate for unsupervised weapons. For this reason, this thesis now considers whether firmware and software can backfill for what otherwise will remain fundamental challenges.

7. Firmware: Embedded process challenges to AWS function

An overarching research question for this thesis is to consider whether an unsupervised weapon system can *feasibly* be reliably compliant and reliably dependable. Firmware is defined as the permanent software already loaded when the weapon's control systems start up.[1] For this purpose, it is considered together with the weapon's middleware, the software that acts as a bridge between the weapon's operating system (or databases) and its several applications, the subject of the following chapter.[2] While middleware is specific to an operating system, firmware is not and together they are used here as a proxy for AWS' architectural capability.[3] It is the combination of firmware and middleware that provides the weapon's *connectivity* enabling multiple processes to run on one or more independent systems and allowing, in theory, interaction across networks. This chapter's purpose is therefore to identify challenges arising from AWS architecture, its firmware and middleware upon which functions the weapon's application software. As noted by Triska, AWS firmware will define its solution processes regardless of weapon tasking.[4] The challenge arises from the tailoring of those processes (and thus weapon firm and middleware) to the capabilities specified by the Design Cohort and, as evidenced from the computer games industry, the difficulties that such subsequent aggregation entails.[5] The chapter is divided into four sections that consider AWS' sources of fragility, specific performance and behaviour consequences arising from the weapon's machine learning (ML) spine and, finally, ramifications arising from how it is expected that AWS will reason, understand and direct its attentions. In this vein, Bostrom highlights engineer practices that are based on remote code libraries which, he suggests, will further compromise AWS' design.[6]

7.1 Sources of technical debt

This chapter's enquiry into AWS feasibility starts by considering how firmware might escalate system fragility. The concept of *technical debt* was first put forward by Cunningham in 1992 in order to quantify costs arising between speed of execution and quality of engineering. Technical debt is therefore a metaphor that links the consequences of poor software design to accumulating a

[1] Jeff Sieracki, 'Machine Learning for Embedded Software is not as hard as you may think', *Reality AI*, (3 August 2016), paras. 6-7 of 12 <https://www.reality.ai/single-post/2016/08/03/5-tips-for-embedded-machine-learning> [accessed 10 October 2017]. Firmware can usefully be understood as that permanent software that is programme into the weapon's read-only memory.

[2] See: Chapter 8 (*Software*), specifically: 8.4 ('*Software processing functions*') and 8.7 ('*Action selection issues*').

[3] Middleware is therefore the software that then mediates between the AWS' assorted application software. Middleware comprises the layer that resides between hardware and application in order to provide critical background services.

[4] Ricardo Triska, 'Artificial Intelligence, classification theory and the uncertainty reduction process', *Federal University of Santa Catherina, Brazil*, (2013), pp. 479-480 <http://www.iskoiberico.org/wp-content/uploads/2014/09/479-486_Triska.pdf>. See also: Kristinn Thorisson and others, 'Why artificial intelligence needs a task theory. And what it might look like', *9th International Conference on AGI*, (2016), pp. 2-3 <http://people.idsia.ch/~steunebrink/Publications/AGI16_task_theory.pdf>.

[5] Cameron Browne and others, *Towards the Adaptive Generation of Bespoke Game Content,* (USA: John Wiley Publishing, 2012), pp. 3-5 <http://ccg.doc.gold.ac.uk/wp-content/uploads/2016/10/browne_ieeechapter14-2.pdf>.

[6] Bostrom, *Superintelligence*, p. 152. See also: Tjorisson, p. 5. and Gerhard Weiss, 'Learning to coordinate actions in multi-agent systems', *Munich Proceedings if the International Joint Conference on Artificial Intelligence*, (1993), p. 311 <http://www.ijcai.org/Proceedings/93-1/Papers/044.pdf>.

'financial debt'.[7] It will be an assertion of this chapter that such 'debt' is particularly relevant to AWS deployment; as a loan must eventually be repaid, with compounding interest, so too hasty design decisions in the removal of human supervision must be paid for with poor weapon reliability, re-factoring, debugging, fragility and complicated testing.[8] In reviewing AWS firmware, an aim is thus to review the premise that machine learning has all of the basic code complexity issues that may be found in normal programming but also 'a larger system-level complexity that can create hidden debt'.[9]

Causes of technical debt will likely include (a combination of) inappropriate AWS architecture[10], shortcuts arising from commercial pressures, a lack of appropriate testing protocols or appropriate whole-system understanding and, as noted by Letovzey and Whelan, lack of ownership, poor technical leadership and pervasive specification changes.[11] Perry also notes that debt arises from 'counterparty development' whereby isolated software routines, once developed, must eventually be merged into a single source base.[12] Finally to this point, Richards highlights the effect of parallelism[13] that will likely exist both within and between a weapon's software releases. Parallelism prevents tidy separations of software into independent work units and decreases both overall awareness and distributed knowledge in the finished product. Scale, furthermore, compounds technical debt arising by increasing both interactions and interdependencies among developers and the Delivery Cohort. Technical debt occurs as projects evolve. In this case, the practice of delayed 'refactoring' of code occurs when specific routines that have become unwieldy must then be largely reworked. Other sources of debt in AWS design will include managing the weapon's configuration and broader integration, resolving semantic conflicts, determining 'logical completeness' and, as inferred from Perry, evaluating iterations to establish the product's unity and fitness for purpose.[14]

The issue here is how the weapon's firmware may add to such debt. First, AWS learning models may subtly erode the weapon's abstraction boundaries.[15] Abstraction is the reduction of AWS processes to their set of essential characteristics. As noted by Hearn, strict abstraction boundaries

---

[7] For a useful primer of technical debt, see: Philippe Kruchtren and others, 'Technical Debt: From Metaphor to Theory and Practice', *IEEE Software*, University of British Columbia, (2012), pp. 18-19 <https://www.computer.org/csdl/mags/so/2012/06/mso2012060018.pdf>.

[8] Zachary Chase Lipton, The high cost of maintaining machine-learning systems', *KD Nuggets*, (January 2015), generally <www.kdnuggests.com/2015/01/high-cost-of-maintaining-machine-learning-technical-debt.html> [accessed 5 March 2017].

[9] Sculley and others, 'Hidden Technical Debt in Machine Learning Systems', *Advances in Neural Information Systems*, (2015), p. 1 <http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf>.

[10] See previous chapter, specifically: 6.2 ('*Architectural approaches*').

[11] Jean-Louis Letovzey and Declan Whelan, 'Introduction to the Technical Debt Concept', *Agile Alliance*, undated, pp. 6-7 <https://www.agilealliance.org/wp-content/uploads/2016/05/IntroductiontotheTechnicalDebtConcept-V-02.pdf>.

[12] Dewayne Perry and others, 'Parallel changes in large-scale software development: An observational case study', *International Conference on Software Engineering*, ICSE98, (2001), p. 309 <https://pdfs.semanticscholar.org/933a/98846a0adc29f2bf8f6557c9c15956562a07.pdf>.

[13] Andrew Richards, 'The Challenges of Delivering Massive Parallelism to Real-World Software', *Codeplay presentation, UKMAC*, (May 2016) <http://conferences.inf.ed.ac.uk/UKMAC2016/slides/Andrew_Richards_The_Challenges_of_Delivering_Massive_Parallelism_to_Real-World_Software.pdf>.

[14] Perry, pp. 311-314.

[15] D Sculley and others, *'Hidden Technical Debt in Machine Learning Systems'*, pp. 2-3.

help define software variation, facilitating later maintenance in the field and simplifying isolated changes and improvements.[16] Abstraction processes are not straightforward and, as inferred separately from Sculley and Dickson, will be particular problematic in AWS deployment.[17] In higher-level models, the abstractions used will often have no relation to how human brains work and will appear illogical. A weapon's low-level models[18] may conversely lack abstraction rules notwithstanding the challenges of memory management, data translation and, notes Zhang, 'parameter tuning' and 'code reuse'.[19] Technical debt in machine-learning models, after all, will be rooted in AWS' reliance on both external and divergent information. Indeed, notes Sculley, 'desired behaviour cannot be effectively implemented in software logic without dependency on external data'.[20] The firmware significance of such tight coupling (between weapon algorithm and remotely retrieved digital information) concerns the degree to which small changes in the AWS' *external* data then alters the way that the algorithm (and weapon) behaves.[21] Any consequences will then be compounded where data acquisition, processing and tuning is managed by different firmware in quite different systems across the weapon. This is not an obvious process and is, note Sun and Giles, one that must be preceded by resolving task appropriateness, data comparison (in multiple domains) as well as the mediation of weapon assumptions where prior knowledge exists.[22]

Frost highlights the computational challenge of ML running hierarchical structures that are based either on long-term priors or on loose dependencies that are tangential and therefore hard to code.[23] The consequence for AWS deployment is that the weapon must then make correlations that are flat (here, non-hierarchical) and represented as simple unstructured lists where every correlated feature on that list is given an inappropriately equal footing which then requires complex proxies (such as the allocating of weightings based only upon the sequence or positioning of relevant data strings).[24] Firmware's reliance, furthermore, upon *labelled* examples leads, notes Marcus, to systemic inefficiency: The choice for the Cohort must either be to incorporate an almost unlimited number of feature detectors on the weapon's grid (that must themselves grow exponentially) or to increase the size of the weapon's training sets in a similarly exponential

[16] Robert Hearn, 'Building Grounded Abstractions for Artificial Intelligence Programming', *MIT Press*, (June 2001), pp. 7-8 <https://groups.csail.mit.edu/mac/users/bob/grounded-abstractions.pdf>.

[17] D Sculley and others, pp. 2-4 ('*Entanglement, correction cascades, undeclared consumers, unstable and underutilised data dependencies and feedback issues*'). See also: Ben Dickson, *'The Limits and Challenges of Deep Learning'*, *TechTalk*, (27 February 2018), paras. 8-9 <https://bdtechtalks.com/2018/02/27/limits-challenges-deep-learning-gary-marcus/> [accessed 6 September 2018].

[18] Rodney Brooks, 'Artificial Intelligence without Representation', *Artificial Intelligence*, MIT AI Labs, Elsevier, 47, (1991), pp. 139-144 <http://www2.denizyuret.com/ref/brooks/brooks.pdf>.

[19] Ce Zhang and others, 'An Overreaction to Broken Machine Learning Abstraction' *HILDA 2017*, Chicago, (14 May 2017) <http://pages.cs.wisc.edu/~wentaowu/papers/hilda17-easeml.pdf>

[20] Sculley and others, p. 2.

[21] The software ramifications of coupling is discussed in Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues*').

[22] Ron Sun and C Lee Giles, 'Sequence Learning: From recognition and prediction to sequential decision-making', *IEEE, Intelligent Systems*, (2001), pp. 4-5 <http://www.sts.rpi.edu/~rsun/sun.expert01.pdf>.

[23] Christopher Frost and others, 'Generalized File System Dependencies', *SOSP*, (14 October 2007), p. 1 and p. 3 <http://featherstitch.cs.ucla.edu/publications/featherstitch-sosp07.pdf>.

[24] Gary Marcus, 'Deep Learning: A Critical Appraisal', *New York University, arXiv*: 1801:00631, (2 January 2018), pp. 9-10 <https://arxiv.org/pdf/1801.00631.pdf>.

manner.[25] The firmware conundrum is that the weapon gets caught in what Marcus terms 'local minima' whereby its system gets stuck on suboptimal solutions with 'no better solution appearing nearby in the space that the weapon's firmware is searching'.[26]

The weapon's firmware is essentially a tool for mixing together and deriving actions from a bank of data sources. Vincent, however, recasts this observation to suggest that AWS firmware is a mechanism for creating *entanglement* whereby the isolation of individual improvements within a self-learning weapon actually becomes impossible as none of the AWS' inputs are properly independent.[27] As above, Sculley terms this trait the CACE principle, whereby *Changing Anything Changes Everything.*[28] In AWS, this relationship will likely be more insidious: While it may be possible for the Delivery Cohort to predict system variation arising from intended developments in code, it can be inferred from Dietterich that deviations in weapon behaviour will likely arise from a broad universe of change agents such as variances in regularization strength, in learning settings, in sampling methods and convergence thresholds.[29] In particular, it is the weapon's prediction sequences that may have nuanced effects on action selection. Here, the weapon's firmware must ensure that output from its learning routines is made accessible to its *other* internal subsystems, either during runtime or by its writing to logs that may later be accessed by those other weapon systems.[30] This is a complex matter that will have hidden consequences. Subsidiary weapon subsystems then become 'undeclared consumers', consuming the *output* of a particular prediction as *input* to another component of that overall sequence.[31] Given its modular composition and limited computational resource, AWS firmware will likely recycle and repurpose exactly these input signals from its sensor banks. As noted by UNIDIR, unintended feedback loops then form between weapon algorithms and the weapon's external world.[32] Such loops may be analogous to filter bubbles in social networks and web search whereby noise suppression mechanisms inadvertently suppress nonconforming data. This feature will also contribute to AWS firmware being treated as a black box, resulting in considerable 'glue code'[33] or, worse still, calibration layers that can lock in assumptions. These ramifications are discussed later in this chapter.[34]

[25] Ibid., pp. 6-7.

[26] Ibid., p. 5.

[27] James Vincent, 'These are Three of the Biggest Problems Facing artificial Intelligence', *The Verge*, 10 October 2016 <https://www.theverge.com/2016/10/10/13224930/ai-deep-learning-limitations-drawbacks> [accessed 6 April 2017], generally.

[28] Sculley and others, pp. 2-5.

[29] Thomas Dietterich, 'Learning and Reasoning', *Department of Computer Science, Oregon State University*, (May 2003), p. 4 <http://web.engr.oregonstate.edu/~tgd/publications/mlsd-ssspr.pdf>.

[30] This issue is sometimes referred to as *visibility debt*. For discussion on action selection, see: Chapter 8 (*Software*), specifically: 8.7 ('*Action selection issues*').

[31] Sculley and others, p. 3.

[32] UNIDIR, 'Safety, Unintentional Risk and Accidents in the Weaponization of Increasingly Autonomous Technologies', *UNIDIR*, 5, (2016), p. 6 and pp. 7-9 <http://www.unidir.org/files/publications/pdfs/safety-unintentional-risk-and-accidents-en-668.pdf>.

[33]Azamat Shakhimardanov and others, 'Best Practice in Robotics', *BRICS Collaborative*, (February 2013), pp. 82-88 ('*Interface* technologies') <http://www.best-of-robotics.org/pages/publications/BRICS_Deliverable_D2.1.pdf>.

[34] See next section: 7.2 ('*Firmware ramifications of machine learning*').

This phenomenon of the 'undeclared consumer' creates additional technical debt with costs arising from the tight coupling of a particular prediction model to other parts of the weapon's sensor and processing stack. Changes within those routines then impact other system processes, usually in ways that are unintended and certainly poorly understood. Empirically, the phenomenon actually makes it problematic to make *any* changes to the weapon's firmware (here, the underlying model of the AWS' platform). Pannu and Moore point also to the secondary costs of these feedback loops that must be created to adjust for such variations.[35] A battlefield example might be a weapon subsystem predicting area-incursion where a sub-routine is then tasked with determining entity-size in possible transgressors. If this entity-size module then starts consuming area-incursion as a discrete input signal and entity-size has an effect on the weapon's propensity to enter the area, then the inclusion of area-incursion in entity-area adds an unwelcome hidden feedback loop. It is easy to imagine a scenario where the AWS uncontrollably recognises increasingly small entity-size transgressors (from an armoured personnel carrier to a rabbit to a flea) in its output generation.

Dependencies comprise a material faultline in AWS firmware.[36] While code dependencies may conceivably be spotted via static analysis, data dependency in the AWS is more complicated to identify and difficult subsequently to untangle. Why is this an issue? Input signals may themselves be unstable, meaning that they qualitatively change behaviour over time.[37] Rolling out 'improvements' to that particular input signal will, however, likely have quite arbitrary effects on weapon performance. A mitigation strategy might then be to create versioned copies of given signals within the weapon's logs but this, notes Sculley, will create inappropriate complexity (through the redundant multiple versions of data) as well as potential staleness within AWS data feeds.[38] A second firmware ramification arises from underutilized data dependencies, mostly unneeded routines within the weapon's programme suites that provide little additional accuracy. Examples might include legacy features, bundled features or additional refinement routines. Each such feature, however, adds to AWS brittleness as the weapon's firmware will assign them some weight notwithstanding their redundancy. Indeed, Sculley highlights that 'the overall system is therefore vulnerable, sometimes catastrophically so, to changes in these unnecessary features'.[39] Subsequent removal of these routines (for instance, a weapon's dead experimental code-paths) may, moreover, be a further source of system error. Vardi points here to the general difficulty of performing third-party static analysis on data dependencies, complicating the tracking of data throughout the weapon system and action sequences.[40] In the interests of efficiency, it may even be

---

[35] Adarsh Pannu and Steve Moore, 'Three Reasons Machine Learning Models Go Out Of Sync*', Inside Machine Learning*, (27 November 2017) <https://medium.com/inside-machine-learning/three-reasons-machine-learning-models-go-out-of-sync-a101b2cdca54> [accessed 12 September 2018].

[36] J David Morgenthaler and others, 'Searching for build debt: experiences managing technical debt at Google'*, Proceedings of the 3rd International Workshop on Managing Technical Debt*, (2012), p. 1 <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/37755.pdf>.

[37] This can happen *implicitly* (a signal input from a separate model updating overtime or from a data-dependent lookup table) or *explicitly* (the input signal is imported or otherwise separate from the weapon system).

[38] YB Kim and Karl Stratos, 'Adversarial Adaption of synthetic or Stale Data', cit. *Proceedings of 55th AGM, Association of Computational Linguistics*, Microsoft AI & Research, (2017), pp. 1297-1298 <http://www.karlstratos.com/publications/acl17adversarial.pdf>.

[39] Sculley and others, p. 4.

[40] Moshe Vardi and others, 'The Implication Problem for Data Dependencies'*, Hebrew University of Jerusalem*, (January 2006), pp. 73-85 <https://www.researchgate.net/publication/226509257_The_implication_problem_for_data_dependencies> [accessed 13 July 2017].

decided to stop computing a particular non-core signal within the weapon's automatic targeting routines. Even if there are no references to that routine in the current version of the weapon's code, there may still be instances with older binaries referencing that input and leading again to unintended outcomes.

A further source of firmware's technical debt is identified by Goeree and arises from the phenomenon of 'correction cascade'.[41] Such cascades typically occur when machine learning models do not learn as predicted requiring an external fix on that model's output. Zavershynskyi notes that 'as the hot fixes pile up you end up with a thick layer of heuristics on the ML model'.[42] In short decision cycles, weapon learning will be reduced as its routines infer nothing new from the weapon's subsequent observations. Seemingly minute variations in a training dataset might justify re-running the earlier training but this time with a small learning correction, the incentive being a fast, low-cost benefit given that the correction should be immaterial to weapon performance. The issue, however, is that any such correction is likely to create its own dependencies within the weapon's initial dataset making it challenging then to attribute on-going improvement in weapon performance. Bella notes that this will be aggravated as correction is applied to 'closely related' (rather than precisely delineated) learning such as calibrating outputs to slightly different test distributions.[43] Within cascading, a further challenge is also for the weapon to factor both the total but also the distribution of this error. Cascading can then produce deadlock whereby the local optimum for that learning system becomes circular and iterative with the result that the component's routine cannot then be improved.[44] While humans can learn relationships from a very a few number of trials, DeepMinds' work on board games and Atari involved billions of such examples.[45] The point for AWS is that such training does not equate to understanding. As noted again by Marcus, '[Atari] has just learnt specific contingencies for particular scenarios'.[46] Similarly, transfer tests (the generalizing of conditions that are different from those encountered during training) demonstrate that machine learning outputs are often extremely superficial.[47] This is unsurprising given the difficulty of evolving machine learning from interpolation (effecting generalization between known examples) to extrapolation (the requirement to advance beyond the space of known training examples).[48]

---

[41] Jacob Goeree, 'Self-correcting Information Cascades', *Review of Economic Studies*, 74.3, (2007), p. 733 <http://pages.wustl.edu/files/pages/imce/brogers/casexp.pdf>.

[42] Maksym Zavershynskyi, 'Technical Debt in Machine Learning', *Towards Data Science*, (1 July 2017) <https://towardsdatascience.com/technical-debt-in-machine-learning-8b0fae938657> [accessed 3 November 2017].

[43] Antonio Bella and others, 'Calibrating of Machine Learning Models', *University of Valencia*, undated, pp. 1-3 <http://users.dsic.upv.es/~flip/papers/BFHRHandbook2010.pdf>.

[44] Data Science blog, 'Neural Networks getting stuck at Local Optima', September 2014 <https://datascience.stackexchange.com/questions/2362/neural-networks-getting-stuck-at-local-optima> [accessed 12 May 2017]. See also: Akarachai Atakulreka and others, 'Avoiding Local Minima in Feed-forward Neural Networks by Simultaneous Learning', *Springer Publications*, Berlin, (2007) <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.2375&rep=rep1&type=pdf>.

[45] Aman Agarwal, 'Explained Simply: How DeepMinds Taught AI to Play Video Games', *Medium.org blog*, (27 August 2017) <https://medium.freecodecamp.org/explained-simply-how-deepmind-taught-ai-to-play-video-games-9eb5f38c89ee> [accessed 12 October 2017].

[46] Marcus, p. 8.

[47] Sebastian Ruder, 'Transfer Learning-Machine Learning's New Frontier', *Ruder blog*, (21 March 2017) <http://ruder.io/transfer-learning/> [accessed 28 July 2017].

[48] Marcus, p. 6.

AWS' firmware will create other sources of technical debt. A further difficulty, notes Marcus, is to pollinate deep learning models with prior knowledge.[49] These models can neither distinguish causation from correlation nor deal with open-ended inference (the difference, say, between 'John promised Mary to leave' and 'John promised to leave Mary').[50] AWS firmware presumes 'a large, stable world' and, as inferred from Marcus, even if ML approximation shows promise, the Delivery Cohort should not generally trust ML's output given that such learning is 'a statistical technique and all statistical techniques suffer from deviation in their assumptions'.[51] It is difficult to guarantee that these weapons will work in alternative circumstances with novel data that, after all, is unlikely to resemble previous data. Bottou compares this challenge to the development of aeroplane engines where design relies upon building complex systems out of simple subsystems; while it may be possible to create secondary guarantees around subsequent performance, the passing down such performance guarantees is not an appropriate path for ML-based AWS.[52]

O'Rourke notes a separate challenge that arises from the procurement of complex systems from disparate commercial parties where the bundling together of several proprietary routines results in a system design that is held together by 'glue' (or 'spaghetti') code.[53] Glue code relates to the quantity of supporting code that must be written to permit data transfer in and out of these specific software packages.[54] This phenomenon can materially contribute to system fragility[55], not least because glue code tends to anchor a system to the peculiarities of those proprietary packages being glued.[56]. Glue code has other ramifications for AWS control and operation. Fundamentally, the feature will entrench the weapon's original construction in 'supporting code' instead of embedding it directly into components that have been designed for specific weapon routines. Such setup discourages experimentation. The magnitude of the problem is illustrated by Sculley's research that mature systems empirically end up being five per cent machine learning code and ninety-five per cent glue code.[57] An adjunct challenge arises from the occurrence in weapon firmware of 'pipeline jungles', a characteristic which Foreman notes often typifies the move from prototype to production manufacture.[58] Pipeline complexity is likely to develop organically as weapon's signals are adjusted and new information routines and sensor inputs are added. Glue code and pipeline

[49] Ibid., pp. 10-11.

[50] Ibid., pp. 11-12.

[51] Ibid., p. 15.

[52] Leon Bottou, 'Two Big Challenges in Machine Learning', *ICML 2015*, presentation papers, (2015), generally <https://icml.cc/2015/invited/LeonBottouICML2015.pdf>.

[53] Ronald O'Rourke and others, 'Multi-year procurement and block buy contracting in Defence Acquisition', *Congressional Research Service*, (8 August 2017), p. 7 <https://fas.org/sgp/crs/natsec/R41909.pdf>.

[54] Neuromorphic Technologies, '"Spaghetti Code": Complexity and Artificial Intelligence', *Admin blog*, (27 March 2018) <http://fernandojimenezmotte.com/mi-articulo/spaghetti-code-complexity-and-artificial-intelligence/> [accessed 8 September 2017].

[55] See: Natali Vlatko, 'The Dangers of Spaghetti Code', *JaxCenter blog*, (5 June 2015) <https://jaxenter.com/the-dangers-of-spaghetti-code-117807.html> [accessed 12 September 2017].

[56] Source: Imperial College London, Department of Computing <http://www.doc.ic.ac.uk/~np2/patterns/scripting/glue-code.html> [accessed 12 September 2017].

[57] D Sculley, 'Hidden Technical Debt in Machine Learning Systems', p. 5.

[58] John Foreman, 'The real world of machine learning for fun and profit; pipeline jungles and hidden feedback loops', *Foreman Business Blog*, 1 May 2015 <http://www.john-foreman.com/blog/the-perilous-world-of-machine-learning-for-fun-and-profit-pipeline-jungles-and-hidden-feedback-loops> [accessed 2 September 2017].

jungles may also arise from overly separated research and engineering roles, a frequent characteristic in government procured weapons hardware.[59] While the resulting firmware may be intended to facilitate data preparation, it may instead be a tangle of scrapes, joins and other sampling steps. Such pipelines may themselves have intermediate file output. A reaction to the difficulty may then be to add both amendments and connecting code as conditional branches within the main system source code. Obsolete and experimental code parts can similarly interact with each other as well as with the weapon's primary system in unpredictable ways.[60] Xie and Engler highlight programmers' use of 'dead flags' as a further source of system fragility.[61] It is the idiosyncracy of these signposts that challenges subsequent intervention by the Delivery Cohort. Dead code that may has been moribund for an extended period can, notes Linders, be awakened by an automatic change to that flag's value with significant (but unpredictable) consequences in the weapon.[62]

In comprising a wide portfolio of engineered capabilities, AWS firmware must also provide end-users with a wide range of 'configurable options' including, inter alia, which weapon features are available to the Cohort, how data is selected as well as a broad gamut of learning settings, logging formats, processing routines and verification methods.[63] The whole process of system configuration is a further source of technical debt. A portfolio of weapon capabilities requires accurate third-party alignment and confirmation. System configuration, moreover, is not an obvious process with the number of lines of 'configuration code' far exceeding the number of instruction lines behind the weapon's learning processes.[64] Each such line has the potential for mistakes given that 'configurations are by their nature ephemeral and less well tested'.[65] Unpredictability may arise in several guises. First, weapon systems that are autonomous still require a myriad of decision thresholds and tradeoff boundaries that must be mediated by humans in order that they can appropriately mirror the Cohort's aims. As noted by Kwang, such thresholds must be manually set as no proven alternative method of intervention is currently available.[66] AWS configuration will also be complex. If a weapon updates on new data, the old manually-set thresholds may be invalid. Updating so many thresholds across different models (each, moreover, in a different learning state) will be time-consuming and brittle. Further fragility then arises when relationships in battlefield data that have been assumed from the design outset no longer exist on the ground in which case the

[59] Inferred from: AIP, 'Defense Department Reorganization Aims To Foster "Culture of Innovation"', *American Institute of Physics*, (10 August 2017) <https://www.aip.org/fyi/2017/defense-department-reorganization-aims-foster-'culture-innovation'> [accessed 13 September 2017].

[60] Sculley highlights the loss by trading shop Knight Capital of $465m in 45 minutes that was attributed to unexpected behaviour from obsolete experimental code paths.

[61] Yichen Xie and Dawson Engler, 'Using Redundancies to find Errors', *IEEE Transactions on Software Engineering*, (2003), p. 1 <https://web.stanford.edu/~engler/tse-redundant.pdf>.

[62] Ben Linders, 'Dead code must be removed', *Info Q*, (9 February 2017) <https://www.infoq.com/news/2017/02/dead-code accessed> [23 September 2017].

[63] See, generally: Holger Hoos and others (eds.), 'Automated Algorithm Selection and Configuration', *Report from Dagstuhl Seminar*, 16412, (2017) <http://drops.dagstuhl.de/opus/volltexte/2017/6956/pdf/dagrep_v006_i010_p033_s16412.pdf>.

[64] See generally: Alejandaro Zacarias and others, 'A Framework to Guide the Selection and Configuration of Machine-Learning-Based Data Analytics Solutions in Manufacturing', *Proceedings CIRP*, 72, (2018), pp. 153-158.

[65] Sculley and others, p. 7.

[66] Kevin Kwang, 'Machine Learning needs Human Helping Hand', *ZDNedt,* (3 April 2014), paras. 8-13 of 13 <http://www.zdnet.com/article/machine-learning-needs-human-helping-hand/> [accessed 27 July 2017].

weapon's prediction and behaviour will perform unpredictably.[67] Monitoring and addressing variation in the associations assumed in AWS deployment models must also occur in real-time across the whole of the weapon system in order for that weapon to be both compliant and valuable. The challenge, however, is *what* metrics the AWS should monitor given that a purpose of its machine-learning is, after all, to adapt overtime. In this way, Osoba and Welser note that it is inappropriate to base weapon monitoring on prediction biases that simply forecast average values of label occurrences without regard to input features.[68] Firmware interventions must therefore be administered manually, they must be fit to each combat scenario and they may require their own feedback loop. Finally to this point, legal, social and political constraints will agitate that such limits are set *conservatively* which, should the action limit unexpectedly trigger, might compromise that weapon's operational usefulness to the Delivery Cohort.[69]

7.2 Firmware ramifications of learning methodologies

The work of Rosenberg and Markoff evidences that ML is pivotal to AWS architecture.[70] The aim of this section is to develop this thesis' analysis of structural challenges that are inherent in the application of ML to unsupervised weapons and assumes, therefore, that AWS architecture will comprise such a learning framework (be it a neural network, enhanced logic or other statistical framework).[71] The aim is to provide a primer in order to assess frailties across learning models notwithstanding that several variants are currently posited for AWS operation. The context for this section is provided by Potember's 2017 study of ML (as it relates to the US DoD) and his conclusion that 'the manifolds whose shape and extent that machine learning is attempting to approximate are almost unnoticeably intricate, leading to failure modes from which there is very little human intuition and even less established engineering practice'.[72]

In this vein, weapon learning likely breaks down into a set of distinct firmware types.[73] *Supervised* learning, as detailed later in this section, attempts to predict an output when given an input; an example might be multiple data points from an AWS' portfolio of sensors being processed

[67] Causal links, for instance, between observed target traits and triggering lethal engagement.

[68] Osonde Osoba and William Welser, 'An Intelligence in our Image: The Risks of Bias and Errors in Artificial Intelligence', Rand, (2017), pp. 7-9 <https://www.rand.org/content/dam/rand/pubs/research_reports/RR1700/RR1744/RAND_RR1744.pdf>.

[69] See, generally: Chapter 5 (*Obstacles*), specifically: 5.6 ('*Behavioural constraints'*) and 5.8 ('*Ethical and accountability constraints'*).

[70] M Rosenberg and J Markoff, 'The Pentagon's 'Terminator Conundrum': Robots that could kill on their own'*, New York Times*, 25 October 2016, generally <https://www.nytimes.com/2016/10/26/us/pentagon-artificial-intelligence-terminator.html> [accessed 3 August 2017].

[71] Economist Magazine Special Report, 'Artificial intelligence', *Economist,* p. 5. For a general discussion on abstraction in machine learning, see: N Bredeche and others, 'Perceptual Learning and Abstraction in Machine Learning: An Application to Autonomous Robots', *IEEE Transactions on Systems, Man and Cybernetics: Part C*, 36, 2, (2006) <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.3559&rep=rep1&type=pdf>. For a commercial overview of this field, see generally: Katrina Wakefield, 'A Guide to Machine Learning Algorithms and their Application', *SAS.com*, undated <https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html> [accessed 30 September 2017].

[72] Richard Potember, 'Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD', *Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to the DoD*, The Mitre Corporation/McLean, (2017), JRR-16-Task-003, p. 2 <https://fas.org/irp/agency/dod/jason/ai-dod.pdf>.

[73] For an introductory overview, see: Anish Talwar and Yogesh Kumar, 'Machine Learning: An Artificial Intelligence Methodology'*, IJECS*, 2, 12, (December 2013), pp. 3400-3402 <http://www.ijecs.in/issue/v2-i12/11%20ijecs.pdf>.

to deliver target selection and engagement. Supervised models may either be regressive (seeking whole numbers and clear outcomes) or classifying (seeking class labels, trends and patterns). *Reinforcement* learning, also discussed below, instead encourages the weapon to select actions that maximize a payoff. Finally, the weapon may learn through an *unsupervised* model whereby the machine will be focused to discover general internal representations from its sensor input.[74] A further assumption for this section is that there is no appropriate role for provisional, probationary or other pilot steps in delivering lethal violence. Indeed, the term 'exploration' appended to reinforcement processes points to a central challenge of learning models: Just as an independent weapon cannot be sure it has found its *best* action for each state until it has tried all possible actions in all states, learning models are too iterative and ambiguous to manage target selection without an appropriately definable end-state.[75] Any error in the weapon's sensing of its current state will, after all, then carry forward in that machine's future learning and future battlefield actions. A further ramification arises should either the AWS' environment or combat task change while the weapon is in a learning phase in which case much of what it has learnt may be invalid. This also poses the challenge of incorporating *un-learning* in the weapon's ML routines should error occur.[76] The trade-off between 'constantly learning' versus employing what is already known to work (at the cost of missing out on further improvement) is a well-discussed conundrum (termed by Cohen the 'exploration/exploitation dilemma').[77] As a theoretical challenge to the process of removing weapon supervision, it creates, however, another intractable set of problems which, suggests Lafrance, is also fraught with agent-principal risk.[78] As noted by Du, 'to what extent such self-learning would be sufficient for AWS to make better decisions is fundamentally unclear'.[79]

This exploration/exploitation dilemma generally informs firmware setup and is the clearest example of the interdependence that will exist between weapon architecture and the software routines that sit on that architecture.[80] Complexity arises from the AWS having to optimize between its *control* policy (the reactive rules that control a weapon for a particular goal) and its use of an appropriate *value* function (the comparative value of being in each state relative to the weapon's overarching goal).[81] This complexity carries over into the AWS' firmware which must facilitate

---

[74] L Busoniu and others, 'Reinforcement Learning and Dynamic Programming using Function Approximation', *Delft Center for Systems and Control, Netherlands*, (November 2009), pp. 2-5 <https://orbi.ulg.ac.be/bitstream/2268/27963/1/book-FA-RL-DP.pdf>.

[75] Michael Horowitz, 'The Promise and Perils of Military Applications of Artificial Intelligence', *Bureau of the Atomic Scientists*, (23 March 2018) <https://thebulletin.org/landing_article/the-promise-and-peril-of-military-applications-of-artificial-intelligence/> [accessed 3 June 2018].

[76] Naveen Joshi, 'Now, What is Machine Unlearning All About?', *Allerin blog*, (12 March 2018) <https://www.allerin.com/blog/now-what-is-machine-unlearning-all-about> [accessed 5 June 2018].

[77] Jonathan Cohen and others, 'Should I stay or should I go? How the Human Brain manages the trade-off between exploitation and exploration', *Royal Society Publishing*, (May 2007) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2430007/> [accessed 12 September 2017].

[78] Adrienne Lafrance, 'Machine Unlearning', *The Atlantic*, 18 March 2016 <https://www.theatlantic.com/technology/archive/2016/03/computers-brains-cybernetics/474273/> [accessed 21 September 2017].

[79] Dr Hongbo Du, School of Computer Science, Buckingham University, in conversation with the author, January 2019.

[80] See: Chapter 8 (*Software*), specifically: 8.3 ('*Utility function;'), 8.5 ('Anchoring and Goal setting issues'*) and 8.6 ('*Value setting issues'*).

[81] For a useful primer, see: Tobias Baer and Vishnu Kamalnath, 'Controlling Machine Learning Algorithms and their Biases', *McKinsey and Company*, (November 2017) <https://www.mckinsey.com/business-functions/risk/our-insights/controlling-machine-learning-algorithms-and-their-biases> [accessed 17 October 2017].

evaluation on how well each such state-action pairing has fared and then absorb the outcome of each resulting trial in running subsequent actions. It is this friction that exists between control tenets (architecture) and value tenets (firmware and software routines) which creates a further conflict for unsupervised weapons. This will be exacerbated in a multi-robot environment (or an environment where several polities come together on a battlefield) where, notes Chopra, identification of *individual* outcomes is particularly difficult and requires unrealistic attribution analysis.[82] The firmware conundrum of allotting credit or blame to feedback-generated actions is termed AWS' 'temporal credit assignment'. In the case of weapons, such assignment will be a particularly intractable task; as autonomous system behaviours change over time (given that weapon's learning patterns), discrepancies will occur between actual system performance and operator expectations, likely leading to system deadlock and certainly preventing useful attribution analysis.[83] It might also lead to inappropriate teammate surprise (and, more important, to subsequent loss of trust) during operations.[84] Soldiers, moreover, do not perceive their environment in terms of a collection of labelled objects. Accordingly, it is not obvious that the Cohort should trust symbolic processing that underpins ML while its battlefield decisions, notes the Army Research Laboratory, must instead be based on the 'incorporation of all opinions and evidence' (here, numeric, contextual and heuristic).[85] The frustration for this section is obviously that the fundamental bases for AWS learning models are evolving very quickly and are not yet set in stone.[86] Progress will likely continue and promising work-rounds will arise. Any analysis must thus be behavioural and subject to the same sense-check that machine output should be based. It is therefore useful to return to machine learning's four primary variants which, for the purposes of assessing their feasibility as a statistical basis for AWS, are now evaluated in order.

First, *supervised* learning may be used to train a weapon with the aid of that same internal labelled set of examples. Here, the weapon must rely upon a control mechanism in order to direct what is being sought. The standard task of supervised learning is to master a decision-making policy for 'one-shot' classification tasks. Under this scenario (and inferred from Dieternich), the weapon will receive an input observation from its sensor portfolio and must then make what will be a *single* decision.[87] In this case, the AWS is undertaking only one evaluation rather than a sequence of evaluations. Using such a database approach, the supervised weapon will work through examples and adjust the weights inside its network to improve accuracy in its tasks. The merit of this approach is that there is no need for a human first to draw up a list of rules or for the Delivery Cohort to implement them through intervention. The weapon system will theoretically learn directly from the labelled data.[88] But supervised training also requires adjunct processes in order to

[82] S Chopra, 'Attribution of Knowledge to Artificial Agents and their Principals'*, International Joint Conference of Artificial Intelligence*, (August 2005), 19, pp. 1175-1176 <http://www.sci.brooklyn.cuny.edu/~schopra/893.pdf>. See also: Chapter 4 (*Deployment*) specifically: '*Human-Machine teaming'.*

[83] Inferred from: Marvin Minsky, 'Steps towards Artificial Intelligence'*, IRE*, (January 1961), pp. 8-9 <http://courses.csail.mit.edu/6.803/pdf/steps.pdf>.

[84] US Department of Defense, 'Summer Study on Autonomy', p. 18. See, also: introduction to Chapter 10 (*Oversight*).

[85] Army Research Laboratory, 'Research Suggests Uncertainty May Be Key to Battlefield Decision Making', *ARL*, (12 July 2018) <https://phys.org/news/2018-07-uncertainty-key-battlefield-decision.html> [accessed 7 September 2018].

[86] The role of context and assumption-building is reviewed in detail in Chapter 2 (specifically: '*Introduction to key concepts'*).

[87] Dietterich, p. 4.

[88] Economist Magazine Special Report, *'Artificial intelligence',* p. 5.

deliver outputs including, notes Preetham, back-propagation algorithms whose purpose is dynamically to adjust the weapon's weightings.[89] This, however, gives rise to possible 'temporal dislocation' as the weapon's filtering routine can only take place having first compared *actual* outputs with *desired* outputs for any given stimulus. Furthermore, Sun points out an unwelcome characteristic of general connectionist learning (here, relating to how humans learn and remember) that such weights are arbitrarily influenced by network connections themselves.[90] A firmware ramification is that considerable variation can remain around how strongly individual modes within the AWS planner might be connected and how this might affect subsequent re-weighting of inputs. The process is unstable and lacks appropriate predictability.

Other learning types highlight the complexity involved. An autonomous weapon relying instead on *unsupervised* learning will, theoretically, train its network by exposing itself to a huge number of examples but without control mechanisms telling its sensor bank what to search. In this model, the idea is that the weapon's network will learn to recognise battlefield features and, suggests Kosko, to cluster similar samples, thus revealing hidden groups, links or patterns within the dataset.[91] Unsupervised learning thus posits a theoretical model that is capable of anomaly search given that the weapon system does not know what it is seeing. But such sequences must still be undertaken within the 'fog of war', of changing conditions, changing priorities and with incomplete knowledge of either friendly or hostile forces and their respective capabilities.[92] Battlefield uncertainty also removes delineation in the weapon's input data features. Without such delineation, the weapon's network must establish (and then rely upon) further probability relationships within what is a dynamic dataset in order to drive meaning from that data, creating additional complexity and unwelcome sources of technical debt. This is a critical observation. The quality of input data into neural network models will, after all, 'strongly influence the results of the data analysis'.[93] A pivotal feature turns out to be the preparation of that input data: Fifty to seventy per cent of time in data analysis is currently taken up with the preparation of that data, a facility that will be unavailable in a battlefield setting if the weapon is still to be valuable.[94] Improperly prepared data, it turns out, can make analysis 'difficult, if not impossible'.[95]

Workarounds exist in an effort to remediate this challenge. The focus, for instance, of *reinforcement* learning, a further possible variant in AWS design, will be to train the weapon's neural network to interact with its environment with only occasional system feedback (and involving therefore less troublesome data handling) which, in this case, will be presented in the

[89] V Preetham, 'Back Propagation – How Neural Networks learn Complex Behaviours'*, Autonomous Agents #AI*, generally <https://medium.com/autonomous-agents/backpropagation-how-neural-networks-learn-complex-behaviors-9572ac161670#.qwz64wcu6> [accessed 4 October 2017].

[90] R Sun, 'Connectionist and Symbolic Approaches'*, University of Missouri-Columbia*, (November 2000), p. 3 <http://www.cogsci.rpi.edu/~rsun/sun.encyc01.pdf>.

[91] Bart Kosko, 'Hidden Patterns in Combines and Adaptive Knowledge Networks'*, Elsevier Science Publishing*, 2, (1988), p. 378 <http://ac.els-cdn.com/0888613X88901119/1-s2.0-0888613X88901119-main.pdf?_tid=245f5116-9b5e-11e7-a4d0-00000aab0f6c&acdnat=1505621609_a7493474593fc9822b5484b3a9589e34> [accessed 3 October 2017].

[92] Zheng, pp. 3-5.

[93] Source: Foreign Exchange Rate forecasting with Artificial Neural Networks, 'Data Preparation in Neural Network Data Analysis'*, International Series on Operations Research Management Science*, pp. 39-62 <https://link.springer.com/chapter/10.1007%2F978-0-387-71720-3_3> [accessed 3 October 2017].

[94] Ibid., p. 39.

[95] Ibid., pp. 39-41.

form of a 'reward system'. The theory is that the weapon's learning algorithm gradually constructs its own internal evaluation function that then assigns values to states, state-action pairs or policies in an engagement sequence.[96] In essence, training will involve adjusting network weights to search for an action strategy that consistently generates higher rewards for the unsupervised weapon.[97] Two shortcomings immediately arise. First, this learning approach underscores the imprecise input/output range that can be expected from such models. A challenge for unsupervised learning is that the weapon's weight-adjusted algorithm may eventually cause similar inputs to have the same output patterns. Any such gradualism is clearly unacceptable for battlefield weapons. As discussed below, all deployment models will require an impractically large set of example inputs as there is no way to forecast output patterns for what is a limitlessly wide (and imprecise) class of battlefield inputs. It should also be inferred from Quinlan that models based on expert decision trees are inappropriate for battlefield deployment as such trees are feed-forward structures.[98] As proposed by Kosko, they have no '*dynamical* behaviour'.[99] Indeed, a decision tree does not represent feedback. A ramification is that search time increases exponentially with tree size; the more rules or branches in an expert system tree, the longer it takes to make an inference or decide upon an action. Real time running is therefore infeasible for large search trees as, Petri notes, models' 'divide and conquer' methods means that they perform poorly when many complex interactions are present and tend towards 'over-sensitivity to training sets, to irrelevant attributes and to noise'.[100] Finally, such trees do not naturally combine.[101] To open a new decision tree will require the weapon to remove an existing edge or subset of an edge in an existing tree. The question arises as to which edge or edges as edge removal is empirically ad hoc.[102] The corollary is that neural networks do not exhibit human cognition[103], their underlying processes are iterative and deliver imprecise outcomes. It can, moreover, be inferred from Kamimura that this will be further exacerbated by the time drag occurring between AWS' sensing of external data and its computation of an executable action[104] requiring additional (and complex) 'stopping' routines in order to sense-check and pause weapon routines.[105]

Hybrid firmware in ML does not address this faultline. In this case, 'transfer learning' looks to build upon previously acquired knowledge rather than the weapon having to be trained from

[96] Bostrom, *Superintelligence*, p. 188.

[97] Economist Magazine Special Report, 'Artificial intelligence', p. 5.

[98] JR Quinlan, 'Induction of Decision Trees', *Kluwer Academic Publishers*, Machine Learning, 1, (1981), pp. 81-85 <http://hunch.net/~coms-4771/quinlan.pdf>.

[99] Kosko, p. 378.

[100] C Petri, 'Decision Trees', *Cluj Napoca*, (2010), p. 8 <http://www.cs.ubbcluj.ro/~gabis/DocDiplome/DT/DecisionTrees.pdf>

[101] Kosko, p. 378.

[102] Bogumil Kaminski and others, 'A Framework for Sensitivity Analysis of Decision Trees', *Central European Journal of Operations Research*, 26, 1, 135-136 <https://link.springer.com/article/10.1007%2Fs10100-017-0479-6> [accessed 18 October 2017].

[103] M Zorzi, 'Self-learning AI emulates the human brain', *European Research Council*, University of Padova, (22 July 2016) <https://erc.europa.eu/projects-figures/stories/self-learning-ai-emulates-human-brain> [accessed 3 August 2017].

[104] R Kamimura, 'Generation of Organised Internal Representations in Recurrent Neural Networks', *Neural Networks*, IJCNN, HJUne (1992), pp. 1-3 <http://ieeexplore.ieee.org/document/287116/> [accessed 2 September 2017].

[105] Lutz Prechelt, 'Early Stopping – But When?', in 'Neural Networks: Tricks of the Trade', *Lecture Notes in Computer Science*, 7700, Springer, Berlin, Heidelberg, undated, p. 53.

scratch each time.[106] 'Multitask learning' promises instead to use the experience of one layer in order to improve efficiencies at another layer. Two further issues arise from such hybrid routes. First, while humans may have numerous representations 'active' at any one time (some originating from our eyes and ears, others evoked from memory or initiated by internal models), a weapon without supervision must seamlessly choose to process *only* relevant associations from the available dataset. Another difficulty is that an AWS' firmware will have to limit the number of active signals such that only relevant associations arise.[107] This model constraint will likely impact unacceptably upon performance unless, perhaps, that platform's tasking is itself suitably narrow and defined.[108] Second, of course, the machine purpose of DNN is to enable the unsupervised weapon to perform sound *reasoning* routines.[109] An adjunct learning basis for machine-learning weapons might therefore be by imitation. Mataric, however, notes that such models, whether driven by imitation or by copying from demonstration, remain technically demanding.[110] Schaal highlights here that it is difficult to separate what is relevant to the task in hand from that which is being taught.[111] As inferred again from Zheng, the enduring challenge for AWS remains the isolation of relevant battlefield data from, for instance, recommendations generated from its internal models or data points that are either divergent or tangential.[112] As part of this balancing, the AWS must then check its *learned* operation parameters for appropriateness in order to ensure conformity and equilibrium with its internal representations. The challenge is then how best to solve for the weapon's design priorities. This is not obvious. To this point, is a weapon's movement *generating* mechanism to be directly employed by or quite separate from its movement *recognition* sequences? An adjunct issue reinforces this point. How might the intention of that movement be recognised within the weapon's ML to be converted into its set of internal goals?[113] To conclude this section, it is useful to emphasise three further aspects of learning. Most such learning is based on what Du terms a 'closed world assumption' whereby one decision out of several predefined possible decisions is made based on received inputs. In real life, however, humans encounter an 'open world' where several scenarios have never before been encountered and where, of course, the trained model is unable to recognise patterns. In AWS design, such an open-world situation should lead to a safe 'default' decision but it is this default that it is so difficult to define. Indeed, the reliability of such training models is systemically questionable when the balance between the number of training examples (rows of a data set) and the number of input features (columns of a data set) is not frictionless.[114] Finally, it should be noted that reinforcement learning can only play a limited

[106] Economist Magazine Special Report, 'Artificial intelligence', p. 6.

[107] Inferred from: Jesse Dunietz, 'The Fundamental Limitations of Machine Learning', *Nautilus*, (20 September 2016) <http://nautil.us/blog/the-fundamental-limits-of-machine-learning> [accessed 12 October 2017].

[108] Analysis on *how* autonomy may be delivered onto the battlefield is the subject of Chapter 4 (*Deployment*).

[109] Source: Microsoft Research, 'From machine learning to machine reasoning', <https://www.microsoft.com/en-us/research/publication/from-machine-learning-to-machine-reasoning/> [accessed 12 October 2017].

[110] Mataric, p. 263.

[111] Stefan Schaal, 'Is imitation learning the route to humanoid annoyed robots', *Trends in Cognitive Science*, 3, 6, (June 1999), p. 283 <http://www.bcp.psych.ualberta.ca/~mike/Pearl_Street/PSYCO354/pdfstuff/Readings/Schaal1.pdf>.

[112] Yaling Zheng, 'Machine Learning with Incomplete Information', *CSE Technical Reports*, 143, (December 2011), pp. 3-5 <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1148&context=csetechreports> [accessed 25 September 2017].

[113] Schaal, p. 248.

[114] Dr Hongbo Du points here to image-based classifications as an example where ISR may have a very limited number of images but with each such image having an extremely large number of features (in the millions if the number of pixel intensity values are taken as features). Dr Hongbo Du, in conversation with the author, January 2109.

role in modifying the weapon's learning model and that regular re-training is unavoidable. It is for this reason that the human's role in verifying the rightness of class labels remains critical.

7.3 Reasoning and cognition methodologies

The balance of this chapter continues to assume the deployment of wide capability, wide task AWS and, therefore, that the weapon's learning routines must facilitate appropriate reasoning. A definition of reasoning is 'the algebraic manipulation of previously acquired knowledge in order to answer a new question'.[115] For AWS to reason, Beautement notes that the weapon must be capable of complex 'value mapping' to establish an appropriate code-based equivalent of what can be sensed, reasoned about and implemented.[116] Norman posits an additional requirement whereby a cognition variant (Norman terms this feature 'affect') must also be present in order to evaluate, judge and then moderate actions resulting from such cognition. This distinction highlights the complexity of capabilities that together (and only together) must comprise appropriate cognition for weapons without supervision. It is 'affect', after all, that will alert the weapon of possible dangers.[117] As a starting point, AWS cognition must be able to accommodate battlefield hypotheses and inferences encompassing 'the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses'.[118] Cognition must thus comprise very many processes (covered in this and subsequent chapters) including knowledge management, attention, memory and working memory[119], judgment and evaluation, reasoning and computation, problem solving and decision making, comprehension and production of language.[120] It is a wide-ranging machine capability that is key to the deployment of broad-task AWS, the assumption of this section.

The difficulty is that cognition, even in its human condition, is challenging to define.[121] Human cognition can be conscious or unconscious. It can be concrete or abstract as well as intuitive ('knowledge' of a language) or conceptual ('model' of a language). Moreover, as noted by Dietterich and Michalski, cognitive processes must use existing knowledge in order to generate new

---

[115] Leon Bottou, 'From machine learning to machine reasoning', *Microsoft Research, arXiv*: 1102.1808, unnumbered, undated, Section 1 ('*Introduction*') and 3 ('*Reasoning revisited*') <https://www.microsoft.com/en-us/research/wp-content/uploads/2011/02/tr-2011-02-08.pdf>.

[116] Patrick Beautement, 'Putting complexity to work; achieving effective human-machine teaming', *The Abaci Partnership LLP*, (2015), p. 15.

[117] A Norman and others, 'Affect and Machine Design: Lessons for the Development of Autonomous Machines', *IBM Systems Journal*, 22, 1, (2003), p. 38 <http://ai2-s2-pdfs.s3.amazonaws.com/244f/65498b2acd07a25416527118a52b8924f6f6.pdf>.

[118] Examples of such decision-making heuristics include effort, fluency, naïve diversification, recognition, scarcity, social proof and simulation heuristics.

[119] Doug Black, 'AI Definitions: Machine Learning vs. Deep Learning vs. Cognitive Computing vs. Robotics vs. Strong AI', *EnterpriseTech*, (19 January 2018) <https://www.enterprisetech.com/2018/01/19/ai-definitions-machine-learning-vs-deep-learning-vs-cognitive-computing-vs-robotics-vs-strong-ai/> [accessed 8 May 2018].

[120] For a brief primer on learning and decision-making, see Max Bezerman, 'Judgement and Decision Making', *Harvard/Noba*, undated <https://nobaproject.com/modules/judgment-and-decision-making#vocabulary-bounded-rationality> [accessed 6 June 2017].

[121] Marisa Tschopp, 'Human Cognition and Artificial Intelligence – A Plea for Science', *Medium.com*, (23 April 2018) <https://medium.com/womeninai/human-cognition-and-artificial-intelligence-a-plea-for-science-21a2388f6e7e> [accessed 8 May 2018].

knowledge.[122] AWS cognition is therefore a general catchall for a broad but poorly demarcated collection of processes that will be pivotal in the weapon's perceiving, thinking and awareness. Indeed, demarcation difficulties are particularly unhelpful to code-based approaches.[123] Four snags may arise. Contamination by indeterminate heuristics in a weapon system's executive processes may directly contribute to this lack of defined boundaries.[124] Given that the weapon must use multiple sensors to acquire multiple layers of information about its environment and physical status, insufficient scripting differentiators will lead to increased system noise in AWS processes.[125] It can also be inferred from Haikonnen that system noise will impair the knitting together of sensed data into a non-contradicting interpretation of the weapon's moment-to-moment situation.[126] Such AWS datasets must be *incrementally* built as not all data is available to the weapon at once. Gil notes the data-priority, allocation and denial challenges that arise in such ML decision routines.[127] Resource bottlenecks are a further ramification to arise from ML. Current syntax and schema-based 'resource matchmakers' are demonstrably an inappropriate model to describe the weapon's likely workflows.[128] In this vein, Gil notes that they are incapable of providing an appropriate mechanism to explore trade-offs in the machine's decision space.[129] Given that data feeds will be arriving *progressively*, the underlying distribution of this data must generally evolve with time. This contradicts the hypothesis of 'identically distributed data'[130] upon which classic data-mining algorithms rely.[131] AWS sensor outputs must all be capable of dependable classification, handing-off, and sense checking.[132] This is non-trivial given that weapon *in*puts will are being acquired from a wide universe of stimuli and, as above, must accommodate performance knock-on arising from data inconsistency and, notes Hensman, training set irregularities.[133] Furthermore, the whole model

[122] See, generally: T Dietterich and RS Michalski, 'A Comparative Review of Selected Methods for Learning from Examples', in *An Artificial Intelligence Approach,* RS Michalski and others (eds), (USA: Tioga Publishing, Paulo Alto, 1983), p. 41 and pp. 43-45.

[123] Venkat Gudivada and others, 'Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleansing and Transformation', *International Journal on Advances in Software*, 10,1, (2017), 1-3 <https://www.researchgate.net/publication/318432363_Data_Quality_Considerations_for_Big_Data_and_Machine_Learning_Going_Beyond_Data_Cleaning_and_Transformations> [accessed 9 June 2018].

[124] CM Teng, 'Dealing with data corruption in remote sensi*ng', Advances in Intelligent Data Analysis,* VI, IDA, Lecture notes in computer science, 3646, Springer, (2005), p. 453 <https://link.springer.com/chapter/10.1007/11552253_41> [accessed 7 July 2017].

[125] Tom O'Haver, 'A Pragmatic Introduction to Signal Processing', Univ*ersity of Maryland at College Park*, Department of Chemistry and Bio-Chemistry, (May 2017) <https://terpconnect.umd.edu/~toh/spectrum/Differentiation.html> [accessed 3 May 2017].

[126] Haikonnen, p. 41.

[127] Yolanda Gil and others, 'Artificial Intelligence and Grids: Workflow Planning and Beyond'*, IEEE Intelligent Systems*, (February 20014), p. 27 <https://scitech.isi.edu/wordpress/wp-content/papercite-data/pdf/gil2004ai.pdf>.

[128] Luo Mai and others, 'Optimizing Network Performance in Distributed Machine Learning', *Hotcloud,* (2015), pp. 1-3 <https://www.usenix.org/system/files/conference/hotcloud15/hotcloud15-mai.pdf>.

[129] Gil and others, p. 28.

[130] Source: Statistics How To blog <http://www.statisticshowto.com/iid-statistics/> [accessed 28 September 2017].

[131] Beatrice Lopez and others, 'Modeling decisions for artificial Intelligence: Ninth International conference'*, MDAI 2012*, Girona, Spain, (November 2012), p. 235.

[132] George Luger, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving,* (UK: Pearson, 6th Edition, 2009), p. 193 <http://iips.icci.edu.iq/images/exam/artificial-intelligence-structures-and-strategies-for--complex-problem-solving.pdf>.

[133] Paulina Hensman and David Masto, 'Impact of Imbalanced Training Data for Convolutional Neural Networks', Degree project, *Kth Royal Institute of Technology*, Stockholm, (May 2015), p. 5 <https://www.kth.se/social/files/588617ebf2765401cfcc478c/PHensmanDMasko_dkand15.pdf>.

remains untested. Indeed, such data folding (here, preparation and handing-off) will require its own intermediate pre-processing, scrubbing and, depending on data provenance and repair, filtering to reduce noise and heuristic error (including availability, representation and base-rate routines).[134] Nor is cognition an exact science. In this vein, a further firmware requirement for AWS deployment is that deduction and reasoning routines factor for information that is *not* directly perceived.[135] This is a critical capability which assumes additional judgement processes be in place in order to garner situational awareness, detect battlefield contradictions and evaluate their significance. Difficulties around defining and balancing such facility demonstrate how human limitations within the Delivery Cohort will constrain development of combat-based machine learning, cognition and awareness.[136] Several obstacles exist. First, it is the Cohort who must define this weapons autonomy. It is the human politician and field commander who must specify its adoption within their battlecraft. It is then the human soldier, naval rating and airman that must implement the technology supporting what is a recurrent theme of this thesis that human strengths likely trump machine capabilities.[137]

As inferred from Grossman and Hart, firmware complexity presents humans with a unique principal-agent control problem.[138] The characteristic arises when a human entity (here, the principal and, in the case of AWS, its Delivery Cohort) appoints another to act in the former's interest. Military procurement and command 'might worry that scientists and programmers implementing a particular project will not act in their best interest'.[139] Recasting Bostrom's principal-agent problem then questions whether an AI (in this case, the unsupervised weapon's control suite) may compromise a project's broader interests (in this case, battlefield objectives or broader strategy). The complexity of agency underlines Pickering's metaphor of a 'mangle', the argument that agency will always be 'temporally evident' in ML processes rather than in either the subjects or objects involved in that process.[140] On this matter, Bostrom and Danahar separately argue that agency may be an intractable challenge as simple observation of AWS' behaviour in its development phase will be insufficient given the ulterior possibility of 'treacherous turn syndrome'.[141] Moreover, Bowden notes that small-scale testing of weapon AI in a laboratory will be markedly different even from limited field study that must precede equipment rollout into the

[134] Source: Study.com <http://study.com/academy/lesson/heuristics.html> [accessed 13 October 2017].

[135] Shashi Phoha, 'Machine Perception and Learning Grand Challenge: Situational Intelligence Using Cross-Sensory Fusion', *Frontiers in Robotics and AI (Sensor Fusion and Machine Perception)*, (6 October 2014) <https://www.frontiersin.org/articles/10.3389/frobt.2014.00007/full> [accessed 18 October 2017].

[136] Don Norman, *Things That Make US Smart: Defending Human Attributes in the Age of Machines,* (USA: Diversion Books, 2 December 2014), generally.

[137] Beautement, p. 12. For a discussion on context, assumptions and the relative roles of social, political, cultural and technical drivers to the debate on AWS deployment, see: Chapter 2 (*Context*).

[138] Grossman and Hart, p. 7.

[139] Inferred from: Bostrom, *Superintelligence,* p. 128.

[140] Andrew Pickering, 'Cybernetics and the Mangle: Ashby, Beer and Pask', *University of Illinois, Department of Sociology,* (March 2002), pp. 1-7. (Alternative source: Social Studies of Science, 32, 3, (2002)).

[141] Bostrom, *Superintelligence*, pp. 128-132. Also: John Danahar, 'Philosophical Disquisitions', *Danahar blog*, (19 May 2014) <http://philosophicaldisquisitions.blogspot.co.uk/2014/07/bostrom-on-superintelligence-3-doom-and.html> [accessed 5 July 2017].

combat environment.[142] In this vein, a further firmware ramification is that the weapon's whole systems must comprise intelligent sub-parts that are themselves capable of agency; each weapon component must therefore be viewed as an autonomous agent in its own right. This is a central finding that raises additional complications. Composite agency may, after all, complicate the weapon's motivational selection. This complexity is unexpectedly fundamental in AWS deployment as the motivations of composite systems depend not only on the impulses of their constituent sub-agents but also on how those sub-agents are organised, an issue that is unsuited to human assessment (both for the commissioning Cohort and those coding its processes).[143]

7.4 Attention methodologies in AWS

A final firmware issue that is relevant to this chapter arises from the *framing* of weapon autonomy. How can the system be structured in a manner that AWS' focus can appropriately be directed? As inferred from Helgason, constructing a suitable architecture within which to manage attention methodologies will be a basic challenge to the removal of weapon supervision.[144] The human brain appears to be free to choose what it looks at, listens to and thinks about. In benign conditions, humans can focus their attention as they please. This drift of information, however, if not limited in any way, would lead to memory overflow, interference and what Haikonnen terms 'contradictory neural cacophony'.[145] Similar to the human brain, the AWS platform must actively select the source and quantity of its information, what to process, what to store and which peripheral information then to attenuate in its decision processes. For humans, this process is termed *attention* which, inter alia, controls sensory information acquisition and subsequent processing. As noted by James as far back as 1890 in his *The Principles of Psychology*, 'attention implies withdrawal from some things in order to deal effectively with others'.[146] It also, he continues, 'has a real opposite in what is the confused, dazed, scatterbrained state which in French is called *distraction* and *Zerstreutheil* in German'. Helgason notes that this is challenging (and possibly intractably so) to mimic in a machine.[147] All such real-world tasks (especially those in a battlefield setting) come with time limits and it is the managing of this attention feature that must therefore be a key constituent of AWS processes.

In the case of AWS control, a ramification of machine firmware is that such attention must be divided into that which is sensory (information acquisition from its senses) and that which relates to inner attention (the selection of relevant inner representations).[148] For instance, AWS attention

[142] Gavin Bowden, 'Real-time Deployment of Artificial Intelligence Network forecasting Models: Understanding the Range of Applicability', Water Resources Research, (31 October 2012), para. 6 of 40 <http://onlinelibrary.wiley.com/doi/10.1029/2012WR011984/full> [accessed 3 October 2017].

[143] Bostrom, *Superintelligence,* p. 23.

[144] Helgo Pall Helgason, 'General Attention Mechanisms for Artificial Intelligence systems', *University of Reykjavik*, PhD paper, (June 2013), generally <https://en.ru.is/media/td/Helgi_Pall_Helgason_PhD_CS_HR.pdf>.

[145] Haikonnen, p. 66.

[146] W James, *Principles of Psychology,* 1, (USA: Henry Holt, NY, 1890), pp. 403-404.

[147] Helgason, p. 2.

[148] Yi-Ling Hwong, 'Attention in Artificial Intelligence Systems', *AGI.io blog*, (22 September 2017) <https://agi.io/2017/09/22/attention-in-artificial-intelligence-systems/> [accessed 28 September 2017]. See also: HP Helgason and others, 'Towards a General Attention Mechanism for Embedded Intelligent Systems', *International Journal of Computer Science and Artificial Intelligence*, 4, 1, (May 2014), 1-7 <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=FC977A50C3C47BAD83BB0BC215F74E40?doi=10.1.1.694.5003&rep=rep1&type=pdf>.

should not exclude non-attended information completely, the so-called 'cocktail party effect' whereby it is possible to shift attention from conversation to conversation in a universe of broad noise.[149] In this manner, the AWS must be able to extract any attended information stream from what is otherwise background noise that comprises the weapon's multiple unattended sensory streams.[150] This core capability must be agnostic both to the weapon's tasking and its degree of independence. It must optimize the weapon's resource allocation. Dillel points out that attentional selection in AWS must be further divided into voluntary and involuntary attention.[151] Involuntary attention takes place when the weapon's attention is captured by strong, novel or significant stimuli.[152] Given their battlefield frequency, this creates an overarching issue of priority. What routine can appropriately mediate stimuli in order to determine that one such sensor input be preferred over others? The issue distils into how the weapon's input intensity should be managed. It would be inappropriate for AWS input strengths simply to be equalized in order to create a level-playing field. This is an important operational conundrum where a solution, Helgason notes, does not lend itself to being embedded into a weapon's machine routines.[153] It is also a pivotal model to AWS deployment given that battlefield information is *dynamic* where its importance fluctuates as events unfold. Martin, Secretary of the AISB[154], observes of these attention routines: 'I don't mean that it's too difficult like "man will never fly" or "man will never land on the moon". I'm saying it's hopelessly misguided like "man will never dig a tunnel to the moon"'.[155]

Weapon attention therefore requires delicate ranking given that it must be subject to update and robust calibration and cannot be determined simply by signal intensity. Battlefield stimuli, moreover, are likely to be quite indistinguishable with often overlapping boundaries. As inferred from Daotis, this complicates coding routines, in particular in those sequences designed to anchor particular attention traits within the weapon's control systems.[156] Examples of inadequate feature definition might include tactical curiosity, battlefield 'memories' as well as system conflicts arising from AWS goal and action selection and the a-priori programmed policies of the Delivery Cohort.[157] In considering its ranking, the weapon's management of such data variables creates intractable challenges: Two variables that may be useless by themselves can, after all, be useful *together*.

---

[149] Barry Arons, 'A Review of the Cocktail Party Effect', *MIT Laboratories*, undated, pp. 1-2 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.446.8514&rep=rep1&type=pdf>.

[150] Denny Britz, 'Attention and Memory in Deep Learning and Natural Language Processing', *WILDML*, (3 January 2016) <http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/> [accessed 28 September 2017].

[151] L Dillel and others, 'Voluntary and in voluntary attention have different consequences; the effects of perceptual difficulty', *US National Library of Medicine*, On-line resource, undated <https://www.ncbi.nlm.nih.gov/pubmed/18609402> [accessed 6 November 2017].

[152] Inferred from: Haikonnen, p. 67.

[153] Helgason, p. 9.

[154] Source: Society for the Study of Artificial Intelligence and Simulation of Behaviour <https://www.gold.ac.uk/news/tungsten-goldsmith-ai/> [accessed 17 November 2017].

[155] Andrew Owen Martin, senior technical analyst at the Tungsten Network, cited in Ben Sullivan, 'Elite scientists have told the Pentagon that AI won't threaten humanity', *Motherboard Magazine*, 19 January 2017 <https://motherboard.vice.com/en_us/> [accessed 17 November 2017].

[156] Marios Daoutis and others, 'Knowledge Representation for Anchoring Symbolic Concepts to Perceptual Data', *Bridges between the methodological and practical work of the robotic and cognitive systems community*, (2012), pp. 2-3 <http://www.aass.oru.se/~sci/chapter-11.pdf>. See also: Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues'*).

[157] Ibid., specifically: 8.5 ('*Anchoring and goal setting issues'*) and 8.7 ('*Action selection* issues').

Similarly, Guyon and Elisseeff note that a single variable that is useless by itself can then be useful with others.[158] In reviewing firmware's handling of variables, Verleysen notes further coding quandaries such as 'variable complimentarity', the 'curse of dimensionality' as well as issues around variable redundancy and variable commensurateness.[159] How, indeed, will the weapon's attention mechanism decide which variable (or combination of variables) to focus upon in any data training?[160] Walczak and Cerpa then point to the additional challenge introduced by attention heuristics being present in the weapon's decision processes (including, inter alia, recency, availability, representative, conjunction, anchoring and adjustment heuristics).[161] This blurring complicates AWS attention processing and adds additional agent-principal challenge.

Two final illustrations serve as a conclusion to this section by highlighting the intractability of this attention coding. First, the precept of *habituation* may lead to programming conflict whereby the weapon system's response is coded to decrease as a stimulus is repeated.[162] For AWS operating in a battlefield environment, it can be inferred from Haikonnen that not all repeating stimuli should lead to such attention habituation.[163] The associated attribute of sensitization actually increases a response to a repeated stimulus.[164] The illustration captures the uncertainty that characterises the coding of these key traits requiring instead the Delivery Cohort to incorporate additional associative learning sequences that mimic links between stimulus, representation and attention reaction. This is unsustainable, not least because of additional require feedback loops required to ensure their appropriate weighting, attribution and function. The issue highlights another firmware impediment. Fletchy notes that attention models must depend upon clearly defined routines separating 'memorization' ('Mary had a little lamb…') and 'memory-making', the act of imprinting episodic recalls into memory.[165] The firmware distinction is fundamental whereby *semantic* learning involves the unsupervised weapon absorbing a definable fact while *procedural* learning will involve the unit learning skill routines, mental or motor sequences. The challenge is that AWS firmware must incorporate sufficient adaptability, generalization and an ability to parse information that has been inductively picked up by the AWS in different situations and from different sources.[166] This integration piece is thus pivotal to weapon system architecture: A modular system that contains all of the parts identified in this chapter must still deliver the promise

[158] Isabelle Guyon and Andre Elisseeff, '*An Introduction to Variable and Feature Selection', Journal of Machine Learning Research*, 3, (2003), 1165 <http://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf>.

[159] M Verleysen and others, 'On the Effects of Dimensionality on Data Analysis', *IWANN*, (2003), pp. 106-108 <https://perso.uclouvain.be/michel.verleysen/papers/iwann03mv.pdf>.

[160] Guyon and Elisseeff, p. 1157.

[161] Steven Walczak and Narciso Cerpa, 'Heuristic Principles for the Design of Artificial Neural Networks', *Information and Software Technology*, 41 (2), (1999), pp. 6-8 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.22.9321&rep=rep1&type=pdf>.

[162] Stephen Marsland, 'Using Habituation in Machine Learning', *Neurobiology of Learning and Memory*, 92,2, (2009,) pp. 260-266 ('*Abstract'* and '*Introduction'*).

[163] Haikonnen, p. 75.

[164] Humans learn not to pay attention to the noise of traffic or other repeating background sound (habituation) but may instead become irritated by a repeating and intense noise (sensitization).

[165] For a general discussion on the issue, see: G Fletchy blog and video, (1 December 2014) <https://gfletchy.com/2014/12/01/from-memory-and-memorization-there-is-a-difference/> [accessed 6 May 2017].

[166] Each such dataset, moreover, will have its own provenance that will require weighting. This might include information from its own immediate sensor streams versus an update from central logistics, peripheral information from colleague robots or changes to its utility function from battlefield command.

that *overall* system performance can be improved by changing out individual modules. This, notes the US DoD, must be possible with little adverse effects on the whole machine and, critically, without the need for repeated full regression testing of the entire system.[167] Although it is intended that AI machines learn from their past experiences or environments, they will not be learning *all* of the time. This will be a complex routine for AWS given the intricacies of battlefield attention and situational awareness. Crucially, for a weapon to be compliant and still valuable, its intended end-state cannot be derived through ad hoc learning; instead, all data gathered by the AWS while on mission might instead be 'sent back home' and used subsequently to improve the unit's underlying model which then has to be redeployed. In this way, it is posited that no individual weapon system can learn 'bad' behavior from a particular environment and then turn rogue. On this basis, it is now possible to investigate AWS *software* components that together will comprise its empirical operation and, to intents and purposes, AWS' capabilities that rest upon the weapon's firmware and middleware.

[167] US Department of Defense, 'Summer Study on Autonomy', p. 25.

8. Software: Coding challenges to AWS function

Having considered infrastructure challenges that arise from the removal of human supervision in lethal weapons, the purpose of this chapter is now to review the principal software routines that will sit on top of this architecture and will likely comprise AWS operation.[1] This section primarily unpicks the capturing and processing of the unsupervised weapon's immediate environment. As noted by Kamimura, this is primarily a translation challenge between the weapon's current sensor data and the a-priori signposts that comprise the machine's initial setup on deployment.[2] The key for this chapter is therefore to determine how such real-time information may be incorporated into subsequent decision and action processes. By way of overview, several challenges exist. While both data capture and decision routines must be current, it is also necessary for the weapon to toggle predictably between old and new information and, within a framework of confidence checks, to ignore and remove stale data. This chapter's crux comprises its discussion on the weapon's likely utility function[3], the conditioning of that function[4], how that function is anchored and then, crucially, how that function interacts with the weapon's current goals, values and behaviours before deciding upon actions.[5] This requires a review of possible models that comprise battlefield engagement sequences (and the restrictions that such processes enforce upon AWS operation). Three factors again require emphasis. Demarcation between the subject matter of chapters Six through Eight (*Wetware, Firmware* and *Software*) is not clear-cut. Moreover, these chapters seek to opine on *cumulative* feasibility in AWS operation. Finally, the sections plot a likely but not definite architecture for the weapon's software routines and do so from a deliberately behavioural rather than technical perspective.

As indicated in the preceding chapter[6] and highlighted by the US DoD, the value of an autonomous weapon is largely reducible to how well its software routines perform.[7] An overview of design issues therefore provides relevant preamble to the chapter. As inferred from Grundspenkis, it is the weapon's primary-level 'representation' that will prove fundamental to compliant deployment whereby the AWS can capture its immediate 'world-state' and then anchor its position relative to its environment.[8] For this to take place, the platform must gather and then process a

[1] For a useful primer on network components and software engineering issues, see: Fabio Beckenkamp, 'A Component Architecture for Artificial Neural Network Systems'*, University of Constance, Software Research Laboratory*, (June 2002), pp. 67-73 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.9621&rep=rep1&type=pdf>.

[2] R Kamimura, 'Generation of Organised Internal Representations', *Neural Networks, IJCNN*, (1992) <http://ieeexplore.ieee.org/document/287116/> [accessed 12 June 2017]. As in previous chapters, unless specified the term 'weapon' is used interchangeably with 'unit', 'platform' and 'machine' to denote an autonomous weapon system.

[3] For explanation on the role of a utility function in AWS operation, see: this chapter, specifically: 8.3 ('*Utility function'*).

[4] EC Kulasekere and others, 'Conditioning and Updating Evidence'*, International Journal of Approximate Reasoning*, 36, (2004), 76.

[5] This chapter, specifically: 8.5 ('*Anchoring and goal setting issues'*), 8.6 ('*Value setting issues'*) ad 8.7 ('*Action selection issues'*).

[6] See: Chapter 7 (*Firmware*), specifically: 7.2 ('*Sources of technical debt'*).

[7] US Department of Defense, 'The Role of Autonomy in DoD Systems', pp. 1-3, pp. 15-17.

[8] For a useful overview on knowledge representation, see: J Grundspenkis, 'Fundamentals of artificial intelligence; knowledge representation and networked schemes', *Department of Systems Theory and Design*, Riga University, Lecture 7, undated, generally <http://stpk.cs.rtu.lv/sites/all/files/stpk/lecture_7.pdf>.

wide variety of data points and types.[9] Stonebroker points out that the routines' complexity arises from the requirement that the weapon poll, categorize, classify and then act upon what must be a dynamic yet poorly correlated set of information.[10] The variety of (and variation in) this sensed data must be counterweighed, moment-by-moment, through compensatory application of confidence levels which, notes Culotta and McCallum generally of ML models, must then be attached to each of the weapon's data strings (dependent, inter alia, upon that data's recency, intensity, relevance, substantiation, completeness, consistency and trustworthiness).[11] This is an intricate requirement and one that must precede any subsequent allocation of processed data output into ensuing routines that will comprise weapon processes such as that platform's goals, values and behaviours. Furthermore, *all* of these routines must be incorporated within code-written instruction notwithstanding the nuanced nature of this list of dependencies.[12] Such coding must also manage the complexities of anchoring this data whereby amendments can then be made to weapon settings in appropriate degree in order to refine the nub of the weapon's operation, its on-going representation.[13]

In order to remove supervision from machine systems, the challenge is that the weapon must maintain a comprehensive (and now independent) understanding of its environment.[14] It is this function that generates the AWS' 'representation' and provides the machine with what Chaudri and others term its 'internal model of its world'.[15] On the battlefield, it is this volatile 'interpretation' that enables the autonomous weapon to apply ML in order to solve problems by considering, deciding upon and actioning possible solutions.[16] Representation is therefore a key capability in AWS' intended function and, as deduced from Mataric, is based on efficient manoeuvre of symbolic information harvested from the weapon's sensors.[17] The immediate challenge, however, arises from data accuracy and fit.[18] Dredze notes the complication that such models are founded upon a frequent (possibly continuous) requirement for sensed (in this case, battlefield) data that must first

[9] ISCA Tutorial, 'Hardware Architecture for Deep Neural Networks', *MIT Press*, nvidia, (24 June 2017) <http://www.rle.mit.edu/eems/wp-content/uploads/2017/06/ISCA-2017-Hardware-Architectures-for-DNN-Tutorial.pdf>.

[10] For discussion on data polling frequency and resolving conflict between data latency and storage, see: M Stonebroker et al, 'The 8 requirements of real-time stream processing', *MIT Press/Brown*, (2008) <http://cs.brown.edu/~ugur/8rulesSigRec.pdf>.

[11] Aron Culotta and Andrew McCallum, 'Confidence Estimation for Information Extraction', *Proceedings of HTL-NAACL*, Association for Computational Linguistics, (2004), pp. 1-2 <https://people.cs.umass.edu/~mccallum/papers/crfcp-hlt04.pdf>.

[12] This chapter, specifically: 8.1 ('*Coding methodologies*') and 8.2 ('*Coding errors*').

[13] This chapter, specifically: 8.5 ('*Anchoring issues*').

[14] See: Chapter 6 (*Wetware*), specifically: 6.1 ('*Software versus intelligence*') and 6.2 ('*Architectural approaches*').

[15] A further useful resource on knowledge representation can be found in slide format at: V Chaudri, 'Knowledge representation and reasoning', *University of Stanford,* Class slides CS227, (Spring 2011) <https://web.stanford.edu/class/cs227/Lectures/lec01.pdf>.

[16] Dietterich, pp. 3-6.

[17] Mataric, p. 13.

[18] Medium.com, 'The AI data Apocalypse' <https://medium.com/@peopleio/the-ai-data-apocalypse-1375ac47ffe4> [accessed 5 July 2017].

be subject to both scaling factors[19] and conviction weightings.[20] These factors must appropriately reflect the confidence level attributable to particular routines (itself a dynamically changing construct) and which will explicitly influence subsequent weapon outcomes. This is a non-trivial and, considering Hartemink, a possibly intractable prerequisite.[21] To this point, Tenk points out that the platform's treatment of sensed data will be prone to corruption.[22] How then might this be managed in an engagement sequence? Empirically, the weapon's sensed data with *least variance* will be given additional weighting in each new polling iteration. This has consequences. Confidence weightings may have the unintended effect of inappropriate data smoothing until that data (in the case of an engagement sequence, information on target signature, classification, positioning, location, movement, threat, obstacles and sensitivity to battlefield clutter) becomes inappropriately scaled to the mean.[23] A further effect is that other sensed data may receive inappropriate weighting, so requiring the introduction of normalization routines[24] in order to force appropriate accounting of those variances in the weapon's routines.[25] Finally, a weapon's sensed data must be comparable (both temporally and structurally) before it can be reliably incorporated in an engagement sequence. The challenge is that smoothing of primary data might also obscure underlying data variations which, in themselves, may be pivotal to ensuring weapon efficiency and compliance.

Allocating a high confidence level to an AWS dataset should then translate a portfolio of specific (and, by definition, vetted) stimuli into a highly allied response action.[26] For this to occur, however, first requires that confidence thresholds change *dynamically* in order appropriately to regulate weapon behaviour.[27] A difficulty is noted by Siddiqi whereby such intervention should apply whether AWS design is top-down (policies that are programmed into the weapon whereby such rules are executed without variation) or bottom-up (weapons based upon machine learning including deliberative planning as well as interaction with humans).[28] There is actually quite little operational flexibility in the scope of weapon routine. This is an important observation as narrow tolerances are themselves a source of inappropriate system brittleness and 'technical debt'.[29]

---

[19] To be applied by the weapon system when a real world set of data needs to be represented on a different scale in order to fit a specific data format.

[20] M Dredze and others, 'Confidence-weighted linear classifiers', *Department of Computer and Information Science,* University of Pennsylvania <https://www.cs.jhu.edu/~mdredze/publications/icml_variance.pdf>.

[21] A Hartemink and others, 'Maximum likelihood estimation of optimal scaling factors', *MIT Laboratory for Computer Science*, p. 1 <https://groups.csail.mit.edu/cgs/pubs/spie.pdf>. The research reports undesirable variation arising from hybridization of sample sets.

[22] CM Teng, 'Dealing with data corruption in remote sensing', *Advances in Intelligent Data Analysis*, VI, IDA, lecture notes in Computer Science, Vol. 3646, Springer, (2005) <https://link.springer.com/chapter/10.1007/11552253_41> [accessed 7 July 2017].

[23] US Field Manual, *Light Cavalry Gunnery: Target Acquisition,* (USA: Field Manual Publications, 17-12-8, February 1999) <http://www.globalsecurity.org/military/library/policy/army/fm/17-12-8/ch3.htm> [accessed 12 August 2017].

[24] Normalization is the weapon's process of reorganizing data so that it meets two basic requirements. First, there may be no redundancy of data (all data is stored in only one place). Second, data dependencies may all be logical (in which case all related data items are stored together).

[25] Hartemink and others, generally.

[26] Source: ISCA Tutorial, 'Hardware Architecture for Deep Neural Networks', generally.

[27] Human judgement will therefore be a critical component in AWS programming with the degree of machines' learning and awareness being determined by how humans have set the filters and weightings.

[28] Abdul Ahad Siddiqi, 'Implications of using Artificial Intelligence Technology in Modern Warfare', *ICCIT*, (2012), p. 33.

[29] See: Chapter 7 (*Firmware*), specifically: 7.2 ('*Machine learning and sources of technical debt*').

Together, after all, these functions will inform the weapon's knowledge representation.[30] It is a fundamentally complex construct.[31] Rather than being a matter of memory management, the issue is *what* and *how* that primary data is represented within the AWS' representation. An example (again relating to AWS land-based navigation) is useful to evidence this intricacy. The unsupervised weapon may be programmed to remember an exact odometric path to a target.[32] Alternatively, Devlin notes that the platform may be programmed to follow a sequence of moves made at particular landmarks in its immediate environment.[33] In both cases, the weapon's topological map must be dynamic and always current if the AWS is to rely upon such an instruction dataset (whether as a complete sequence or as a series of movements triggered at each landmark). In inferred from Cipar, the weapon's representation is thus structurally prone to obsolescence.[34] It may also be susceptible to camouflage, start-state confusion and other muddling noise.[35] Distortion may arise in data points from elevation, obstacles and other fast-changing battlefield metrics.[36] Furthermore, an odometric path is only useful if the AWS' environment is itself static and understood. It similarly relies on precise tracking of distances *and* turn angles. As confirmed by Aquel, the procedure is non-obvious and challenging.[37]

Once so deployed, the AWS will move around an environment, dynamically polling new information from its sensors.[38] The challenge is that the weapon must use this refreshing routine to refine its own probability distributions using conditionalization.[39] The theory is that the weapon is thus computing a new probability for its immediate world that is consistent with new information being polled from its sensors. Any inconsistent probabilities will be set back to zero and then 'renormalized' over the remaining possible outcomes. The weapon's model for conditionalization, however, is that it must reliably calculate conditional probabilities for *each* set of possible causes for *each* of its given observed outcomes. The intended routine is for the weapon to construct a complex composite comprised of the *received* probability of each such cause and the *conditional*

[30] Chaudri, generally.

[31] Andreas Engel and Wolf Springer, 'Temporal Binding and the Neural Correlates of Situational Awareness', *Trends in Cognitive Science*, 15, 1, (January 2001), pp. 16-17 <http://ieeexplore.ieee.org/document/287116/> [accessed 12 September 2017].

[32] The use of data from motion sensors to estimate change in position over time, usually in relation to a starting position.

[33] For a useful discussion of the issues, see: H Devlin, 'Google creates and artificial intelligence program that uses reasoning to navigate the London tube', *Guardian Newspaper*, 12 October 2016 <https://www.theguardian.com/technology/2016/oct/12/google-creates-ai-program-that-uses-reasoning-to-navigate-the-london-tube> [accessed 4 February 2017].

[34] This feature presents an adjunct challenge whereby every slowdown of a given data thread can delay all other data threads and paralyse weapon processing. Inferred from: James Cipar and others, 'Solving the Straggler Problem with Bounded Staleness', *HotOS*, 13, (2013), pp. 1-3 <http://www.cs.cmu.edu/~seunghak/hotOS-13-cipar.pdf>.

[35] Although written in 1998, see: James Llinas and others, 'Studies and Analyses of Aided Adversarial Decision Making: Phase 2 – Research on Human Trust in Automation', *US Air Force Research Laboratory*, (April 1998), generally <https://pdfs.semanticscholar.org/1424/5ae4ec038f3a3b9c737e40b9d289dc79a612.pdf>.

[36] DY Kim and others, 'Data filtering system to avoid total data distortion in IOT networks', *MDPI,* (2017) <www.mdpi.com/journal/symmetry> [accessed 2 March 2017].

[37] M Aquel et al, 'Review of odometry: Types, approaches, challenges and applications', *Springer*, (28 October 2016) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5084145/> [accessed 3 March 2017].

[38] J Hendler and others, 'Handbook of Knowledge Representation', in *Foundations of Artificial Intelligence*, 978-0-444-52211-5, (USA: Elsevier Publishing, 2008) <http://dai.fmph.uniba.sk/~sefranek/kri/handbook/handbook_of_kr.pdf>.

[39] Bostrom, *Superintelligence',* p. 10.

probability of the outcome of those causes. Challenges emerge from conditionalized data streams[40] and engineering reliable practical application of a conjectural theorem remains enduringly complex.[41] Furthermore, the process must be driven by an initial set of error-free (and yet subjective) decision weightings. The process involves architectural, language, symbol and semantic complications, each component of which may likely prove intractable with any hiccup in these operational inputs grossly affecting platform performance.[42] Furthermore, as the weapon makes additional observations under this model, its list of derived probabilities may become ever more focused on a *shrinking* set of possible worlds that nevertheless remain consistent with the evidence being provided by its external sensors: This condition appears an intractable feature that cannot be designed away. The weapon, after all, is generating a series of posterior probability distributions that the machine uses as its *new* prior in every *next-time* step. Finally to this point, a weapon system that captures data from its immediate environment is dependent upon appropriate feedback loops.[43] Notwithstanding their importance to the AWS model, Norman notes that such pointer mechanisms are empirically difficult to manage and that such underlying routines can themselves create hidden loops whereby gradual changes may not be apparent from quick experimentation and testing by the weapon's software routines.[44] Indeed, hidden feedback complicates projected system changes and obstructs even very simple correction to the unsupervised platform.[45]

Within this general framework, it is then possible to identify software shortcomings in AWS mapping. It is relevant to revisit first principles. Mapping models, notes Wallgrun, generally require the unsupervised machine to remember a passage by creating a metric map of its immediate environment.[46] Each instance, however, requires precisely accurate measurement for *all* of that passage. Mararic notes the substantial storage, management and processing requirements of such datasets.[47] The model type, moreover, is particularly prone to obsolescence and poor stability.[48] This arises from the unavoidable requirement that *all* relevant navigable space (which is accessible to the weapon) must first be identified, then processed and then made 'map-ready' for each of the machine's internal representations. The resulting representation must then be searched in real

[40] Saad Mohamad, 'Active Learning for Data Streams', *Bournemouth University/IMT Lille Douai*, (October 2017), Abstract <http://eprints.bournemouth.ac.uk/29901/1/MOHAMAD%2C%20Saad_Ph.D._2017.pdf>.

[41] Although outside the remit of this analysis, see, inter alia: narrative on intractability of the Barman-Hartmanis conjecture, the P versus NP conjecture, Hodge's conjecture and Poincare's conjecture.

[42] See: previous chapter (*Firmware*), specifically: 7.2 ('*Machine learning and technical debt'*), 7.3 ('*Firmware ramifications of learning methodologies'*) and 7.4 ('*Attention methodologies'*).

[43] Y Gatt, 'Space Mapping and Navigation for a behaviour-based Robot'*, University of Neuchatel*, PhD thesis, (1994), pp. 12-13 and pp. 23-24
<https://www.cs.cmu.edu/~motionplanning/papers/sbp_papers/integrated2/muller_mapping.pdf>.

[44] D Norman, 'The Problem of Automation; Inappropriate feedback and interaction'*, ICS Report 8904*, Institute for Cognitive Science, University of California, (1989)
<https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19900004678.pdf>.

[45] Sculley and others, p. 3.

[46] A useful primer on robot navigation is found at: JO Wallgrun, 'Spatial Representation and Reasoning for Mobile Robots'*, Springer Publishing*, DOI 10.1007/978-3-642-10345-2_2, (2010), p. 12.

[47] Mataric, p. 147.

[48] James Vincent, 'The Biggest Headache in Machine Learning? Cleaning Dirty Data off the Spreadsheets', *The Verge*, 1 November 2017, generally <https://www.theverge.com/2017/11/1/16589246/machine-learning-data-science-dirty-data-kaggle-survey-2017> [accessed 12 January 2018].

time to establish available paths that are appropriate[49] before dynamically computing a *best* path for the AWS and then reframing the weapon's goals given the revised fit of that identified path.[50] This is a complicated set of consecutive routines that, given sensor and environmental noise, will acutely be prone to error.[51] Navigation (the example used throughout this analysis to illustrate model complexity) is, after all, a testing process for a human operator to master.[52]

Davis and colleagues point to a second software challenge that arises in representation-based models.[53] Expected unit function requires that comprehensive information is available to the weapon's internal routines. Its scope must include, inter alia, data on the AWS' self (stored proprioception on its goals, intentions, plans and self-limitations), its battlefield environment (navigable spaces, obstructions, limitations and deviations), its objects, actions, outcomes and its combat tasks (objectives, priorities, alternatives, goals). These components themselves comprise individual sub-models. Warwick points out that they are highly elaborate, variably correlated and likely conflicting.[54] Not all of these routines are intrinsic to core operation. Some may be relatively quickly constructed, briefly used and, crucially, must then be subject to appropriate 'forgetting routines' as they are likely quickly discarded.[55] This, however, adds material complexity as it will be intractably challenging to classify, process and then migrate sensed data between types and stages of the representation model.[56] As pointed out by Hunicke, dynamic balancing complicates the programming, maintenance and operational anchoring of autonomous systems once deployed.[57] This point is worth further explanation. In the first instance, it is unclear which constituent of the Delivery Cohort will both govern and then manage such weightings.[58] It is unclear what happens if

[49] For a useful discussion on decision waterfalls for autonomous vehicle motion, see: B Paden and others, 'A Survey of Motion Planning and Control Techniques for Self-driving Urban Vehicles', *MIT Press*, (25 April 2016), generally <https://arxiv.org/pdf/1604.07446.pdf>.

[50] Although written for a financial audience, see: S Brassia, 'Artificial Intelligence in the path planning of mobile agent navigation', *SciVerse ScienceDirect*, Elservier, (2012) <http://ac.els-cdn.com/S2212567112001475/1-s2.0-S2212567112001475-main.pdf?_tid=6d536560-6611-11e7-9d94-00000aab0f6c&acdnat=1499761249_fe59e466dd607005fe4ebe2da5722507> [accessed 3 August 2017].

[51] Andreas Birk, 'A Quantitative Assessment of Structural Errors in Grid Maps', *Autonomous Robots,* Jacobs University, 28, (2010) ,pp. 187-196 <https://pdfs.semanticscholar.org/aa32/bdb2fe20faf7585c2763bb70c3fb42f54196.pdf>.

[52] Directorate Land Warfare, 'Operation HERRICK Campaign Study', *AD LXC*, Warminster, (2015), generally. See again: Chapter 9 (Hardware), specifically: 9.3 ('*Navigation issues*').

[53] R Davis and others, 'What is Knowledge Representation?', <http://groups.csail.mit.edu/medg/ftp/psz/k-rep.html> [accessed 3 May 2017].

[54] For a general discussion on AI conflict resolution, see: Kevin Warwick, *Artificial intelligence; the basics,* (UK: Routledge, 2012), p. 34.

[55] The complicated role of *forgetting* in AWS control routines is considered later in this chapter. Mapping, for instance, typically entails incremental adjustment of an internal model followed immediately by the discarding of that sensed observation.

[56] S Dash, 'A comparative study of moving averages: Single, weighted and exponential', *Trade Station Labs*, (9 May 2012) <https://www.tradestation.com/~/media/Files/TradeStation/Education/Labs/Analysis%20Concepts/A%20Comparative%20Study%20of%20Moving%20Averages/Moving%20Averages.ashx> [accessed 6 April 2017]. For detailed discussion on hardware sensors and ramifications arising from sensed data, see: Chapter 9 (Hardware), specifically: 9.1 ('*Hardware and sensor fusion for the AWS*'). The process will require data sensing, data processing, back testing, memory management and data cleansing. See: this chapter, specifically: 8.1 ('*Coding methodologies*').

[57] A useful comparator here comes from the gaming industry. For a discussion on challenges to dynamic adjustment of computer programmes, see: R Hunicke et al, 'Artificial Intelligence for dynamic difficulty adjustment in games', *Northwestern University*, undated <http://www.cs.northwestern.edu/~hunicke/pubs/Hamlet.pdf>.

[58] For the purposes of this thesis, the broad Delivery Cohort is used throughout the thesis to describe the several parties responsible for the design, implementation and deployment of AWS.

some but not all deployed machines receive such intervention.[59] External revision of scaling factors will, after all, be more difficult in a battlefield environment where communication is problematic and data collection is compromised. Different levels of interpolation will presumably lead to weapons with subtly different software composition, subtly different learning traits and, over time, ever less identical profiles. A complication might even arise when these weapon routines result in AWS protocols ignoring attempts by parties to alter these weightings.[60] Similarly, difficulties will arise when the sizing (and revising) of scaling factors implies directly the preferences of the user.[61]

8.1 Coding methodologies

A key software faultline relates to *how* the AWS will be programmed.[62] A purpose of this section is therefore to review the efficacy of computer code as an appropriate means of replacing human supervision over weapons through unambiguous machine instruction.[63] Knight highlights the gap between capabilities that can be expressed through software script and the tasking required of underlying machine models; software written in high-level languages, he notes, increasingly 'crushes' transparency and subsequent serviceability in these models.[64] The challenge to AWS deployment is that anticipating the full execution costs of weapon code is increasingly difficult. Wren identifies, for example, a related trend in the use of inappropriate higher-order coding tools, in particular 'object-orientated methods and iterators that cannot easily be described and manipulated in weapon-specific low-level terms'.[65] Nor is coding efficiency a linear construct; as noted by Selman, some quite involved tasks readily break down into programmable kernels while other apparently light routines appear to defy expression. It is well documented that the efficiency curve of coding appears very variable with certain routines becoming unexpectedly iterative and unsolvable.[66] This should not be unexpected: Godel, after all, evidenced as early as the 1930s that not all arithmetic truths that exist are provable, demonstrating that mathematical statements occur

[59] The issue is well documented (in particular by service providers). See: D Moran, 'Tackling RF-Denied Environments', *Harris Corporation*, (9 March 2017) <https://www.harris.com/perspectives/innovation/tackling-rf-denied-environments> [accessed 23 June 2017].

[60] This is less to do with Hawkins' concerns on nefarious robot sentience (see: Appendix Two: '*The issue of singularity in AWS*') and more to do with individual AI agents creating overarching software priorities that cut out, for instance, external updating.

[61] Peter Vas, *Artificial-intelligence based Electrical Machines and Drives,* (Oxford: Oxford Science Publications, 1999), p. 579. See also: Sanford Grossman and Oliver Hart, '*An Analysis of the Principal-Agent Problem',* Econometrica, Vol. 51, Issue 1, January 1983, p. 10 <http://brousseau.info/pdf/cours/grossman_hart_83.pdf>.

[62] Jason Tanz, 'Soon we won't program computers. We'll train them like dogs: The End of Code', *Wired Magazine*, Business, 17 May 2016 <https://www.wired.com/2016/05/the-end-of-code/> [accessed 23 September 2017].

[63] Robert Harper, 'Structure and Efficiency of Computer Programming', *Carnegie Mellon School of Computer Science*, (23 July 2014), p. 2 <http://repository.cmu.edu/cgi/viewcontent.cgi?article=3673&context=compsci> [accessed 18 October 2017].

[64] Will Knight, 'AI's Language Problem: Machines that truly understand human language would be incredibly useful. But we don't know how to build them', *MIT Technology Review*, (9 August 2016) <https://www.technologyreview.com/s/602094/ais-language-problem/> [accessed 3 September 2017].

[65] Alisdair Wren, 'Relationships for Object-Oriented Programme Language', *University of Cambridge Computer Laboratories*, Technical Report Number 702, (November 2007) <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-702.pdf>. This is considered further later in this chapter.

[66] Bart Selman, 'Challenge Problems for Artificial Intelligence', *13th National Conference on AI*, AAAI-96, undated <http://erichorvitz.com/selman.htm> [accessed 16 October 2017].

that cannot be derived from arithmetic axioms by arithmetic rules alone.[67] The effect is that situations exist that *cannot* be captured through code-based scripting. Andersson, moreover, points to the breadth of (currently) unsolved coding problems including several capabilities that are prerequisites to AWS deployment such as 'closed domain solutions', 'common sense reasoning', 'strong generalization', 'essential learning', as well as counter-factual reasoning, intuitive physics, intuitive psychology and sensorimotor problems.[68] Andersson's list thus encapsulates what is a systemic challenge to the removal of human supervision from battlefield weapons.[69]

A central matter is therefore to understand how the Delivery Cohort can express *intention* in unsupervised weapon systems.[70] An etymological primer provides useful context to these challenges. For coding purposes, there are two kinds of 'directive construct', *concrete* and *abstract*.[71] The 'concrete' has a meaning that can be directly perceived and indicated such as a seen object (a tank) or an action (engaging the tank). Basili notes that concrete words may be taught by associating the word with the entity concerned.[72] In AWS instruction, the meaning of a concrete word (here, the Cohort's intention) must be linked vertically within the weapon controller to a representation on one of the AWS model's learning planes. This linkage must then be coded in a manner that one such associatively connected entities can evoke another in particular sequences. In this way, the meaning of an instruction will not be restricted to just one association. The challenge arises around coding's capture of abstracts, an important conjecture that is reviewed below.[73] Haikonnen is highlighting a key issue for general coding when he notes that most words are 'no more or less abstract, they do not have a meaning that can simply be pointed out; the meanings of abstract words are defined by other words'.[74] It is the complexity of a command's construction that complicates coding. For the unsupervised weapon, the information contained within a command must also be coupled to previously given information as well as to information that is to follow. This is rarely obvious. In particular, associatively connected words are likely to create ambiguity in the weapon's control routines.[75] Nested structures and, notes Danniels, conditionals that are common

[67] Source: Kurt Godel, 'Incompleteness Theorem', <https://www.britannica.com/topic/Godels-first-incompleteness-theorem> [accessed 12 October 2017].

[68] Simon Andersson, 'Unsolved Problems in Artificial Intelligence', *AI Roadmap Institution, @goodAI blog*, (3 February 2017) <https://medium.com/ai-roadmap-institute/unsolved-problems-in-ai-38f4ce18921d> [accessed 3 September 2017].

[69] Other parties prioritise slightly different lists of outstanding capabilities. See therefore: The Great Debate, 'Determinism and Free will in Science and Philosophy', generally <http://thegreatdebate.org.uk/determinismandfreewill.html>) [accessed 28 August 2017].

[70] See: Chapter 6 (Wetware), Footnote 123: For the purposes of this thesis, the broad Delivery Cohort describes the several parties responsible for the design, implementation and deployment of AWS.

[71] Rohan Paul and others, 'Grounding Spatial Concepts for Language Interaction with Robots', *Proceedings of 26th International Joint Conference of Artificial Intelligence*, (2017), p. 4929 <https://www.ijcai.org/proceedings/2017/0696.pdf>.

[72] R Basili and others, 'Using Word Association for Syntactic Disambiguation', *Trends in Artificial Intelligence*, Springer Verlag, 549, (30 May 2005) <https://link.springer.com/chapter/10.1007/3-540-54712-6_237> [accessed 12 October 2017].

[73] See, generally: John Launchbury, '*A DARPA perspective on Artificial Intelligence', DARPA slides*, undated <https://www.darpa.mil/attachments/AIFull.pdf>. Launchbury uses the examples of 'existence' and 'truth' to illustrate this coding constraint with capturing abstracts.

[74] Haikonnen, p. 135.

[75] Inferred from: Haikonnen, p. 233.

to complex instructions will likely create similarly challenging syntactic issues.[76] It is, after all, peculiarly human practice to be able to understand what has been directed without having to figure out exactly the meaning of the words. This will be particularly difficult for weapon coding where trusted outcomes are vital if the Delivery Cohort is properly to exploit AWS' deployment; while Henry points out that context-free, rule-based parsing may be able to establish who-is-doing-what-to-whom, this is insufficiently comprehensive as either an allocation or a tasking basis upon which to deploy AWS.[77]

Coding, furthermore, cannot capture context. Here, context may encapsulate the information in adjacent instructions, from non-associated coding routines and from other information sources being polled from the battlefield.[78] As pointed out by Suchman, coding must facilitate the AWS' broad situational awareness in order to be LOAC complaint but also in order for the independent weapon actually to be accretive.[79] Code cannot simply be a mechanism to parse text and, in order to enable empirically autonomous operation, it must empower exactly those challenging processes identified by Andersson such as perception, memory making, subjective interpretation, appreciation of temporal order and the emotional evaluation of significance.[80] In this vein, rules of engagement are not binary and do not lend themselves to binary coding; given then that these routines are generally comprised from knitted together sub-routines, Crammer notes that this property has material potential for whole-machine corruption.[81] Moss notes that conflicts of interest similarly complicate AWS coding.[82] Several examples exist. It will be necessary to incorporate different categories of facts within AWS syntax including indexical facts, normative facts, strong convictions, observations, hints, clarifications, reinforcements as well as basic ontological statements. The challenge is that such categorization will itself be volatile and will change unpredictably according to sources of subsequently gleaned intelligence, feedback and input from the AWS' primary sensors. Nevertheless, the management of such 'facts' will directly impact AWS priorities and actions. Conflicts may then arise from the weapon's subsequent processes, in particular its interpretative routines in what Tanz highlights is already a crowded set of system processes.[83] This requires further comment. Yao notes that the role of intermediate processing is to update machine states by revising dynamically that weapon's existing set of

[76] Chip Danniels and others, 'Harnessing Initiative and Innovation: A Process for Mission Command'*, Military Review*, (September-October 2012), p. 18 <http://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MilitaryReview_20121031_art006.pdf>. On nested statements, see: Microsoft Help functions <https://www.techonthenet.com/excel/formulas/if_nested.php> [accessed 28 October 2017].

[77] Winston Henry, *Artificial intelligence,* (USA: Addison-Wesley Publishing, Reading, MASS, 1984), pp. 295-314.

[78] For a discussion on the role of context in lethal engagement see: Chapter 2 (*Context*), specifically: 2.6 ('*The role of situational awareness and uncertainty'*).

[79] Suchman, 'Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy', pp. 6-7.

[80] Andersson, pp. 1-2.

[81] Inferred from: Koby Crammer and others, 'Learning from Data of Variable Quality'*, Advances in Neural Information Processing Systems*, (2006), pp. 219-220 and pp. 223-224 <https://www.cis.upenn.edu/~mkearns/papers/vardata.pdf>. For discussion on battlecraft and how AWS deployment might undertake broad combat routines see: Chapter 4 (*Deployment*), specifically: 4.3 ('*Human-Machine teaming'*) and Chapter 10 (*Operations*), generally.

[82] Richard Moss, 'Software writing Software and the broader challenge of computational creativity'*, New Atlas*, (3 March 2015) <https://newatlas.com/creative-ai-computational-creativity-challenges-future/36353/> [accessed 12 October 2017].

[83] Tanz, 'Soon we won't program computers. We'll train them like dogs: The End of Code', generally.

representations.[84] As inferred from Longa, an unavoidable consequence of the weapon's ML is likely conflict arising between, inter alia, the weapon's deployment parameters, its representational model on deployment, the weapon's intended end-state and, finally, pointers arising from sensed data[85], a consequence of which will therefore require precise arbitration and moderating routines.[86] Arbitration based on a simple recency heuristic is unlikely to be the appropriate tool with which to adjudicate sensor inputs in an engagement sequence given the battlefield's complex temporal considerations. What constitutes 'acceptable' delay? What represents urgency? Other arbitration strategies are similarly challenging. Averaging protocols that select weapon actions according to where the most conditions have been satisfied (also referred to as 'longest matching') are inappropriate as, by definition, they must reduce data precision (indeed, what data is being lost?). It can be inferred from Bullinaria that 'context limiting' rules will compromise efficacy; the activity of certain grouped conditions (for instance, in the engagement sequence) may be turned on or off according to the perceived fit of data to that intended action.[87]

Embedded conflicts create a second software challenge. As inferred from Kon, the unsupervised weapon must first be coded with appropriate *initial* beliefs.[88] The AWS' initial deployment represents a definite point in time after which its actions will be determined by that initial setting as amended by the weapon's subsequent cognitive processes.[89] Two challenges arise. The first comprises the degree of stepped change (the increment of learning termed 'anchoring') that each follow-on process then exerts on the weapon's immediately prior set of beliefs.[90] A second challenge is that weapon actions must reliably compose the appropriate reaction to every battlefield stimulus. To this point, weapon 'curiosity' might be a code-based composite of the two states of 'novelty' and 'attraction' (with each state occasioning a specific and often conflicting action routine in the weapon).[91] 'Astonishment' might be the combination of system reactions' 'attraction', 'withdrawal' and 'curiosity'. Fundamentally, this is currently how the weapon's Delivery Cohort will action weapon responses.[92] Similarly, within the AWS' reward system, 'aspiration' might be a combination of 'inclination', 'attraction' and 'arousal' (again, each with its own pre-determined portfolio of actions which, notes Knight, are in theory then blended with reference to machine-

[84] M Yao, 'Four approaches to natural language processing and understanding', *Topbots*, (31 March 2017) <http://www.topbots.com/4-different-approaches-natural-language-processing-understanding/> [accessed 29 September 2017]. See also: this chapter, specifically: 8.4 ('*Software processing functions'*).

[85] Luca Longa, 'Argumentation of Knowledge Representation, Conflict Resolution, Defeasible Inference and its Integration with Machine Learning', *Machine for Health Informatics, Computer Science*, 9605, (10 December 2016) <http://www.topbots.com/4-different-approaches-natural-language-processing-understanding/> [accessed 12 July 2017].

[86] Source: Study.com, 'Problem Solving Methods; Definitions and Types', *Study.com*, Lesson 4, <http://study.com/academy/lesson/problem-solving-methods-definition-types.html> [accessed 17 June 2017].

[87] John Bullinaria, 'Biases and Variances, Under-fitting and over-fitting', *Birmingham University School of Computer Science,* (2004) <http://www.cs.bham.ac.uk/~jxb/NN/l9.pdf>. See also: Kevin Warwick, *Artificial Intelligence; the basics*, p. 34.

[88] Mark Kon and others, 'Statistical Representations of Prior `Knowledge in Machine Learning', *Artificial Intelligence Applications*, (2005), pp. 1-2 <http://math.bu.edu/people/mkon/B5Final.pdf>. See also: Vincent Vanhoucke and others, 'Improving the Speed of Neural Networks on CPUs', *Processor Deep Learning and Unsupervised Feature Learning, NIPS Workshop*, 1, (2001) <http://andrewsenior.com/papers/VanhouckeNIPS11.pdf>.

[89] Vanhoucke and others, generally.

[90] See: this chapter, specifically: 8.5 ('*Anchoring and goal setting issues'*).

[91] See also: Chapter 11 (*Conclusion*), specifically: 11.1 ('*The nature of deployment challenges'*).

[92] Knight, generally.

generated weightings that are appended to each feature).[93] The point in this case is to illustrate the approximating nature of the methodology that underlies programming combinations which are intended digitally to synthesize a weapon's broad spectrum of reactions. Knight's analysis points to additional difficulties.[94] It will be intractable, for example, to capture any reliably temporal dimension in this procedure: Reactions such as 'aggression' and 'submission' require the AWS to initiate an immediate course of action. How then does weapon coding provide for the notion of a delayed response, a measured response, a slight deferral while additional information is sought and processed or, more complicated, a variable response? These aggregated responses must also be refined in the AWS over time (in line with lessons learned on the battlefield). As noted by Moyes, this is unlikely to happen in lock-step across colleague machines leading instead to unexpected idiosyncrasies and 'emergent' behaviours between one weapon and the next.[95] The general model requires, furthermore, that these 'emotional ratings' be accurately tagged within the weapon's internal set of 'memories' in order to ascribe appropriate significance (and immediately subsequent weighting) to relevant happenings received from its sensor array.[96] Only by such tagging (which itself will require appropriate feedback and calibration routines), notes Boros, might machine behaviour be appropriately evolutionary.[97] This is a further software conundrum; a model that reliably allows machines to *allocate* worth and values to sensor-derived experiences has yet to emerge.[98]

Conventional programming does not lend itself either to capturing or explaining ambiguity. Human communication is not only about sending a message that can be recovered and enacted upon by the receiver. There is usually a gulf between the written word and the intended message. Communication empirically has the following complicating phases: Translating the intended message into an appropriate form of expression, subsequent transmission, reception and decoding, interpretation and understanding. Three primary ambiguities may therefore confound the weapon's coding process.[99] Lexical ambiguities tend to concern omitted, imprecise and simply error-prone script. Syntactic and semantic ambiguities concern interpretative uncertainty. Finally to this point, pragmatic ambiguity characterises communicating parties with quite separate contextual bases. All three areas of uncertainty are directly relevant to AWS programming given the requirement that unsupervised weapons be situationally aware of their broadest ecosystem.[100] Moreover, as the coding process is affected by noise and nuance, Gosavi notes that supplementary

[93] Knight, 'AI's Language Problem', generally. See also: this chapter, specifically: 8.3 ('*Utility functions*'). See also: Daniel Dewey, 'Reinforcement Learning and the Reward Engineering Principle'*, AAAI Spring Symposium Series*, (2014), pp. 1-4 <http://www.danieldewey.net/reward-engineering-principle.pdf>.

[94] Haikonnen, p. 116.

[95] Emergent behaviour is dealt with later in this chapter, specifically: 8.8 ('*Behaviour setting and coordination'*). See also: Richard Moyes, 'Emergent behaviour and Risk – a sketch for a risk management approach'*, Article 36*, (April 2017), p. 344.

[96] Stephan Edelkamp, 'Memory Limitations in Artificial Intelligence'*, Algorithms for memory Hierarchies*, LNCS 2625, pp. 233-234 <http://www.csd.uoc.gr/~hy460/pdf/AMH/11.pdf>.

[97] Tiberiu Boros and others, 'Large Tagset Labeling in Feed Forward Neural Networks'*, Researchgate,* 9 August 2013, pp. 692-693 <https://aclweb.org/anthology/P/P13/P13-1068.pdf>.

[98] Inferred from: Haikonnen, p. 116.

[99] For a general discussion on coding ambiguities see: SW Developers, Tripod.com <http://swdevelopers.tripod.com/english/language/chap4.html> [accessed 27 February 2017].

[100] For a discussion on situational awareness in AWS see: Chapter 5 (*Constraints*), specifically: 5.1 ('*Geneva Convention and the Laws of Armed Combat'*) and Chapter 10 (*Operations*), generally.

routines will be necessary to enable the AWS to repeat certain message sections as well as to seek affirmation and feedback within that messaging.[101] Such procedures will be complex to initiate, complex to manage and will add to the weapon's overall technical debt as disambiguating either instructions or symbols cannot be undertaken without extensive, potentially unlimited, knowledge of the weapon's *real* world.[102] In the AWS, however, only its internal state models are available for this important purpose which must, by coding (and mechanical) definition, be partial and limited. As inferred from Liang and Potts, it will not be possible for the weapon to analyse meaning from an instruction in natural language syntactically until such a potential source of ambiguity has been resolved semantically.[103] In natural language, moreover, the boundaries of meaning are inherently indistinct. The margin, for instance, between day-time and night-time is unclear and can be arbitrarily set according to particular purposes. Kassan notes that this conflicts with the computationalists' assumption (here, underlying AWS programming) that 'the world consists of unambiguous facts that can be manipulated algorithmically'.[104] In several circumstances, therefore, an AWS' output may only be *partially* true, again requiring the control suite to apply that complicating confidence level to outputs.

Ambiguity will have material bearing in the weapon's coding process. Lapin notes that decision rules based on a utility function but where values are compromised by imperfect (and thus ambiguous) data degrade machine performance.[105] Given that lethal action must be initiated on a *priority* basis (whereby decisions are made according to the AWS' highest expected utility), this posits a further weakness. The weapon's dataset (or, more complicatedly, the melding of several battlefield datasets) enabling this decision may, notes Khemlani and Trafton, be insufficiently precise to enable any such workable utility function.[106] An adjunct challenge her is then that the possible range of battlefield outcomes is too broad for an appropriate utility to be calculated. Instead, the independent weapon must work to an *approximate* normative model that may or may not be sufficiently close to the original intention of the Delivery Cohort. Does the weapon pause if a computed outcome is too wide of this normative ideal and how might this relationship be dynamically managed?[107] A further challenge can be inferred from the work of Mikolajczyk and Schmid noting the inadequacy of sufficiently rich *descriptors* to capture the Delivery Cohort's

---

[101] Abhijit Gosavi, 'The Effect of Noise on Artificial Intelligence and Meta-heuristic Techniques', *Proceedings of the Artificial Neural Network in Engineering Conference*, 12, (2002) p. 7.

[102] See: Percy Liang and Christopher Potts, 'Bringing Machine Learning and Compositional Semantics Together', *Annual Review of Linguistics*, 1.1, (13 April 2014), pp. 1-2 <https://web.stanford.edu/~cgpotts/manuscripts/liang-potts-semantics.pdf>. For a discussion of technical debt and its ramifications to AWS operation, see: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of Technical Debt*').

[103] Ibid.

[104] Kassan, p. 4.

[105] Maksim Lapin, 'Image Classification with Limited Training Data and Class Ambiguity', *Saarland University*, PhD title, (June 2017), p. 2 <http://scidok.sulb.uni-saarland.de/volltexte/2017/6909/pdf/lapin17phd.pdf. For a discussion on the role of utility functions, see: this chapter, specifically: 8.3 ('*Utility functions*').

[106] Sangeet Khemlani and JG Trafton, 'Percentile Analysis for Goodness-of-Fit Comparisons of Models to Data', *Navy Center for Applied Research into Artificial Intelligence*, Proceedings of 36th Annual Conference of the Cognitive Science Society, (July 2014), pp. 737-738 <https://www.nrl.navy.mil/itd/aic/content/percentile-analysis-goodness-fit-comparisons-models-data> [accessed 2 October 2017].

[107] The concept of 'pause' gives rise to a further intractable conflict whereby the battlefield performance of AWS is directly affected by the weighting strength applied to any such intermission routine. This introduces several considerations. Can, for instance, the battlefield commander rely on AWS to execute if arbitrary late-stage routines exist to break that weapon's engagement sequence?

purpose.[108] Nor can these weapon descriptors be described in terms of high-level human concepts or other philosophical paraphrase: As Bostrom points out, all action definitions must eventually bottom out in terms that are actually defined in the AI.[109] Such coding definitions must be scripted in mathematical *primitives*, each with addresses that point to relevant memory registers in the AWS.[110]

Such software challenges may be summed up by Sharkey's aphorism that 'the AWS is nine-tenths code, one-tenth a portfolio of hardware bits. To paraphrase, it's all about the code, stupid'.[111] The breadth of tasking necessary for both productive and yet still compliant weapon deployment will, moreover, be very wide-ranging. This creates several challenges with Giordana and Serra pointing out that a machine that acts illegally must also 'learn' from its mistakes.[112] While this feature will require its own coding, it must be a routine that immediately and appropriately influences *all* other weapon actions. Feedback loops will similarly be required to prevent that mistaken action becoming part of the weapon's updated set of operating procedures.[113] Russell observes that the importance of such routines is likely to be understated by the Design Cohort and to be lost amongst other machine priorities.[114] The trait raises basic ethical issues. Are 'occasional mistakes' acceptable on the basis that AWS are similarly (and legally?) susceptible to ethical deficiencies as human soldiers?[115] Second, is a lower legal bar appropriate for machines with *less* lethality? Several coding difficulties arise from the precepts of ethics and morality.[116] In this matter, Haidt has tried to build a classification of those moral traits relevant to weapons-directing AI.[117] That his framework (or anyone else's framework) has yet to be settled points to the complexity of this task. Nor does the framework currently under discussion incorporate emotions such as guilt, embarrassment, shame, anger, disgust or deceit, all of which would appear to be ethical

[108] Krystian Mikolajczyk and Cordelia Schmid, 'A Performance Evaluation of Local Descriptors', *IEEE Transactions on Pattern Analysis and Machine Learning*, 27, 10, (October 2005), p. 1615 and pp. 1616-1620 <https://www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/mikolajczyk_pami2004.pdf>.

[109] Inferred from: Bostrom, *Superintelligence,* p. 187.

[110] A coding adjunct is how humans will be able physically to *liaise* with their unsupervised machines. Human behaviour means that the spoken word must remain a key means. In practical terms, AWS may need to communicate through spoken or hastily written descriptors with humans on the battlefield. This too is complicated. In order to understand what is being said, the weapon will need to evoke *relevant* (and current) connections, concepts and ideas. Even then, battlefield communication must be hierarchic; new concepts cannot be learnt if the AWS has nothing with which they may be associated. Moreover, these verbal descriptors rely heavily on limited capacity short-term memory. Such routines will also require suitable feedback loops in order to ensure both fitness for purpose and compliance.

[111] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[112] Attilio Giordana and Alessandro Serra, 'Learning from Mistakes', *Human Machine Perception,* Springer, (2001), pp. 89-90 <https://link.springer.com/chapter/10.1007/978-1-4615-1361-2_7> [accessed 18 August 2017].

[113] Isaac Caswell and others, 'Loopy Neural Nets: Imitating Feedback Loops in the Human Brain', *Stanford Publishing*, (2016), p. 1 <http://cs231n.stanford.edu/reports/2016/pdfs/110_Report.pdf>.

[114] Inferred from: Stuart Russell and others, 'Research Priorities for Robust and Beneficial Artificial Intelligence', *Association for the Advancement of Artificial Intelligence*, (Winter 2015), pp. 107-111 <https://futureoflife.org/data/documents/research_priorities.pdf>.

[115] Tonkens, p. 155.

[116] See: Chapter 5 (*Constraints*), specifically: commentary on Article 57 and the protection of civilian populations in a combat zone: 5.1 ('*Geneva Convention and the Laws of Armed Conflict*').

[117] Inferred from: Jonathan Haidt, 'The Moral Emotions', in R Davidson et al, eds., *Handbook of Affective Sciences*, (Oxford: Oxford University Press, 2005), generally.

components of the AWS' lethal engagement.[118] Current frameworks also ignore the precept of gratitude. Why is this important? ICRC argues that the responses of an AI agent should be appropriately broad in order to execute upon an engagement decision that is still compliant under LOAC.[119] In this vein, a 'morals portfolio' (for that is what it must be) should, notes Cowie, include such precepts as responsibility, appraisal, norm violation appraisal, negative self-evaluation, worry and motivation.[120] Without this, tasking of an unsupervised weapon must be correspondingly narrow if it is still to remain compliant.[121] AWS coding, moreover, should enable 'restoration' given the consequences of collateral damage (in particular in public relations) arising from AWS error.[122] Might, for instance, the AWS' learning routines be unwittingly discriminatory against particular parties? Notwithstanding a comprehensive test programme, Google had to apologise when the automatic tagging system in its photos app identified certain individual traits as 'gorillas'.[123]

In particular, it is the coding of 'guilt' that demonstrates the coding complexity of compliant AWS.[124] The programming, for example, of Arkin's 'Ethical Adaptor' (theoretically working in tandem with an 'Ethical Governor', an 'Ethical Behaviour Control', a 'Responsibility Advisor', a set of constraints and reasons, a Commander and an Operator) requires the seamless incorporation of an extraordinary number of inputs.[125] Several non-obvious inputs must be in place. These should include, inter alia, current data on friendly casualties, non-combatant casualties and the amount of civilian structural damage. Siddiqi notes that this last input requires complex calculation based

[118] Professor Noel Sharkey, in conversation with the author, 12 January 2016.

[119] International Committee of the Red Cross, 'Law of Armed Combat: Basic Knowledge'*, ICRC*, (June 2002), pp. 3-4 <https://www.icrc.org/eng/assets/files/other/law1_final.pdf>. A useful summary of this argument can be found at: Moral Foundations, 23 March 2016 <http://www.moralfoundations.org/> [accessed 4 June 2017].

[120] Roddy Cowie and others, 'Beyond Emotion Archetypes: Databases for Emotion Modelling using Neural Networks', *Neural Networks,* Elsevier Publishing, (19 May 2005), pp. 2-3 <https://www.researchgate.net/profile/Roddy_Cowie/publication/7782589_Beyond_emotion_archetypes_Databases_for_emotion_modelling_using_neural_networks/links/00b4951b63a63c205f000000/Beyond-emotion-archetypes-Databases-for-emotion-modelling-using-neural-networks.pdf>.

[121] The purpose here is not to impose any value judgement on the scope of tasks that a machine is suited. Instead, it should inform the extent to which such tasking must be limited by the platform's capability to integrate emotional precepts. Narrow emotional capability in otherwise autonomous weapons must equate to narrow taking. For a discussion on deployment 'degrees', see: Chapter 4 (*Deployment*).

[122] Jonathan Haidt, *The Moral Emotion,* (USA: University of Virginia, Oxford University Press, 2003), generally. See also: Siddiqi, p. 33.

[123] Economist Magazine Special Report, 'Artificial Intelligence', p. 15. See also: M Hutson, 'Even Artificial Intelligence can acquire biases against race and gender', *Science Magazine*, Science AAAM, 13 April 2017, paras. 4-7 and 10 of 12 <http://www.sciencemag.org/news/2017/04/even-artificial-intelligence-can-acquire-biases-against-race-and-gender> [accessed 26 May 2017].

[124] For a discussion on the ethics of autonomous weapons, see: Chapter 5 (*Obstacles*)*,* specifically: 5.8 ('*Ethical and accountability constraints'*)*.*

[125] Ronald Arkin and others, 'An Ethical Governor for Constraining Lethal Action in an Autonomous System'*, Georgia Institute of Technology Robot Lab*, Technical Report, GIT-GVU-09-02, (2009), pp. 1-2 <https://www.cc.gatech.edu/ai/robot-lab/online-publications/GIT-GVU-09-02.pdf>. Under Arkin's model for ethical control, robotic behaviour is expressed as an equation that includes all interpretable stimuli coming from the machines on-board sensors as well as from externally received information, a limitless number of possible responses both in terms of their strength and direction of action and an ability to set thresholds above which a response will be generated. Additionally, mapping must be established between the stimuli and the overall response range that then defines the behavioural function to be triggered; See also: 'Governing Lethal Behaviour: 'Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture'*, Georgia Institute of Technology Robot Lab*, Technical Report GIT-GVU-07-11, (2011), pp. 14-19.

upon accurate situational awareness.[126] He notes that it is a prerequisite in calculating the degree that this exceeds what is allowed under the test of Military Necessity in lethal engagement. At the very least, testing any such Ethical Governor in battlefield conditions will involve a presumably unacceptable process of trial and error as well as material processing delay, undermining Arkins' notion of a real-time and independent simulator running *in parallel* in order to ensure that the unsupervised weapon is working within statutory rules.[127]

Feasibility is not, however, simply a matter of appropriately coded instructions. A further difficulty arises from the *firing sequence* of those rules that will comprise AWS' instructions. Each rule (and the pattern created by those rules) will result in quite different values being taken forward at any given point of a sequence by the AWS. In order to provide a single end-value upon which to generate an appropriate battlefield action, the challenge is that the unsupervised weapon must first aggregate these values. As noted by Osmeyer and Cowell, this will comprise a complex weighted average ('centre of gravity') routine.[128] Such models increase AWS' 'technical debt' given ML's inherent limitation of reading one symbol at a time (with each such symbol being processed on information collected from previous symbols).[129] The difficulty is that weightings must be time-dependent in order to influence which rule the AWS fires first. A more manageable challenge exists given that the measures of quantity for the weapons' key data inputs may each be different (the weapon's voltages, temperatures and flow rates all measure quite differently and in quite different units). Each such weapon value may therefore require real-time translation requiring additional complexity (and what Warwick terms 'learning latitude').[130]

A recurring software challenge relates to the quality of AWS underlying data. O'Kane notes that algorithmic tools demonstrably become less useful as uncertainty grows around scenarios (whether through incomplete data, inappropriate coding routines or simple enemy feint).[131] Uncertainty, of course, is everywhere: The wheels of weapons may spin and battlefield obstacles move unpredictably. Rasmussen notes two relevant types of uncertainty which can be extrapolated to AWS. First, prediction uncertainty arises when the effects of actions are not fully predictable (here, uncertainty about the weapon's future states). Second, sensing uncertainty relates to ambiguity about the weapon's current state.[132] How then might AWS' software optimize outcomes given these constraints? How the Delivery Cohort tasks AWS unsurprisingly starts with as assessment of the weapon's skills-based behaviours (well-understood sensory-motor actions that,

[126] Siddiqi, p. 34.

[127] Arkin, 'An Ethical Governor for Constraining Lethal Action in an Autonomous', pp. 1-2. In this way, Atkins envisages incoming command sequences first being run on this simulator to ensure that any outcome meets an established rule-set *prior* to any autonomous action sequence. See also: US Department of Defense, 'Summer Study on Autonomy', p. 34. An Ethical override might theoretically provide for appropriate in-situ ghosting to filter malicious commands, theoretical erasure of key data should the machine be compromised and a block on malign reverse engineering.

[128] Jared Osmeyer and Lindsay Cowell, 'Machine Learning on Sequential Data using a Recurrent Weighted Average', *Cornell University Library*, (4 May 2017) <https://arxiv.org/abs/1703.01253> [accessed 4 October 2017].

[129] Osmeyer and Cowell, 'Machine Learning on Sequential Data using a Recurrent Weighted Average', generally.

[130] Warwick, pp. 43-44.

[131] Jason O'Kane and others, 'Algorithms for Planning under Uncertainty in Prediction and Sensing', *Autonomous Mobile Robots: Series in Control Engineering,* University of Illinois, (2005), 501-547, (pp. 501-504) <http://msl.cs.illinois.edu/~lavalle/papers/OkaTovCheLav06.pdf>.

[132] J Rasmussen, 'Skills, Rules and Knowledge: Signals, signs and symbols, and other distinctions in human performance models', *IEEE Transactions on Systems, Man and Cybernetics*, 13, 3, (1983), pp. 257-266.

for soldiers, become highly automatic after a period of appropriate training).[133] In demonstrating battlefield skills, AWS' software challenge is to incorporate a sufficiently robust classifier that can correctly predict classes of 'new' objects (and tasks) given the weapon's prior training undertaken on datasets of 'old' objects.[134] Such classifiers must overcome the battlefield problem of significant sequential correlation whereby nearby x and y values are almost certain to be closely related to each other.[135] As task complexity increases, the treatment of such classifiers must change from rules-based routines to knowledge-based routines and finally to sequences that require palpable expertise.[136]

How then might this tasking continuum create software challenges in AWS deployment? While O'Kane notes that those tasks which are 'highly repetitive with inherent feedback loops that can be controlled through mathematical statement' may lend themselves to coding, additional confirmatory routines become necessary as task complexity intensifies.[137] In a battlefield context, there are too many solutions to too many possible problems. Cummings also notes that more interpretation is required as the complexity of battlefield tasks increases, 'especially in cases of multiple and compound problems'.[138] The challenge is for coding routines to determine which rule type may best apply as uncertainty mounts in a sequence.[139] Why does this create an important crossroad? In situations that are characterised either by incomplete data or by ambiguous sensor input, algorithms are unlikely to understand sufficiently the weapon's solution space.[140] In this case, Smith points to code's inability to generalize and the resulting obligation of AWS' Delivery Cohort to base its weapon's decision processes on what are imperfect variables.[141] This notion of 'mounting difficulty' is supported by Gal who notes that such complication unsurprisingly peaks at the point of 'maximum' uncertainty in the AWS' data set.[142] This represents the point where AWS' ML spine is at its most inappropriate and when expert behaviours are required that instead leverage

---

[133] Thomas Dietterich, 'Machine Learning for Sequential Data: A Review', Structural, Syntactic and Statistical Pattern Recognition', *Oregon State University,* (2002), pp. 1-2 <http://web.engr.oregonstate.edu/~tgd/publications/mlsd-ssspr.pdf>.

[134] Here broadly relating to new events, sequences, actions, target types, order configuration, new parties and behaviours.

[135] Dietterich, 'Machine Learning for Sequential Data: A Review', generally.

[136] Missy Cummings uses human pilots as an example of such skills-based activity; human pilots train to interpret their cockpit dials before adjusting aircraft controls appropriately to ensure the aircraft's actual state matches the intended state. See: Missy Cummings, 'Artificial intelligence and the future of warfare', *Chatham House,* Research Paper draft, (January 2017), pp. 5-6.

[137] Jason O'Kane and others, 'Algorithms for Planning under Uncertainty in Prediction and Sensing', p. 3.

[138] Cummings, p. 6.

[139] Leslie Kaelbling and others, 'Planning and Acting in Partially Observable Stochastic Domains', *Artificial Intelligence*, Elsevier, 101, (May 1998), pp. 100-101 <https://ac.els-cdn.com/S000437029800023X/1-s2.0-S000437029800023X-main.pdf?_tid=c7163328-a68c-11e7-b7f0-00000aacb35e&acdnat=1506851102_e13f50110d2bfa5bfd9cb5e6bab83d35> [accessed 12 October 2017].

[140] Cummings, p. 7.

[141] Inferred from: PJ Smith and others, 'Brittleness in the design of cooperative problem-solving systems: The effects on user performance', *IEEE Transactions on Systems, Man and Cybernetics*, Part A: Systems and Humans, 27, 3, (1997), pp. 360-371.

[142] Yarin Gal, 'Uncertainty in Deep Learning', *Department of Engineering, Cambridge*, PhD submission, (September 2016), p. 7 <http://mlg.eng.cam.ac.uk/yarin/thesis/thesis.pdf>. Uncertainty here is a cumulative feature arising from, inter alia, out-of-distribution test data, 'aleatoric' factors (measurement imprecision), uncertainty in model parameters as well as from structural uncertainty (termed by Gal, model or epistemic uncertainty).

judgement, intuition and time-critical quick assessment of that situation. It is at this point where Cummings notes that algorithms fall short of the human expert who is instead able to make difficult decisions 'in a fast and frugal manner, since comparing all possible plan alternatives is time-intensive, especially in the face of uncertainty'.[143] The point may be that at such higher levels of expertise, battlefield commanders do not even recognise that they are making decisions; rather, they are fluidly interacting with a changing situation and responding to patterns that they recognise. Training, experience, subjectivity, biases, personality and a wide grasp of context are (among several factors) critical attributes that, as noted by Wang, cannot reliably be captured by code.[144] It is, after all, these same characteristics that best express the practical application of situational awareness. The crux is software's inherent 'symbol grounding problem' whereby each miscellaneous symbol that is employed to instruct the weapon is itself defined using *other* such symbols, the consequence being that it is difficult to relate meaning to these real world situations.[145] This points again to the circular challenge to such code-based models: How does the weapon determine *which* of its internal representations is relevant to the scenario that is immediately unfolding? It is this general issue of having to work to an appropriate context that is collectively termed AI's 'frame problem'.[146]

8.2 Coding errors

The preceding analysis suggests that coding limitations, whether in a class or cumulative basis, create intractable challenges to the removal of weapon supervision. Such analysis assumes acceptable coding accuracy throughout these weapon routines. A further attribute to compliant deployment, however, must be the successful management of coding errors, the subject of this section. The matter is not straightforward as coding accuracy, automatic bug repair as well as automated machine programming continues to evolve and relevant precedent is thus difficult to establish.[147] Certain observations on AWS' coding can, however, be made. Empirical experience of software development suggests that even veteran programmers unknowingly write one mistake into every ten lines of code.[148] How might this datum relate to battlefield systems? The F-35 fighter jet has more than twenty million lines of code, eight million of which relate solely to its missile and threat management systems.[149] Error rates occur regardless of current quality controls that seek to mitigate system risk, provide redundancy and deliver reliability. Error risk, moreover, is unlikely to diminish given that weapon platforms (and, by extension, AWS) will increasingly be comprised of multiple sub-systems produced by multiple manufacturers, each with different testing regimes,

[143] Cummings, p. 6.

[144] Yingxu Wang, 'On Abstract Intelligence: Towards a Unifying Theory of Natural, Artificial, Machinable and Computational Intelligence', *International Journal of Software Science and Computational Intelligence*, 1(1), (January 2009), 1-3 <http://www.ucalgary.ca/icic/files/icic/24-IJSSCI-1101-AbstractInt.pdf>.

[145] Kassan, p. 4.

[146] Murray Shanahan, 'The frame problem', *MIT Press*, (February 1997), '*Abstract*'.

[147] Larry Hardesty, 'Automatic bug repair: System fixes bugs by importing functionality from other programs without access to source code', *MIT News*, (29 June 2015), paras. 2-4 and 9-10 and 17 <http://news.mit.edu/2015/automatic-code-bug-repair-0629> [accessed 24 June 2017].

[148] Dan Mayer, 'Continuously Deployed', *Blog*, (11 November 2012) <https://www.mayerdan.com/ruby/2012/11/11/bugs-per-line-of-code-ratio> [accessed 5 October 2017]. Although written in 2002, see also: Andy Chou and others, 'An Empirical Study of Operating System Errors', *ACM SIGOPS Operating System Review*, 35, 5, (2002), Abstract and generally <https://pdos.csail.mit.edu/archive/6.097/readings/osbugs.pdf>.

[149] Source: Drone Mag, (April/May 2016), p. 62 <www.drone360mag.com>.

verification and control standards. Collaboration outcomes in the AWS' supply chain will only be as good as individual routines that knit together the weapon's overall processes.[150] The US Department of Defense recognises this issue of coding error in its software procurement; 'in algorithms, patterns-of-life [integrated routine practices] are critical and must be managed properly to ensure accuracy and correctness in subsequent decision-making process'.[151] Singer notes that such military-civilian collaboration (with each party's different incentives in this process) adds fragility to the procurement chain; currently seventy-five per cent of the maintenance and weapons loading for drone systems has been outsourced to private contractors 'with patchy results at best'.[152] These error rates are quite broadly corroborated. Other sources judge the minimum number of software errors observed at some 'two-and-a-half errors per function point'; at this rate, notes Kassan, a software program large enough to simulate the human brain would contain some twenty trillion errors.[153]

This analysis ignores the ramifications arising from coding error and coding brittleness[154], a characteristic of ML systems identified by Cox and Perlis where there is a tendency for coding to break when confronted with input deviation from situations anticipated by their designers.[155] AWS will require robust diagnostics (and malleable action selection) if it is to identify that designed functions are either operating erroneously or, as inferred from Gange, no longer being performed.[156] As noted by Prior, a further error dynamic will likely arise from the adding of software patches or other repair sequences: 'Patched software is impenetrable software' and will likely make code verification more challenging.[157] The attraction of AI may be that it is 'greater than the sum of its parts'[158] but such hidden interactions among what is a myriad of sub-systems manifestly increase the likelihood of coding error.[159] Finally, such analysis ignores the human context of AWS coding errors where service-level understanding of autonomous and other 'go-

[150] Walker, pp. 22-31.

[151] Department of Defense, 'Unmanned Systems Integrated Roadmap FY 2013-2038', p. 67. The matter is further discussed in Chapter 9 (*Hardware*), specifically: 9.2 ('*Calibration issues'*) and Chapter 10 (*Deployment)*, specifically: 10.2 ('*Validation and testing*').

[152] Peter Singer, 'Statement to US House of Representatives Committee on Oversight and Governmental Reform', cit. *Rise of the Drones, Unmanned Systems and the Future of War,* (USA: Congressional Research Service, Nimble Books LLC, March 2010), p. 2; Army systems operating in Iraq have been described as 'government-owned-contract-operated'.

[153] Kassan, p. 3.

[154] For a detailed discussion on brittleness in AWS routines, see: Chapter 7 (Firmware), specifically: 7.1 ('*Sources of technical debt'*) and 7.2 ('*Firmware ramifications of learning methodologies'*).

[155] Michael Cox and Don Perlis, 'Self-adjusting autonomous systems*', Awareness magazine*, 10.22417, 2011, pp. 1-2 <http://awareness-mag.eu/pdf/003951/003951.pdf>.

[156] Graeme Gange and others, 'Interval Analysis and Machine Arithmetic: Why 'Signedness' Ignorance is Bliss*', ACM Transactions on Programming Languages and Systems*, 37, 1, (January 2015), p. 1 and p. 3 <http://cliplab.org/~jorge/docs/ACM-TOPLAS-wrapped.pdf>. For general discussion of the issue (and potential solutions), see: Matthew Schmill and others, 'The Role of Metacognition in Robust AI Systems'*, aaai.org*, (2008) <http://www.aaai.org/Papers/Workshops/2008/WS-08-07/WS08-07-026.pdf>.

[157] For a discussion on AWS verification and validation procedures, see: Chapter 10 (*Oversight*), specifically: 10.2 ('*Verification and testing'*). See also: Dark Reading, <http://www.darkreading.com/vulnerability/microsoft-patch-problems-underline-trade/240160244> [accessed 1 March 2017].

[158] Fil Macias, 'The test and evaluation of unmanned and autonomous systems'*, ITEA Journal,* 29, (2008) 388. See also: Benjamin Alan Pryor, 'Assessing the Army's Software Patch Management Process', *Defense Acquisition University*, Aberdeen MD, (4 March 2016), p. 3 ('*Problem Statement'*) <http://www.dtic.mil/dtic/tr/fulltext/u2/1040604.pdf>.

[159] Richard Selby, 'Analyzing Error-Prone System Structure'*, IEEE Transaction of Software Engineering*, 17, 2, (February 1991), 141 <http://cgis.cs.umd.edu/~basili/publications/journals/J42.pdf>.

through' processes is rarely reducible to a single level of explanation. Coding error creates one further material conflict. Battlefield AI must throughout prioritise avoiding the possibility of high-regret outcomes, more likely in a contested environment where an adversary is intent on neutralizing assets either through deception, through direct force or through meddling to cause collateral damage.[160]

8.3 Utility function

How then might such software processes manage the unsupervised weapon's action sequence? In order to understand challenges arising, it is again useful to consider a brief process narrative. AWS' action sequences will be arbitrated according to an 'optimality notion', each unique to a specific weapon class and set by the Delivery Cohort as an initial set of decision rules.[161] It is this internal 'aide memoire' that will dictate, inter alia, the confidence premium (here, 'weighting') to be attributed to each of the weapon's possible worlds as set out above. Bartak notes the instability of this function which must rank the desirability of possible outcomes in order to establish a set of basic preferences for the weapon.[162] In this way and at each step, the weapon will select an action with the *highest expected* utility. Various challenges arise from this utility model. The margin for error is considerable and, in the case of an autonomous and lethal engagement sequence, comes with the possibility of high-regret outcomes.[163] To find the action with the highest expected utility, the AWS must first run an internal computation on *all* possible actions, a considerable task given the almost limitless number of battlefield parameters including target status and risks arising from poor execution. The utility function must also regulate what constitutes appropriate use of force, the involvement of colleague assets, consideration of next steps, a data audit ahead of an action sequence as well as post-event communication of each engagement. Assuming satisfaction of these prerequisites (and after iterative conditionalizing as set out above), only then can the weapon identify its suitable (and legal) priority by calculating the expected value of an action as the sum of the value of *each* possible world multiplied by the conditional probability of that world given the action. Humans do this innately.[164] A challenge for AWS is usefully expressed by machine action in the event of a tie in that function after which the weapon then resorts to picking a random action in order to achieve its expected utility which is likely to be unacceptable both from a compliance and utility perspective.[165]

---

[160] Richard Moyes, 'Emergent behaviour and risk – A sketch for a risk management approach', *Article 36 article*, (April 2017), p. 13. See also: US Department of Defense, 'Summer Study on Autonomy', p. 14.

[161] Paul Christiano, 'The Reward Engineering Problem', *AI Alignment*, (31 May 2016), pp. 3-5 <https://ai-alignment.com/the-reward-engineering-problem-30285c779450> [accessed 24 July 2017].

[162] R Bartak, 'Artificial intelligence', *KTIML course*, lecture slides, (2016), generally <http://ktiml.mff.cuni.cz/~bartak/ui2/lectures/lecture05eng.pdf>) [accessed 5 July 2017].

[163] Dennis Gallilandand others, 'A Note of Confidence Interval Estimation and Margin of Error', *Journal of Statistics Education*, 18,1, (2010) <https://amstat.tandfonline.com/doi/pdf/10.1080/10691898.2010.11889474?needAccess=true> [accessed 29 July 2017].

[164] Diana Gitig, 'Humans have an innate ability to assess probability but odds are we're bit off', *Genetic Literacy Project,* (19 May 2016), generally <https://geneticliteracyproject.org/2016/05/19/humans-have-innate-ability-to-assess-probability-but-odds-are-were-a-bit-off/> [accessed 18 March 2017].

[165] D Poole, 'Probabilistic Conflicts in a Search Algorithm for Estimating Posterior Probability Problems in Bayesian Networks', *Department of Computer Studies*, University of BC, Vancouver, (May 1996) <http://www.cs.ubc.ca/~poole/papers/seaalg.pdf>.

There is risk that the Delivery Cohort's epistemological specifications for this function may be plain wrong (or even slightly wrong or, more likely, may subsequently drift away from the day's battlefield priorities). Nor does the process lend itself to arbitration or adhoc intervention. It wrongly assumes whole-system predictability while, as inferred from Lopez-Paz, subtle disparities in AWS' prior probabilities will grossly affect how the weapon behaves.[166] The AWS might work to a prior that assigns zero probability to enemy forces being delivered by land transport: No matter how much the battlefield evidence it accrues to the contrary, that weapon may stubbornly rejected any sensor-based intelligence to the contrary and make dangerous, unpredictable choices as a consequence.

8.4 Software processing functions

AWS feasibility fundamentally depends on its ability to process information. A basic challenge therefore arises from what Sun terms 'the absence of attached meaning' in machine routines.[167] Instead, weapon sequences will likely be based on the *lowest* possible level of symbology to both external and internal entities, actions and relationships. Why might this be unacceptable in AWS design? In order for that unsupervised weapon to be compliant and valuable, sensor input should trigger wide and *unpredictable* associations between the weapon's external environment and that weapon's future actions.[168] This requires the machine to make wide extrapolation of outcomes from available data sets. For this reason, the richest possible symbology is therefore important if the weapon is to operate effectively but still within the boundaries (be they goal-based, value-based, territorial, action-based or peripheral) as defined by the Delivery Cohort. In this vein, a complicated environment, a conflicting list of tasks as well as mildly clashing goals or priorities must compromise weapon function. As Haikonen concludes, 'purposeful operation calls for order and priorities as everything cannot be attended to at once'[169] and, as inferred from Dietterich in the case of AWS, this is further complicated given that the only relevant product of its processes is as intermediate results that must themselves then be presentable for further processing.[170] A key challenge is posited by the weapon's control processes, by definition, being dynamic and unstable as there is no obvious *end-state* for the independent weapon. Jones' inquiry on how often the machine should poll its sensors highlights a similarly fundamental challenge to AWS deployment given operational requirements to refresh that end-state and undertake update routines.[171] Is this a millisecond occurrence (in order to approximate the human brain) or is it necessary to introduce

[166] David Lopez-Paz, 'Towards a Learning Theory of Cause-Effect Inference', *arXiv*.1502.02398, (9 February 2015), p. 1 <https://arxiv.org/pdf/1502.02398.pdf>.

[167] R Sun, 'Connectionist and Symbolic Approaches', *University of Missouri-Columbia*, (November 2000), generally <http://www.cogsci.rpi.edu/~rsun/sun.encyc01.pdf>. For the purposes of this section, 'information' here refers to data provided by, inter alia, range sensors, state sensors as well as sensors monitoring environmental, storage, hardware and operational parameters.

[168] In contrast to narrowly defined factory applications. Actions available to the weapon should reflect an unconstrained range of goal-directed tasks that are only then tempered by limitations imposed by its environment as well as boundaries imposed by the Delivery Cohort.

[169] Haikonnen, p. 168.

[170] Dietterich, '*Machine Learning for Sequential Data: A Review*', pp. 230-233.

[171] This is a complex procedure given that sensors may be passive/active, mobile/static/remote and vary in complexity, power and modes of operation. Its complexity may require that game theory is needed to ensure 'best group performance' in an AWS' sensor portfolio. See: P Jones, 'An Iterative Algorithm for Autonomous Tasking in Sensor Networks', *Decision and Control, IEEE Conference paper*, (2006) <http://ieeexplore.ieee.org/document/4177379/> [accessed 8 July 2017].

buffering to prevent data overload and allow more extensive processing but thereby increase the risk of performance stutter? Given AWS' twin requirements of compliance and utility, the decision mechanism to determine when the weapon's *intermediate* results become *final* results is a further complexity.

Similarly intractable are the set of sub-routines necessary then to support the weapon's processing functions. While a simple switch sensor might indicate its state by, perhaps, voltage variation in its control circuit, AWS' actions will depend on extensive processing of data drawn from multiple sensors, from internal memory, from external feeds and all across what must remain an undefined time continuum.[172] The weapon must incorporate subroutines in order, notes Housien, to amplify (or deny), filter, scrub, classify and manage such data before progressing towards decision and execution.[173] This creates sequencing issues. Middleware processes are particularly processing-intensive.[174] As pointed out by Abdulhafiz, scrubbing routines generate material complexity given data noise, missing values (as well as the dynamic need to bridge data in order 'to fill holes'), data inconsistencies, uncertain data, ambiguous and conflicting data and duplicated data.[175] The weapon's data-handling (termed ETL or Extraction, Transformation and Loading) can be further corrupted during data merging as well as process effects from representing what may be very similar and therefore unexpectedly undifferentiated information. Similarly, Yadav notes that currently available scrubbing methodologies (Nearest Neighbour, Clustering, Greedy and Rules-based routines[176]) each introduce specific deficiencies into machine data preparation including memory constraints, degraded run-time performance, certain 'common variant problems' that cannot yet be solved, premature and local convergence of data, random handling of cluster data and even the requirement for manual intervention.[177] To this point, Zhao questions whether this is yet a properly automated process.[178] Furthermore, the imprecise and heterogeneous nature of such primary data makes machine output particularly prone to error.[179] The slowest and most general phase of the signal-to-symbol process (especially given the scope of competing datasets generated either on-board, from colleague machines or received externally), data scrubbing is likely to be an enduring constraint in the development of unsupervised weapons.

---

[172] See: Chapter 9 (*Hardware*'), specifically 9.1 ('*Hardware and sensor fusion issues for AWS*').

[173] HI Housien and others, 'A comparison study of data scrubbing algorithms and frameworks in data warehousing', *International Journal of computer applications,* Central Southern University, Changsa, China, (0975-8887), 68, 25, (April 2013) <http://research.ijcaonline.org/volume68/number25/pxc3887406.pdf>.

[174] Economist Magazine, 'Intel on the Outside: The rise of artificial intelligence is creating variety in the chip market', *Economist*, 25 February 2017 <http://www.economist.com/news/business/21717430-success-nvidia-and-its-new-computing-chip-signals-rapid-change-it-architecture> [accessed 2 August 2017]. See: introduction to Chapter 7 (*Firmware*). Middleware here is defined as the software mediating between the weapon's permanent and application software.

[175] W Abdulhafiz, 'Handling Data Uncertainty and Inconsistency Using Multi-sensor Data Fusion', *Siemens, Cairo*, Academic Paper, (27 May 2013), generally <https://www.hindawi.com/journals/aai/2013/241260/> [accessed 6 July 2017].

[176] Pankaj Yadav and others, 'Nearest Neighbour-based Clustering Algorithms for Large Datasets', *arXiv*.1505.05962, (22 May 2015), p. 2 <https://arxiv.org/pdf/1505.05962.pdf>.

[177] HI Housien and others, pp. 1-5.

[178] S Zhao, 'Manual versus Automated Data Validation', *Siemens commercial blog*, (9 February 2016), para. 5-7 of 11 <https://www.edq.com/blog/manual-vs.-automated-data-validation/> [accessed 6 July 2017].

[179] Osoba and Welser, generally. For a military perspective, see: S Russell and others, 'Human Information Interaction, Artificial Intelligence and Errors', *US Army Research Laboratories*, (2016), 'Association for the advancement of AI'.

This balance between processing speed and data accuracy is similarly a constraint to their deployment. What additional costs arise from this balance? Given that the weapon must work in real-time, speed may be gained by dividing the machine's sensed datasets into specific subroutines that can each be processed at a lower and thus faster level. This too has a validation and calibration cost. As noted by Levy (and Borrie), it also increases the likelihood of weapon failure through 'coupling contagion' requiring additional mediation routines to arbitrate any sewing together of data sources.[180] Weapon processes might instead require division into smaller-packet commissions to be performed by sensors that are best suited for that task. AWS might be configured to limit dynamic polling to relevant sensory information (action-oriented protocols, termed 'active sensing') in order to refine (although likely compromise) the weapon's sensory inputs.[181] In this case, weapon sensors would be directed in a direction where information is most needed or available (termed task-driven attention). The model, however, remains theoretical. Nor does the model mitigate the fundamental complexity of data scrubbing discussed above. While partition of the weapon's immediate world into perceptual categories might aid data conversion, it will require the Delivery Cohort to introduce suitable negotiation and sense-testing routines. A battlefield equivalent might involve basing people detection on the use, for example, of simple temperature, movement, colour or distance sensing: Such sensors might be materially simpler than comprehensive vision technology (and require substantially less processing) but such paring down risks compromising overall system efficacy. It is, after all, the combination of sensor inputs that will facilitate autonomous operation by providing appropriately granular information for the AWS to 'understand' its immediate environment. This process concept is termed sensor fusion to which this thesis now turns.[182]

The set of routines that comprise a weapon's state representation cannot materially be simplified.[183] As pointed out by Suchman, the terms 'world' and 'representation' are 'a very general gloss for an open horizon of potentially relevant circumstances'.[184] Data feeds, furthermore, may be complicated by human heuristics and biases unwittingly incorporated by the broad Delivery Cohort into the weapon's AI routines.[185] Termed the *Paradox of Artificial Intelligence*, this contradiction increases system intricacy.[186] First, Wang notes the complexity of data patterns arising from the scale of available data from multiple sensors will tend to encourage routine compartmentalization

[180] A Levy, 'Combining Artificial Intelligence and Databases in Data Integration', *AI Today*, Lecture notes in Computer Science, 1600, Springer, (1999), pp. 249-268 <https://link.springer.com/chapter/10.1007%2F3-540-48317-9_10> [accessed 2 April 2017]. A consequence of increasing complexity is the emergence of tight, little understood and non-intuitive system associations that are likely to take place between sensor inputs, processing routines and subsequent action output in AWS platform. See: John Borrie, 'Security, unintentional risk and system accidents', Panel presentation, *United Nations Institute for Disarmament Research*, Geneva, (15 April 2016). This is discussed in Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt*').

[181] W Matthews and A Gheorghu, 'Repetition, expectation and the perception of time', *ScienceDirect*, (16 February 2016), pp. 110-11 <http://www.sciencedirect.com/science/article/pii/S2352154616300420> [accessed 12 February 2017].

[182] Inferred from: Mataric, p. 79.

[183] Osoba and Welser, pp. 17-21.

[184] Suchman and Weber, 'Human-machine autonomies', p. 92.

[185] There are several methods of identifying such biases. WEAT (Word Embedded Association tests) and IAT (Implicit Association Tests) are recognised tools for isolating otherwise hidden inferences. A useful example is symbology that fails to differentiate between steam and ice given their proximity (but semantic opposite) from plain-state water.

[186] Brian Bergstein, 'The Great AI Paradox', *MIT Technology Review*, (12 December 2017) <https://www.technologyreview.com/s/609318/the-great-ai-paradox/> [accessed 3 March 2018].

in an effort to create usable data structure.[187] Second, the smoothing practice of removing 'confounding covariates' in weapon datasets introduces inappropriate hazard by filtering out statistically volatile data points in order to regulate whole-dataset sensitivity.[188] Other biases will likely arise from data commingling, from sample size disparity (one sensor dataset versus another), from reward functions in particular AWS sensor routines being incongruent and, finally, straightforward cultural differences in interpreting different output categories arising from this sensor fusion.[189] Examples abound, after all, of risk-estimating algorithms being based on plainly incorrect probabilistic models.[190] Additional challenges arise. Different sensors, for instance, will collect quite separate types of battlefield information requiring pre-processing in order to re-work datasets into homogenous, comparable file types that are then serviceable.[191]

It can be inferred from Lansner that 'associative processing' will be necessary for appropriate identification, selection and engagement of targets if the weapon is to integrate 'associative evocations' of battlefield representations with other representations, even incomplete ones.[192] In this way, the weapon must directly attribute its *sensed* battlefield experience with a portfolio of a-priori signals that depict and then define properties of that entity or association.[193] Garfield defines such transient linking of representations (here, battlefield shapes, characteristics and identities) as the central matter of machine attention. It is certainly complicated but also a pivotal capability if datasets are to be classified by the AWS according to common features. The challenge is that these representations are not equipped to represent properties numerically and can only generally inform whether that designated property (here, a battlefield feature) is either present or not present. This gives rise to difficulties. Should a particular representation be volatile (and thus forgotten after processing) or is it essential to the AWS' current goal set and therefore to be written

[187] R Wang, 'Active Sensing Data Collection with Autonomous Mobile Robots', *International Conference, Robotics and Automation, IEEE,* Carnegie Mellon University, (2016) <http://www.contrib.andrew.cmu.edu/~rpw/MyBioWebpage/ICRA16.pdf>.

[188] Osoba and Welser, p. 21.

[189] Osoba and Welser, p. 17. See also: M Hurston, 'Even Artificial Intelligence can acquire Biases against Race and Gender', *Science Magazine*, Science AAAS, 13 April 2017, generally.

[190] S Ashby, 'The 2007-2009 Financial Crisis: Learning the Risk Management Lessons', *University of Nottingham*, (January 2010) <https://www.nottingham.ac.uk/business/businesscentres/crbfs/documents/researchreports/paper65.pdf>.

[191] See: Chapter 9 (*Hardware*), specifically: 9.1 ('*Hardware and sensor fusion issues in AWS*'). A battlefield example is useful. Using reflected light might appear to offer a simple sequencing solution to detecting the presence of a target object, the distance to that target, some detail on that target's surface and to recognise other embedded features. A weapon's reflectance sensor, however, is unexpectedly complicated. Light reflectivity is affected by the target's colour, smooth or rough texture and other surface properties. Light reflection depends upon surface colour and is therefore less reliable in detecting dark objects. Similarly, the reflectance sensor must ignore ambient light in order to be sensitive only to its own emitter's reflected light. While this single subroutine may be undertaken by the weapon using multiple sensor readings, the level of processing then required (for what, after all, is a single component of a very complex sequence) demonstrates how difficult it will be to process even quite basic datasets into a useable form. This would be effected by undertaking one pass with the emitter on and one with it off and then subtracting one from the other having first adjusted the data in order not to conclude with a negative light. A useful primer on reflectance can be found at: 'Basic Principles of Surface Reflectance', generally <https://www.cs.cmu.edu/afs/cs/academic/class/15462-f09/www/lec/lec8.pdf>.

[192] A Lansner, 'Associative Processing in Brain Theory and Artificial Intelligence', *Springer Link,* Conference paper, (1986) <https://link.springer.com/chapter/10.1007/978-3-642-70911-1_12> [accessed 4 May 2017].

[193] J Garforth, 'Executive Attention, Task Selection and Attention-Based Training in a Neurally Controlled Simulated Robot', *Neurocomputing*, 69.16, (2006), pp. 1923-1945.

to the weapon's updated action list? The challenge is that the detail and intensity of such representations are just not consistent between (or, actually, throughout) engagement routines.

AWS planning routines must look ahead to the *outcomes* of possible actions, prioritising and incorporating this analysis in order to create a sequence of actions that moves the weapon towards the desired (yet still compliant) goal of the Delivery Cohort. The complicating software feature of autonomous planning routines is noted by McNaughton: They must of course be parallel-running as well as being reliably complimentary to the weapon's currently set goals.[194] This planning cycle is not simple. To this point, a search routine by the weapon's navigation kernel may prioritise the *shortest* path for AWS movement given time and battery power considerations. But other action criteria such as safety, isolation, proximity to friendly forces as well as other topographical advantages (concealment and feint) must also form part of that planning and optimization routine.[195] Any such search must look for multiple *and* parallel solutions but, as the number of possible available states becomes ever larger, planning becomes inappropriately slower and less reliable. In this vein, a planning sequence that takes longer to solve might then be based on a weapon's dataset that may itself already be outdated. For humans, this is instinctively dealt with by experience. There will similarly exist incongruous incentive for the weapon to exercise planning sequences as infrequently as possible (in order to reduce lag and error) which is likely to compromise further the weapon's IHL compliance. Furthermore, AWS' decision routines assume that the weapon's immediate environment remains in tolerance with set probabilities and does not change during mid-sequence in a way that affects outputs. It can similarly be inferred from Simpson that AWS routines will need to include detailed rules for all possible scenario combinations in order to ensure that no such mutually exclusive conditions are created, the combinatorial effects of which will generate unmanageable complexity.[196]

8.5 Anchoring and goal setting issues

The software challenges identified thus far stem primarily from the weapon's data capture and processing sequences rather than subsequent action or feedback routines. One such function relates to the 'anchoring', the degree by which the weapon's current representation is amended to reflect recently polled data. Anchoring is also susceptible to a cognitive bias whereby the independent weapon relies too heavily upon an initial piece of information when making engagement decisions.[197] An adjunct of the bias is that the weapon attaches too much importance to one aspect of a circumstance (here, a battlefield episode) causing error in that weapon's

[194] For an example of parallel decision processes in robotics, see: M McNaughton, 'Planning algorithms for real-time motion planning', *Carnegie Mellon University*, Dissertations, 7-2011, Paper 179, undated <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1180&context=dissertations> [accessed 9 May 2017].

[195] Such processes must be themselves be undertaken by specific software routines that will each require mediation before integration into decision and action sequences.

[196] E Simpson et al, 'Dynamic Bayesian Combination of Multiple Imperfect Classifiers', *Decision Making and Imperfection,* 474, (2013), pp. 1-35 <http://www.robots.ox.ac.uk/~sjrob/Pubs/galaxyZooSN_simpson_etal.pdf>. An example of such conflict might arise from basing AWS control routines upon design-time rather than run-time whereby a *complete* set of rules must be pre-programmed into the weapon notwithstanding the difficulties of ensuring that this set of rules is both appropriate and current.

[197] Amos Tversky and Daniel Kahneman, 'Judgement under Uncertainty: Heuristics and Biases', *Science*, New Series, 185, 4157, (27 September 1974), pp. 1124-1131 <http://www.its.caltech.edu/~camerer/Ec101/JudgementUncertainty.pdf>.

subsequent prediction of utility of a future outcome.[198] For the purposes of this section, such anchoring processes must execute the appropriate degree of change between what is the current weapon state and an amended updated weapon state that is suggested by the platform's application layer.[199] The heuristic, notes Kahneman, has ancillary effects, several of which will be directly relevant to the Delivery Cohort's prescriptions for AWS control routines: In order (again) to be valuable, weapon processing must be appropriately sensitive to sample size, to prior probabilities of outcomes (here, 'base rate frequency') as well as the roles of chance and relevant probability distribution.[200] Similarly, the unsupervised machine must be *in*sensitive to illusory correlation and biases of 'retrievability' and 'imaginability' (termed by Kahneman 'the availability heuristic').

Anchoring in AWS will be complicated by a requirement that the weapon 'knows' how and where it is located at any (and every) point in its decision process.[201] In the human brain, the soldier's vantage point arises from his ability to attribute sensations to a point of origination. This capability is complicated, however, to mimic in a weapon as it requires the machine to detect visual and auditory stimuli (including directions and distances) and then incorporate that data as dynamic digital features into its routines.[202] It will involve the weapon managing mental maps and, inferred from Haikonnen, an ability to toggle between externally and internally generated sensations.[203] The common challenge is to fix the weapon into what Dragan terms 'a localized observer' in order to ensure that all routines treat the implied locus of the weapon in a homogenous manner.[204] This construct raises fundamental questions. Bonarini notes that it is computationally challenging for the weapon to perceive the origination point of such stimuli as being somewhere *other* than the point of that particular on-board sensor.[205] Inappropriate adjustment of the weapon's anchoring routines will also impact how the AWS fixes itself within those surroundings as well as likely creating iterative conflict.[206] Challenge in this instance is evidenced by the required extent of processing during which any of these heuristics can interfere with the Delivery Cohort's intentions.[207] A further complication, identified by Sculley and discussed in the previous chapter, is

198 This aspect of the heuristic is termed the 'focusing effect'. See: Daniel Kahneman and others, 'Would You Be Happier if You Were Richer? A Focusing Illusion', *CEPA Working Paper*, 125, (May 2006), generally <http://www.morgenkommichspaeterrein.de/resources/download/125krueger.pdf>.

199 Edward Teach, 'Avoiding Decision Traps', *CFO*, (17 June 2004), generally <http://ww2.cfo.com/human-capital-careers/2004/06/avoiding-decision-traps/> [accessed 12 May 2017].

200 Tversky and Kahneman, 'Judgement under Uncertainty', generally.

201 This aspect of the heuristic is termed the 'adjustment effect', and empirically suggests such incremental adjustments are usually insufficient. See: Tversky and Kahneman, 'Judgement under Uncertainty', p. 1128.

202 Gretchen Chapman and Eric Johnson, 'Anchoring, Activation and the Construction of Values', *Organizational Behaviour and Human Decision Processes*, 79, 2, (August 1999), p. 115 ('*Abstract*').

203 Inferred from: Haikonnen, p. 74.

204 A Dragan and others, 'Integrating Human Observer Inferences into Robot Motion Planning', *Robotic Institute*, Carnegie Mellon University, in Autonomous Robots, Springer 37.3, (2014) <http://www.ri.cmu.edu/pub_files/2014/7/legibility_AURO14.pdf>.

205 Andrea Bonarini and others, 'Concepts for Anchoring in robots', *Congress of the Italian Association for artificial intelligence*, conference paper, (September 2001) <https://link.springer.com/chapter/10.1007/3-540-45411-X_34> [accessed 23 May 2017].

206 Hanan Shteingart and others, 'The Role of First Impressions in Operant Learning', *Journal of Experimental Psychology*, 142, 2, (2013), 476-477 <https://elsc.huji.ac.il/sites/default/files/476.pdf>. Here, weapon routines may be compromised by a 'primacy effect' whereby recall of primary (initial) information is more

207 Haikonnen, p. 71. This primarily comprises weapon distinction between precepts dating to its deployment-day settings, between precepts subsequently acquired from sensor feedback, from its immediate external world and, finally, from output that is either generated by subsequent ML or received from external sources. Finally to this point,

introduced by ramifications arising from any such changes (here, the CACE Principle whereby *Changing Anything Changes Everything*).[208] The conclusion is that this breadth requires human intervention in order to provide a governable means whereby these changes in weapon state can be regulated, under-scrutinized data dependencies can be mediated and erosion of the Cohort's intended boundaries can be managed.

Similarly intractable is the Cohort's need to set *goals* to impose the weapon's priorities, responses and action selection.[209] Pollock distinguishes here between values and goals in machine autonomy: Goals, he argues, prompt an intelligent system to develop plans of action while values enable it to assess the comparative merits of such plans.[210] If this process is inappropriately undertaken, the weapon will either be illegal or useless. The difficulty is to ensure that goal satisfaction mirrors the intentions of the defining Delivery Cohort. This and the following section now consider two such dynamics in configuring artificial intelligence-directed weapon systems: First, the issue of setting and updating goals for such machines and, second, the issue of setting and maintaining values under which such weapons should operate. AWS tasking, after all, will generally be determined by a generated priority order that governs what must be undertaken at once, what should be undertaken next, the resumption of a task that was previously discontinued and, more complex for the weapon, what actions should *subsequently* take place based on sensor feedback in order to capitalize on battlefield opportunities.[211] As noted by Verschule, goal setting is not simply a process of tuning the machine's reward signal as several issues exist to complicate this relationship.[212] A primer on the scope of weapon goals can be seen in the set up required in configuring friendly-forces in an on-line video war-game.[213] Configuration steps are complicated by the cascading and interrelated nature of that process. Furthermore, errors in goal setting may have unforeseen battlefield ramifications including logistical bottlenecks whereby 'more ammunition requires more logistics requires more infrastructure requires more safe areas'.[214] Inappropriate goal setting may also compromise weapon priorities with similar operational implications ('wider engagement parameters will require more intelligence will require more data processing which will

anchoring must also take into account the effects of data entanglement, correction cascades and undeclared data consumption.

[208] Sculley and others, pp. 2-4. See also: discussion on CACE, Chapter 7 (*Firmware*), specifically: 7.1 ('*Technical debt'*).

[209] P King and others, '*Int*elligence is turning out to be a computational problem. What about goal-setting? Is that a uniquely human endeavour?'*, Quora magazine*, 6 May 2016 <https://www.quora.com/Intelligence-is-turning-out-to-be-a-computational-problem-What-about-goal-setting-Is-that-a-uniquely-human-endeavor> [accessed 2 July 2017]. For a discussion on human-robot decision waterfalls, see: P Schermerhorn and others, 'Dynamic robot autonomy; investigating the effect of robot decision making in a human-robot task team*', ICMI-MLMI,* (2009) <https://hrilab.tufts.edu/publications/schermerhornscheutz09icmi.pdf>.

[210] JL Pollock, *Thinking about acting: Logical foundations for rational decision-making,* (Oxford: Oxford University Press, 2006), generally. See also: Angie Hunt, 'Are 'Machine Values' Replacing our Principles?', *Futurity blog*, (19 April 2017), generally <https://www.futurity.org/technology-machine-values-1406692-2/> [accessed 15 January 2019].

[211] Scott Drew Pendleton and others, 'Perception, Planning, Control, and Coordination for Autonomous Vehicles', *MDPI*, Machines, 5.1, 6, (2017), pp. 16-13 ('*Section Three, Planning'*).

[212] Inferred from: P Verschule and others, 'The why, what, where, when and how of goal-directed choice: Neuronal and computational principles', *Philos Trans Royal Society London*, 1655-20130483, (5 November 2014) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4186236/> [accessed 12 July 2017].

[213] See, for example: Mihir Sheth, *'16 Expert Tips for Conquering God of War's Difficulty Mode'*, *PlayStation Blog*, (16 May 2018) <https://blog.eu.playstation.com/2018/05/16/16-expert-tips-for-conquering-god-of-wars-brutal-give-me-god-of-war-difficulty-mode/> [accessed 2 October 2018].

[214] Derived from Sheth. The analysis clearly assumes the deployment of very broad task weapon autonomy.

require, counter intuitively, more human intervention and intermediate engagement steps'). 'Infrastructure profusion', for instance, relates to a machine unexpectedly transforming disproportionately large parts of its reachable resources into the service of an inappropriate internal goal.[215] In the case of AWS, such requisitioning behaviour would have the side-effect of unbalancing resource allocation across the battlefield and is best illustrated by Bostrom's paperclip analogy.[216] Such error has consequences given that there is likely no reason (or, indeed, mechanism) for the AWS to *cease* activity on achieving that flawed goal. On the contrary, it can be inferred from Bostrom that the weapon, if it is a logical Bayesian agent, will 'never assign exactly zero probability to the hypothesis that it has not yet achieved its goal'.[217]

Russell notes that complexity in the weapon's goal-setting is also evidenced by current workarounds that are intended to solve the issue.[218] It may be that a goal definition directs the AWS to look no further once it has identified a course of action that it gives a probability of success that exceeds, perhaps, a threshold of ninety-five per cent. This *satisficing* approach is plainly inappropriate. It fails, notes Stimpson, to ensure that the weapon will select any humanly intuitive or empirically sensible way of achieving that goal.[219] Omohundro observes that ML goal generation which is based on probability thresholds is 'more likely to build in suboptimal behaviour'.[220] Grandon-Gill even conjectures that the machine's drive towards self-improvement will tend towards inherent instability in its goal setting as all of its objects are increasingly represented as 'economic' utility functions that are inappropriate for AWS' compliant deployment.[221] It will also be challenging to allocate reliable thresholds in what is likely to be a set of unfolding events: Should then a slightly diluted set of actions being coded for, say, a ninety per cent threshold? How too can the *whole* scope of an engagement sequence be captured in that threshold?[222] Finally to this point, a learning weapon platform might prioritise internal processes that have 'moral status' (notwithstanding the coding and ambiguity ramifications of such precepts) leading Yudkowsky to

[215] For a discussion on possible failure modes in unsupervised robots, see: Lesswrong.com, 'Superintelligence 12: Malignant Failure Modes', (2 December 2014) <http://lesswrong.com/lw/l9t/superintelligence_12_malignant_failure_modes/> [accessed 3 June 2017].

[216] Inferred from: Bostrom, *Superintelligence,* p. 123. Again, this analysis assumes the deployment of very broad task weapon autonomy. Here, an AI, designed to manage production in a factory, is given the final goal of maximizing the manufacture of paperclips and then proceeds to convert the whole Earth into paperclips.

[217] Ibid., p. 125.

[218] For a useful discussion on the role of goal-setting in AI control, see: S Russell, 'Rationality and Intelligence: A brief update', *Berkeley University School of Computer Science*, undated <https://people.eecs.berkeley.edu/~russell/papers/ptai13-intelligence.pdf>.

[219] Inferred from: J Stimpson and others, 'Learning to cooperate in a social dilemma; a satisficing approach to bargaining', *Proceedings of the 20th International Conference on machine-learning,* Brigham Young University, ICML-03, (2003) <https://www.aaai.org/Papers/ICML/2003/ICML03-095.pdf>.

[220] S Omohundro, 'The Basic AI Drives', *Self-aware Systems*, Paulo Alto, California, pp. 1-4 <https://selfawaresystems.files.wordpress.com/2008/01/ai_drives_final.pdf>. Omohundro's work on AI agents protecting their economic utility function is developed further in his paper to the 2007 Singularity Summit: 'The Nature of Self-Improving Artificial Intelligence' <https://selfawaresystems.com/2007/10/05/paper-on-the-nature-of-self-improving-artificial-intelligence/> [accessed 6 April 2017].

[221] T Grandon Gill, 'A Psychologically Plausible Goal-Based Utility Function', In*forming Science: The International Journal of an Emerging Transdiscipline*, 11, (2008), 228. Inferred from: Omohundro, p. 1 and pp. 3-5.

[222] Source: 'UK Tactical Aide Memoire (TAM) Part 2', Issue 3.0, (1998); in particular, fire discipline parameters including arcs of fire, authority routines before opening fire, STAP, priority of targets, controlled rates of fire, ammunition conservation and target indication.

highlight that such routines may be particularly prone to value-loading.[223] This is complicated on several levels, not by an operational imperative that weapon goals be both dynamic and emergent. Indeed, it will be impossible to test a-priori the validity, utility or entirety of such weapon goals.[224]

Similarly, goal-setting will have broad deployment ramifications. While a soldier's actions are influenced by complex and often impossible-to-define tenets such as leadership, empathy, experience, training and team dynamics, the weapon's behaviour must instead be directed through code-based machine learning. In this vein, Konidaris specifies a 'motivational framework' based on a numerically comparable reward mechanism but such goal setting models remain theoretical.[225] AWS, for instance, will have complicating short term and long-term goals with each such object calving sub-goals, each with their own hierarchical targets that require moderation and prioritising. For this reason, Bikakis notes the challenging prerequisite of an appropriate governor to mediate between goal conflicts.[226] Finally to this point, a reliable 'fatigue process' is then necessary which allows the unsupervised weapon to recognise setbacks, to re-prioritise data feeds and adjust subsequent weightings, introduce appropriate inhibition routines as well as inform (without compromising) the weapon's overall goal regime.[227]

8.6 Value setting issues

Developing a model that integrates values into decision sequences will be similarly challenging, particularly in a manner that allows the weapon to learn and dynamically refine those values, a key component (inferred from Cohn and others) of AWS function in a battlefield setting.[228] In order to appreciate its challenge, it is useful to rely on Cooper's value assessment for the deployed AWS: At one end is a values-based framework based on general probabilistic inference while at the other end is a portfolio of special-case, average-case and other approximation algorithms.[229] Neither,

---

[223] For a discussion on value loading in AI agents see: E Yudkowsky, 'The Value Loading Problem', *Machine Intelligence Research Institute*, undated <https://www.edge.org/response-detail/26198> [accessed 17 June 2017). A second more existential challenge emerges: In running a huge array of simulations, the weapon might then discard earlier programmed iterations, 'gloriously' altering its initial setup under what Bostrom calls 'mind crime'. See: Bostrom, *Superintelligence,* p. 139. On ML ramifications to this point, see: Chapter 7 (*Firmware*), specifically: 7.2 ('*Firmware ramifications of learning methodologies*').

[224] Although written in 1987, see: P Ranky, 'A summary of Robot Test Methods with Examples', *University of Michigan*, RSD-1-87, 1987, p. 7. An adjunct here is that goal setting must incorporate a reliable *temporal* dynamic. Furthermore, the construct of 'perverse instantiation' posits a possible conflict between a weapon's theoretical rush to achieve its goals notwithstanding its Delivery Cohort's a-priori intentions.

[225] G Konidaris, 'An Adaptive Robot Motivational System', *University of Massachusetts at Amherst*, undated <http://www-anw.cs.umass.edu/pubs/2006/konidaris_b_SAB06.pdf>.

[226] A Bikakis, 'Alternative strategies of conflict resolution in multi-context systems', *International Federation for Information Processing,* Conference paper, 296, Springer, Boston MA, (2009), p. 4 <https://link.springer.com/chapter/10.1007/978-1-4419-0221-4_6> [accessed 12 May 2017]. See also: M Muraven, 'Goal Conflict in Designing Autonomous Artificial Systems', *University of Albany*, (18 March 2017), pp. 1-5 <https://arxiv.org/pdf/1703.06354.pdf>.

[227] A Wigfield, 'Expectancy Value Theory of Achievement Motivation; a developmental perspective', *Educational Psychology Review*, 6, 1, (1994), p. 50. Wigfield discusses the relative acceptability to a party (inferred as the AWS) of succeeding or failing at that task.

[228] For a useful discussion on values in AI, see: A Cohn, 'How do we align artificial intelligence with human values?, *Future of Life Institute*, (3 February 2017) <https://futureoflife.org/2017/02/03/align-artificial-intelligence-with-human-values/> [accessed 13 December 2017].

[229] G Cooper, 'The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks', *Knowledge System Laboratories*, Stanford University, (1990), p. 1 <http://www2.stat.duke.edu/~sayan/npcomplete.pdf>.

however, provides an appropriate value framework for AWS deployment. The first is intractably complicated (if the battlefield asset is to retain utility and trust within the Delivery Cohort) while the second relies on approximation procedures that are ill-suited to ensuring LOAC compliance.[230]

Current value-setting for autonomous agents is predicated either upon a simple scaffolding approach (based upon a series of interim values en-route to the weapon 'landing' on a final set of values) or, more likely, a model that instead retains an *unchanging* set of values throughout the weapon's learning and then operational phases. In this way, subsequent experience does not change the AWS' final goal with learning routines instead conditioning the weapon's *beliefs* about the Cohort's values and goal.[231] This, however, creates challenge. Again, human context is important. As noted by *Wired*, '[w]hen engineers peer into a deep neural network, what they see is an ocean of math: a massive, multilayer set of calculus problems that—by constantly deriving the relationship between billions of data points—generate guesses about the world'.[232] How can diverging goals and value settings be understood by the battlefield commander or Delivery Cohort given the complexity of determining which experiences should update goal and weapon values?[233] Moreover, how are AWS values and goals to be calibrated when individual weapons may be at materially different stages of development? Similarly, what mechanisms are available to bring the AWS back into line should erroneous goal development lead to general (even slight) malfunction? A trivial aberration (perhaps a wrinkle en route to an updated goal) might otherwise develop into a material divergence from the Delivery Cohort's intentions for its weapon.

Value setting creates its own idiosyncratic issues. As inferred from Nof, it is empirically implausible that front-line mechanics can reliably reconfigure the AWS' value calibration, especially if unable to communicate with that platform.[234] Adjustment to weapon value settings will therefore require several unlikely routines.[235] Adjustment, after all, must account for *all* of the weapon's conditional probabilities and all expected utility outcomes with the mechanic then repeating his procedures for every possible course of action facing the AWS.[236] This is currently infeasible with no visibility on when such capabilities might be possible.[237] It questions how the learning AWS will

---

[230] L Busoniu and others, 'Reinforcement Learning and Dynamic Programming using Function Approximation', *Delft Center for Systems and Control,* Netherlands, (November 2009), p. 2 <https://orbi.ulg.ac.be/bitstream/2268/27963/1/book-FA-RL-DP.pdf>. See also: Chapter 7 (*Firmware,* specifically: '*Machine learning and sources of technical debt*' and '*Firmware ramifications of learning methodologies* and *Reasoning and Cognition Methodologies*').

[231] Inferred from: Bostrom, *Superintelligence,* p. 192.

[232] Tanz, pp. 1-5.

[233] Laurent Orseau and Mark Ring, 'Self-Modification and Mortality in Artificial Agents', *International Conference on Artificial General Intelligence*, Lecture Notes on Computer Science, 6830, (2011), Springer, Berlin, '*Abstract*' and generally.

[234] Maintenance challenges are well illustrated by: S Nof, *Handbook of Industrial Robotics,* (USA: John Wiley & Sons, 1999, generally. For discussion of command and control in a communications-denied environment, see: Ramon Ahrens, 'Mission Control in a Communications Denied Environment', *Air War College*, (16 February 2011), generally <http://www.dtic.mil/dtic/tr/fulltext/u2/1036912.pdf>.

[235] Considerations here arise from on-line and off-line repair processes, field repairs and servicing, replenishment and restocking, routine maintenance, system updating and unit retrieval processes.

[236] Inferred from: Nick Bostrom, 'Hail Mary, Value Porosity, and Utility Diversification', *nickbostrom.com*, (19 December 2014), p. 4 <https://nickbostrom.com/papers/porosity.pdf>.

[237] James Somers, 'The Coming Software Apocalypse', *Atlantic*, (26 September 2017), generally <https://www.theatlantic.com/technology/archive/2017/09/saving-the-world-from-code/540393/> [accessed 6 December 2018].

pursue exactly those values intended by its Delivery Cohort. The weapon's values 'must act according to an appropriate incentive system as the repository of an appropriate final value'.[238] This, however, posits a circular argument; in retrospect, the weapon's values framework must similarly be revisable as the weapon refines its own representation on the basis of amended intelligence about its world and environment.[239] This learning characteristic is pivotal and presents the weapon with at least three further goal and value quandaries. First, it will be difficult for the unsupervised weapon to incorporate new data if it must remain slaved to an unchanging final value that is coded from its initial deployment. Second, it is unfeasible for the Delivery Cohort to specify reliably a set of value actions for *every* such scenario.[240] Finally to this point, the impact on the broad value-loading problem from these descriptors is *cumulative*.[241]

Christiano notes that 'reward maximisation' is an inadequate basis for defining these values.[242] In this case, the intention is for a reward framework to screen the weapon's battlefield and environmental experiences prior to selective update of that weapon's value system. The model is challenging.[243] How, notes Song, can sufficient priority be coded for *complementary* stimuli in the process?[244] Its weaknesses are clear-cut: While humans' ability to effect a workable goal-acquiring is innate, it is not reducible to coding. Similarly, while the Delivery Cohort may be able to articulate on paper a values waterfall to guide AWS operation on deployment, the untested challenge (as inferred from Alami) is its reliable translation into lines of code in a manner that is not immediately lost in that weapons' subsequent learning.[245] LeCerra and Bingham conclude that such learning models are 'so closely tailored to human neurocognitive architecture that it is not transferable across machine intelligences except where those platforms would one day be based on whole brain emulation'.[246]

It is also unlikely that the Delivery Cohort would want to construct weapon-controlling AI with exactly the same value disposition as a human. Zawieska notes, after all, that flawed human nature is subject to arbitrary (and even evil) traits inappropriate to LOAC-compliance.[247] The Cohort's

---

[238] Daniel Dewey, 'Learning What to Value', *MIRI*, (2011), p. 3 <http://www.danieldewey.net/learning-what-to-value.pdf>.

[239] Eliezer Yudkowsky, 'Complex Value Systems in Friendly AI', *International Conference on Artificial General Intelligence*, Lecture Notes on Computer Science, 6830, (2011), Springer, Berlin, Heidelberg, '*Abstract*' and generally.

[240] Bostrom, *Superintelligence,* p. 185.

[241] Eliezer Yudkowsky, '2015: What do you think about machines that think?', *The Edge*, 2015, generally <https://www.edge.org/response-detail/26198> [accessed 5 May 2017].

[242] Paul Christiano, 'The Reward Engineering Problem', *AI Alignment*, (30 May 2016) https://ai-alignment.com/the-reward-engineering-problem-30285c779450> [accessed 6 May 2017].

[243] Ibid.

[244] HF Song and others, 'Reward-based Training of Recurrent Neural Networks for Cognition and Value-based Tasks', *ELife*, (13 January 2017), generally <https://elifesciences.org/articles/21492> [accessed 4 May 2017].

[245] R Alami and others, 'Toward Human-Aware Robot Task Planning', *LAAS-CNRS, American Association for Artificial Intelligence*, (2006) <http://www.aaai.org/Papers/Symposia/Spring/2006/SS-06-07/SS06-07-006.pdf>. For a detailed discussion on coding and issues, see: this chapter, specifically: 8.1 (*'Coding methodologies')* and 8.2 (*Coding errors')*.

[246] Peggy LaCerra and Roger Bingham, 'The Adaptive Nature of Human Neurocognitive Architecture: An Alternative Approach', *PNAS*, (September 1998) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC21635/> [accessed 18 August 2017].

[247] K Zawieska, 'Do Robots Equal Humans? Anthropomorphic Terms in LAWS', *Industrial Research Institute for Automation and Measurement, PIAP*, undated

choice of weapon values must instead navigate through a thicket of philosophical problems. This is particularly true when the weapon's decision context is unfamiliar. In deciding upon action paths, should the AWS use causal decision theory, evidential decision theory, 'updateless' decision theory or something quite else? The consequence of giving a lethal autonomous weapon either flawed or inappropriate decision rules is likely to be significant. Similarly, the Cohort's value definition must be consistent with 'understood' and normative battlefield standards in order to avoid heterogeneity and maverick action.[248] Getting this wrong might occasion irrevocably bad decisions including, inter alia, the weapon rewriting itself to run thereafter on an unforeseen basis. Harman notes that coding values are difficult precisely because human goal representations are so complex, imprecise and evolutionary.[249] Nor is the model complimentary to how humans work: Human goals, after all, may or may not develop from prior actions with values then acting (or not) as triggers and with goal upgrades happening either in real-time of after an indeterminate delay. Value-setting in machines, notes Bostrom, is enduringly inexact and not at all reducible to code.[250]

8.7 Action selection issues

How then might such goals and values be *calibrated* within AWS' tasking? The desired state of the weapon, also known as its goal state, is the benchmark to which the Delivery Cohort will set its weapon system. The challenge is that the AWS will require elaborate feedback in order to achieve and then maintain that set-point by continuously comparing its current battlefield state with this desired state.[251] Furthermore, once an action task is completed, the AWS must then conclude that particular delivery routine by integrating appropriate lessons and, as necessary, feedback to the Cohort. Maintenance tasks, conversely, require ongoing active work (including arbitration and resource allocation) by that weapon. Such clear delineation between these actions is unlikely to be obvious, requiring complicating arbitration in order to mediate the process.[252] The difference between the current and desired states will be the weapon's observed error.[253] The goal of the AWS' action selection will then be to *minimize* that error. A sequence of feedback controls, each an individual routine that must be integrated into weapon goal-setting and action selection, will then

<http://www.unog.ch/80256EDD006B8954/(httpAssets)/369A75B470A5A368C1257E290041E20B/$file/23+Karolina+Zawieska+SS.pdf>.

[248] Larry Lewis, 'Redefining Human Control: Lessons from the Battlefield for Autonomous Weapons', *CNA Washington*, (March 2018), pp. i-vi <https://www.cna.org/cna_files/pdf/DOP-2018-U-017258-Final.pdf>.

[249] For a useful primer on philosophical pitfalls, see: G Harman, 'Artificial Intelligence and some Philosophical Issues', *University of Princeton*, (4 October 2005) <https://www.cs.princeton.edu/courses/archive/fall05/cos402/readings/harman.pdf>.

[250] Bostrom, *Superintelligence,* p. 186; Bostrom's narrative is useful to evidence this infeasibility: 'From a noisy time series of two-dimensional patterns of nerve findings, the visual cortex must work backwards to reconstruct and interpret three-dimensional representations of external space. A sizeable portion of our precious one square meter of cortical real estate is required to process visual information, billions of neurons of working ceaselessly to accomplish this task… how could our programmer transfer this complexity into a utility function?'

[251] See: Hado van Hasselt, 'Reinforcement Learning in Continuous State and Action spaces', *Reinforcement Learning,* Springer, Berlin, Heidelberg, (2012), pp. 1-2 and p. 6 ('*Function Approximation*') <https://www.researchgate.net/profile/Hado_Van_Hasselt2/publication/239843999_Reinforcement_Learning_in_Continuous_State_and_Action_Spaces/links/0c9605220e5949a8a4000000/Reinforcement-Learning-in-Continuous-State-and-Action-Spaces.pdf>. Also see: Mataric, p. 121.

[252] See: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt'*) and 7.2 ('*Firmware ramifications of machine learning*).

[253] For a useful discussion on error types and their ramification on function in AI, see: D Allchin, 'Error types', *Perspectives on Science,* 9: 38-59, undated <http://members.tcq.net/allchin/papers/e-types.pdf>.

compute an output range in order to assist the unit in maintaining goal state.[254] Two issues arise. The pace of such error correction is not obvious and depends on how often the error is computed and how much correction is then made on each feedback loop.[255] While feedback control may play a role at low level tasking (continually moving actuators, for example) it is also significantly less adroit at modifying the weapon's higher-level action selection such as navigation, longer-term coordination, environmental interaction, collaboration and human-robot interaction.[256] Finally to this point, in selecting actions, the weapon controller must be front-facing and determine, therefore, its system state based on sub-goals set for itself *ahead of time*.[257] This too appears an intractable challenge given that closed-loop feedback systems rely upon homogeneity in in the weapon's immediate environment: As that environment becomes less anticipatable, weapon performance will certainly degrade.[258]

A further key to AWS feasibility is therefore its ability to determine action selection. Its controller must, after all, select *one* such output action or a defined, well-understood combination of actions. The AWS' internal routines to decide this are termed fusion and arbitration.[259] Fusion is the weapon's combining multiple candidate actions into a single action output. Kam notes, however, that the procedure remains an enduring constraint in robotics.[260] The AWS must then monitor possible action sets concurrently in order always to be ready to respond to, or modify, its actions. Several hurdles exist as changes to the weapon's immediate environment may occur at any time. Process roadblocks, notes Lewis, may arise at several junctures in an AWS' action sequence including degradation in the weapon's primary information through spoofing and camouflage, system gridlock, sensor overload and other sources of data congestion.[261] Weapon systems must therefore be able to support parallelism, the complex ability to monitor and execute multiple actions at once.[262] Sequential processing would otherwise risk missing events that might be critical to compliant combat operation. For this reason, other action sequence models must exist for the AWS.[263] It might then be a hybrid control model that appears to offer the best of both worlds; as observed by Abraham, the AWS would thus benefit from 'the speed of reactive control and the

[254] Jessica Taylor and others, 'Alignment for Advanced Machine Learning Systems', *MIRI*, (2016), p. 3 ('*Motivations*') <https://intelligence.org/files/AlignmentMachineLearning.pdf>.

[255] Source: University of Salzburg Center for Human-Computer Interaction, *'*To err is robot: How robots learn to recognise error*'*, <https://hci.sbg.ac.at/outputs/hri-error-situations/> [accessed 10 July 2017].

[256] Inferred from: Mataric, p. 129.

[257] A technical analysis on *feed-forward* control theory is set out by: Control Guru <http://controlguru.com/the-feed-forward-controller/> [accessed 4 December 2016].

[258] Ion Stoica and others, ' A Berkeley View of System Challenges for AI', *arXiv*.1712: 05855v1, (12 December 2017), para. 3 ('*Trends and Challenges*') <https://arxiv.org/pdf/1712.05855.pdf>. See also: Paragraph 3.4 ('*AI Demands Outpacing the Moore's Law*') and 4.2 ('*Secure AI*').

[259] Mataric, p. 167. See also: Monica Nicolescu and others, 'Learning Behaviour Fusion from Demonstration', *Interaction Studies*, 9.2, (2008), pp. 319-320.

[260] M Kam, 'Sensor Fusion for mobile robot navigation*', Proceedings of IEEE,* 85, 1, undated, generally.

[261] Lewis, 'Redefining Human Control: Lessons from the Battlefield for Autonomous Weapons', pp. 21-22.

[262] Nuno Amando and others, 'Exploiting Parallelism in Decision Tree Induction', *Proceedings for the ECML/PKDD Workshop on Parallelism and Distributed Computing for Machine Learning*, LIACC, (2003), pp. 1-2 <https://www.dcc.fc.up.pt/~fds/FdsPapers/w2003_ECMLW7_namado.pdf>.

[263] Warwick, p. 111.

brains of deliberate control'.[264] In terms of LOAC compliance and practical feasibility, however, hybrid control is challenging to achieve. It requires the amalgamation of fundamentally different control protocols, different time-scales (short for reactive, long for deliberative) and different representational models (none for reactive, explicit and elaborate world models for deliberate).[265] Additionally, Flach points out that not all relevant information in AWS routines necessarily flows from the 'bottom up' (that is, from the reactive layer to the deliberative layer) with often-repeated situations likely stored away in the weapon's internal 'contingency tables'.[266] Such action models are, by definition, parsimonious in order to avoid strong a-priori assumptions introducing bias.[267] The issue is that it remains uncertain how a weapon without supervision can reconcile off-line planning with dynamically generated in-line planning.

Finally, in reviewing AWS' action selection, it is relevant to consider bottlenecks in a weapon's *planning routines*. A theoretical model might be as follows: Pre-compiled battle plans are incorporated into the AWS intended for certain narrow engagement tasks to be undertaken without supervision; such domain knowledge should require neither internal reasoning nor expensive real-time processing. What then are the relevant issues? Practically, the state space required under this methodology is simply far too large. It is unworkable to keep such plans appropriately current given a shifting battlefield (the state space, after all, of the AWS). Similarly, Konecnik notes that any change in the AWS' goal set will empirically require real-time modification of the platform's *entire* internal rule set.[268] Unsupervised toggling, moreover, between goal-determined processed actions and pre-defined rules will be precarious and, as inferred from Vershule, must anyway be compromised in a communications-denied environment.[269] Furthermore, Bengio states that any such alternating middle layer within a machine's control suite is hard to design and harder to implement.[270] It would tend towards being very special-purpose and designed for specific platform architectures thereby requiring infeasible reengineering for almost every new robotic weapon and mission, an example perhaps of Harari's adage that 'designing artificial intelligence to be the best of all worlds can end up being the worst of all worlds'.[271]

Multisensory integration might also appear to be a straightforward task of binding together battlefield datasets but their integration into a weapon's control routines (first by data amalgamation and then by data consolidation) is problematic, an unsurprising characteristic given that each such dataset must influence the perception processes of all other modalities in that

---

[264] A Abraham, 'Hybrid Intelligent Systems: Evolving Intelligence in Hierarchical Layers', *Soft Computing.net*: *Do Smart Adaptive Systems Exist?,* (2005), pp. 159-179 <http://www.softcomputing.net/gabrys.pdf>.

[265] Mataric, p. 177.

[266] Peter Flach, 'Machine Learning: The Art and Science of Algorithms that Make Sense of Data', *University of Bristol*, (25 August 2012), generally <http://www.cs.bris.ac.uk/~flach/mlbook/materials/mlbook-beamer.pdf>.

[267] L Kenal and others, 'Uncertainty in Artificial Intelligence', Elsevier, (28 June 2014), p. 384 and generally.

[268] Inferred from: K Konecnik, 'Pre-programming Artificial Intelligence is a risky business', *Daily Kos*, 10 November 2016 <https://www.dailykos.com/blog/Kenneth%20Konecnik> [accessed 12 August 2017].

[269] P Verschule and others, 'The why, what, where, when and how of goal-directed choice: Neuronal and computational principles', *Philos Trans Royal Society London*, 1655-20130483, (5 November 2014) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4186236/> [accessed 12 July 2017].

[270] Yoshua Bengio, 'Learning Deep Architecture for AI', pp. 30-34.

[271] Mataric, p. 183. See also: Yuval Noah Harari, 'Homo Sapiens as we know them will disappear in a century or so', *Observer*, 19 March 2017 <https://www.theguardian.com/culture/2017/mar/19/yuval-harari-sapiens-readers-questions-lucy-prebble-arianna-huffington-future-of-humanity> [accessed 15 June 2017].

weapon's sequencing.[272] Given data conflicts, however, it is not simply a matter of summing data inputs. A review then of how the weapon might then manage such divergence signposts the challenges arising from this final combinatory stage of the AWS' processing routine. Here, resolution strategies will likely be based on rule matching routines: Bikakis notes that these will either be *forward*-chaining (using deduction routines) or *backward*-chaining (using expert premises as additional sub goals that will then determine data fit).[273] The complexity of these rules (the responsibility of the Delivery Cohort) is a further challenge to appropriate weapon function. The weapon's rule set must either be deterministic (in an ideal but unrealistic world where, at most, no more than one rule is matched in each resolution cycle) or, more likely, non-deterministic (where multiple rule matching occurs in the engagement sequence, requiring the intricate introduction of an 'inference engine' in order to provide interpretation before determining a decision path that is based on up-to-the-second battlefield data).[274] As Yao points out, comparing one data string to another will lead to widely different outcomes.[275] The challenge is underlined by Haikonen who similarly concludes that repeated combining of sensed inputs soon means that meaningful resemblance to primary battlefield data is lost as amalgamated data is blunted by the summing of numerous prior combinations.[276] Combinatory routines introduce one additional (and potentially pivotal) challenge: As particular properties of the weapon change, the complex requirement to anchor other properties has already been noted in order to prevent wholesale alteration of the weapon's function.[277]

8.8 Behaviour setting and coordination

This thesis' analysis into AWS' deployment challenges demonstrates the prerequisite of dependable behaviour if independent weapons are to be both compliant and valuable.[278] In this vein, system behaviour is how the deployed AWS acts and reacts. The complexity is that behaviours are time-extended and not instantaneous. Their review, therefore, is appropriate as a final section on AWS software challenges precisely because they will be a complex amalgam of weapon values, goals, utility function and learning processes that are covered above. A relevant precept to illustrate their complexity, borrowed from Neukart, is by reviewing AWS' likely abstraction, the handling and transformation of available data in that weapon's given representation (whether captured directly

[272] Defense Technology Information Center, 'Artificial Intelligence and Sensor Fusion'*, International Conference on Integration of Knowledge Intensive Multi-Agent Systems ADP021440*, unclassified, (October 2003), p. 595. The paper concludes that the definition and development of multi-sensor ontology remains entirely work-in-progress.

[273] Bikakis, 'Alternative strategies of conflict resolution in multi-context systems', generally.

[274] For a primer on conflict resolution in AI systems, see: Study.com, 'Problem Solving Methods; Definitions and Types'*, Study.com*, lesson 4 <http://study.com/academy/lesson/problem-solving-methods-definition-types.html> [accessed 12 June 2017].

[275] M Yao, 'Four approaches to natural language processing and understanding', *Topbots*, (31 March 2017), pp. 3-5 <http://www.topbots.com/4-different-approaches-natural-language-processing-understanding/> [accessed 23 June 2017) The balance of the article usefully covers the complexity of instruction language.

[276] Inferred from: Haikonnen, p. 57.

[277] S Coradeschi and A Saffiotti, 'An Introduction to the Anchoring Problem'*, Article for Robotics and Autonomous Systems,* Department of Technology, Orebro University, Sweden, (2003) pp. 89-94 <https://www.cs.utexas.edu/~kuipers/readings/Coradeschi-ras-03.pdf>. See also: This chapter, specifically: 8.5 ('*Anchoring and goal-setting* issues') and its discussion of the phenomenon of *Changing Anything Changes Everything*. See also: Carson Khan, Uber's head of Machine Learning, Twitter <https://twitter.com/carsonkahn/status/735526311305699328> [accessed 3 June 2017].

[278] See, generally: Chapter 4 (*Deployment*),

from the weapon's sensors or the product of a weapon's subsequent action routine) prior to that data becoming the basis for the software routines outlined above.[279]

As inferred from Holte and Fan, it is abstraction that will cement (or not) the weapon's ability to hop between these different data sources while still preserving that data's useful properties in an action sequence.[280] Abstraction also comprises the weapon's ability to deal with 'ideas' rather than battlefield events and, for the purposes of this final section, is a convenient proxy to distill the foregoing software challenges into a general behavioural consequence for AWS deployment. Dorigo notes that it remains untested how to organise such abstraction within an AI agent (here, the unsupervised weapon) and how then to arbitrate its weightings within moment-by-moment control routines.[281] This uncertainty has several consequences for AWS behaviour. Individual behaviours must operate on compatible timescales: Incorrect calibration may compromise a weapon's sequences when, for instance, hopping between fast and slow behaviours is required. Weapon behaviour, after all, will be a direct product of AWS inputs (either from internally generated actions or from external validated sources) and, as such, is contingent upon the same processes of data efficacy, mediation and anchoring called into question above. In the AWS, abstraction and behaviours will also be complicated by the platform never reaching an 'immutable' or final state.[282] The consequence is that individual weapon behaviours may be in conflict with other behaviours and, certainly, colleague machine behaviours. As highlighted by Bonarini, machine behaviours will likely remain fuzzy, hard to isolate and, on account of such different timescales and learning paths, intractably difficult to coordinate across multiple machine platforms.[283] The same instabilities inherent in AWS' ML model will be exactly evidenced across weapon behaviour.[284] Similarly, subsequent modification of AWS behaviour will reflect those same complexities evidenced in the amending of headline goals and values once the weapon is deployed in the battlefield. A counterfactual is noted by Scharre whereby independent weaponry would otherwise be susceptible to hacking by adversaries should AWS behaviour be entirely predictable.[285]

Finally to this point, weapon behaviour requires a workable process of 'data forgetting' if the platform behaviour is not to be compromised either by information overflow or the retention of

---

[279] F Neukart, 'A Machine Learning Approach for Abstraction based on the idea of Deep Learning Belief Artificial Neural Networks', *24th DAAAM International Symposium of International Manufacturing and Automation*, (2013), pp. 1499-1508.

[280] R Holte and G Fan, 'State Space Abstraction in Artificial Intelligence and Operations Research', *AAAI Workshop,* University of Alberta, (2015), p. 55 <https://www.aaai.org/ocs/index.php/WS/AAAIW15/paper/download/10134/10234> [accessed 3 May 2017].

[281] Marco Dorigo and others, 'Evolving Self-Organizing Behaviours for a' Swarm-Bot'', *Autonomous robots,* 17.1, (2004), pp. 3-4 <http://people.idsia.ch/~luca/swarmbot-control.pdf>.

[282] For a useful discussion on shared control protocols and hardware immutability, see: T McConaghy, 'Blockchains for Artificial intelligence', *Bigchain,* undated <https://blog.bigchaindb.com/blockchains-for-artificial-intelligence-ec63b0284984> [accessed 13 June 2017].

[283] A Bonarini, 'Learning to compose fuzzy behaviours for Autonomous Agents', *International Journal for Approximate Reasoning*, 17, 4, (1997), 409-432 <http://www.sciencedirect.com/science/article/pii/S0888613X97000029> [accessed 3 May 2017].

[284] See, generally: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt*') and 7.2 ('*Firmware ramifications of learning methodologies*').

[285] Paul Scharre, 'Autonomous weapons and operational risk', *Centre for a New American Security*, (2016), p. 36 <https://www.files.ethz.ch/isn/196288/CNAS_Autonomous-weapons-operational-risk.pdf>.

sub-optimal (or wrong) data.[286] Although a main competency for feasible AWS, Foster questions whether there is an obvious model for this.[287] Several challenges exist. It is enduringly difficult to create filtering or significance criteria to manage a 'delete sequence' that is still appropriate to the embedded deployment of unsupervised weapons.[288] It is complex to validate such binary routines (durable data versus erased data?) and for the weapon to decide what constitutes primary visual evidence versus, for example, peripheral (and therefore contextual) information. Ishikawa notes one final hitch exists. While forgetting is a necessary capability in behaviour setting, the AWS should *not* generally forget acquired skills, a conundrum noted by Fratto that she terms 'catastrophic forgetting, one of the fundamental limitations of neural networks'.[289] The issue, moreover, has several levels. How should data age conflate with data redundancy?[290] The frequency of such routines' data polling will impact the weapon's redundancy calculations. Under what battlefield circumstances should the AWS be able to retrieve 'forgotten' information?

The inference used here is that machine behaviour is a useful proxy through which to evidence the *cumulative* effects of AWS' software complexity. As Rahwan points out 'we cannot certify that an AI is ethical by looking at its source code any more than we can certify that humans are good by scanning their brains'.[291] This is relevant given, as above, weapon behaviours derive *directly* from the software components and controls that comprise such machines. As autonomy becomes more pervasive, Nikerson and Reilly note the verso where machine output might conflict with output from human decision-makers.[292] Unsurprisingly, therefore, the challenge throughout is contextual and distills down to the relative confidence assigned by the Delivery Cohort to each such conflicting source of advice. The conclusion for this chapter is that intricacy is an inherent and unavoidable property of AI agents that together will be tasked with enabling weapon independence: Gershgorn notes that emergent behaviour and the inescapable consequences of AWS' ML basis leads to system behaviour that 'will actually be impossible to predict even by its own programmers'.[293] More

[286] For a general primer on unlearning issues, see: Yinzhi Cao and Junfeng Yang, 'Towards Making Systems Forget with Machine Unlearning', *IEEE Symposium on Security and Privacy*, (20 July 2015), generally <https://ieeexplore.ieee.org/Xplorehelp/#/ieee-xplore-training/working-with-documents#interactive-html> [accessed 7 May 2017].

[287] N Foster and others, 'Is awareness all the ability to forget (and to remember) critical for demonstrating directed forgetting?', *University of Illinois-Champaign*, (November 2016), generally <https://experts.illinois.edu/en/publications/is-awareness-of-the-ability-to-forget-or-to-remember-critical-for> [accessed 2 March 2017].

[288] Natalie Fratto, 'Machine Un-learning: Why Forgetting Might be Key to AI', *Hackernoon.co*m, (31 May 2018) <https://hackernoon.com/machine-un-learning-why-forgetting-might-be-the-key-to-ai-406445177a80> [accessed 6 June 2018]. In reviewing the use in unlearning of Long Short Term Memory Networks (LSMN), Elastic Weight Consolidation (EWD) and the use of Bottleneck Theory (the squeezing of data through code bottlenecks in order to retain only those features most relevant to general concepts), Fratto notes the inappropriate fitting and compressing of data that underlines such techniques.

[289] Matsumi Ishikawa, 'Structural Learning with Forgetting', *Neural Networks*, 9, 3, (April 1996), pp. 509-521.

[290] Ibid., pp. 509-511.

[291] Iyad Rahwan and Manuel Celbrian, 'Machine Behaviour Needs to be and Academic Discipline', *Nautilus*, (29 March 2018) <http://nautil.us/issue/58/self/machine-behavior-needs-to-be-an-academic-discipline> [accessed 7 October 2018].

[292] Jeffrey Nikerson and Richard Reilly, 'A Model for Investigating the Effects of Machine Autonomy on Human Behaviour', *Proceedings of the 37th International Conference in Security Science*, (2004), generally <https://web.stevens.edu/jnickerson/ETSIB01.PDF>.

[293] Dave Gershgorn, 'AI is now so Complex its Creators can't Trust why it Makes Decisions', *Quartz*, 7 December 2017 <https://qz.com/1146753/ai-is-now-so-complex-its-creators-cant-trust-why-it-makes-decisions/> [accessed 9 October 2018].

importantly, shortcomings in AWS' individual componentry (the capture, perhaps, of battlefield behaviour's basic tenets such as goals, values and utility) means that currently posited software, in the balance of probability, is likely incapable of generating appropriate actions to an extent that can win the trust of its procuring Delivery Cohort.[294] It is against this background that this thesis can now consider the hardware upon which those routines must operate and certain of the challenges that physical configurations create.

[294] See: this chapter, specifically: 8.5 ('*Anchoring and goal setting issues*') and 8.6 ('*Value setting issues'*).

9. Hardware: Build challenges to AWS function

The foregoing analysis demonstrates that preconditions must be met before a machine can function autonomously.[1] The conclusion from previous chapters is also that it will be AWS design that determines the weapon's physical or 'affect' characteristics.[2] This, however, is not a fixed relationship.[3] Nor is there a distinct division between hardware and software.[4] Instead, Impagliazzo predicts that hardware complexity (sensor capabilities and the management of multiple systems) will grow exponentially as software capabilities appear and weapon tasking moves towards autonomy including abilities that encompass 'reflection' and the sensor fusion that this entails.[5] Notwithstanding, then, what may be an increasingly artificial divide between wetware and firmware and now between hardware and software (they obviously work in tandem), the aim of this chapter is to review distinct *hardware* constraints that challenge AWS deployment, in particular those complexities that arise from posited combinations of physical equipment required for compliant (yet still expedient) weapon operation. The intention is to highlight cumulative bottlenecks that stem from machine deployment in the harshest of conditions.[6] In this vein, there is a long-established trade-off, identified by Ferrell in 1994 but nevertheless still pertinent to AWS deployment:

> Having many sensors and actuators is a double edge sword. Multiple sensors provide for reliable sensing and a richer view of the world. More actuators provide more degrees of freedom. However, more components also mean there is more that can fail and subsequently degrade performance… mechanical failure, and electrical failure or sensor failure.[7]

Such causes of failure have changed little in the intervening quarter century (as evidenced in the case of Ferrel's *Hannibal* robot by sensor signal drift, 'graceful degradation' in its hardware

[1] Haikonnen, p. 169. See: Chapter 1 (*Introduction*), generally.

[2] Specifically: Chapters 6 (*Wetware*), 7 (*Firmware*) and 8 (*Software*).

[3] Donald Norman, Andrew Ortony and Daniel Russell, 'Effect and Machine Design: Lessons for the Development of Autonomous Weapons'*, paper prepared for IBM Systems Journal,* Northwestern University, (22 July 2002), 2-4 <https://s3.amazonaws.com/academia.edu.documents/30792296/10_22_Norman5.8F.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1526204286&Signature=LJnPY02sH0nXttxfQLPiox1WBt8%3D&response-content-disposition=inline%3B%20filename%3DAffect_and_machine_design_Lessons_for_th.pdf>. Norman questions here where AWS will sit on Norman's *Reaction, Routine, Reflection* continuum.

[4] P Niranjan, 'Software and Hardware for Autonomous Robots Using Distributed Embedded System', *International Journal of Computer Applications,* 55, 11, (October 2012), 31-32. See also: Rick Osterloh, 'The Best Hardware, Software and AI –Together'*, The Keyword, Google Company Blog*, (4 October 2017) <https://www.blog.google/technology/ai/the-best-hardware-software-and-ai-together/> [accessed 15 May 2017]. Osterloh is a Senior Vice President, Hardware at Google.

[5] Russell Impagliazzo and others, 'Which Problems Have Strong Exponential Complexity?'*, Journal of Computer and System Science*, 63, (2001), 512-513 <https://cseweb.ucsd.edu/~russell/ipz.pdf>.

[6] For useful parallels in the deployment of autonomous cars, see, for instance: Darrel Etherington, 'Ford Details Some of the Big Hardware Challenges to Overcome in Self-Driving'*, TechCrunch*, (9 March 2017), para. 5 of 6 <https://techcrunch.com/2017/03/09/ford-details-some-of-the-big-hardware-challenges-to-overcome-in-self-driving/> [accessed 7 March 2018].

[7] Cynthia Ferrel, 'Fault Recognition and Fault Tolerance of an Autonomous Robot'*, Adaptive Behaviour*, 2.4, (1994), pp. 4-5 <http://web.media.mit.edu/~cynthiab/Papers/Breazeal-AB94.pdf>.

performance and challenges arising from subsequent patches in efforts to confine those errors).[8] By way of context, Gent relegates overall hardware challenges of robotics behind those of software, control, of deployment and ethical constraints and, to a degree, this chapter's relative brevity acknowledges the possible transience of hardware challenges as technologies mature, certainly compared to the systemic software constraints identified in preceding chapters.[9] Particular technologies, however, pose enduring challenge. In this vein, *National Instruments* identifies, for instance, progress in battery and actuator components as key constraints holding back delivery of fully autonomous machines.[10] Bourque further suggests that technical progress is not uniform across such hardware challenges.[11] The aim of this section is therefore to identify specific hurdles within what is a broad portfolio of hardware componentry that will be required for reliable deployment of AWS.

While Jaeger characterises the opportunity of machine autonomy as a revolution in machines' ability to predict, the hardware componentry for AWS to achieve such competence is substantial and includes a wide variety of on-board assets that must sense, garner and process the data behind such prediction.[12] These processes remain fundamentally a hardware matter based upon apposite physical sensors, the nature of which will depend on the deployment model (and hence tasks) to be carried out by that weapon.[13] It is AWS sensors, after all, that will allow the robot to know its state, the general physical notion describing itself at any point in time notwithstanding, notes Mataric, that such states may be visible, partially hidden or hidden (unobservable).[14] This points to a general constraint. As noted by Degutis, Director of Product Management at Bosch, 'all of these sensors have strengths and weaknesses', a general work-round being that the Delivery Cohort will opt to incorporate *multiple* such sensors in an effort to design out these weaknesses but, in so doing, building in technical debt and the need for systemic mediation whereby those states may then be discreet (up, down, blue, red) or continuous (a thousand miles).[15]

---

[8] Ferrel, pp. 5-6.

[9] Ed Gent, 'The Ten Grand Challenges Facing Robotics in the Next Decade', *SingularityHum.com*, (6 February 2018) <https://singularityhub.com/2018/02/06/the-10-grand-challenges-facing-robotics-in-the-next-decade/#sm.001hc0z0a1azad5etv91d6at6w0df> [accessed 10 May 2018]. Hardware considerations concern power and energy efficiencies, progress in new materials and fabrication schemes. See also: Matt Simon, 'Want Awesome Robots? You'll Have to Best These Challenges', *Wired Science*, (2 May 2018) <https://www.wired.com/story/want-awesome-robots-youll-have-to-best-these-challenges/> [accessed 19 May 2018]. See also: Economist, 'After Moore's Law: The Future of Computing – The Era of Predictable Improvement in Computer Hardware is Ending. What Comes Next?', *Economist Magazine*, (12 March 2016) <https://www.economist.com/leaders/2016/03/12/the-future-of-computing> [accessed 12 May 2018].

[10] Source: National Instruments, 27 July 2017 <http://www.ni.com/newsletter/50878/en/> [accessed 10 May 2018].

[11] Brad Bourque, 'The Tables Have Turned. Hardware Finally has to Catch up with Software', *Digital Trends*, (7 January 2017) <https://www.digitaltrends.com/computing/hardware-vs-software-ces-2017/> [accessed 12 May 2018].

[12] Herbert Jaeger, Jacobs University in Bremen, cit. Natalie Wolchover, 'Machine Learning's 'Amazing' Ability to Predict Chaos', *Wired Science*, (21 April 2018), para. 3 of 10 <https://www.wired.com/story/machine-learnings-amazing-ability-to-predict-chaos/> [accessed 23 April 2018].

[13] See, generally: Chapter 4 (*Deployment*).

[14] Mataric, p. 22.

[15] Charles Degutis, Product Director, Bosch, cit. Charles Pickering, 'How AI is Paving the Way for Autonomous Cars', *Engineer*, (15 August 2017) <https://www.theengineer.co.uk/ai-autonomous-cars/> [accessed 12 May 2018]. 'All of these sensors have strengths and weaknesses. Radar can bounce of tunnels and bridges and can struggle to differentiate small and closely spaced objects. Video can be blinded by glare. Lidar can degrade in high moisture situations'. For a primer on this relationship, see also: Niranjan, 'Software and Hardware for Autonomous Robots Using Distributed Embedded System', p. 33. On technical debt, see Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt*').

Sensors, of course, are just one part of the hardware portfolio enabling machine autonomy.[16] Effectors, for instance, enable the weapon to undertake physical actions including locomotion and manipulation.[17] A complexity is that manipulators (robot arms and grippers), in broad terms, must move in one or more dimensions. This characteristic, notes Mataric, introduces a further robotic difficulty around 'degrees of freedom' (DOF), the minimum number of coordinates required to specify completely the motion of a mechanical system.[18] Challenges arising from DOF are discussed in this chapter's next section but, while simple actuators such as motors control a single motion (up-down, left-right) of the weapon's effector, Correll points out that more complex effectors, such as a weapon's robotic arms, have exponentially more DOF that then require more complicated actuators with exponentially more complex control mechanisms.[19] It would be ideal if a robot included an actuator for every DOF but this is rarely the case and would result in unacceptable machine complexity. A further effector challenge arises from the weapon's power requirements without it being loaded down with heavy batteries.[20] Its electronics, moreover, should be isolated from its sensors and effectors and steps taken to prevent loss of performance as power levels drop or there is a sudden spike in power demand. In this vein and depending upon tasking, it can be inferred from Ieropoulis that AWS will need to replenish that power in a similarly autonomous fashion.[21] In order to satisfy this constraint, the weapon's Delivery Cohort must enable its machine to get to a particular location following a particular path: As evidenced by Fanucchi and others, motion planning (here, a combined effector and control issue) will be a computationally complex process involving a search and evaluation through *all* possible trajectories in order to decide upon a path that satisfies all requirements.[22] The relevance of the example becomes apparent: Depending on that task, Hachoun notes that several components will be required to finding the very best route (shortest, safest, most efficient, least challenging) that also takes into account the machine's own geometry (shape, turning radius) and steering mechanism (the AWS' holonomic properties).[23]

All of these issues may eventually be solvable but, as evidenced by Sitte and Winzer, hardware issues nevertheless place clear constraints upon deployment. Much of AWS' hardware challenge relates to the enduring complexity of 'the last meter' whereby hardware components must manage

[16] Aku Pietikainen and others, 'Design of the Mechanics and Sensor System of an All-Terrain Robot Platform', *International Conference on Robotics and Automation*, IEEE, (2008), pp. 1-2 <https://pdfs.semanticscholar.org/eca8/532658ae4864c3e9a4a264af107e3cc0ed69.pdf>.

[17] For a primer on robotic parts, see: CCSI, 'Parts of a Robot', <http://www.mind.ilstu.edu/curriculum/medical_robotics/parts_of_robots.php> [accessed 12 May 2018].

[18] Mataric, p. 39.

[19] For a useful primer on robots' physical challenges, see: Niklaus Correll and others, 'Analysis and Observations on the Frist Amazon Picking Challenge', *arXiv*.1601.05484v3, (22 September 2017), pp. 2-5 and generally <https://arxiv.org/pdf/1601.05484.pdf>.

[20] For a general discussion on battery development, see: Chapter 3 (*Drivers*), specifically: 3.2 ('*Technology creep and dual use drivers*').

[21] For a discussion on developments in autonomous power replenishment in robots, see: Ioannis Ieropoulos and others, 'Energetically Autonomous Robots', *University of West of England*, Bristol, undated <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.739&rep=rep1&type=pdf>.

[22] D Fanucchi and others, 'An Overview and Ideas on Autonomous Robot Path Planning Algorithms', *Wits University*, (2010), pp. 53-56 (Section 2.4: '*Overview of Methods*') <https://www.wits.ac.za/media/migration/files/cs-38933-fix/migrated-pdf/pdfs-2/2009AutonomousRobotpathplanning.pdf>.

[23] O Hachoun, 'Path Planning of an Autonomous Robot', *International Journal of Systems Application, Engineering and Developmen*t, 4, 2, (2008), 178-181 <http://www.wseas.us/journals/saed/saed-45.pdf>.

the difficult interaction between host weapon and its immediate environment. Lida attributes this challenge to that interaction's characteristics of poor predictability, low programmability, the plethora of emergent battlefield scenarios as well as the frequently delicate and idiosyncratic tasks required of that process.[24] Trajectory planning also provides a relevant proxy. Calculating a best path, after all, becomes exponentially more complex in three-dimensions (the case with a weapon's robotic arms).[25] Manipulation is particularly challenging given the AWS' requirement to compute in real time the *free* space of each manipulator (the space in which movement is possible) in order then to model that space for a particular action sequence or combat task.[26] All such routines, moreover, must be preceded by precursor sequences including, for instance, whether each such effector action is both goal and value-compliant?[27]

Johnson points out that sensors and effectors may have *separate* controllers requiring further levels of internal coordination and feedback on the weapon's location and state.[28] As above, an effector is the hardware device on the autonomous platform that has most physical effect, impact and influence on the weapon's immediate environment. Just as sensors must correspond to the weapon's task, so must the effectors be similarly matched, each with a conforming actuator that enables the effector to execute that action or movement. Yin and others note the complication that such actuating will likely be undertaken in several quite different non-conforming modes, whether by electric motors, hydraulics, pneumatics or by using dissimilar materials which may be photo-reactive, chemically reactive, thermally reactive or piezoelectric.[29] A deployment consequence is the amount of idiosyncratic management and security (individual programming languages, bespoke instruction routines, custom feedback loops) required for *each* component's appropriate

[24] Dr Fumiya Lida, Department of Engineering, Cambridge University, *Prowler.io Decision Summit*, (15 November 2018) and in conversation with the author.

[25] Joaquin Sitte and Petra Winzer, 'Mastering Complexity in Robot Design'*, Proceedings of 2004 IEEE International Conference on Intelligent Robots and Systems,* (2004), 1815-1816 <https://groups.csail.mit.edu/drl/journal_club/papers/robot-design-iros2004.pdf>.

[26] Mataric, p. 61.

[27] IEEE Global Initiative for Ethical Consideration in Artificial Intelligence and Autonomous Systems, 'Embedding Values into Autonomous Intelligent Systems'*, IEEE*, undated, pp. 22-23 <https://standards.ieee.org/develop/indconn/ec/ead_embedding_values.pdf>. The document provides a key primer to the issues of moral overload and value conflicts (p. 25), algorithmic biases (p. 26), building empirical norms into machine architecture (pp. 29-32) and third-party evaluation of alignment in machine values (p. 33).

[28] Matthew Johnson and others, 'Team IHMC's Lessons Learned from the DARPA Robotics Challenge Trials', *Journal of Robotics*, (March 2015), 4-5 and generally <https://s3.amazonaws.com/academia.edu.documents/41980899/Team_IHMCs_Lessons_Learned_from_the_DAR20160203-30232-1p7o2um.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1540061344&Signature=kTtjZhdj2UozAahn7Cbwujww3X0%3D&response-content-disposition=inline%3B%20filename%3DTeam_IHMCs_Lessons_Learned_from_the_DARP.pdf>.

[29] Mark Yin and others, 'Modular Self-Reconfigurable Robotic Systems: Grand Challenges of Robotics'*, Robotics and Automation Magazine*, IEEE, 14, 1, (2007), pp. 4-52 <http://ieeexplore.ieee.org/abstract/document/4141032/> [accessed 20 May 2018].

operation.[30] A recurring theme in this thesis (and noted by Yatamanchili) is that system heterogeneity will invariably add complexity and fragility to AWS' overall function.[31]

Nesnas similarly highlights build challenges relating to such hardware configuration, especially around component organisation.[32] An AWS' vision package requires a camera. In order to be fit for purpose, however, it empirically requires a hardware portfolio that enables stereo processing, visual odometry, structure-from-motion, visual tracking, object finding and template matching.[33] The issue is therefore achieving a level of hardware generalization that is appropriate. Nesnas notes that this will be complicated by quite different design paths, from modules intended to achieve specific functionality to componentry that must be organised instead around more generic use-based classes.[34] Challenges also arise from the definition of a common system vocabulary, from mediation between different programming models (declarative programming as opposed to procedural programming), as well as from a requirement to reconcile different representations of outwardly similar input information.[35] An adjunct complexity is of course the ensuring of clear governance when setting priorities in the case of *shared* hardware resources given that multiple devices comprising AWS hardware may be logically decoupled while still remaining physically coupled.[36] In this vein, conflicts must be mediated between opposing hardware architectures (for instance, an AWS model based upon a central processor versus a model where control is migrated throughout the weapon firmware in distributed nodes).

Hardware selection requires recurrent design compromise.[37] Haikonnen identifies a complicating denominator here to be the machine's expression of sensory information that is based entirely on numeric values.[38] The general hardware challenge is that resulting performance

[30] Lois Batson and Donald Wimmer, 'Unmanned Tactical Autonomous Control and Collaboration Threat and Vulnerability Assessment', *Calhoun NPS Institutional Archive*, (June 2015), pp. 8-11 ('*Five Pillars of Information Assurance'*) and 25-30 ('*Breakdown of Threat Template'*) <https://calhoun.nps.edu/bitstream/handle/10945/45738/15Jun_Batson_Wimmer.pdf?sequence=1&isAllowed=y> [accessed 23 May 2018].

[31] For a discussion on ramifications of system heterogeneity, see: Sudhakar Yatamanchili, 'Software Challenges of Heterogeneity*', School of Electrical and Computer Engineering*, Georgia Institute of Technology, undated <http://www.socforhpc.org/wp-content/uploads/2015/07/yalamanchili-SW-Challenges.pdf>.

[32] Issa Nesnas, 'The CLARAty Project: Coping with Hardware and Software Heterogeneity'*, Software Engineering for Experimental Robotics*, Springer, (2007), p. 5 <http://gias720.dis.ulpgc.es/Gias/asignaturas/master-siani-isspe/Bibliografia/CLARAty/06_nesnas_starbook.pdf>.

[33]Ibid.*,* p. 5 and generally.

[34] Ibid., p. 2.

[35] Dirk Fahland and others, 'Declarative Versus Imperative Process Modelling: The Issue of Maintainability', *International Conference on Business Process Management*, Springer, Berlin, Heidelberg, (2009), pp. 6-7 and generally <https://www.matthiasweidlich.com/paper/declarative_vs_imperative_maintainability_ERBPM_2009.pdf>.

[36] Nesnas, p. 7 and p. 23.

[37] For a detailed discussion on conflict resolution methodologies, see: Chapter 8 (*Software*), specifically: 8.4 ('*Software processing functions'*) and 8.6 ('*Action selection issues'*).

[38] Haikonnen, p. 169. See also: B Selam (moderator) and others, 'Challenge problems for Artificial Intelligence'*, 13th National Conference on AI,* AAAI-96, <http://erichorvitz.com/selman.htm> [accessed 16 June 2017]. 'We lack the equivalent of a Perceptron Book (Minsky and Papert, 1969)'. Selam defines system brittleness as the difficulty of porting coding solutions across routines. 'It is hard to know how to take things from successes and apply them to new problems'.

variation is deeply coupled, is incapable of clear attribution[39] and will anyway be brought about by exogenous factors such as sensor noise[40] and what Pendleton terms as the dynamic 'change state of the machine's immediate environment'.[41] In a similar vein, it can be inferred from Guszcza that the relationship between the weapon's software and the hardware that it is driving is itself dynamic, difficult to model and largely dependent upon that weapon's receptiveness to local prediction and the weapon's ability to learn.[42] In order, then, to judge the complexities arising from such hardware selection, it is necessary to consider the ramifications that arise from fusing together what is heterogeneous componentry.

9.1 Hardware and sensor fusion issues for AWS

An effector is any device on the AWS that touches the robot's environment. Some narrative is again useful to identify constraints arising from design issues. An action sequence will typically commence with the weapon controller initiating a command to its effectors in order to produce a particular outcome on the environment that is based on the weapon's current task. Linked to weapon tasking, it is important to state that AWS actuator types must be very broad and encompass electric motors, hydraulic or pneumatic cylinders as well as photo-reactive, temperature sensitive, piezoelectric or chemically sensitive materials that can actuate wheels, tracks, arms, grippers and other effectors on the AWS.[43] In this vein, it is relevant to consider challenges created by an effector's degrees of freedom (DOF).[44] Current robotic hands may have thirty DOF.[45] For the purpose of this analysis, a helicopter has six DOF and moves in three dimensions. Considering first the configuration of an autonomous land-based vehicle, only two dimensions of movement are required. In this case, the AWS' movement governor can control only two things: forward/reverse and rotation. Although that vehicle therefore has three DOF, only two of them are controllable. Since there are always more DOF than are controllable, Mataric highlights that there will be motions that cannot be undertaken by that machine such as moving sideways.[46] Why is this important? However effective the two DOF, an unsupervised land-based machine must likely generate a complicated path in order to carry out its motion task. The hardware challenge is that

[39] For a discussion of sensor decisions, see: R Luo and M Kay, 'Multi-sensor Integration and fusion for Intelligent Machines and Systems', *Ablex Publishing,* North Carolina University, (1995), pp. 5-9.

[40] R Brooks and others, 'Automatic correlation and calibration of noisy sensor readings using elite genetic algorithms', *Artificial Intelligence,* 84, Elsevier, (1996), pp. 339-354 <https://pdfs.semanticscholar.org/cef2/9a6a615d9875d9d538c02cef71a7d29df190.pdf>. Unsurprisingly, machine calibration becomes more difficult the further sensor noise moves from standard deviation limits.

[41] Pendleton and others, 'Perception, Planning, Control, and Coordination for Autonomous Vehicles', pp. 1-4 and generally.

[42] J Guszcza and N Maddirala, 'Minds and Machines: The Art of Forecasting in the Age of Artificial Intelligence', *Deloitte Review*, 19, University Press, (25 July 2016) <https://dupress.deloitte.com/dup-us-en/deloitte-review/issue-19/art-of-forecasting-human-in-the-loop-machine-learning.html> [accessed 12 June 2017].

[43] FC Park and KM Lynch, 'Introduction to Robotics: Mechanics, Planning and Control', *North Western University Publishing*, (20 September 2016), pp. 9-14 ('*Degrees of Freedom'*) and 19-24 ('*Configuration space'*) <http://hades.mech.northwestern.edu/images/2/2a/Park-lynch.pdf>.

[44] D Lowe, 'Characterising complexity by the degrees of freedom in a radical basis function network', *Neurocomputing*, 19, April (1998), p. 199 <http://www.sciencedirect.com/science/article/pii/S0925231297000659> [accessed 2 March 2017].

[45] Dr Fumiya Iida, Department of Engineering, Cambridge University, *Prowler.io Decision Summit*, (15 November 2018) and in conversation with the author.

[46] Mataric, p. 41.

continuous trajectory must be achieved with *discontinuous* velocity: the weapon must stop and start in order to reach each destination, so creating deep complexity. Two complicating relationships emerge. First, the non-holonomic robot has more DOF than it can control.[47] Second, Sporn and Edelman evidence that the more DOF a robotic weapon exhibits, the more complicated it is to control.[48]

A further physical complication in land-based AWS architecture relates to unit stability.[49] Current iterations of military robots tend to be based on four-legged mobility in order better to deal with the issues of centre of gravity (COG) and balance.[50] While Bottcher may note that four-legged construction provides an 'optimal polygon of support', the design also introduces significant additional hardware challenge.[51] Human COG is quite high on our bodies and keeping stability does not happen without experience and training. For this reason, Waibel is able to confirm that any two-legged AWS will have a small such polygon and computational routines will be required to keep COG stably aligned in order to keep the weapon unit upright.[52] For the purposes of evaluating such challenge, desirable robot gaits have five required collaborative properties. 'Stability' insures the robotic weapon does not fall over, 'speed' allows the weapon to move quickly while 'energy efficiency' ensures the robot can exhibit durability and 'robustness' allow it to recover from various failure modes. Finally, overall 'simplicity' will insure that the unsupervised weapon's gait and operation is not unwieldy.[53] These characteristics must therefore comprise the weapon's physical design fundamentals which, given the importance of free passage across what will be a contested combat environment, leads Shurkin to highlight the confounding issue of error (whether human or machine) in the empirical operation of robots.[54]

A further source of hardware error arises from inaccuracies arising from AWS' configuration and, noted by Hofman, from errors arising in unsupervised measurement systems.[55] While hardware error might appear relatively 'primitive' next to challenges that originate in that weapon's software control and governance routines, such challenges cumulatively contribute to AWS infeasibility. Stability routines cannot exist in isolation and must be integrated into the

[47] *Holonomic* refers to the relationship between controllable and total degrees of freedom of a robot. If the controllable degree of freedom is equal to total degrees of freedom, then the robot is said to be holonomic.

[48] For a discussion of the principles involved, see: Olaf Sporns and Gerald Edelman, 'Solving Bernstein's Problem: A Proposal for the Development of Coordinated Movement by Selection', *Child Development*, 64,4, (1993), generally <http://e.guigon.free.fr/rsc/article/SpornsEdelman93.pdf>.

[49] For a useful general discussion on robotic stability, S Bottcher, 'Principles of Robot Locomotion' <http://www2.cs.siu.edu/~hexmoor/classes/CS404-S09/RobotLocomotion.pdf>.

[50] Boston Dynamics, 'Big Dog robot' <https://www.youtube.com/watch?v=mpBG-nSRcrQ> [accessed 17 April 2017].

[51] Bottcher, 'Principles of Robot Locomotion', p. 2 and p. 9.

[52] B Waibel and others, 'Theory and Experiments on the stability of robot compliance models', *Transactions on Robotics*, IEEE, 7, 1, generally.

[53] Mataric, p. 52.

[54] J Shurkin, 'When Driver Error becomes Programming Error', *Inside Science*, (18 February 2015), generally <https://www.insidescience.org/news/when-driver-error-becomes-programming-error> [accessed 12 June 2017].

[55] D Hofman, 'Common sources of errors in measurement systems', *Steinbeis Transfer Centre for Quality Insurance*, Handbook of Measuring System Design, (2005) <http://eu.wiley.com/legacy/wileychi/hbmsd/pdfs/mm154.pdf>. Hofman's research provides a useful aide memoire on common error sources relevant to AWS deployment including input, sensor, signal transmission, conversion, cumulative, gain, dataset, materials, drift, load, thermal, operator, degradation, communication, mapping and other software errors.

weapon's control and decision functions in order to mediate wobble, lean and deviation.[56] Rasmussen's work on reinforcement learning (in this case, its application to riding a unicycle) demonstrates the fundamental instability of certain hardware relationships.[57] To this point, stability exists in two states; the AWS must be statically stable as well as dynamically stable. In general, an AWS with more legs (or ground points) can maintain better static stability given its raised centre of gravity and broad polygon of support. Physical componentry also gives rise to error. Backlash and inaccuracy arising from gear mechanisms are likely to be catalysts for error.[58] Similarly, keeping the weapon's centre of gravity over an efficiently small contact-point with the ground requires active and trained effort and is an obvious source of weapon instability, especially over contested ground. Four-legged robots, for instance, can only lift one leg at a time as three legs must remain grounded in order to remain statically stable. There are, therefore, several compromises to be made between a weapon's stability, its speed of movement, energy conservation, robustness and simplicity.[59]

A further hardware challenge relates to motion and, notes Lewis, arises generally from the requirement that AWS' manipulators (a grabber, perhaps, or other subsidiary system used to handle objects in its immediate environment) must move relevant to a three-dimensional orientation.[60] Complex computational processes must triangulate dynamically the weapon's body state, its manipulators and the task in hand. This must take place at all times and be accurate at all times. In particular, it must involve understanding the 'free space' of each weapon module that must account for all space in which machine movement is possible. These routines must take place in advance of (but also in tandem with) AWS' calculations and risk-assessment on this free space. The AWS' consequent output must then be dynamically integrated into the weapon's utility function and action selection routines[61], the more so as modules with multiple DOF will likely require the weapon to support significant additional physical machinery in order to achieve that movement.[62] This too has follow-on implications on weapon feasibility. Power for each such motor may require additional platform weight in turn needing stronger motors to lift the platform manipulators. In this vein, the work of Duran and Thill demonstrate that traditional ball-and-socket

[56] E Nebot and others, 'Navigational Algorithms for Autonomous Machines in Off-Road Applications', *Journal of Autonomous Robots*, 14, (2000) , Paragraphs 1 ('*Introduction'*) and 3 ('*Non Model Based Navigation'*) <http://www8.cs.umu.se/research/ifor/dl/LOCALIZATION-NAVIGATION/Navigation%20Algorithms%20for%20Autonomous%20Machines%20in%20Off-Road%20Applications.pdf>.

[57] Professor Carl Edward Rasmussen, Department of Engineering, Cambridge University, *Prowler.io Decision Summit*, (15 November 2018) and in conversation with the author.

[58] Machine design, 'Methods to Minimise Gear Backlash'*, Machine Design*, (23 December 2015) <http://www.machinedesign.com/datasheet/methods-minimize-gear-backlash-pdf-download> [accessed 12 March 2018].

[59] Mataric, pp. 50-57.

[60] For a useful primer on robot manipulators, see: Frank Lewis and others, *Robot Manipulator Control: Theory and Practice,* (USA: Marcel Dekker Publishing, 2004) <https://pdfs.semanticscholar.org/d1f0/2a7db3294ddf775555bd1f26610b5df2e467.pdf>.

[61] See also: Chapter 8 (*Software*), specifically: 8.7 ('*Action Selection Issues'*).

[62] Sources: Quora.com, 'Degrees of Freedom' <https://www.quora.com/How-does-one-calculate-a-robots-DOF-degrees-of-freedom-in-the-strict-sense-of-mobility> [accessed 12 December 2016]. See also: Whatis.com, 'Degrees of Freedom', <http://whatis.techtarget.com/definition/degrees-of-freedom> [accessed 12 December 2016].

joints are particularly difficult to incorporate in artificial systems.[63] By comparison, muscles in animal rotary joints are linear actuators (relatively lighter, springier, more flexible and stronger). Whereas humans use their hands as general-purpose manipulators to work specific tools (knives, screwdrivers), the AWS' robot manipulators will likely be specialized with dedicated lethal tools at the endpoint. Yao argues that endpoint engineering is intractably complex requiring computationally intense inverse kinematics to manage the weapon's end-effector to a desired point, involving real-time conversion and management of each weapon's endpoint tools from a Cartesian (x, y, z) position.[64] None of these hardware sequences are straightforward and must account for manipulator travel relative to the weapon's centre of gravity, problems caused by friction and unexpected obstacles as well as background hand-off, priority and tasking issues.[65]

An architectural analysis of likely AWS *sensor* inventory is also relevant. AWS' deployment depends, as above, on reliable capability to sense the condition of its systems as well as predicting the states of its immediate and far environments. As noted by Martinelli and others, proprioceptive sensors are required to administer individual elements of the robot's internal state (the position of its wheels, the joint angles of its arms) while exteroceptive sensors must process the platform's external world (light levels, distances to objects, sound).[66] Taken together, these sensors constitute the weapon's perceptual system. As identified in previous chapters, the challenge is that its efficacy is based upon sensor inputs that must by definition be tangential, fragmented and the product of multiple units.[67] Sensed input must be processed ex-ante to manage limitation such as effector and actuator noise, as well as constraints arising from data sources being derived from either hidden or partially observable states.[68] Other factors compromise effective proprioception including the *quantity* of information that multiple sensors are returning dynamically to a controller: While a simple contact switch may provide just one single bit of information (on or off), Mataric and others note that a vision sensor will be stunningly rich and similarly complicated in the amount of information captured.[69]

Such AWS' dependence on hardware sensors creates other deployment challenges. AWS' sensor stimuli will have properties that are not separately divisible. While significant progress may continue to be made in artificial recognition, it can be inferred from Horvitz that battlefield shapes and sizes empirically do not appear 'alone' and often are themselves properties of other larger

[63] B Duran and S Thill, 'Rob's Robot: Current and Future Challenges for Humanoid Robots', *Intech*, The Future of Humanoid Robot Research and Application, (2012), p. 280 ('*Introduction'*) and pp. 282-288 ('*Mechanical Requirements and Engineering Challenges*').

[64] Ming Yao, 'Mathematics for Inverse Kinematics', undated tutorial <http://www.cs.cmu.edu/~15464-s13/lectures/lecture6/IK.pdf>. See also: Mataric, p. 65.

[65] This is empirically complex and must be achieved algebraically (through the use of matrix equations), geometrically (by combining knowledge of the weapon arm with dynamic trigonometry) or numerically (by using 'guess work' and incremental adjustment in order to minimise local error(. See: Applied Go Tutorials, 'Inverse kinematics; how to move a robotic arm (and why this is a harder then it seems)', *Applied Go Tutorials*, (16 June 2016) <https://appliedgo.net/roboticarm/> [accessed 5 June 2017].

[66] Agostino Martinelli and others, 'Multi-Robot Localization Using Relative Observations', *IEEE International Conference on Robotics and Automation*, (2005) <https://infoscience.epfl.ch/record/97559/files/a1093.pdf>.

[67] See: Chapter 7 (Firmware), specifically: discussion in 7.1 ('*Sources of technical debt'*) on data efficacy.

[68] See: Chapter 8 (*Software*), specifically: 8.4 ('*Software processing functions'*).

[69] Mataric, p. 71. See also: Eurostat, 'Big Data Conversion Techniques including their Main Features and Characteristics: 2017 Edition', *Eurostat*, (2017), p. 17 and p. 18 <http://ec.europa.eu/eurostat/documents/3888793/8123371/KS-TC-17-003-EN-N.pdf/ad617aaa-6d34-4f05-a341-fa8db6043045> [accessed 14 May 2018].

objects that have conflicting properties.[70] Adversarial feint will grossly complicate machine attribution. Second to this point, the output from such sensor hardware (whether through filtering, parsing, normalisation, smoothing) will clearly impact data granularity[71] and add, therefore, to the platform's technical debt.[72] Hardware constraints directly create governance issues. Should, for instance, consistency, uniformity and accuracy tests be undertaken interactively or selectively on hardware sensors? Rahm's study of data cleaning similarly highlights challenges arising from overlapping, data conflict as well as the stitching together of different hardware sensor types.[73] While management of weapon sensors may illustrate an overlap between hardware and software assets, it is complicated by file matching challenges (the 'object identify problem') as well as data duplication and purging challenges arising from hardware shortfall.[74] An adjunct constraint is noted by Cai whereby data improvement will be enduringly difficult to implement at scale and, by inference, is not simply a matter of adding further hardware capability to the AWS.[75] As noted by Steinruecken, the role still required of humans in the building, understanding and interpretation of probabilistic models suggests that AWS will require an automatic statistician (akin, perhaps, to Arkin's Ethical Governor[76]) if the model is to be appropriately scalable.[77] A further adjunct complexity arises then from 'sensor scheduling', the handshaking process that decides which weapon sensor (or mode of operation) should dynamically be chosen to provide the *next* relevant measurement. By inference, Krishnamurthy apportions this challenge to weapon overload, weapon energy constraints and ensuing data ambiguity.[78] A quite separate hardware challenge arises from the requirement that the weapon possesses broad interface-processing capabilities.[79] An example is

[70] Eric Horvitz, 'Artificial Intelligence and Life in 2030: One Hundred Year Study of Artificial Intelligence*', AAAI*, Stanford University, (September 2016), pp. 4-6 <https://ai100.stanford.edu/sites/default/files/ai_100_report_0901fnlc_single.pdf>.

[71] Michael Ohler and others, 'Proper Data Granularity Allows for Stronger Analysis*', Six Sigma blog*, (22 May 2018) <https://www.isixsigma.com/tools-templates/measurement-systems-analysis-msa-gage-rr/proper-data-granularity-allows-stronger-analysis/> [accessed 22 May 2018].

[72] For a discussion on process complexity as it relates to AWS, see: Chapter 7 *Firmware)*, specifically: 7.1 ('*Sources of technical debt'*).

[73] Erhard Rahm and Hong Hai Do, 'Data Cleaning: Problems and Current Approaches*', University of Leipzig*, p. 2 <http://www.betterevaluation.org/sites/default/files/data_cleaning.pdf>.

[74] Leonidas Guibas, 'The Identification Management Problem – A Short Survey*', Information Fusion*, 11th International Conference, IEEE, (2008), pp. 1-2 <https://geometry.stanford.edu/papers/g-impss-08/g-impss-08.pdf>.

[75] Li Cai and others, 'The Challenge of Data Quality and Data Quality Assessment in the Big Data Era*', Data Science Journal*, 14, 2, (2015), 1-10. Data remediation will rely on rules-based combing routines which complicate the subsequent management of the weapon's primary data given their dependence on transitive closure of source files See: M Harnandez and S Stolph, *'*Real World Data is Dirty: Data Cleansing and the Merge/Purge Problem*', Data Mining and Knowledge Discovery*, 1998, section 2:9.

[76] See Chapter 5 (*Obstacles*), specifically: 5.6 ('*Behavioural constraints'*).

[77] Christian Steinrueken, Department of Engineering, Cambridge University, *Prowler.io Decision Summit*, (15 November 2018) and founder of 'Automatic Statistician' at <https://www.automaticstatistician.com/index/> [accessed 18 January 2019].

[78] V Krishnamurthy, 'Algorithms for Optimal Scheduling and Management of hidden Markov mode Sensors', *IEEE Transactions on Signal Processing*, 50, 6, (June 2002), p. 1382 <http://ece.ubc.ca/~vikramk/Kri02.pdf>.

[79] Lamont Wood, 'Service Robots: The Next Big Productivity Platform*', PWC*, (8 September 2016) <http://usblogs.pwc.com/emerging-technology/service-robots-the-next-big-productivity-platform/> [accessed 25 March 2018].

identified by Wood in the sensor-input of third-party speech with its important nuances in rate, volume, pitch and other indicators of personality.[80]

Practical difficulties arise from linking hardware combinations to weapon tasks, a further design constraint. A weapon's visual systems will tend to be mounted on platforms that are usually moving. Similarly, the target under evaluation may be moving, complexity arising when other battlefield objects are moving independent of the AWS or intended target. Movement on a battlefield, moreover, is a conflicting combination of, on the one hand, purpose and, on the other, reaction to a series of unsystematic, chaotic drivers.[81] Does the weapon's movement sequence (its reaction, for instance, to an unexpected environmental hazard) trump its attack sequences? Lumelsky highlights that such movement occasions unique hardware challenges and requires specific hardware permutation.[82] Weapon balance must be maintained on uneven surfaces and the weapon's speed of execution must be adjusted constantly in order to suit such changing physical conditions.[83] Correctional sub-routines are then necessary to smooth for motion-generated interference with such adjustments compensating for the weapon's starting, turning, climbing, descending and stopping. Furthermore, the AWS must trigger such motor actions slightly in advance of execution and cannot be based on fixed sequences or sequenced timing charts.[84] Most motor acts, note Hoffman and Ju, will involve serial combinations of quite different motor sequences with each act being factored appropriately into that machine's processes.[85] It is, reckons Haikonnen, 'like having a strange remote control with buttons but no markings… there is no inherent connection between the inner image of an action and [situational awareness routines] that could cause the desired action'.[86]

Hardware components are generally task-specific, each with idiosyncratic challenges to AWS deployment. Ultrasound sensors are one such component that may be key to AWS function, measuring distance to an object using sound waves. A second component set is sonar and attendant power issues given that significant current is required to emit each ping.[87] Indeed, the range of

[80] Natural language processing and its hardware pose further complications. Despite progress in processing written language, it will remain difficult for an AWS to understand *spoken* work unless that platform is addressed with a few chosen words that are already familiar, without disruption or interference and delivered without accent.

[81] Small Wars Journal, 'An Advanced Engagement Battlespace: Tactical. Operational and Strategic Implications for the Future Operational Environment*', SWJ*, undated <http://smallwarsjournal.com/jrnl/art/advanced-engagement-battlespace-tactical-operational-and-strategic-implications-future> [accessed 12 May 2018]. The publication interestingly separate battlefield assets (and their likely movement traits) into 'finders versus hiders', 'strikers versus shielders', dispersed assets as well as assets that are deliberately dormant.

[82] Vladimir Lumelsky, 'Algorithmic and Complexity Issues of Robot Motion in an Uncertain Environment*', Journal of Complexity*, 13, 2, (1987), 146-150 <https://ac.els-cdn.com/0885064X87900252/1-s2.0-0885064X87900252-main.pdf?_tid=383b9a49-fb51-4072-8a52-ed8b58717958&acdnat=1526910324_3de58822542e77ff1958bcd94192fbaf> pp. 146-150. See also: Chapter 8 (*Software*), specifically: 8.4 ('*Navigation issues*').

[83] Haikonnen, p. 205.

[84] Ibid., p. 204.

[85] Guy Hoffman and Wendy Ju, 'Designing Robots with Movement in Mind', *Journal of Human-Robot Interaction*, 1, 1, (2012), 3 and 5 <http://guyhoffman.com/publications/HoffmanJuJHRI14s.pdf>. The way that motor neurons might be connected to the rest of the cognitive machine is not yet clear in order for planning and control of motor acts by inner imagery to take place.

[86] Haikonnen, p. 206.

[87] Mataric, p. 99.

these sensors is determined entirely by the signal strength (and therefore power use) of the emitter. Ping emission by an AWS also betrays that weapon's state and position. A further hardware complication arises because sound waves do not necessarily bounce off the nearest surface and return as expected; instead, notes Hamalainen and MacIsaac, the process is hampered by blind spots, multiple or unwanted reflections, obstructions, noise and lack of consistency.[88] The direction of reflection depends on several factors including gameable surface properties and the incident angle of the sound beam. Moreover, a disadvantage of ultrasound sensing is its susceptibility to specular reflection, the reflection from the outer surface of the bounce-back object (here, the AWS' target). The smoother that target, the bouncing sound generates a false far-away reading. In contrast, rough target surfaces produce irregular reflections. In cases of specular recollection, sound bounces around the target environment and may not return to the detector. A deployment issue is that the AWS may thus be fooled into concluding there is no object or that it is at a great distance. The property also facilitates sensor spoofing by the adversarial designing of feints into target surfaces that are then not anticipated by the AWS which must instead trust sensor readings regardless of how unpredictable they may be. Laser sensors come with other trade-offs. Using phase-shift rather than time-of-flight principles, lasers involve much higher-power electronics. Furthermore, laser units are likely to remain large relative to other AWS components and, while the laser's narrow beam is adept at detecting distance to a particular point, a weapon's laser must sweep in order to cover its area, requiring significant instruction, processing activity, substantial battery power and management of further large information datasets.[89]

9.2 Calibration issues

Calibration is then the process of adjusting the unsupervised weapon so as to maximize hardware performance.[90] Rivera notes that increasingly automated calibration routines may be able to correct sensor offset, gain variation and reduce compromises caused by poor data integrity.[91] The issue for AWS deployment, however, is that sensors may require multiple (even dynamic) calibrations. Beautement points out that such adjustment will be needed even to calibrate responsibility levels in the weapon's decision-making (matched here to unfolding battlefield circumstances) as well to validate the weapon's ability to choose new actions or make policy changes. Calibration will also be required to prevent the AWS from attempting the unachievable, preventing its obligations from exceeding its scope of permissions, initiating reductions in current obligations in order to take on new tasks and, consequently, modifying permission levels to

[88] A Hamalainen and D MacIsaac, 'Using Ultrasonic Sonar Rangers: Some Practical Problems and how to Solve Them', *Proceedings from the XXXVI Annual Conference of the Finnish Physical Society*, (2002), p. 1 <http://physicsed.buffalostate.edu/pubs/TPT/TPTJan02SONAR/poster.pdf>.

[89] For a useful primer on hardware challenges to autonomous robots, see: 'Overview of Challenges in the Development of Autonomous Mobile Robots', (23 August 2011) <http://web.eecs.utk.edu/~leparker/Courses/CS494-529-fall11/Lectures/Aug-23-Development-Challenges.pdf>.

[90] A Elatta, 'An Overview of Robot Calibration', *Information and Technology Journal,* 74-78 <http://docsdrive.com/pdfs/ansinet/itj/2004/74-78.pdf>.

[91] J Rivera, 'Self-calibration and optimal response in intelligent sensors design based on artificial neural networks', *Sensors, Basel*, 7, 8, (August 2007), pp. 1509-1529, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3814866/> [accessed 12 December 2016]. Ambient light levels change throughout the day requiring, for instance, vision sensors to be recalibrated repeatedly in order to stay accurate and useful.

accommodate new capabilities.[92] These are complex constructs that will require fine (and changing) balance. Programming additional actions into a routine is generally regarded to be an inefficient method of dealing with such uncertainty.[93]

How then might dynamic calibration affect hardware performance in AWS? In control parameters and sampling rates, 'proportional' tuning will make the weapon respond to the detected error using both the direction and magnitude of that error.[94] Determining how to size this calibration response, is termed gain.[95] It is non-obvious and, crucially for both AWS compliance and utility, remediation will require inappropriate trial and error.[96] Furthermore, Wilhelm points out that system response to calibration routines must itself be adjusted such that AWS' actions are appropriately dampened to avoid oscillation.[97] The complication is that when an AWS is close to its desired state, the means for its control will be materially different than when it is far from it. As noted by Roth, the momentum generated by the controller's response to error, its own error correction, will otherwise carry the overall weapon system *beyond* a weapon state that is both LOAC-compliant and satisfactory to the deploying commander.[98] This requires complicated management whereby an amount that is proportional to that weapon's velocity (regardless of modality) is first subtracted from the AWS' current error momentum in order to achieve suitable correction.[99] Further complexity arises given that calibration will generally require that the weapon account for and tracks its own errors, in particular any repeatable or fixed errors (steady state errors) that it displays but that may subsequently change over time.[100] Calibration must similarly inform the weapon's 'state estimation', the weapon's process of reckoning its system state from measurements.[101] This, too, posits a challenging problem in unsupervised weaponry. Any estimation process is, after all, indirect. The AWS will measure *what* it can and *when* it can and establishes its state based on that particular dataset. In this case, calibration must mirror the noise of real-world sensors and real-world properties.[102] Measurements are unlikely to be available to

[92] For a primer on calibration challenges, see: Vijay Pradeep and others, 'Calibrating a multi-arm multi-sensor Robot', *Willow Grange Inc*, undated, pp. 1-2 and p. 10 ('*Choosing what to calibrate*') <http://www.willowgarage.com/sites/default/files/calibration.pdf>.

[93] Mataric, p. 104.

[94] Source: Indian Institute of Technology, Madras, 'Performance Enhancement of Robotics using Calibration Data', <https://ed.iitm.ac.in/~robotics_lab/files/Calibration.ppt>, undated [accessed 3 January 2017]. Error here may arise from design tolerances or from variances in the unit's assembly as well as from those operational challenges noted above.

[95] Ahmed Joubair, 'How Can Industrial Robots be Calibrated*?*', *Robotiq.com*, (16 November 2014) <https://blog.robotiq.com/bid/73064/How-Can-an-Industrial-Robot-Be-Calibrated> [accessed 9 December 2017].

[96] *Robotics beta blog*, Stack Exchange, (7 June 2016) <http://robotics.stackexchange.com/questions/10029/damping-vs-friction> [accessed 9 September 2016].

[97] Lisa Wilhelm and others, 'Oscillation Analysis in Behavioural-Based Robot Architecture'*, Autonome Mobile Systeme*, Springer, Berlin, (2009), pp. 121-122.

[98] Stephen Roth, 'Evaluating path tracker performance in outdoor mobile robots'*, National Robotics Engineering Consortium*, Pittsburgh, undated <http://www.nrec.ri.cmu.edu/projects/toro/tech/evaluating_tracker.pdf>.

[99] Ferrell*,* pp. 4-5 ('*Issues*') and p. 6 ('*Confinement of errors*').

[100] Although dated, for a primer on fault monitoring principles, see: M Gini and R Smith, 'Monitoring Robot Action for Error Detection and Recovery'*, NASA publications,* University of Minneapolis, (1987), pp. 67-68 <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19890017177.pdf>.

[101] Timothy Barfoot, *State Estimation for Robotics,* (Cambridge: Cambridge University Press, 2018), pp. 3-4 and p. 9 <http://asrl.utias.utoronto.ca/~tdb/bib/barfoot_ser17.pdf>.

[102] Mataric, p. 226.

AWS systems all of the time requiring that processing be carried out in batches and only when sufficient data has been accumulated: Such intermittency will add further to calibration inefficiency and hardware inaccuracy.

9.3 **Case study: navigation issues**

As the previous section focuses on specific hardware components, it is relevant to focus on the feasibility of specific hardware *routines*. In this vein, autonomous navigation merits additional review. Unsurprisingly, further challenge arises from the model's deep uncertainty. Interim destinations for the AWS are likely to be outside its immediate sensory range. Navigational parameters are complicated by moment-to-moment changes in that platform's immediate environment as well as its dependence on volatile mapping representation that is, notes Bacchus, likely out of date.[103] AWS navigation differs fundamentally from emerging models of driverless cars which will operate solely on prescribed roadways that will be networked to enable dynamic information sharing in what is an uncontested, coordinated and geo-fenced setting. Similarly, AWS' path planning will not be framed by a single-dimension geographical obligation ('get me to a location', as in the autonomous car) but is instead just one part of a complex goal-based process generated as part of the weapon's overall tasking.[104] Similarly, AWS cannot rely upon neat surface descriptors, a visualization capability and a static inventory of major objects positioned along given mapped corridors. As inferred from Joshi, however, some characteristics are common to both car and weapon models. Much of the hardware portfolio may be shared across platforms and there is similar accent on 'negative' objects (here, unexpected obstacles and efficient path planning).[105] Additionally, the AWS must proactively search through waypoints and landmarks but also goals, tasks as well as subsidiary and third party priorities that relate to its current navigational objective (here, what Frazzoli terms the 'universal coverage problem').[106] The challenge for AWS deployment is that such localization and mapping must occur at the same time, each capability requiring a complex marriage of hardware and software that empirically becomes less accurate the further the AWS travels.[107] AWS navigation thus has several additional layers of complexity. Following an arbitrary path to an intended destination is considerably more challenging than having to move to

[103] Arif Bacchus, 'Microsoft Admits to 'Stale' Mapping Data, Working on a Fix', *OnMSFT blog*, (March 2018) <https://www.onmsft.com/news/microsoft-admits-to-stale-maps-data-working-on-a-fix> [accessed 8 May 2018].

[104] See Chapter 8 (*Software*), specifically: 8.7 ('*Action selection issues*').

[105] Sourabh Joshi and others, 'Going Driverless with Sensors', *International Journal of Science, Engineering and Technology*, 2, 5, (24 June 2014), 299-301 <http://ijset.in/wp-content/uploads/2014/06/ijset.0620140074.1011.1806_Geet_298-305.pdf>. See also: Frederic Large and others, 'Navigation Among Moving Obstacles Using the NVLO; Principles and Applications to Intelligent Vehicles', *Autonomous Robots*, 19, (2005), pp. 159-160 <https://s3.amazonaws.com/academia.edu.documents/46032891/Navigation_Among_Moving_Obstacles_Using_20160528-30355-1hvecvz.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1527024240&Signature=RCzv%2F5VX0w7BFUs2c3cmiUdDxuQ%3D&response-content-disposition=inline%3B%20filename%3DNavigation_Among_Moving_Obstacles_Using.pdf>.

[106] Emilio Frazzoli, 'Real-Time Motion Planning for Agile Autonomous Vehicles', *Journal of Guidance, Control and Dynamics*, 25, 1, (January 2002), 116-118 <https://s3.amazonaws.com/academia.edu.documents/42379352/frazzoli_gcd_02.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1527024703&Signature=DNfi%2BLDhjiIjfEq2uZtAusWtuGk%3D&response-content-disposition=inline%3B%20filename%3DReal-Time_Motion_Planning_for_Agile_Auto.pdf>.

[107] Erik Gomez, 'Map-building and Planning for Autonomous Navigation of a Mobile Robot', *Center for Research and Advanced Studies*, National Polytechnic Institute, Mexico, (January 2015), pp. 26-28 and generally, *('Problems').*

that place using a route that has been specified.[108] AWS' motion planning will be computationally complex given the imperative now to search and evaluate *all* possible permutations in what is likely a fast changing and, as noted by Ahren, a likely communications-denied environment that is characterised by competing task priorities.[109] Depending on a given task, it may be necessary for the AWS to find the best or the shortest or safest or quickest or most efficient route in order for the weapon to determine an optimal trajectory while maximizing its utility function. It is for this reason that autonomous route calculation must incorporate situational awareness.[110] Furthermore, given that an AWS platform is not itself simply a single point, the weapon's geometry (shape and turning radius) as well as its steering properties must all be taken into account in arriving at appropriate motor commands.

Sensor heterogeneity, discussed above, also creates specific problems for AWS navigation. Consider, for instance, an unsupervised weapon that has a range sensor and a workable internal map of its environment. The platform can readily take range measurements and compare these with its map representation but, in practice, several features in that environment may look identical to particular sensors; corners, corridors, featureless topographies and other undifferentiated descriptors empirically lead, notes Toth, to widespread confusion in robot navigation outside the bounded environment that characterise the world of driverless cars.[111] The issue for AWS is to weight priorities that will determine that optimal path. Which element of the AWS engagement process should influence route, speed across the ground, feint or timing? Should path planning be based on distance covered, friendly asset disposition or on danger and safety criteria? The AWS' priority matrix will, after all, vary moment-to-moment depending, presumably, on its own state, mission timeline and the battlefield. Likewise, and as noted by Mataric, finding that optimal path requires searching *all* available paths in order not to miss the very best one.[112] This is computationally complex, potentially slow and open to intractable conflict.

9.4 Operational hardware issues

The foregoing analysis suggests empirical AWS deployment will diverge from the operational expectations of the Delivery Cohort. Despite the evident success of unmanned aircraft (as demonstrated by growing multi-service demand for these systems), Jacobsen notes that cost overruns and programme delays evidence that complex systems are difficult to deliver.[113] This is not a recent phenomenon. Even by July 2009, the US Government Accountability Office (GAO) had reported that six unmanned programmes were exhibiting 'cost growth' ranging from sixty per cent

[108] Brad Plumer, 'Five Big Challenges that Self-Driving Cars Still Have to Overcome'*, Vox.com blog*, (21 April 2016) <https://www.vox.com/2016/4/21/11447838/self-driving-cars-challenges-obstacles> [accessed 9 March 2018].

[109] Ramon Ahren, 'Mission Control in a Communications Denied Environment', *Air War College*, Montgomery, (16 February 2017), generally.

[110] See: Chapter 2 (*Context*), specifically: 2.6 ('*The role of situational awareness and uncertainty*').

[111] CK Toth and others, 'Mobile Mapping and Autonomous Vehicle Navigation', *Revue Francaise Photogramm*, Teledetection 185, (2007), pp. 57-61 <http://www.isprs.org/proceedings/XXXVI/part1/Papers/T08-36.pdf>.

[112] Mataric, p. 227.

[113] Mark Jacobsen, 'The Promise of Drones'*, Harvard International Review*, (3 November 2016) <http://hir.harvard.edu/article/?a=13949> [accessed 7 May 2018]. Jacobsen's paper also sets out the statistical basis of UAV's recent take-up. See also: Walker, *Killer Robots?,* pp. 99-100 ('*Economic considerations*'). Also: Chapter 1 (*Introduction*), generally.

to more than two hundred and fifty per cent.[114] Four of the UAS programmes were reported by GOA in 2015 to have experienced delays of between one to four years, mainly as a result of hardware development and testing problems.[115] Such schedule breaches inform generally on AWS feasibility including, inter alia, the 'high level of concurrency between development, production and testing; poor contractor performance; developmental and technical problems; system failures; and bad weather'.[116] Furthermore, AWS hardware economics have been impacted by the lack of commonality between the various systems, payloads, sub-systems and even ground control stations.[117] Hardware heterogeneity may lead to performance degradation: After months of continuous operation in 1991's Operation Desert Storm, the efficacy of the Patriot Missile System degraded significantly, its radar prone to 'drift from its prescribed search fan leading to a significant miss ratio'.[118] Such limitations clearly have ramifications. In this case, accuracy issues caused the Patriot's control systems to recalibrate using this miscalculated fire data requiring a system-wide refit with updated software.[119] As noted by Sculley (and discussed in earlier chapters), such ramifications are increased by AWS' general entanglement which 'is innate to machine learning. In practice, this all too often means that shipping the first version of the system is easy but making subsequent improvements is unexpectedly difficult'.[120]

In judging hardware aspects to AWS feasibility, specific challenges arise from the targeting apparatus of an unsupervised weapon. As set out in the US Department of the Army's *Targeting Process*, targeting is an involved process.[121] An understanding of its facets provides useful context to hardware assets where targeting must be undertaken without human supervision. Targeting comprises three phases ('Decide', 'Detect' and 'Deliver).[122] Removed from human supervision, weapon hardware must instead enable autonomous apportionment of value to targets as well as determine engagement effects on each such target. AWS componentry must facilitate the analysis (and then execution) of when and how to attack having first integrated restrictions relating to that attack. Its hardware must enable appropriate battle damage assessment ahead of engagement.[123] Without oversight, it is AWS hardware that must now determine engagement responsibilities, execution of target tracking, liaison with friendly assets, establishing common datum and ensuring

---

[114] US Department of Defence, 'Unmanned System Roadmap 2007-2032', cit. United States Government Accountability Office ('GAO'), *Testimony before Subcommittee on National Security and Foreign Affairs*, (March 2010) p. 5.

[115] Alice Ross, 'Watchkeepers: Boxed Up, Barely Used and Four Years Late', *Bureau of Investigative Journalism*, (2 October 2015) <https://www.thebureauinvestigates.com/stories/2015-10-02/boxed-up-barely-used-and-4-years-late-watchkeeper-the-armys-affordable-1-2bn-drone-programme> [accessed 21 May 2018].

[116] US Department of Defence, '*Unmanned System Roadmap 2007-2032',* p. 6.

[117] Ibid., p. 14.

[118] General Accounting Office, 'Patriot Missile Defence; software problems led to system failure at Dhahran', *GAO Office Report*, (2002) <http://www.fas.org/spp/starwars/gao/im92026.htm> [accessed 2 November 2017].

[119] John Hawley, 'Patriot Wars: Automation and the Patriot Air and Missile Defense System', *Center for a New American Security*, (25 January 2017), generally <https://www.cnas.org/publications/reports/patriot-wars> [accessed 26 May 2018].

[120] Sculley and others, p. 2. See also: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt'*) and 7.2 ('*Firmware ramifications of learning methodologies'*).

[121] US Department of the Army, *The Targeting Process,* (USA: Field Manual Publications, 3-60, November 2010) <https://www.globalsecurity.org/military/library/policy/army/fm/3-60/fm3-60.pdf>.

[122] Ibid., Appendix E-1.

[123] US Department of the Army, *Battle Damage Assessment and Repair,* (US Army, Doctrinal Guide Publications, March 2012) <http://asktop.net/wp/download/GTA/gta01-14-001.pdf>.

overall synchronisation in each engagement. Other hardware targeting tasks will include validation, coordination of other relevant assets and mediation of potential fratricide.[124] In this case, Reiner notes that hardware considerations make the automation of this routine significantly challenging, especially around the practice of machine detection and autonomous identification of targets, a central capability for AWS.[125]

A further challenge is that proximity, heat effects and other target monikers will tend to introduce cueing bias into the weapon's automatic target detection (ATD).[126] ATD has several challenging characteristics (and differs subtly from the AWS ATR, the subject of Appendix One to this thesis). Targets that are missed by ATD routines may have a much higher regret possibility than a weapon system throwing up false alarms. As noted by Verly in 1989, the unchanging difficulty is that the weapon's hardware must dependably handle a variety of targets under a variety of conditions.[127] Furthermore, target detection is materially different from target identification and, as noted by Lipton and others, requires incorporation of additional (and challenging) dissection tools.[128] Surprisingly, this may be complicated by a general mismatch that continues between weapon system requirements and available computational power.[129] Woods highlights the amount of data that will be generated by these machine visual perception routines[130]; were the images provided by human eyes to be stored, more than one hundred gigabits of data each day would be created from this single modality at even a low sample rate of one picture per second, again demonstrating the pervasive link that exists between AWS hardware and software. In this case, the hardware issue is that 'meaning' lies in *relationships* within that data and not in the data itself (the basis behind machine learning's 'context sensitivity problem'[131]). The ICRC notes that the proprietary natures of commercial ATR and ATD modules generally create bottlenecks that complicate hardware collaboration and knowledge-sharing between colleague weapons and agencies.[132]

Military hardware has generally been deployed in an environment *as encountered* with limited opportunity to edit in advance, structure or map that environment (all seeming preconditions for

[124] US Department of the Army, 'The Targeting Process', Appendix E-1 and F-1.

[125] See Chapter 4 (*Deployment*), specifically: 4.3 ('*Machine and human teaming models*').

[126] Andrew Reiner, 'Effects of Automatic Target Detection on Detection and Identification Performance'*, Department of Industrial Engineering*, University of Toronto, (2016), p. 3 <https://tspace.library.utoronto.ca/bitstream/1807/70554/1/Reiner_Adam_J_201511_MAS_thesis.pdf>.

[127] Jacques G Verly and others, 'Machine intelligence technology for automatic target recognition', generally.

[128] Alan Lipton and others, 'Moving Target Classification and Tracking from Real-Time Video'*, Application of Computer Vision*, Fourth IEEE Worksop, undated, pp. 1-2 <http://www.vision.cs.chubu.ac.jp/MPRG/C_group/C001_Lipton1998.pdf>.

[129] Some context is useful here. Even a bee has more than a gigabyte of storage and performs several tera-operations per second (See: DARPA Neural Study, *AFCEA International Press*, November 2008, generally).

[130] David Woods and others, 'Can we ever escape from data overload? A cognitive system diagnosis'*, Cognition*, Technology and Work, April 2002, 4, 1, p. 22 <https://link.springer.com/article/10.1007/s101110200002 [accessed 6 June 2016].

[131] Derek Ball, 'What Are We Doing When We Theorize bout Context Sensitivity?'*, St Andrews University*, undated, pp. 1-2 <https://www.st-andrews.ac.uk/~db71/ball_context.pdf>.

[132] International Committee of the Red Cross, 'New Technologies and Warfare', *ICRC*, 94, 886, (Summer 2012), pp. 457-458.

efficient AWS deployment).[133] This might compromise function against intelligent adversaries seeking to defeat the technology in use, requiring 'broad robustness' in design in order to deal with deception, assault and counter-autonomy tools.[134] Critics, moreover, of autonomous hardware abound. As noted by USAF General Hostage, 'Predators and Reapers are useless in a contested environment. Pick the smallest, or weakest country with the most minimal air force, [it] can deal with a Predator'.[135] There are various ramifications that arise from his view. Predator landings are empirically complex as the unit's large wing-area makes it sensitive to wind gusts.[136] A further challenge arises from existing time lag that is caused by current satellite communications, exacerbated by operation in communications-denied battlespace.[137] Indeed, UAVs are already acutely vulnerable over urban areas to radio-frequency interference, in particular from Warlock jammers used to prevent remote control of IEDs. As highlighted by Hambling, there is currently a frequent and unhelpful sound signature heard at the target *prior* to a Predator's engagement.[138]

Two conclusions therefore arise from this chapter. The first is the gulf between, on the one hand, emergent hardware (that might individually suggest sufficient advance has been made to warrant removing human supervision from weapon systems) and, on the other, the appropriate knitting together of such technologies to create an independent weapon that is appealing to the Deployment Cohort and still compliant under LOAC.[139] Second, hardware and software challenges to compliant AWS deployment are inextricably linked and should be considered in tandem. In this way, an analysis of issues facing AWS hardware folds into the purpose of this thesis' later chapters, the identification of shortcomings that *cumulatively* undermine the case for removing a human from the loop and the pinpointing of likely frailties in AWS deployment (the constituents of 'technical debt'). It also supports the conclusion that contextual considerations trump technical considerations in deploying such assets.[140] It is only within this basis that behavioural and technical constraints can now be evaluated against the central issue of AWS oversight.

[133] Inferred from: UK House of Commons Defence Committee, 'Gambling on 'Efficiency': Defence Acquisition and Procurement', *First Report of Session*, 2017-2019, (2017), p. 6 and p. 9 <https://publications.parliament.uk/pa/cm201719/cmselect/cmdfence/431/431.pdf>.

[134] US Department of Defense, 'Summer Study on Autonomy', p. 13.

[135] USAF General Mike Hostage, 'Drone combat missions may be scaled back eventually, Air Force chief says', *cit. Washington Post,* 13 November 2013 <http://wapNd> [accessed 16 March 2018].

[136] John Hawley, 'Automation and the Patriot air and missile defence system', Sections 5 ('*Ineffective human-automation integration*') and 6 ('*Observations, lessons and cautions*').

[137] Ramon Ahren, 'Mission Control in a Communications Denied Environment', *Air War College*, Montgomery, (16 February 2017), generally.

[138] Hambling, p. 48.

[139] Chapter 6 (*Wetware*), specifically: 6.3 ('*The AWS Delivery Cohort*').

[140] As set out in: Sections 4.7 ('*Operations and causes of failure*') and 7.1 ('*Sources of technical debt*').

10. Oversight: Command and control constraints to AWS deployment

Significant transformation must take place in how armies fight if AWS are practically to be deployed. An assumption for this chapter is that removing supervision in weapons must materially transform *how* combat is undertaken. Regardless of deployment model[1], several well-tried concepts that have long comprised battlecraft[2] will require fundamental reexamination as autonomy is introduced throughout combat practices.[3] Such reappraisal, however, must account for the human factors that form part of that battlecraft. Accordingly, the aim of this chapter is to review the role of the human in combat's command and control following adoption of unsupervised weapons. It is also to consider the weighting that should be applied to the human dimension of AWS deployment. A starting point is provided by Warne whereby '[t]echnological materials (sic), while valuable in their own right, are not as valuable as human life, in every (sic) operational scenario'.[4] The imperative under review is that human involvement comes with obligations that outweigh any 'value' of physical equipment.

The chapter is comprised of four sections. An analysis of 'command' across its several levels leads first to a review of weapon targeting in order to gauge whether such processes' obligations can feasibly be captured by code. The chapter then considers the scope of new behavioural competencies that will be needed as AWS are deployed. This requires a review of skill sets, the development of individual and team capabilities, innovative proficiencies and benchmarks as well as new models of leadership. The chapter's final section, this thesis' synthesis, is only then able to review Meaningful Human Control (MHC) as an appropriate (and statutory) benchmark with which to frame adoption of autonomy across battlefield practices. For the purposes of this chapter, the command chain is defined as the line of authority *and* responsibility along which orders are passed within units and between units. It is the will of that commander expressed for the purpose of bringing about a particular action. Control is then that command (which might be less than full authority) exercised by the local commander over part of the activities of subordinate or other organisations.[5] This chapter primarily concerns control from the perspective of the Delivery Cohort

[1] See, generally: Chapter 5 (*Deployment*).

[2] Samuel Bendett, 'Russia Poised to Surprise the US in Battlefield Robotics', *Defense One*, (25 January 2018), paras. 5-9 of 16 and generally <https://www.defenseone.com/ideas/2018/01/russia-poised-surprise-us-battlefield-robotics/145439/> [accessed 18 October 2018]. For a discussion on battlecraft in AWS deployment, see also: Michael Guetlein, 'Lethal Autonomous Weapons – Ethical and Doctrinal Implications', *Researchgate*, (February 2005), Abstract.

[3] John Govern, 'The Importance of Distance in Modern Warfare*', Modern Warfare Institute*, West Point, (16 May 2016) <https://mwi.usma.edu/reexamination-distance-modern-warfare/> [accessed 29 July 2017].

[4] Leoni Warne and others, *'The Human Dimension of Future Warfighting'*, *Australian Department of Defence* (Defence Science and Technology Organisation), (September 2004) <http://www.dodccrp.org/events/9th_ICCRTS/CD/presentations/7/162.pdf>. The Australian study is useful in highlighting human traits that challenge AWS deployment, specifically the behaviour divide between the warrior (*discipline, decisiveness, loyalty, confidence*) and the peace-keeper in his responses (*patience, empathy, responsibility, rapport, lesson-learning*) to combat scenarios. The study also emphasises the increasing role of trust (*devolving responsibility to lower levels, disseminating information to ever wider audiences*) and context (*antidote to volume, presentation, testing reliability in battlefield data*) that remain likely in future warfighting and yet incompatible with models of AWS deployment.

[5] Source: US DOD, 'Dictionary of Military and Associated Terms', *US DoD Publications,* (November 2018) <https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/dictionary.pdf>.

and, in particular, the human commander deciding upon AWS deployment.[6] It also touches on control from the deploying States' standpoint by reviewing consequences that might arise from statutory constraint of AWS.

Several headings in the UK's 2011 *Army Doctrine Primer* would seem to conflict with AWS deployment, especially around 'shaping tasks', 'the Decisive Act' and that primer's definition of the general nature of battlefield tasks.[7] In this case, doctrine is the expression of how military forces contribute to war (from campaigns down to individual engagements). It acts as a guide-to-action rather than a set of defined rules. It is also, notes Spencer, a common frame of reference that 'reflects [an] Army's views about what works in war based on past experience'.[8] In considering the risks of removing human oversight, command is then the appropriate meld of control, authority, and permissions as well as the power to influence or direct behaviour within courses of events.[9] It is difficult to see how the Primer can remain fit for purpose in an environment where certain engagements are to be undertaken without human oversight. Regardless of technical feasibility, a deployment issue will be how command can be exercised given the radical recasting that must first take place if areas of battlecraft are to be undertaken without humans in-the-loop.[10] Much of this required transformation relates, notes Cordingley, to broad control structures and the fit that must be maintained between weapons-directing AI, weapons' battlefield tasks and adopted rules of engagement.[11] While the trials of command are well documented[12], Thompson points to new challenges created by exponential growth in the *dimension* of battlefield activities.[13] UK Army doctrine again provides an appropriate starting point for this analysis. Command formulae (as set out in the UK Army's *Land Operations* publication) include themes such as 'unity of effort', 'freedom

[6] See: Chapter 6 (*Wetware*), specifically: 6.3 ('*The Delivery Cohort*').

[7] UK Army, 'Army Doctrine Primer', AC 71954, (UK Army Doctrine Publications, May 2011), section 4.7 ('*Military Activities in the Land Environment*'), <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/33693/20110519ADP_Army_Doctrine_Primerpdf.pdf>.

[8] John Spencer, 'What is Army Doctrine?', *Modern War Institute*, (21 Marcb 2016), para. 8 of 19 <https://mwi.usma.edu/what-is-army-doctrine/> [accessed 9 March 2018].

[9] Department of US Army Headquarters, 'The US Army Functional Concept for Battle Command 2015-2024', *TRADOC Pamphlet*, 525-3-3, Version 1.0, (30 April 2007), p. 19 ('*Mission Command*') <http://www.tradoc.army.mil/tpubs/pams/p525-3-3.pdf>. This also encompasses legal responsibility as set out in Chapter 5 (*Obstacles*).

[10] Ibid., section 4.9 ('*Tactical Actions in the Land Environment*'). Also: Major-General Patrick Cordingley, Commander, 7th Armoured Brigade, Gulf War, 1991, in conversation with the author, January 2019.

[11] Source: UK Government, 'Rules of Engagement', *Wikileaks*, UK and Danish ROE, (June 2008) <https://file.wikileaks.org/file/uk-danish-roe-iraq-2006.pdf>. Although outside the remit of this analysis, see also: UK Government, 'The Joint Service Manual of the Law of Armed Conflict', (Joint Services Publication 383, 2004 Edition), pp. 21-26, pp. 51-100 and pp. 101-104.

[12] David Johnson and others, 'Preparing and training for the full spectrum of military challenges: Insights from the experiences of China, France, the United kingdom, India and Israel', *National Defence Research Institute*, Rand Corporation Publishing, (2009), pp. 15-16, pp. 236-247 and pp. 257-275 <https://www.rand.org/content/dam/rand/pubs/monographs/2009/RAND_MG836.pdf>.

[13] Loren Thompson, 'Five Reasons the Army's New Battlefield Networking Strategy Won't Work', *Forbes* Magazine, (20 November 2017), generally <https://www.forbes.com/sites/lorenthompson/2017/11/20/five-reasons-why-the-armys-new-battlefield-networking-strategy-wont-work/> [accessed 12 July 2017]. For a review of associated literature, see: Australian Army Occasional Papers, 'Command and Control in Modern Warfare', *Command and Leadership,* Series 001, (September 2017) <https://www.army.gov.au/sites/g/files/net1846/f/publications/command_control_b5.pdf>.

of action', 'building of trust', 'timely and effective decision-making' and 'mutual understanding'.[14] The framework is thus inappropriately conditional *and* inappropriately contextual for those AWS' deployment models earlier discussed. It is conditional because it relies upon complex sets of requirements first being met (authority, permissions, trust and processes).[15] While such conditions may not be new (Van Creveld was highlighting the vulnerability arising from the dispersion of modern armed forces in the 1980s[16]), removing human supervision *disruptively* crosses technical, legal and ethical boundaries thereby making such conditionality unworkable.[17] It is also contextual because it concerns intangibles and processes where, as above, basic challenges exist even to capturing their meaning in computer code.[18] Put simply, it is command's basic processes that complicate deployment including routine estimating, routine decision-making, the assigning and executing of tasks, deriving missions and formulating concepts, acquiring and processing feedback as well as communicating intent.[19]

This irreducibility of command into code is noted by General Sir Rupert Smith in his framework on leadership, in particular his 'trust test' whereby subordinates will follow their commander based on 'intangible principles of comradeship, respect, endurance and sacrifice regardless of their situation'.[20] Smith usefully conflates these traits with 'an enduring bond' underpinning, he notes, *all* battlefield activities.[21] Similarly, Smith continues, trust is based on 'character' and 'competence' and must be laid down 'over decades' in advance of battle. What then is this framework's relevance to AWS deployment? To Smith, the impact of leadership on outcomes arises from its bedrock of moral and physical courage.[22] Leadership can 'crisis-proof' a battle-plan and, appropriately executed, is one component in fostering willingness to delegate, to innovate in adversity and cement the army's chain of command.[23] Cordingley notes that introducing a means of lethal engagement which wholly falls outside this framework must clearly overturn current practice; even if AWS are to be adopted by States at the margins, piecemeal adoption must still

---

[14] UK Army, 'Land Operations', *Land Warfare Development Centre*, (Army doctrine publication, AC 71940, undated) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/605298/Army_Field_Manual__AFM__A5_Master_ADP_Interactive_Gov_Web.pdf> Specifically: 6.12 ('*Unity of effort*'), 6.13 ('*Freedom of action*'), 6.14 ('*Trust*'), 6.15 ('*Mutual understanding*'), 6.16 ('*Timely and effective decision-making*').

[15] Justin Lynch and Lauren Fish, 'Soldier Swarm: New Ground Combat Tactics for the Era of Multi-Domain Battle', *Modern War Institute*, West Point, (5 April 2018) <https://mwi.usma.edu/soldier-swarm-new-ground-combat-tactics-era-multi-domain-battle/> [accessed 11 June 2018].

[16] Martin van Creveld, *Command in War*, (Harvard University Press, 1985), p. 2.

[17] Chapter 5 (*Obstacles*) and Chapter 11 (*Conclusion*).

[18] See: Chapters 7 (*Firmware*) and 8 (*Software*), specifically: 8.1 ('*Coding methodologies'*).

[19] US Army Field Manuals, 'Battle Command', *FM 7-30, The Infantry Brigade*, Chapter 3, (1995 and revisions) <https://www.globalsecurity.org/military/library/policy/army/fm/7-30/Ch3.htm> [accessed 2 February 2019].

[20] Previously Deputy Supreme Allied Commander Europe. Author of: 'Utility of Force: The Art of War in the Modern World', *Vantage*, (2008). Here, Buckingham University, Masters in Modern War Studies lecture and subsequently in conversation with the author, 11 January 2017.

[21] Ibid.

[22] See also: Sgt Nicholas Holmes, 'Leadership and Resiliency Training from a Soldier's Perspective', *US Army, www.army.mil*, (19 December 2017) <https://www.army.mil/article/198421/leadership_and_resiliency_training_from_a_soldiers_perspective>, [accessed 14 July 2018].

[23] Buckingham University, Masters in Modern War Studies lecture and subsequently in conversation with the author, 11 January 2017.

compromise function in what is a tested eco-system.[24] While previous deployment of new weaponry has certainly brought about changes to battlefield processes, the conclusion here is that AWS deployment will generate unprecedented disruption[25] that will be evident in force projection, the acceleration of obsolescence in battlefield assets and operational concepts and, suggest Davis and Wilson, in an 'unusual inadequacy should parties react merely incrementally to the removal of human supervision in the use of force'.[26] It is not by accident that this thesis continually links AWS compliance and AWS utility: if AWS efficacy is so poor as to deplete users' trust then its Delivery Cohort will presumably simply ignore the system.

Army Leadership Insights, published by the UK's Centre for Army Leadership, provides further pointers to why AWS' coding for command and control is enduringly infeasible.[27] Clark's paper on *The Intelligently Disobedient Soldier* highlights lasting battlefield benefits of curiosity, critical thinking, imagination and the open-mindedness that arises from intelligent challenge within military frameworks. AWS architecture (based on its strict ML spine, training sets and defined learning) is unsuited to enable such flexibility.[28] Wilson's paper, also from the Centre and titled *What the Hell Do We Do Now?*, points instead to the frequency whereby *no* solutions are available in battlefield scenarios. In this case, it must also be impossible to capture every such scenario in code (apart, in extremis, from ensuring the AWS closes down or otherwise renders itself useless to the Delivery Cohort).[29] Cooper's work on *Empowerment: Beyond Delegation* similarly notes the advantages in combat situations of free thinking, bottom-up generation of ideas, innovation and collaboration, none of which can be captured in AWS routines that are expressed in code.[30] Skinner's *Learning to Change* highlights the empirical premium of parties (here, the human soldier) who are able constantly to learn. There is, moreover, an important differential between, on the one hand, 'poor learners' (rules-based, rote learning, a proxy here for AWS) versus innately human mechanisms that are better adapted to uncertain situations where previous experience is repurposed in order improve outcomes in subsequent scenarios.[31] Finally to this point, Grodecki

[24] Major-General Patrick Cordingley, in conversation with the author, January 2019. As with the majority of this thesis' review, the analysis focuses on State rather than non-state deployment of unsupervised weaponry. See: Chapter 1 (*Introduction*), specifically: 1.2 ('*Introduction to key concepts'*).

[25] Examples here might include precision weapons, advanced air-defence, advanced anti-ship, long-range delivery weapons, space and cyber munitions.

[26] Paul Davis and Peter Wilson, 'Looming Discontinuities in US Military Strategy and Defense Planning: Colliding RMAs Necessitate a New Strategy', *National Defense Research Institute*, RAND, (2011) <https://www.rand.org/content/dam/rand/pubs/occasional_papers/2011/RAND_OP326.pdf>, p. 3 and generally. For analysis of likely disruption to current models, see: Chapter 4 (*Deployment*), specifically: 4.7 ('*Operations and causes of failure*').

[27] Souce: Centre for Army Leadership, <https://www.army.mod.uk/who-we-are/our-schools-and-colleges/centre-for-army-leadership/army-leadership-insights/> [accessed 7 November 2018].

[28] See Chapters 7 (*Firmware*), specifically: 7.2 ('*Firmware ramifications of learning methodologies'*) and 7.3 ('*Reasoning and cognition methodologies'*) and 8 (*Software*), specifically: 8.8 ('*Behaviour setting and coordination'*). Professor Lloyd Clark, 'The Intelligently Disobedient Soldier', *Centre for Army Leadership*, (March 2017).

[29] Ibid. See also: Luke Wilson, 'What the Hell Do We Do Now?', (August 2017). By inference, Wilson highlights the useful coding distinction between 'what to do in a particular situation' versus 'how to prepare a team for what to do in a particular situation'. He also points to a bias in general leadership protocols as well as the transactional nature of task allocation under pressure as opposed to transformational styles of action selection for outcomes requiring long term development.

[30] Paul Cooper, 'Empowerment: Beyond Delegation', (April 2018), generally.

[31] Ibid. See also: Kirsty Skinner, 'Learning to Change', (July 2018), generally.

and Turner borrow from Shelley to highlight that 'nothing is so painful… as a great and sudden change'.[32] A danger for the Delivery Cohort is also that group-think arises (within the Cohort or, more likely, within AWS' code strings) either from unsuitable hierarchies or from group dynamics within parties constructing weapon routines.[33] This faultline is likely to be accelerated without the new education, training, development and command relationships that removal of supervision must occasion.[34]

Isolating the tenets of battlefield command provides a pointer to whether each can be incorporated into the routines that are posited for AWS deployment. This challenge is exacerbated by command's complexity given the plethora of specialized troops, units, functions and equipment (and attendant command structures) that comprise a modern army.[35] In this context, the coordination and control of force has several characteristics that cannot be overlooked by AWS' Delivery Cohort. The speed and range of modern weapons have already reduced the time in which to exercise such control.[36] Machines' capacity for fast, accurate calculation has now exceeded that of the human commander and it is therefore technical advance (rather than operational thinking) which is determining the role for those weapons' autonomy.[37] In order to protect hardware assets, it is also these same capabilities that will see armies spread out over considerable areas, further complicating the process of command.[38] Similarly, command and control complications will arise in the *division* of tasks between human commander, human subordinate and autonomous componentry. It is difficult to foresee how these same control complexities (dynamic authority and permissions, character and competence) should be organized in AWS. They require comprehensive fact collection (second nature for the human soldier but, in the case of AWS, requiring complicated real-time processing and, notes both Kim and Doare, likely interpretation error).[39] Introducing

[32] Adam Grodecki and Ruth Turner (together, The Forward Trust), 'Leading Responsibly Through Change: A Call for Creative Conflict', (September 2018).

[33] They also point to a phenomenon of general 'busyness' (here, a proxy for AWS operation) to evidence inertia and a reluctance to embrace change.

[34] Professor Lloyd Clark, in conversation with the author, September 2018. Also: Major-General Patrick Cordingley, in conversation with the author, January 2019.

[35] Air and Space Power Mentoring Guide, *Three Levels of War,* 1, (Air University Press, 1997) <https://www.cc.gatech.edu/~tpilsch/INTA4803TP/Articles/Three%20Levels%20of%20War=CADRE-excerpt.pdf>.

[36] The example of command and control constraints is usefully illustrated in emerging cyber weapons. See: Joseph Gardiner and others, 'Command and Control: Understanding, Denying and Detecting', *University of Birmingham*, (February 2014), pp. 6-8 ('*The Command and Control Problem'*) <https://arxiv.org/pdf/1408.1136.pdf>.

[37] US Department of Defense, 'The Role of Autonomy in DoD Systems', *Task Force Report*, (2012), p. 21 <https://fas.org/irp/agency/dod/dsb/autonomy.pdf>, pp. 7-10 ('*Autonomous Systems Pose Unique Acquisition Challenges'*).

[38] In so doing, this paradoxically reinforces an attraction of independent weaponry. See: Canadian Department of National Defence, 'Adaptive Dispersed Operations: The Force Employment Concept for Canada's Army of Tomorrow', *Directorate of Land Concepts and Designs*, (2007), pp. 16-22 ('*The Adaptive Dispersion Operation Concept'*) and pp. 28-29 ('*Command'*) <http://publications.gc.ca/collections/collection_2009/forces/D2-188-2007E.pdf>. The study usefully defines command as a 'human endeavour [that] depends on culture, the need to accept risk and instil trust. It is the creative expression of human will necessary to accomplish a mission'. See also: TN Dupuy, *The evolution of weapons and of warfare,* (Indianapolis, 1980), p. 312; cit. van Creveld, *Command in War*, p. 277.

[39] See: Kim, 'Enhanced Battlefield Visualisation for Situational Awareness'*, Computer and Graphics*, 27.6, (2003), pp. 873-885 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.87.6406&rep=rep1&type=pdf>. See also: Chapter 7 (*Firmware*), specifically 7.4 ('*Attention methodologies'*).

weapon independence clearly weakens this already fragile and subjective arrangement.[40] Nor are the arguments straightforward: The dependence, for instance, of command systems on electronically transmitted and encrypted data has made them disproportionately open to electronic warfare designed to interrupt their flow and, notes Wilgenbusch, paradoxically encouraging further independent processes and the deployment of unsupervised weaponry. [41]

It is useful to consider weapon control and its consequences from the perspectives of AWS' procurer and operator. Currently, the user axiom for weaponry may broadly be a *threshold for liability* and an *obligation to prevent harm.*[42] This is relevant as it suggests MHC be right in the middle of those routines that identify, select and apply force to targets[43] (identified, after all, by ICRC as the 'critical functions' of a weapon).[44] A crux then to any removal of supervision becomes the level, nature and primacy of human control over specific weapon *functions* rather than any link between such control and specific technologies (given, after all, the speed with which those technologies will change as weapon systems evolve).[45] Article 36 refines further this relationship to include 'when, where and how weapons are used; what or who they are used against; and the effects of their use'.[46] Notwithstanding Garcia's note that several ambiguities still require address through negotiation[47], the adoption of MHC must distinguish between autonomous weapons that are covered by such a ban and the many existing weapons that already have autonomous functions. It would be similarly unworkable to base any restrictions on what is an artificial separation between offensive and defensive tasking.[48]

Regardless of framework, the policing of both compliance and verification (a further oversight process) remains an outstanding constraint in AWS deployment as it will be challenging to monitor, inter alia, either machine intent or whether a contentious engagement was carried out with or

[40] Ronan Doare and others, 'Robots on the Battlefield: Contemporary perspectives and implications for the future'*, Army Combined Arms Center*, Fort Leavenworth, KS Combined Studies Institute, (2014), generally.

[41] Ronald Wilgenbusch and Alan Heisig, 'Command and control of vulnerabilities to communications jamming'*, JFQ*, ndupress.ndu.edu, 69, (June 2013) <http://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-69/JFQ-69_56-63_Wilgenbusch-Heisig.pdf>.

[42] Human Rights Watch, 'Killer Robots and the concept of meaningful human control'*, Memorandum to Convention on Conventional Weapons*, (April 2016), p. 2.

[43] Article 36, 'Key elements of meaningful human control'*, Memorandum to Convention on Conventional Weapons*, April 2016, p. 1.

[44] Source: Statement of the International Committee of the Red Cross, CCW Meeting of Experts on Lethal Autonomous Systems, Geneva, 13 April 2015.

[45] UK MOD, 'Human Machine Touchpoints: The United Kingdom's perspective on human control over weapon development and targeting cycles', *UK submission to CCW GGE on LAWS*, (August 2018), generally.

[46] Article 36, 'Killing by a machine: Key issues for understanding meaningful human control', cit. Human Rights Watch, 'Killer Robots and the Concept of Meaningful Human Control'*, Memorandum to Convention on Conventional Weapons CCW Delegates*, (April 2016), p. 2.

[47] Denise Garcia, 'Governing Lethal Autonomous Weapon Syst*ems*'*, Ethics and International Affairs*, Carnegie Council, (December 2017) <https://www.ethicsandinternationalaffairs.org/2017/governing-lethal-autonomous-weapon-systems/> [accessed 13 August 2018].

[48] These would fail to capture AWS based, for instance, on mobile robotic vehicles. Other distinctions have been suggested from this user/procurer perspective that might better be policed such as the difference between fixed and mobile weapons or between recoverable robotic vehicles and non-recoverable munitions.

without human authorization.[49] To this point, a different control avenue might instead be to consider a ban on AWS that specifically target people under the separate axiom of 'let machines target machines and let people target people'.[50] A further regulatory option might then be to create a *non-legal* code of conduct on users and procurers of AWS that is centered on simple, self-enforcing rules such as 'robotic vehicles should not fire unless fired upon' and 'returned fire should be limited, proportionate and discriminating'. While such oversight models might check escalation and ensure predictable reactions from participating States, it should also be assumed that such rules will likely collapse in war. It is for such real-politik reasons that this chapter instead focuses upon enshrining human judgement (here, MHC) as the basis of a legal framework to govern AWS deployment.

Before considering MHC, two further considerations arise from a review of control. Any prohibition framework should neither anticipate nor legislate for capabilities that might emerge in due course. Likewise, the analysis implies that certain AWS characteristics (technical infeasibility, issues of trust and reliability as well as contextual drivers) will combine to ensure that human agency *does* remain enduringly present in lethal engagement. It is therefore this thesis' contention that AWS control is best achieved by articulating what is a *positive* requirement for human oversight in the use of force, an obligation that irreducible human control (and therefore human judgement) must precede a machine's initiation of violence, regardless of whether that violence is lethal. This accords, after all, with the general observation (here, set out by the Holy See) that 'prudential judgement *cannot* be put into algorithms'[51] and where exercise of judgement depends on more than numeric analysis of data. It is, notes the Holy See, too difficult for AWS, no matter how much data is processed, to exercise required levels of judgement.[52]

10.1 Meaningful Human Control

Given the human factor of battlefield command, this final section reviews MHC as a mechanism to retain human participation in otherwise independent weaponry. This thesis' recommendation is that consequential human supervision be enshrined in all situations involving force. Where, however, should this intervention 'sit'? Is it within the broad act of employing violence or should it be around specific identification of a target as legitimately hostile? Robinson notes that command responsibility (like that of State responsibility) considers control to be a prerequisite for assigning liability.[53] The circumstance of AWS deployment as an autonomous 'vehicle of judgement'[54]

[49] The concept, for instance, of 'plausible deniability' is discussed in the introduction to Chapter 4 (*Deployment*). See also: Chapter 8 (*Software*), specifically: 8.7 ('*Action selection issues*'). Oversight of AWS is considered in Chapter 10 (*Oversight*).

[50] Scharre, *Army of None,* Norton Publishing, 2018, p. 355.

[51] Statement of the Holy See, *CCW Meeting of experts on lethal autonomous systems*, Geneva, (16 April 2015), p. 4.

[52] Ibid., pp. 5-6.

[53] Darryl Robinson, 'How Command Responsibility Got So Complicated: A Culpability Contradiction, its Obfuscation, and a Simple Solution', *Melbourne Journal of International* Law, 13, 1, (2012), 2-6 ('Terminology') and 7-9 <https://law.unimelb.edu.au/__data/assets/pdf_file/0003/1687242/Robinson.pdf>. This is variously covered in Chapter 5 (*Obstacles*), specifically: 5.1 ('*The Geneva Convention and Laws of Armed Combat*').

[54] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

assumes, moreover, several other broad obligations.[55] In order to ensure compliance, involved parties must each understand *how* AWS systems will operate such that controlling individuals can make informed (and thus legally compliant) decisions regarding the *use* of those weapons.[56] Each deployed weapon must also generate evidence of its reliability and performance in order for human decision-making to be appropriately accountable.[57] It is this 'obligation on the person' that is articulated by the most recent US DoD *Law of War Manual*.[58] LOAC also obliges AWS-deploying parties to assume a portfolio of obligations including, inter alia, those tests of distinction and calculations around proportionality discussed above.[59] Neither Delivery Cohort nor battlefield commander can meet that legal obligation unless proper information on the context of *each* individual attack (and, indeed, its expected effects) is reasonably understood at the point of decision.[60] This requirement, moreover, de facto precludes AWS from operating 'without strict bounds in space and time', a further material challenge to their deployment.[61] AWS, then, that operate without communication are unlikely to be able to fulfill a commander's obligation to undertake these timely LOAC calculations. The inference must be that an unsupervised weapon, out-of-touch with the human responsible for using that weapon, cannot appropriately undertake proper situational assessment given AWS' prima facie reliance on timely, fresh and appropriate information. A communication link would appear vital under this same analysis in order that the AWS receive authorization for *all* individual attacks.[62] The conflict here is that such ratification (effectively the subjection of AI to human review *before* it is put into effect) also creates disadvantage.[63] Complexity, after all, is being added by the degree of exercisable modification/veto. Temporal complications will also arise from command bottlenecks and technical snags. In this case it will be easier for the Delivery Cohort to minimise non-compliant engagement rather than devote resources trying to optimise the machine's every detail in what is a fully autonomous engagement. This characteristic would point to continued adherence of MHC in lethal engagements.

---

[55] Jack Beard, 'Autonomous Weapons and Human Responsibility', *University of Nebraska-Lincoln*, College of Law Faculty Publications, 196, (2014) <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=http://scholar.google.co.uk/&httpsredir=1&article=1196&context=lawfacpub>, pp. 622-625.

[56] See: Chapter 5 (*Obstacles*), specifically: 5.1 ('*The Geneva Convention and laws of armed combat'*) and 5.5 ('*Article 36 and LOAC-complaint weaponry'*).

[57] Inferred from: Roff and Moyes, p. 3. See also: Chapter 2 (*Context*), specifically: 2.6 ('*The role of situational awareness and uncertainty*').

[58] United States Department of Defense, 'Department of Defense Law of War Manual', (June 2015), P.6.5.9.3; 'LOAC obligations of distinction and proportionality apply to *persons* rather than the *weapons* of themselves... as these rules do not impose obligations on the weapons… and of course an *inanimate* object could *not* assume an obligation in any event'.

[59] See Chapter 5 (*Obstacles*), specifically 5.1 ('*The Geneva Convention and Laws of Armed Conflict'*).

[60] Darryl Robinson, pp. 20-23 ('The Problem of the Successor Commander'). This also raises the issue of that commander's selection, training, education and development and the type of individual able to undertake this tasking. See this chapter, specifically: 10.2 ('*Required new competencies in human resources'*). Also: Professor Lloyd Clark, in conversation with the author, September 2018.

[61] Roff and Moyes, p. 4.

[62] Chapter 5 (*Obstacles*), specifically: 5.1 ('*The Geneva Convention and Laws of Armed Combat'*) and 5.5 ('*Article 36 and LOAC compliant weaponry'*) charts that it is beholden on the human decision-maker to ensure proper situational awareness for each such attach in order responsibly to grant that authorization.

[63] Bostrom, *Superintelligence,* p. 226.

At its most basic level this thesis' support of MHC is based on two premises.[64] First, MHC conforms with ICRAC's contention that it is inherently wrong for a weapon to be fired if the human being at the point of firing has not properly aimed it. Without MHC, after all, a human is fundamentally not making the decision to initiate violence. [65] There remain, however, several ways to frame the principle of MHC. A human simply pressing a fire button (in response, for instance, to indications from a computer) without cognitive awareness is clearly insufficient to be considered 'human control' in any substantive sense. The term 'meaningful' therefore presents substantial space for diverging opinions on where the boundaries of necessary human control might lie.[66] A key is also whether MHC is applied to the technology itself or to the wider action within which the technology might be applied. It is noteworthy that the efficacy of human control is already patchy in existing military systems, questioning thereby the extent to which current practice should shape normative expectations for future weapon systems.[67] A second premise for this thesis' support of MHC is that any regulatory framework must be practical; the drafting of statutory MHC should be appropriately broad in order to obviate the need for pre-assessment of each new emerging weapons technology and to limit subsequent legal and political meddling. As argued generally in this thesis' contextual review, future deployment decisions will likely have overtly political contexts with different actors having quite different interpretations based on local considerations, priorities and interests. MHC, moreover, can also infer *negative* control and the 'prevention of any unauthorized use of such weapons', a clear departure from traditional *positive* control, 'the assurance that authoritative instructions to perform military missions will be carried out'.[68]

The main challenge to the application of MHC is the weapon's targeting cycle. NATO's process for *Joint Targeting* applies both to deliberate as well as dynamic targeting and, as written, currently comprises five contiguous phases as detailed in the footnote.[69] Only one component of that delineated targeting cycle (here, Phase 5, comprising '*mission planning and force execution'*[70]) is currently considered appropriate for automation if compliance is to be retained and is reviewed below in detail.[71] The remaining phases (in particular, commander's intent but also capabilities around analysis, decision and assignment) nevertheless demonstrate the extent of challenge facing broad capability deployment of independent weapons.[72] Each of these components empirically

[64] Article 36, 'Lethal autonomous weapons, artificial intelligence and meaningful human control'*, Briefing document with ASU Global Security Initiative*, (February 2016), p. 2.

[65] Source: International Committee for Robot Arms Control mission statement <http://icrac.net/statements/>.

[66] Meaningful, significant, consequential, material, worthwhile, relevant, appreciable, substantial…

[67] Thomas Adams, *Future Warfare and the decline in human decision making,* 2, (Parameters, 2001), generally <http://ssi.armywarcollege.edu/pubs/parameters/articles/2011winter/adams.pdf>.

[68] Roff and Moyes, p. 5.

[69] NATO Standardisation Office, '*Allied Joint Publication 3.9: Allied Joint Doctrine for Joint Targeting'*, (NSO, Edition A, Version 1, April 2106) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/628215/20160505-nato_targeting_ajp_3_9.pdf>, p. 2-2. Phase 1: '*Commander's intent, objectives and guidance',* Phase 2: '*Target development*', Phase 3: '*Capabilities analysis*', Phase 4: '*Commander's decision, force planning and assignment',* Phase 5: '*Mission Planning and force execution*', Phase 6: '*Assessment*'.

[70] Ibid.

[71] UK MOD, 'Human Machine Touchpoints', p. 3. Specifically, actions around 'find, fix, track, target, engage, exploit and assess' and, as above, only then subject to appropriate regulation, specification, design, verification and adherence to operating processes and ROE.

[72] Chapter 1 (*Introduction*), specifically: 1.2 ('*Introduction to key concepts'*).

comprises just one part of a targeting continuum which is insufficiently delineated for capture by machine code. Similarly, these sub-routines will vary between engagement type and situation. They also vary in risk; generally, as amount of time available for a targeting decision decreases, the amount of risk in that routine increases but this is demonstrably not a reliably linear relationship. Further targeting complexity will arise in cases of close combat as well as from fluid levels of authority in engagements where authority may be dynamically delegated between friendly forces.[73]

Targeting gives rise to other control-related difficulties that require on-going human supervision. Specifically, the nomination and prioritisation of targets must be guided by goals and values despite what may be a dynamically changing utility function. This is not a fixed relationship and becomes a key requirement for coordination.[74] Similarly, targeting by unsupervised systems must contemporaneously factor for time sensitivity, apt payoff calculation, apt target development, appropriate toggling between lethal and non-lethal outcomes, the handling of restricted targets and no-strike entities as well as targeting's 'decide phase'.[75] Any deficiencies in the process will, moreover, be magnified when local communication is lost and battlefield data is compromised.[76] Even Phase 5 of NATO's *Joint Targeting Cycle* (where it is currently judged that automation might be possible) is comprised of seven individual sub-components (the 'fit, fix, track, target, engage, exploit and fail-safe' of '*Mission planning and force execution*' discussed above).[77] This presents considerable scope for imprecision and it is this layering that requires on-going and intimate human management, both to retain legal compliance but also to retain best possible efficiency in each use of battlefield force.[78]

Review of targeting's Phase 5 is useful to support the notion of statutory MHC. The component that determines 'fit' must, for instance, be theatre-specific and derive from precise data which fuses intelligence, surveillance and reconnaissance (ISR) assets if it is to be appropriate. Targeting's 'fixing' routine is then responsible for each engagement's risk assessment (in tandem with the weapon confirming compliance with local ROE) in order to decide upon levels of appropriate force in each engagement such that operational success can be achieved having factored for possible collateral costs. This is technically complex and will be inappropriately prone to error.[79] Similarly, targeting's 'tracking' and 'engaging' phases require that the AWS balance for broad situational awareness as part of compliant targeting, a routine that is repeatedly shown above to be problematic.[80] Finally to this point, targeting must always consider the broader picture of colleague

[73] The end-point of this continuum would therefore be instances of autonomous self-defence. The issue becomes both the setting by a human agent of the weapon's parameters and, second, the means of subsequent human intervention in that engagement. Furthermore, parameter setting (and the engagement's ensuing riskiness) depends on the degree of trust that the human agent can place on the whole targeting cycle. It also rests on any subsequent ability for those parameters, dependent after all on political and strategic considerations, to be changed *by the machine* after that engagement is launched.

[74] See: Chapter 8 (*Software*), specifically: 8.3 ('*Utility function'*), 8.5 ('*Anchoring and goal setting issues'*) and 8.6 ('*Value setting'*).

[75] NATO Standardisation Office, 'Allied Joint Publication 3.9', Section II, p. 5-4 ('*Decide, detect, deliver, assess'*).

[76] See: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt'*).

[77] NATO Standardisation Office, 'Allied Joint Publication 3.9', generally.

[78] NATO Standardisation Office, 'Allied Joint Publication 3.9', Section III, p. 5-7 ('*Targeting at component level').*

[79] Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of technical debt'*).

[80] For a discussion on situational awareness in combat engagements, see: chapter 9 (*Hardware*), specifically: 9.3 ('*Navigational issues'*), and Chapter 11 (*Conclusion*), specifically: 11.1 ('*Nature of deployment challenges'*).

assets, joint operations and dovetail in as part of the local commander's overall (and dynamically evolving) picture. Absent MHC, for example, the weapon must undertake relevant battle damage assessment without third party assistance in order to action accurate feedback following every engagement.

There are also contextual considerations to the definition and implementation of MHC. Oversight of human control within military systems cannot be limited to the weapon's targeting cycle. Instead, weapon control must be present throughout a broad range of scale, intensity and task complexity as well as be able to operate reliably in 'cluttered, congested, complex and contested operational environments'.[81] Contextually, moreover, 'assessment, evaluation and revision of the system and, notes the UK MOD in *Human Machine Touchpoints*, the surrounding political, strategic and operational "wrap" (sic) is a *continuous* cycle rather than a linear, sequential process and a combination of these factors are required to ensure human control at the appropriate points'.[82] It is within this context that discussions to ban AWS have arisen which, unusually, are being led largely by NGOs and not by States-members to the Geneva Convention.[83] It is similarly unusual that AWS deployment is currently positioned as a concern around civilian wellbeing rather than as a strategic issue. As noted by Scharre, 'bans that are motivated by concern about excessive civilian casualties pit [what is] an incidental concern for militaries against [what is] a fundamental priority: military necessity'.[84] Pivoting this consideration of control from the weapon now to the user (here, the Delivery Cohort), the dynamic then becomes the *degree* of control to be ceded and how State signatories might be prejudiced if they comply with existing law versus if they support to a new ban on AWS. As noted by Etzioni, the prevailing assumption is that weapon autonomy will most benefit advanced militaries (and, that weaker parties are not in practical terms giving up anything should weapon autonomy be peremptorily banned).[85] Analysis, however, of deployment models suggests that once autonomous technologies diffuse across borders and tasks, AWS may actually benefit exactly those weaker parties given the incongruent costs (economic and operational) of maintaining communications in contested environments versus deploying weapons that are capable of independent targeting.[86]

In the case of a statutory MHC framework two further challenges arise. There is no robust answer to how signatories can defend themselves against parties who subsequently deploy AWS in contravention of such a ban. It is also difficult to dismiss the argument that casualty-centric (rather than strategic) initiatives appear merely to disarm those States agreeing not to field AWS. This, notes Scharre, 'would be the worst of all possible outcomes, empowering the most odious regimes with potentially dangerous weapons while leaving nations who care about international law at a

[81] UK MOD, '*Human Machine Touchpoints*, p. 2. Also: Major-General Patrick Cordingley, in conversation with the author, January 2019.

[82] Ibid. These include national regulation and law, adherence to specifications, appropriate human-machine design, appropriate verification and validation as well as observance of operating processes and ROE.

[83] See: Stop Killer Robots Coalition, <https://www.stopkillerrobots.org> [accessed 12 September 2018].

[84] Scharre, *Army of None,* Norton Publishing, 2018, p. 348.

[85] Amitai Etzioni and Oren Etzioni, 'Pros and Cons of Autonomous Weapon Systems', 72-3 and generally. Alternative source: <http://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/> [accessed 12 August 2018].

[86] Introduction to Chapter 4 (*Deployment*).

disadvantage'.[87] Tackling AWS control through MHC is nevertheless supported by IHL's requirement that human control over *individual attacks* should be the relevant unit of legal management in tactical actions.[88] The challenge, after all, is that AWS compliance in targeting must otherwise depend on 'an innate capacity to detect and interpret subtle cues' that must be particular to each engagement.[89] This must presumably include features such as voice tone and body language.[90]

10.2 Validation and testing

Appropriate validation and testing is a central control component of weapon oversight in order to ensure both AWS utility and LOAC compliance. This is complicated on several levels. Given the pace of technical innovation and a likely lack of structure to AWS deployment, Grush argues that deployment models cannot rely upon definite 'baselines' (or, germane to this chapter, a standard battlefield scenario) against which the responsible commander can deploy AWS with appropriate confidence.[91] Conversely, a concern is that possible battlefield gain may encourage deployment notwithstanding that testing is expensive and time consuming when humanitarian failure might appear to have little tangible cost. Lonsdale's point is also that reliance on such baselines may prove impossible either after first engagement, whether through combat or electronic interference[92] in a process that is likely to follow Moltke's maxim that 'no battle plan survives first contact with the enemy'.[93] While deployment models must also factor empirical difficulties arising in AWS testing, the danger is that inadequate practice then sets the norm. To this point, Knight questions who in the Delivery Cohort can authorise AWS sub-components procured from an array of weapons manufacturers?[94]

The US Department of Defense's directive, *Autonomy in Weapon Systems,* appears quite unambiguous on the matter of testing in stating that AWS must 'go through rigorous hardware and software verification and validation and realistic system developmental and operational test and evaluation'.[95] The issue, however, is whether this is realistic. Within the one hundred and twenty-nine US Marine Corps MV-22 Ospreys that have entered service by early 2018, Freedberg highlights that there are *seventy* different configurations, identical to the untrained eye but all subtly

[87] Ibid., p. 351.

[88] Article 36, 'Key elements of meaningful human control', p. 1.

[89] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[90] Robert Sloane, 'Puzzles of Proportion and the "*Reasonable Military Commander"; Reflections on the Law, Ethics and Geo-politics of Proportionality', Harvard* National Security Journal, 16, (2015), 301-304 <http://harvardnsj.org/wp-content/uploads/2015/06/Sloane.pdf> generally.

[91] Bern Grush, 'The rise of autonomous vehicles: planning for deployment and not just development', *R&D Lab Design*, (24 January 2018) <https://www.rdmag.com/article/2018/01/rise-autonomous-vehicles-planning-deployment-not-just-development> [accessed 10 February 2018], paras. 8-9 of 23.

[92] David Lonsdale highlights the role of deception and enemy attacks to degrade information systems that will be critical to AWS. See: David Lonsdale, *Clausewitz and Information Warfare,* (Oxford: Oxford University Press, 2007), p. 248.

[93] Source: Lexician.com, <http://lexician.com/lexblog/2010/11/no-battle-plan-survives-contact-with-the-enemy/>.

[94] W Knight, 'The US Military wants its Autonomous Machines to explain themselves', *MIT Technology Review*, (14 March 2017) <https://www.technologyreview.com/s/603795/the-us-military-wants-its-autonomous-machines-to-explain-themselves/>, paras. 4-9 of 12.

[95] Department of Defense Directive, *Autonomy in Weapon Systems,* Para 4, Policy 2.

dissimilar and requiring different flight checklists, maintenance procedures and spare parts.[96] Validation and verification procedures (V&V) will be key to AWS deployment precisely because they are intended to provide oversight and corroboration that the deployed systems meets user expectations, that they comply to norms and that they work to specification.[97] Roske suggests that two validating methods will be used for AWS deployment.[98] *Formal* validation methods will be based upon a deductive review of the new system to establish a mathematical proof that the system is working. The deployment challenge here is the complex requirement for complex (and accurate) translation of the AWS' entire properties into a formal mathematical language. *Testing* validation methods will instead use inductive processes to *infer* that the AWS is working based on a representative number sample of test cases. Du notes that AWS testing is hamstrung by the limited amount of available test data and appropriately broad 'simulated input scenarios'. He also identifies that 'formal' methods of validation rely upon the embedding of sufficient logical conditions into AWS' state specifications: An example is a Czech Airbus overrunning the runway in 1992 because its logical condition was that its brakes should be released when its wheels were not turning, caused in this case by a layer of runway ice, and leading Du to conclude that 'no magic solutions exist yet to tackle such an underlying problem behind such software development'.[99] The oversight challenge is a weak proof of each model's correctness.

Additional V&V challenges will impact AWS deployment. In particular, Kot highlights what is termed the 'state-space explosion problem', the AWS' ever-larger decision space and the corresponding increase of options at its disposal.[100] In this case, it becomes impossible to foresee possible event combinations that could lead to system failure. A second challenge is noted by Boulanin and Verbruggen, authors of the SIPRI report, whereby AWS' ML basis will presumably require 'automatic reparameterisation and partial reprogramming' of the entire weapon system after every learning iteration:[101] Each time the AWS learns something new, its performance and correctness will need to be revalidated. A further testing issue across all deployment models relates then to weapon specification. While testing conventional software involves confirmation that behaviour matches the manufacturer's descriptors in the case of every possible input, the likely 'connectionistic' software that will comprise AWS processes, notes Kassan, 'comes with no [ready] specifications and instead is expected to learn patterns or act like a natural system'.[102] Validation is

[96] Sydney Freedberg, 'Streamlined MV-SS Maintenance', *Breaking Defense*, 5 February 2018 <https://breakingdefense.com/2018/02/streamlined-mv-22-maintenance-from-70-osprey-types-down-to-5/?utm_source=hs_email&utm_medium=email&utm_content=60470967&_hsenc=p2ANqtz--Xu8INgREBr-YhTJmbRVeXy27_N9SZ9JPZQr4grwHsYyP--GM_lxTQHRDrX5AM1UrpLsLF8NPRcVjPO4KBvlHvzR9F4w&_hsmi=60470967> [accessed 6 February 2018].

[97] S Russell and others, 'Research priorities for robust and beneficial artificial intelligence', *Future of Life Institute*, Boston, (2015) <https://futureoflife.org/data/documents/research_priorities.pdf>, pp. 108-10.

[98] Vincent Roske and others, 'Autonomous System Challenges to Testing and Evaluation', *National Defense Industry Association test and evaluation conference*, (March 2012) <https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2012/TEST/13782_Roske.pdf>, '*Conference pack*'.

[99] Dr Hongbo Du, School of Computer Science, Buckingham University, in conversation with the author, January 2019.

[100] Martin Kot, 'The State Explosion Problem', unpublished thesis, 2003, p. 1 <http://www.cs.vsb.cz/kot/down/Texts/StateSpace.pdf>.

[101] Boulanin and Verbruggen, p. 70.

[102] Inferred from: Kassan, 'AI gone awry: futile quest for artificial intelligence', *The Skeptics Society and Skeptic Magazine*, undated, p. 3 <https://www.skeptic.com/reading_room/artificial-intelligence-gone-awry/> [accessed 14 September 2016].

therefore challenging exactly because of the circular networks that underpin AWS behaviour.[103] SIPRI's conclusion is similarly clear. As autonomous systems become 'more intelligent, interactive and capable of adapting to complex and dynamic environments, it becomes, practically and financially, infeasible to continue to test all ranges of imports to, and possible states of, the system'.[104] Anderson and Waxman concur, concluding that the number of variables comprising AWS scripting, prototyping and testing, its refining, version control, updates, patching and distribution makes *any* testing of AWS software problematic.[105] On this basis, foreseeable V&V methodology only appears fit for purpose for non-learning machine systems operating in understood and static environments where the *overall* design of the weapon system is understood by technicians.

Mindful of potential high-regret outcomes, AWS testing and validation must be 'red-teamed' whereby an independent, properly resourced group with an adversary's mindset challenges the AWS at all phases of concept evaluation, development, simulation and deployment.[106] In AWS, V&V is also complicated given that rules of engagement must be common across weapon variants, cross-border and standardized. It must also be in place in advance of that deployment and not 'learned' by experience. AWS' programmers must consider policies and regulations arising from a myriad of 'appropriate authorities' that are complex, lack incentives to act quickly and, being multi-jurisdictional, are frequently in conflict.[107] Brat and Jonsonn thus question the practical value of AWS 'planners' (the weapon's mission and flight plan) given that certain deployed systems may need to lodge in advance (and have approved) detailed battle-space intentions in what is a very large 'state space'.[108] Nguyen similarly posits that appropriate V&V techniques need to be in place just to verify for inconsistencies, ambiguities and incompleteness, a dynamic exercise that will need to be undertaken at multiple levels (including output from the AWS' executive, functional, integration and decision layers as well output arising from the interaction and collaboration of these layers).[109] That autonomous agents will rarely, if ever, base action selection on exact values from their sensors[110] further complicates the V&V process.[111] Finally to this validation point, AWS will require sophisticated and real-time V&V in order to search internally for possible electronic or mechanical failure that might impair performance. The challenge is that commercial parties in the AWS' procurement chain have little incentive to devote resources to such testing. In this vein,

[103] See, generally: Chapter 8 (*Software*).

[104] Boulanin and Verbruggen, p. 69. It should be noted that this problem is not exclusive to autonomous weapon systems but it is applicable across all machine autonomy.

[105] Simple financial metrics are likely to limit innovation in this field; in 2006 the US DoD had completed its Sniper Return platform but software verification and testing has prevented the technology from being deployed (cit. Noel Sharkey, in conversation with the author, 25 July 2017).

[106] US Department of Defense, Defense Science Board, p. 19.

[107] Department of Defense, 'Unmanned Systems Integrated Roadmap FY2013-2038', p. 82; For maritime autonomous systems alone see: US Coast Guard, Navigation Safety Advisory Council, 1972 International Regulations for the Prevention of Collisions at Sea. For systems that fly, see FAR, FAA, and the ICAO.

[108] Guillaume Brat and Ari Jonsonn, 'Challenges in verification and validation of autonomous systems', *USRA/RIACS,* p. 5 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.1234&rep=rep1&type=pdf>.

[109] Cu Nguyen and others, 'Evolutionary testing of autonomous software agents', *Autonomous Agents and Multi-agent Systems*, 25.2, (2012) <https://nms.kcl.ac.uk/michael.luck/resources/aamas09d.pdf> p. 7.

[110] Michael Fisher and others, 'Verifying Autonomous Systems', *Communications of the ACM*, 56, 9, (September 2013), pp. 84-93.

[111] US Department of Defense, Defense Science Board, p. 30.

commentators note[112] a clear gap in technical literacy to deliver an appropriately robust regime.[113] Given that unsupervised weapons are complex systems with emergent practices and variable outcomes ('what we must test) rather than traditional systems ('what we test').[114]

The conclusion for this thesis is that MHC is required to bridge what will be an enduring gap between verifiable execution of AWS' specifications (the correctness, reliability and, as above, verification of the weapon's controls) and that software's ability to assess battlefield conditions (generally, the context underlying each engagement) in which these specifications must apply.[115] While simulated testing of AWS may attempt to assess such 'correctness', the fact of CACE demonstrates that AWS testing cannot assure reliability under each condition of use, requiring instead the fail-safe of MHC in all uses of machine violence.[116] This is, moreover, all but recognised by the US Department of Defense's current *Directive* and its stipulation that facility exist for 'terminating engagements or seeking additional human operator input before continuing the engagement'.[117] As AWS are deployed in adversarial environments, the challenge is that these weapons will encounter situations that their designers and others from the AWS Delivery Cohort would never have considered.[118] It is for this reason that MHC is required to mitigate such circumstances in conjunction, of course, with a deployment condition-precedent that AWS 'gracefully vitiate' in cases of malfunction.[119]

[112] Chatham House conference, Autonomous Weapons, February 2014.

[113] Alan Hobbs and others, 'Human Challenges in the Maintenance of Unmanned Aerial Systems', *FAA and NASA report*, (May 2006) <https://humansystems.arc.nasa.gov/publications/UAV_interimreport_Hobbs_Herwitz.pdf>, pp. 9-10, pp. 10-16 and pp. 16-18.

[114] Fil Macias, 'The Test and Evaluation of Unmanned and Autonomous Systems', *International Test and Evaluation Association, ITEA Journal*, 29-4, (2008), 388.

[115] Suchman, 'Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy', p. 5.

[116] CACE refers to *Change Anything, Change Everything*. See: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of Technical Debt'*).

[117] Department of Defense Directive, 'Autonomy in Weapon Systems', para. 4. Unhelpfully, the Directive is not clear in its definition of autonomous and semi-autonomous systems but at least the document's thrust is unambiguous.

[118] US Air Force, 'Autonomous horizons; system autonomy in the air force – a path to the future – human-autonomy teaming', *Office of the Chief Scientist,* AF/ST TR 15-01, (June 2015), p. 6.

[119] See: Chapter 4 (*Deployment*), specifically: 4.7 ('*Operations and causes of failure'*).

11. Conclusion

This work has sought to identify and then assess challenges to the deployment of autonomous weapons. In so doing, the thesis' structure is divided equally between analysis of non-technical and technical impedimenta.[1] For the purposes of this conclusion, these constraints are now classified broadly as soft challenges (the contextual and the behavioural) and hard challenges (the technical and the systemic). This is an important distinction as two divergent characteristics drive this thesis' conclusions. The first is the breakneck pace of battlefield change[2] while the second is the constancy of man's role in battlefield activities.[3] As observed, after all, by Liang and Xiangsui in *Unrestricted Warfare*, the relative roles of soldiers and their weaponry are currently in 'unprecedented flux'.[4] This inference is central to this thesis. As noted by Brooks (and setting aside definitions of 'war'), the processes of war and violence still require human participation notwithstanding that the attacks of 9/11 may demonstrate that 'you don't need soldiers to start a war' (irregular combattants simply hijacked civilian airliners 'with nothing more lethal than boxcutters').[5] The assumption that such ostensibly opposing characteristics – breakneck technical development versus the enduring involvement of man in war's base acts – fuse together is supported by General Sir Richard Barrons, former Commander of Joint Forces Command, in his October 2017 op-ed for *Wired Magazine*: 'A peaceful generation in Europe does not change what war is, yet the character of conflict – how war is fought – always changes as thinking and technology advances'.[6] It is therefore a key premise that the use of technology in battlespaces actually remains a basic human endevour and one that is based upon basic human engagement. As such, it follows that any core of *future* military advantage (here, the adoption of autonomy across battlefield processes) will occasion a step-change in what is already a long-standing trend of integration between humans *and* machines in appropriate war fighting systems, each bespoked to outperform the opponent but still requiring enduring human coordination in that process.[7] As Scharre points out, 'the winner of the robotics revolution will not be who develops this technology first or even who has the best technology, but who figures out how best to use it'.[8] While technical challenges will undoubedly temper AWS

[1] First, non-technical impedimenta: Chapter 2 (*Context*); Chapter 3 (*Drivers*) on drivers to adoption and deployment; Chapter 4 (*Deployment*) on current practices and likely pathways to the removal of human supervision in engagements (Chapter 4); Chapter 5 (*Obstacles*) on legal and other obstacles in front of such adoption Second, technical impedimenta: Chapters 6 and 7 (*Wetware* and *Firmware*) on architectural challenges, Chapter 8 (*Software*) on control issues; Chapter 9 (*Hardware*) on likely equipment deficiencies.

[2] This is covered in Chapter 3 (*Drivers*), in particular: Section 3.2 ('*Technology creep and dual-use technology trends'*).

[3] Chapter 2 (*Context*), specifically: 2.2: ('*The role of Context in AWS' argument'*) and 2.4 ('*Defence Planning'*).

[4] Qiau Liang and Wang Xiangsui, *Unrestricted Warfare,* (USA: PLA Literature and Arts Publishing, 1999) p. 15 <http://www.c4i.org/unrestricted.pdf>

[5] Rosa Brooks, 'Can There Be War Without Soldiers?*', Foreign Policy*, (15 March 2016) <http://foreignpolicy.com/2016/03/15/can-there-be-war-without-soldiers-weapons-cyberwarfare/> [accessed 12 December 2017].

[6] Richard Barrons, 'The nature of war is changing. It's time governments caught up', *Wired Magazine*, 14 October 2017 <http://www.wired.co.uk/article/innovation-will-win-the-coming-cybersecurity-war-richard-barrons-opinion> [accessed 12 June 2018].

[7] Chapter 4 (*Deployment*), specifically: 4.3 ('*Machine and Human-Teaming models'*).

[8] Paul Scharre, 'Robotics on the Battlefield Part I: Range, Persistence and Daring'*, Centre For A New American Security*, (May 2014) p. 9 <https://s3.amazonaws.com/legacy.cnas.org/publications-pdf/CNAS_RoboticsOnTheBattlefield_Scharre.pdf>.

deployment, this is less to do with the difficulties of LOAC compliance and more around human anxieties about combat optimisation, about battlefield efficiency, practical priorities and how *humans* decide war is best prosecuted.[9]

This conclusion is divided into two sections. The first rehearses the broad context of AWS deployment in order now to evaluate the challenges that emerge in the thesis' nine chapters. The second reviews the *nature* of these challenges in order to identify common threads before knitting these constraints together in order to establish a broad finding from the overall thesis.[10] Such structure (and this concluding chapter) therefore requires certain restatement of critical findings in order to weight evidence. It is also based on a set of assumptions that require reaffirmation. In this vein, a tipping point throughout is the engagement limits that must determine AWS behaviour and an 'ability to err when confronted by situations outwith their originally-intended design parameters'.[11] Taken to its conclusion, this requires Meaningful Human Control (MHC) to remain in the engagement sequence, thereby recognising what is that class of 'broadly' autonomous weaponry where individual componentry can of course act without human supervision but subject always to ultimate human fail-safe and, in targeting, human permission in the authorisation of target selection. Such definition acknowledges that when autonomous componentry fails it will tend either to fail catastrophically or will require human attention at points of highest stress. Indeed, Cordingley's embrace of MHC is based upon two premises. First, it is difficult, he notes, to foresee in practical terms why the Design Cohort would rule out a proven overrule mechanism in otherwise autonomous weapons. Second, he points to legal and operational imperatives which make delegation of significant tasks to AWS by the local commander thoroughly improbable without that same override mechanism being in place.[12] But this too has unexpected consequences. In such circumstances, that human operator may either be insufficiently engaged or inadequately trained to meet the legal, performance and trust thresholds identified above.[13] It is a generally accepted heuristic that operators skills which go unpracticed tend to wither.[14] While this condition is a fact of machine operation, statutory implementation of MHC (as discussed below) will de-risk this by ensuring consequential oversight.

The analysis also finds that human oversight must be retained in order to monitor emergent effects of autonomous componentry, to intervene when circumstances exceed machine capabilities (and thence for the unsupervised weapon to avoid inappropriate action) as well as, of course, to take over in situations where human capabilities empirically trump machine weaknesses.[15] MHC would also ensure that AWS conforms to this analysis' conclusion that *discriminatory* capacity is a legal precondition for initiating violence and that the judgement that this entails is intrinsic to (and

[9] Chapter 5 (*Obstacles*), specifically: 5.6 ('*Behavioural constraints'*).

[10] See below: Section 10.1 ('*The nature of Deployment Challenges'*).

[11] Ministry of Defence, 'Human-Machine Teaming', p. 32.

[12] Major-General Patrick Cordingley, Commander, 7th Armoured Brigade, Gulf War, 1991, in conversation with the author, January 2019.

[13] Chapter 5 (*Constraints*), specifically: 5.6 ('*Behavioural constraints'*).

[14] Ministry of Defence, '*Human-Machine Teaming'*, p. 32.

[15] Michael Hanlon, ''Super Solders': The Quest for the Ultimate Human Killing Machine', *The Independent,* (17 November 2011), paras. 1-3 and 5 of 15 <https://www.independent.co.uk/news/science/super-soldiers-the-quest-for-the-ultimate-human-killing-machine-6263279.html> [accessed 10 December 2017].

must remain with) human battlefield commanders.[16] Several arguments, both behavioural and technical, inform this deduction. At its most basic, appropriate human involvement will (in the near and medium-term of this thesis from the present until 2040) be required simply to manage the array of potential forms as well, of course, as the *degree* of autonomy employed by each such remote system, to supervise the wide permutations in which these weapons will perform in human-machine teams and, crucially, to oversee the *dynamic conditions* of each such use.[17] AWS, after all, may be deployed in many forms from large inter-continental bombers to undersea vehicles, from swarming micro-planes to small ground robots.[18] Similarly, use of AWS might involve long or short duration with static, attacking, defensive or loitering battle roles in a variety of sea, land, air or cyberspace environments, all executable with variable payloads that require coordination and optimisation. In addition, human oversight must manage operational variability arising through compound target selection processes, the speed of those processes as well as the *degree* of weapon autonomy given that not all autonomous tasks are equal in their significance, in their complexity or their risk.[19] It is thus barely useful to denote a weapon as 'autonomous' without referring (and understanding) the specific battlefield routine that is being made autonomous and the human soldier's relationship to that routine.[20]

Several characteristics form this thesis' conclusions on AWS' technical infeasibility and the ensuing requirement for human oversight in the delivery of *all* force. First is the unexpected endurance of context as a constraint to AWS deployment. This is variously evidenced.[21] An accelerating 'art of the possible', whereby societal *and* procurement expectations around battlefield technologies move further away from what is achievable, continues to influence the landscape in which unsupervised weapons are being considered.[22] This is labelled by Sabin as 'a revolution in expectation'.[23] The thesis' technical analysis suggests, however, that such belief is not convincing.[24] Instead, more important are those behavioural, organizational and structural considerations which must first align before human supervision can practically be removed from lethal engagements. A finding is that such alignment remains absent. To exploit properly developments in military AI and robotics (and in order then to embed new autonomous componentry into battlefield practice), State-parties must first adopt aggressive experimentation, concept development and organisational refinement. This is unlikely to occur given forecast procurement and training practices.[25] AWS

[16] Chapter 5 (*Constraints*), specifically: 5.1 ('*The Geneva Convention and Laws of Armed* Combat') and 5.5 ('*Article 36 and LOAC-complaint* weaponry'). See also: Lucy Suchman, *'Situational awareness'*, p. 8.

[17] Ministry of Defence, 'Human-Machine Teaming', p.vi. For deployment models, see: Chapter 4 (*Deployment*) specifically: 4.3 ('*Machine and Human Teaming Models'*).

[18] This is explored in detail in Chapter 4 (*Deployment*).

[19] For a detailed discussion on targeting's ramifications to AWS feasibility, see: Chapter 10 (*Oversight*), specifically: 10.1 ('*Nature of deployment challenges'*).

[20] Indeed, several applications of military autonomy are non-controversial including unmanned logistics and reconnaissance.

[21] See: Chapter 2 (*Context*).

[22] Peter Lee and Steve Wright, 'Killer Drones: Be Afraid or Ignore the Hype?', *Dleano.lu blog*, (4 December 2017) <http://delano.lu/d/detail/news/killer-drones-be-afraid-or-ignore-hype/163173>. [accessed 12 August 2018].

[23] Professor Sabin, Professor of Strategic Studies at KCL, in discussion with the author, 29 July 2017.

[24] See: Chapters 6-9 (*Wetware, Software, Firmware* and *Hardware*), specifically: 6.5 ('*Missing pieces'*) and 7.1: ('*Sources of technical debt'*).

[25] See: Chapter 4 (*Deployment*), specifically: 4.2 ('*Planning tools'*) and 4.7 ('*Operations and causes of failure'*).

deployment, moreover, may be particularly sensitive to such prerequisite; extensive trialling will be unusually significant in order to build operators' trust in what is unpredicatable autonomous componentry, to establish best teaming models, to anchor the technologies' attraction within its users and to provide datasets that are central to subsequent performance improvement. Finally to this point, a weapon which is networked should always be more valuable than one which is independent and, in the case of AWS, operating on its own.[26] By overtly connecting each weapon to his network, the commander can ensure that the munition 'becomes part of a broader system that can harness sensor data from other ships, aircrat and satellites to assist its targeting'.[27] Under that commander's direct control it is less likely to be wasted, redundant or mis-tasked. Such benefits are broadly entrenched: As emphasised by US Deputy Secretary of Defense Work, only by exercising direct control over military assets will humans 'inside the battle network have tactical and operational overmatch against their enemies'.[28]

Such analysis does not, however, imply that military affairs (and, specifically, military procurement) are immune from disruption. In this vein, pace of change has clear deployment ramifications.[29] The nature of this change, specifically the advent of creditable unmanned systems, means that fitness for purpose in States' current battlefield capabilities can no longer be assumed.[30] Low-cost, swarming UAVs provide an obvious example[31]: The UK's two aircraft carriers currently under construction may each cost some £3.5 billion but the risks posed to them by swarming low-cost drones lead commentators to suggest that they may only be usable within a comprehensive and therefore multi-national taskforce.[32] The example highlights a further anomaly given the seeming asymmetry between autonomous defenders (who must negate every vulnerability and challenge) and autonomous attackers (who just have to land one hit). This is particularly the case in the case of autonomous cyber attack.[33] Technical transformation, moreover, has other imprecise (but nevertheless still enduring) consequences for AWS deployment. Empirically, analysis suggests that it encourages reflexive policy-making, both militarily and governmental, as parties stuggle to keep up with the pace of change evidenced above.[34] In the case of AWS, leaving engagement decisions to a machine removes a time buffer that had previously existed as a brake on the impulsive making of choices. But this same speed of change also fosters institutional paralysis. To this point, autonomy, its battlefield uses and likely configuration remain difficult to define[35] and it is this same imprecision that confounds discussion on a statutory framework (perhaps around MHC)

[26] This is also relevant when considering groupings of allies. The US, notes Cordingley, would expect to be a key player in such alliances in which it participates.

[27] Paul Scharre, *Army of None,* Norton Publishing, 2018, p. 55.

[28] Ibid., p. 99.

[29] Sputnik News, 'Missile Defence Systems Could be Made Obsolete by Small Powerful Laser Weapons', *Sputnik*, (8 April 2017) <https://sputniknews.com/military/201704081052451679-pentagon-lasers-make-missiles-obsolete/>.

[30] This flux characteristic is evidenced in Chapter 3 (*Drivers*), specifically its introduction, and Chapter 4 (*Deployment*), also in that chapter's introduction.

[31] See: Chapter 4 (*Deployment*), specifically: 4.6 ('*Swarming models'*).

[32] Steve Hawkes, 'Britain's two new aircraft carriers may never be able to go to war alone as 'UK forces unable to support them'', *Sun newspaper*, 1 May 2018 <https://www.thesun.co.uk/news/6189612/aircraft-carriers-unable-defend-unassisted/> [accessed 12 June 2018].

[33] Autonomous cyber is considered in 3.5 ('*Operational drivers'*).

[34] See: Chapter 2 (*Context*), specifically: 2.2 ('*The role of context in AWS' argument'*).

[35] For a discussion on definitions and challenges arising see: Chapter 1 (*Introduction*).

which has been taking place at the UN's CCW since 2012.[36] Also to this point, the analysis confirms a general difficulty around the prediction of *how* AWS might be deployed, under what conditions and with what effect. The analysis is similarly clear that policy paralysis around the issue of AWS is sustained by existing legal frameworks, both IHL and IHRL, being empirically unfit for purpose in the face of such technical innovation[37]; it is unsurprisingly difficult to 'shoehorn independent weaponry into laws of armed conflict that were framed in the 1940s' and this fosters friction and inertia.[38] As noted by the ICRC's President in his June 2018 address to HRW, that legal framework is already under unprecedented strain with 'international law increasingly being viewed by parties as a matter of transaction to be undertaken transitorily by belligerent parties as a tool of barter'.[39] It is within such volatile context that AWS deployment must be considered.

While it is arguable that most difficult challenges identified in the thesis' technical analysis may, by and large, be solvable over time on an individual basis[40], it is rather the cumulative and portfolio nature of required innovation (which, after all, must *together* be available) that primarily constrains AWS deloyment. This thesis piggybacks on Beard's conclusion that AWS constituencies (here, the lawyer and activist, the politician and voter, the businessman and the soldier) require that unsupervised weapons be straightaway reliable.[41] In common with most precedent weapon systems, Monagham notes that AWS should work as expected from first deployment.[42] In particular, it is unacceptable that independent weapons be adopted on any trial or approximating basis.[43] It is the enduring challenge of predictability that best highlights what is a cross-over existing between AWS' hard and soft deployment constraints: Empirically, a conclusion might even be that technical challenges always distill down into contextual, soft constraints on AWS deployment. The decision to deploy unsupervised weaponry will, after all, be itself a fundamentally human judgment (and one, therefore, which is owned by that specific human agency) resolving how, when and to what extent

---

[36] UN Office at Geneva, 'The Convention for Certain Conventional Weapons', *UNOG* <https://www.unog.ch/80256EE600585943/(httpPages)/4F0DEF093B4860B4C1257180004B1B30?OpenDocument> [accessed 12 January 2018]. See also: Ariel Conn, 'The Problem of Defining Autonomous Weapons', *The Future of Life Institute*, (30 November 2016) <https://futureoflife.org/2016/11/30/problem-defining-autonomous-weapons/> [accessed 17 April 2018].

[37] Geneva Academy, 'Autonomous Weapon Systems under International Law', *Geneva Academy Briefing,* 8, (November 2014) <https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf> p. 27.

[38] ICRC, 'ICRC, IHL and the Challenges of Contemporary Armed Conflicts', *28th International Conference of the Red Cross and Red Crescent*, (December 2003) <https://casebook.icrc.org/case-study/icrc-ihl-and-challenges-contemporary-armed-conflicts> [accessed 13 July 2018].

[39] Peter Maurer, President of ICRC, Human Rights Watch Annual Forum, Swiss Re, Zurich, 8 June 2018 and subsequently in conversation with the author.

[40] See: ARM Holdings Publication, 'AI Today, AI Tomorrow: awareness, acceptance and anticipation of AI: A global consumer perspective', *ARM Northstar*, (2018) <https://pages.arm.com/rs/312-SAX-488/images/arm-ai-survey-report.pdf>.

[41] Jack Beard, 'Autonomous Weapons and Human Responsibility', *University of Nebraska-Lincoln*, College of Law Faculty Publications, 196, (2014) <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=http://scholar.google.co.uk/&httpsredir=1&article=1196&context=lawfacpub>, pp. 622-625.

[42] Timothy Monagham, 'Military Fault Tolerant Requirements', *Foundation of Dependable Computing (Office of Naval Research Advanced Book Series*', 283 (1994), Abstract.

[43] See: Chapter 7 (*Firmware*), specifically: 7.2 ('*Firmware ramifications of learning methodologies*'). Also: General Sir Richard Barrons, Commander Joint Forces Command (Retd.) in conversation with the author, 23 June 2016.

autonomy is employed in individual lethal engagements. This human factor is as true for polities seeking legal compliance (State signatories to international laws) as it is for the non-state actor whose motivations may have much less to do with established LOAC.

11.1 The nature of deployment challenges

While analysis of AWS ise complicated by it being a 'future-oriented argument'[44], the thesis nevertheless recognises a portfolio of drivers accelerating the deployment of faster, less expensive, more numerous independent weapons. This is a broadly-held phenomenon and the analysis evidences its wide appeal.[45] Such catalysts are facilitated by autonomy's dual uses with several agencies, predominantly commercial, funding research into technologies that are immediately relevant to AWS development.[46] The introduction of autonomous componentry into weapons and then an ensuing path for those weapons becoming properly autonomous also fits with established models on how disruptive adoption of technology takes place.[47] Etzioni and Etzioni note a broadly held view that embrace of weapon autonomy will tend to give combat advantage and, after achieving a tipping point in its adoption, it will rapidly transfigure what are currently manned battlefield practices.[48] Indeed, the academic concern might instead be the *absence* of parties arguing against the promise of AWS' improved performance (specifically, unsupervised weapons' potential of broad role extension and the Delivery Cohort's prospect of moving from a focus on mission outcome to mission performance), the promise of enhanced ethical function as well as significent cost reduction (while still achieving force multiplication). As noted, by Singer, 'the focus on military robotics is to use robots as a replacement for human losses'.[49] This thesis' analysis suggests, however, that not all of these drivers are evidence-based. While soldiers may not be the perfect fighting unit, published data on their ethical behaviour while fighting is, in particular, clearly problematic.[50] Nor, note Reeves and Johnson, are States' pursuit of autonomous technologies as irrefutable as often reported.[51] Scharre highlights what he sees as a 'muddled picture' in militaries' adoption of robotic weaponry, underscoring instead the US' 'intensive cultural resistance… to handing over combat jobs to uninhabited systems'.[52] Plotting cause and effect, moreover, is

[44] Richard Moyes, Article 36, 'War without oversight; challenges to the deployment of autonomous weapons', Buckingham University Humanities Research Institute Seminar, 13 May 2018.

[45] Sean Gallagher, 'The Air Force Wants Weapons Faster, Cheaper as it Sees Writing on the Wall', *ARS Technica*, 31 July 2014, generally <https://arstechnica.com/tech-policy/2014/07/air-force-wants-weapons-faster-cheaper-as-it-sees-writing-on-wall/> [accessed 2 August 2018]. See generally: Chapter 3 (*Drivers*).

[46] Boulanin and Verbruggen, pp. 20-21 and pp. 105-111.

[47] See: Chapter 4 (*Deployment*), specifically: 4.3 ('*Machine and human-teaming models'*). See also: Tony Seba, '*Clean Disruption'.*

[48] Amitai Etzioni and Oren Etzioni, 'Pros and Cons of Autonomous Weapon Systems', pp. 72-74. See also: Chapter 3 (Drivers), specifically: 3.2 ('*Technology Creep and Dual-Use Technology Trends'*) and 3.3 (*'Structural and procurement drivers'*).

[49] Singer, *Wired for War,* p. 418.

[50] Megan Thompson and Rakesh Jetly, 'Battlefield Ethics Training: Integrating Ethical Scenarios in High-Intensity Military Field Exercises', *European Journal of Psychotraumatology*, (August 2014) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4138704/>.

[51] Shane Reeves and William Johnson, 'Autonomous Weapons: Are you sure that these are killer robots – Can we talk about it', *The Army Lawyer*, 25, (2014), p. 25. See also: John Brock, 'Why the United States Must Adopt Lethal Autonomous Weapon Systems', *School of Advanced Military Studies, Leavensworth*, (2017), pp. 1-12 <http://www.dtic.mil/dtic/tr/fulltext/u2/1038884.pdf>.

[52] Scharre, *Army of None,* p. 61. Scharre also highlights a US disconnect between 'ambitious dreams for robots in a variety of roles' and the budgetary realities within those departments. 'Without funding, these visions are more

particularly complex in AWS deployment given that relevant precedents do not yet exist. Favourite measurement metrics, moreover, often have multiple interpretations. In the case of cost drivers, the US Army's outlay on personnel may appear statistically bloated in comparison to the manpower budgets of Russia or China[53] but this may ignore skills and other qualitative advantages arising from such allocation.[54]

Given such AWS drivers and general difficulties of prediction, a danger for this analysis remains being blindsided by the future.[55] Dystopian representations such as *Slaughterbots*[56] and *As Much Death As You Want*[57] suggest, after all, a potential nearness to broad AWS deployment.[58] This unexpectedly complicates the role of AWS' Delivery Cohort. States' procurement of autonomy may be influenced by unfounded developments in neighbouring nations' arsenals, by the pyschological impact promised by quite unfeasible autonomous technologies, by potential escalation that AWS deployment in a bordering State might create[59] as well as by a more nebulous 'fear-of-missing-out'.[60] The analysis therefore concludes that two field models are particularly relevant for AWS deployment. First, there is on-going (but often individually imperceptable) replacement by machines of specific tasks that were previously undertaken by humans leading to erosion of human supervision through *incremental* delegation of battlefield tasks and engagement routines to those machines. The second model is evidenced by the fit between adoption of weapons autonomy and Seba's disruptive S-curve model[61]; the methodology of his model (specifically, tipping points leading to non-linear adoption) looks an applicable archetype for AWS deployment notwithstanding the prerequisite that *all* AWS componentry must first be available if human supervision is feasibily to be removed.[62] It is in this vein that the thesis' technical analysis focuses upon the enduring nature (and significance) of 'technology holes' that remain including, inter alia,

hallucinations than reality. They articulate goals and aspirations but do not necessarily represent the most likely future path'.

[53] Sydney Freedberg, 'US Defense Budget Not That Much Bigger Than China, Russia', *Breaking Defense*, (22 May 2018) <https://breakingdefense.com/2018/05/us-defense-budget-not-that-much-bigger-than-china-russia-gen-milley/> [accessed 3 June 2018].

[54] Helena Careiras and Celso Castro (eds.), 'Qualitative Methods in Military Studies; Research Experiences and Challenges', Routledge, (2013), generally.

[55] Notwithstanding AWS' infancy, a Google search on the term 'autonomous weapon' returns around 9,650,000 references [accessed 17 July 2018].

[56] Source: <https://www.youtube.com/watch?reload=9&v=9CO6M2HsoIA> [accessed 12 January 2018].

[57] Lucien Crowder, 'As Much Death as you Want', *Bulletin of the Atomic Scientists*, (2 December 2017) <https://thebulletin.org/2017/12/as-much-death-as-you-want-uc-berkeleys-stuart-russell-on-slaughterbots/> [accessed 12 January 2018]. See introduction to Chapter 2 ('Context') for a detailed discussion of this theme.

[58] For discussion of possible timelines, see: introduction to Chapter 2 (*Context*).

[59] Economist, 'Autonomous Weapons are a Game-changer', Economist Magazine, (25 January 2018) <https://www.economist.com/special-report/2018/01/25/autonomous-weapons-are-a-game-changer> [accessed 23 July 2018].

[60] Andreas Kirsch, 'Autonomous Weapons will be Tireless, Efficient Killing Machines – and there is no way to stop them', *Quartz News*, 23 July 2018 <https://qz.com/1332214/autonomous-weapons-will-be-tireless-efficient-killing-machines-and-there-is-no-way-to-stop-them/> [accessed 2 August 2018].

[61] Introduction to Chapter 3 (*Drivers*), specifically the discussion of Seba's adoption model.

[62] Chapter 3 (*Drivers*, specifically that chapter's introduction) and Chapter 5 (*Obstacles*, specifically: 5.7 '*Proliferation constraints*').

capabilities around goal setting, value setting, ATR and achoring.[63]

It is also its focus on deployment's outwardly *soft* constraints that underpins this thesis' conclusions, in particular its emphasis on existing legal frameworks (and how unsupervised weapons might fit within these structures) and whether the Delivery Cohort can empirically 'risk such deployment'.[64] This analysis identifies several difficulties.[65] The legal corpus largely dates from before unsupervised weapons had been practically conceived and, if not repeatedly unfit for purpose, the law at least requires unhelpfully contentious interpretation around AWS deployment. The analysis also highlights that international statutory bodies dealing with this framework are generally ineffective and without appropriate authority either to decide or enact decisions. This creates uncertainty and impunity. Legal review of AWS deployment has two elements. The analysis assumes that relevant law is addressed to humans not machines. Weapons, after all, are incapable of agency in respect of the laws of armed combat. The issue therefore becomes the dilution of human agency as attack characteristics get ever wider.[66] Although a truism, battlefield complexity (in the case of AWS, unprecedented elements of speed, innovation and multiple agency) means that conditions of use can (for purposes of coding) never be appropriately absolute.[67] Just as this analysis evidences the difficulty of writing machine code that can deal with ambiguity[68], it is also problematic to link weapon coding to what is a clearly imperfect legal framework. Accountability (especially in instances of weapon error and unattributable use, herein 'plausible deniability') is particularly problematic as is the *degree* of compliance with which such weapons can adhere to LOAC.[69] To this point, the analysis identifies several continua that individually impact AWS deployment (defence versus offence, manpower versus firepower, history-repeats versus history-is-change, politics versus technology et al). It is possible to construct one further continuum with, at one end, State-signatories to international obligations and, at the other, non-State actors with their quite different drivers to adoption. The deployment equation is thus a question of balance between ethical and moral obstacles versus legitimate uses of autonomous technologies and, the conclusion of this thesis, effective human intervention to ensure proper compliance. Within this equation, the matter of what happens when machines out-perform people in elements of the engagement sequence (which is already evident) is no longer necessarily the correct question. The issue is less whether that weapon component performs better than a human operator, but rather the level of risk arising for the Delivery Cohort in situations when (not if) that autonomous weapon fails.[70]

---

[63] These are key sections in the analysis' technical consideration of AWS feasibility. In particular, see: 8.3 ('*Utility function'*), 8.5 ('*Anchoring and goal setting issues'*), 8.6 ('*Value setting issues'*) and 8.7 ('*Action selection issues'*). '

[64] Major-General Patrick Cordingley, in conversatrion with the author, January 2019. Also: Paul Scharre, 'Autonomous weapons and operational risk', *Centre for a New American Security*, (2016), pp. 8-18 <https://www.files.ethz.ch/isn/196288/CNAS_Autonomous-weapons-operational-risk.pdf>.

[65] Chapter 5 (*Constraints*), specifically: 5.1 ('*Geneva Convention and the Laws of Armed Combat'*) and 5.5 ('*Article 36 and LOAC-compliant weaponry'*).

[66] Chapter 5 (*Constraints*), specifically: 5.6 ('*Behavioural Constraints'*) and 5.8 ('*Ethical and accountability constraints'*).

[67] Chapter 9 (*Hardware*), specifically: 9.1 ('*Hardware and Sensor Fusion Issues for AWS'*).

[68] Chapter 8 (*Software*), specifically: 8.1 ('*Coding methodologies'*).

[69] See: Chapter 5 (*Constraints*), specifically 5.8 ('*Ethical and accountability constraints'*).

[70] For discussion on the role of the Delivery Cohort, see: Chapter 6 (Wetware), specifically: 6.3 ('*The AWS' Delivery Cohort'*).

In assessing the nature of these challenges, the thesis concludes that outsize weighting should be allocated to the role of context in considering AWS deployment. Is it better to 'field good soldiers and excellent kit' or instead to field excellent soldiers with merely good equipment?[71] Cordingley, for instance, is quite clear that good soldiers are the enduringly key asset.[72] This point of balance is nevertheless the subject of Chapters Four (*Deployment*), Five (*Obstacles*) and Ten (*Oversight*) in order to mediate between the relative value of technology and that of human supervision. The analysis agrees with Cordingley and finds that the excellent soldier remains the Cohort's best battlefield asset, uniquely placed to counter unexpected adversarial activity on the battlefield, to capitalise upon weaknesses and to neutralise newly unsupervised platforms (Boot's concept of nullification), possibly undertaking this task through 'quite low-tech responses'.[73] A contribution of this analysis is then to identify and frame AWS' deployment challenges within this context. In this case, AWS' invariable ML spine means that adversarial feint and 'non-cooperative targets'[74] will likely obfuscate machine attribution as well as degrade the ATR effectiveness, a prerequisite to AWS deployment.[75] Also important to the Cohort's deployment decision is the finding that battlefield innovation rarely confers lasting advantage and, in instances of deploying 'low tech that works well enough', adds further uncertainty to the decision to deploy independent weapons.[76] In deciding battlefield priorities, it is human endevour, human appraisal and, notes Clark, soldiers' broad knowledge, skills and experience (KSEs) that must empirically shape AWS deployment.[77] On the one hand, politicians may be drawn to autonomous solutions that offer bloodless and remote engagement (minimizing friendly casualties and and their media costs).[78] Conversely, that same politican will likely be knocked back by a 'Tesla moment'[79], by structural impediments (the unanimity model, for example, that underpins NATO decision-making and, in the case of the UN's CCW, tends to dilute decisions to an inappropriately low common denominator in order to achieve consensus) as well as, crucially, by popular pressures (the restrictions on bombing raids that were applied by politicians in Kosovo after public opprobrium).[80] It is also far from certain that strict weapon autonomy will empirically remove per se human participation from combat sequences; current pre-cursors (such as Predator and Reaper drone operations) already require some ten operators to staff each drone, a further twenty operators needed to manage the unit's sensors and

[71] General Sir Richard Barrons, Commander Joint Forces Command (Retd.) in conversation with the author, 23 June 2016.

[72] Major-General Patrick Cordingley, in conversation with the author, September 2018.

[73] Chapter 2 (*Context*), specifically: 2.6 ('*The role of situational awareness and uncertainty*).

[74] Here, those targets that do not broadcast their location and require, therefore, active sensing from engaging munitions in order to find those targets. This is complex. In particular, building algorithms that can automatically decipher SAR output (synthetic aperture radar analysis) will likely be inappropriately error-prone.

[75] Chapter 9 (*Hardware*), generally. Also, Appendix One; '*Case study: Automatic Target Recognition'*.

[76] Chapter 2 (*Context*), specifically: 2.4 ('*Defence Planning'*).

[77] Professor Lloyd Clark, in conversation with the author, June 2018. Clark here notes that the development of both individuals and teams will require reassessment under AWS deployment. Similarly, those soldiers will require new competencies, attributes and very different models of leadership. See: Chapter 10 (*Oversight*), generally.

[78] Chapter 3 (*Drivers*), specifically: 3.5 ('*Operational Drivers'*).

[79] Danny Yadron and others, 'Tesla Driver Dies in First Fatal Crash While Using Autopilot Mode', *Guardian news*paper, 1 July 2016 <https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk> [accessed 12 January 2018]. In this instance, research into autonomous cars was postponed after public clamour following a fatal accident involving that company's technology.

[80] Benjamin Lambeth, *NATO's Air War for Kosovo: A Strategic and Operational Assessment,* (USA: Santa Monica, CA, 2001), p. 185.

scores of intelligence analysts to sift through resulting sensor data.[81] As noted by Scharre, 'it's a cumbersome way to operate [and] not a cost-effective strategy if they require ever larger numbers of highly trained (and expensive) people to operate them'.[82]

Analysis of AWS deployment must distinguish between the nature of war and the character of war. This is similarly not clear-cut. Mewett's definition of war ('[w]ar's nature is violent, interactive and fundamentally political. Absent any of these elements, what you're looking at is not war but something else') now appears rooted in traditional views on old fashioned inter-State conflict.[83] Just as a key dynamic is human endevour, this requires that decision-making around AWS deployment must be correspondingly 'human', bringing together the broadest selection of parties that includes politicians, commanders, those with commercial interests as well as those advocating for the third sector and faith organisations. Mewett similarly distinguishes between war and warfare. If warfare is merely the way that war is made, then AWS deployment can properly be treated as simply another means of warfare.[84] A conclusion from this behavioural review is that changes in the character of war, here defined by AWS deployment, will actually be incremental and shaped primarily by the manner in which the *deploying* party is organised. Such synthesis accords with Ricks' *Future of War* and his argument that emerging technologies continue to change war's character by a gradual blurring of lines. In the case of AWS, this may be the removal of weapon supervision compromising existing and previously stable legal structures, upending pockets of influence between public and private, between military and the intelligence community and, as occasioned by AWS' 'plausible deniability', even by an 'eroding [of] traditional conceptions of sovereignty'.[85] Such complexity is then reflected in the difficulties experienced by defence planners whose commission is to ensure minimum regret in times of rapid technical change and the rapid erosion of these previously well-understood relationships.[86]

The analysis places considerable weight on the complex role of the AWS' Delivery Cohort, an artifice used throughout the thesis to describe the layers of interested parties involved in the decisions and implementation of AWS.[87] This Cohort (a vortex, after all, of human consituents in

---

[81] See: Chapter 5 (*Obstacles*), specifically: 5.6 ('*Behavioural constraints'*).

[82] Scharre, *Army of None,* p. 16.

[83] Christopher Mewett, 'Understanding War's Enduring Nature Alongside its Changing Character'*, War on the rocks*, Texas National Security Network, (21 January 2014) https://warontherocks.com/2014/01/understanding-wars-enduring-nature-alongside-its-changing-character/, generally. Mewett links this point to: Clausewitz 'Spirit of the Age; Understanding War's Enduring Nature Alongside its Changing Character', para. 3 of 9. See also: Chapter 1 (*Introduction*), specifically: 1.2 ('*Introduction to key concepts'*).

[84] For a review of RMA, see introductions: Chapter 1 (*Introduction*) and Chapter 2 (*Context*).

[85] Thomas Ricks, '*T*he Future of War (II): As the Nature of War Changes, the Familiar Dividing Lines of our World are Blurring across the Board'*, Best Defense, Foreign Policy Magazine*, 15 January 2015 <http://foreignpolicy.com/2014/01/13/the-future-of-war-i-a-new-america-project-looking-at-21st-century-conflict/>. Suarez' doctrine of 'plausible deniability' is a case here in point.

[86] Introduction to Chapter 2 (*Context*).

[87] As above, the term Delivery Cohort is used as a device to convey the parties involved in delivering the deployment of AWS and will include, inter alia, the following taskings: neurophysiologists to coordinate AWS networks, psychologists to coordinate learning and cognition, biologists for adaption strategies, engineers for control routines, logisticians, roboticists, electrical specialists, behaviorists, politicians, NGOs, sociologists, lawyers, company directors, weaponists, military tacticians, manufacturers, professionals involved in miniaturization, simulation, configuration, coding, power supply and modularity, specialists in sensors, in distributed and decentralized routines, ethicists, specialists in tooling and calibration.

AWS processes) faces several challenges, not least on accountibility, on audit, testing and validation, on delivering on expectations as well as ensuring ongoing process-improvement throughout AWS' adoption. Cohort challenges are both fundamental (for instance, which normative theory should underpin AWS deployment?) and operational (how might seventeen hundred pages of NATO's recent Rules of Engagement be captured in AWS' framework?). These soft considerations are difficult to weigh. The nature of AWS challenges requires, after all, that the Cohort factor in proliferation concerns as well as methods to counter escalation. It must account for economic considerations, especially given allocation issues that arise from States' assignment of scarce combat resources.[88] The hardware comprising AWS will clearly be expensive.[89] The likely bespoke nature of individual AWS will require small production runs of unique parts, few of which can be bought off the shelf.[90] Similarly, Nesnas points out that costs cannot therefore be recovered over long manufacturing runs. AWS will also suffer from supply chain constraints:[91] States cannot save money by acquiring AWS components globally and having their weapons assembled in lower-cost but possibly adversarial neighbours. Other matters confronting the Cohort simply resist definition. In reviewing of deployment obstacles, these factors include navigating a likely lower threshold to parties' initiation of violence once AWS are fielded as well as the potential phenomenon of 'ubiquitous engagement'.[92] As argued by Suarez, implementing AWS risks accelerating 'destabilizingly [and] unattributable violence' given the ease with which deploying parties can fall back on 'plausible denial' when tasking independent and remote weaponry.[93]

The nature of AWS challenges is clearly shaped by the role of history and by lessons that arise from other deployment precedents. The appeal of disruptive weaponry, after all, is to break out of the development cycle that is typical of mature technologies whereby even material investment in legacy systems leads only to incremental improvement. Indeed, the analysis does not disagree that modest investment in autonomous technologies will deliver disproportionate advantage in military capability. In so doing, it concurs with the UK MOD's 2018 Joint Concept note which unequivocally states the 'utility of AI and robotics already outstrips that of the many mature technologies which are often in many orders of magnitude more expensive to incrementally improve'.[94] Again, however, this relationship is ambiguous. The statement ignores palpable legal, ethical, operational and human constraints to 'whole-weapon' adoption of AWS. Setting aside specific context, instances abound of illogical (and unsustainble) use of battlefield technology. In March 2017, US Army General Perkins revealed that the US had used a thyree million dollar Patriot missile against a quadcopter that cost two hundred dollar from Amazon.[95] Shortly after, it emerged that Houthi

[88] Ibid.

[89] RR Hoffman and others, 'The Myths and Costs of Autonomous weapon Systems', *Bulletin of the Atomic Scientists*, 72, 4, (2016), Abstract and generally.

[90] Issa Nesnas, 'CLARAty: Challenges and Steps Towards Reusable Robotic Software', *International Journal of Advanced Robotic Systems*, 3, 1, (2012), 24-27 ('*Challenges*') and generally <http://journals.sagepub.com/doi/pdf/10.5772/5766> [accessed 18 January 2018].

[91] Chapter 4 (*Deployment*).

[92] Chapter 5 (*Constraints*), specifically: 5.6 ('*Behavioural Constraints*').

[93] Ibid.

[94] Ministry of Defence, 'Human-Machine Teaming', *UK MOD*, Joint Concept Note 1/18, p. 15.

[95] Samuel Osbourne, 'Small Drone 'Worth $300' Shot Down by Patriot Missile Worth $3m, Says US General', *Independent Newspaper*, 13 March 2017 <https://www.independent.co.uk/news/world/americas/small-drone-quadcopter-patriot-missile-shot-down-us-general-david-perkins-army-a7631466.html> [accessed 12 December 2017].

rebels in the Yemen had used low-cost drones to disable state-of-the-art Saudi Patriot missile systems.[96] Such anlysis reiterates that technical advance (here, the facility for remote and independent engagement) is rarely the enduring preserve of wealthy nations as evidenced by the use of sophisticated UAV by *all* sides, both State and non-state, in recent conflicts in Syria, Iraq and Ukraine.[97] Regardless of its precise definition and quite separate from the adoption of autonomy, disruption is clearly evident across *all* aspects of battlefield practice[98] with lessons from recent conflicts involving UAS suggesting that certain assumptions around battlecraft and battlefield assets are already hollow.[99] An example might be air supremacy and developments across aerial hardware: 'Even where enemy aircraft have been neutralised, being observed and targeted by remote and automated systems must be continually treated as a risk'.[100]

A conclusion is also that continued human involvement in lethal force is necessary from a *moral* perspective. After all, no written agreement can prevent parties from deploying AWS if they desire it. Indeed, Strohn notes the degree to which this is essentially 'a debate of choice' for 'soft, Western States' that is likely to fall away, for instance, in any war for genuine survival.[101] The argument, however, is also that machines can display neither empathy nor remorse. In a truism that is argued by the Holy See, only humans can ever feel 'the emotional weight and psychological burden of choosing to kill another human being'.[102] The analysis therefore agrees with Heyns' conclusion that weapons autonomy 'precludes a moment of deliberation in those cases where it may be feasible'.[103] Empathy, concludes Human Rights Watch, can act as a check on killing but only if humans have control over who to target and when to fire.[104] Although contentious, additional context on the matter's moral angle is provided by Krishnan whereby 'taking away the inhibition to kill by using robots for the job could weaken the most powerful psychological (sic) and ethical restraint in war; war would be inhumanely efficient and would no longer be constrained by the natural urge of soldiers not to kill'.[105] This thesis' conclusion that MHC remain a prerequisite in lethal engagement therefore accords with Wylie in that battlefield control fundamentally concerns people[106]; only boots on the ground (or, practically, their equivalent) empirically provide a

[96] Christopher Diamond 'Report: Houthi rebels flying Iranian-made kamikaze drones into surveillance radars', *DefenseNews*, (27 March 2017), https://www.defensenews.com/global/mideast-africa/2017/03/28/report-houthi-rebels-flying-iranian-made-kamikaze-drones-into-surveillance-radars/> [accessed 12 December 2017].

[97] Ministry of Defence, 'Human-Machine Teaming', UK MOD, Joint Concept Note 1/18, p. 23.

[98] Thomas Macauley and others, 'The Future of Technology in Warfare: From AI Robots to VR Torture', *Techworld*, (13 January 2017) <https://www.techworld.com/security/future-of-technology-in-warfare-3652885/> [accessed 15 December 2017].

[99] David Hambling, 'Introduction: Weapons Technology', *New Scientist*, (4 September 2006), generally <https://www.newscientist.com/article/dn9980-introduction-weapons-technology/> [accessed 12 December 2017].

[100] Ministry of Defence, 'Human-Machine Teaming', p. 30.

[101] Dr Matthias Strohn, in conversation with the author, January 2019.

[102] Ministry of Defence, 'Human-Machine Teaming', p. 6.

[103] Christopher Heyns, 'Report of the UN Special Rapporteur on extrajudicial, summary and arbitrary executions to the Human Rights Council', A/HRC/23/47, (9 April 2013), para. 94. Heyns also reasons that machines lack morality and mortality and should as a result not have life and death powers over humans.

[104] See: Chapter 5 (*Constraints*), specifically: 5.8 ('*Ethical and Accountability Constraints*').

[105] Krishnan, '*Killer Robots*', p. 130.

[106] Colin S Gray, *Modern Strategy,* (Oxford: Oxford, 1999), p. 26 <www.dtic.mil/doctrine/jfq-22.pdf>.

consistent and sustainable authority to 'exert control over… a populace or other critical resource'.[107]

This moral imperative provides a potent component to the argument for MHC. Borenstein, Director for *The Centre of Ethics and Technology* at Georgia Tech University, posits that the mere prospect of fighting wars without military fatalities removes an important deterrent to waging war.[108] This dynamic is incorporated in the current call for debate by the UK's MoD to 'ensure that we do not risk losing our controlling humanity and make war more likely'.[109] This thesis' analysis arrives at a similar conclusion but from different angles. The long-held adage that a State may be more inclined to wage war if it calculates that the threat to its own troops has been reduced may be materially accelerated by AWS deployment.[110] In that case, 'States with roboticised forces might behave more aggressively [whereby] robotic weapons alter the political calculation for war'.[111] In this vein, removing weapon supervision would tend to posit particularly poor outcomes on civilians.[112] The commander who is able to deploy AWS (and who also believes his own forces are less prone to attrition) is likely, argues Sharkey, to place insufficient weight on threats to civilian life.[113] Absent MHC, an upshot is that the burden of armed conflict is shifted from soldiers to civilians given military personnel, replaced by machines, are no longer physically on the ground making decisions and controlling lethality.[114] In engagements, after all, casualties remain inevitable whoever is located in a battlespace, the more so if AWS deployment will lead, as noted by HRW, to 'disproportionate civilian suffering'.[115] Can, asks Clark, war actually be 'won' without man-on-man engagement and what may be the psychological impact of AWS deployment?[116] This highlights a separate challenge concerning the weighting of AWS' ethics and morality, specifically the difficulty of untangling criticisms that are aimed at weapons autonomy versus those which are really directed at the basic act of war. As noted by Scharre, 'what does it mean to say that someone has the right to life in war when killing is the essence [sic] of war?'[117] Empirically, it is humans who kill in war, whether using unsupervised weapons, remote weapons from a distance or up close and personally.

The nature of AWS' deployment challenge thus becomes increasingly behavioural, based upon the constancy and trustworthiness of individual weapon componentry as well as the *combination* of

[107] JC Wylie, *Military Strategy: A General Theory of Power Control,* (USA: Annapolis, MD, 1967), p. 89.

[108] Jason Borenstein, 'The Ethics of Autonomous Military Robots'*, Studies in Ethics, Law and Technology*, 2, 1, (2008), p. 8. As noted by Strohn, this is also key in light of recent conflicts as evidenced by the rise of private security agencies (Dr Matthias Strohn, in conversation with the author, January 2019).

[109] UK Ministry of Defence, 'The UK Approach to Unmanned Aircraft Systems'*,* pp. 5-9.

[110] RJ Rummel, *Understanding Conflict and War: War, Power and Peace*, 4, 16, 'Causes and Conditions of International Conflict and War' (USA: Beverley Hills, Sage, 1979), generally.

[111] Singer, *Robots at war: the new battlefield,* p. 48.

[112] Bonnie Docherty, 'We're Running Out of Time to Stop Killer Robot Weapons'*, Guardian newspaper*, 11 April 2018 <https://www.theguardian.com/commentisfree/2018/apr/11/killer-robot-weapons-autonomous-ai-warfare-un> [accessed 23 June 2018].

[113] Sharkey, 'Saying No to Lethal Autonomous Targeting'*,* generally.

[114] Human Rights Watch, 'Losing humanity', p. 39.

[115] Ibid., p. 41.

[116] Professor Lloyd Clark, thesis supervisor, in conversation with the author, October 2018.

[117] Scharre, *Army of None,* p. 294.

such autonomous componentry. For this reason, much of this thesis' analysis focuses on why removing human supervision will materially degrade weapon predictability.[118] As noted by Wallach, '[i]n the evaluation of weaponry, predictability means that within the task limits for which the system is designed, the anticipated behaviour will be realised, yielding the intended result'.[119] An upshot for AWS deployment is captured in Wendell's conclusion whereby 'an unanticipated event, force or resistance can alter the behaviour of even highly predictable systems'.[120] In systems theory, complex adaptive systems (here, unsupervised weapons) 'have tipping points that lead to fundamental reorganisation' which will materially complicate the processes of the battlefield commander.[121] An inescapable characteristic confronting the Cohort (that AWS' emergent properties will be intrinsically difficult to predict and difficult to explain) is underpinned by Wirth's rule that machine complexity increases several measures quicker through software development than it does through hardware development.[122] The thesis' technical analysis is important precisely because it evidences the link between AWS' tight coupling and system brittleness and the likely impact of these features on operational predictability. The issue here is clarified by Du. ML techniques either induce 'explainable' classifiers or 'blackbox-type' style classifiers when grading sensed data. The former provides a decision outcome but also the reason behind that decision (examples here being decision trees, nearest neighbour and rule-based classifiers). The latter also provides a decision outcome but without any attendant reasoning (here, the neural network).[123] It therefore fails on any 'duty to explain' that is a component of properly 'intelligence based systems' and a further driver for MHC. Indeed, without MHC, this will also be visible in unintended interaction between weapon components given the general absence of system slack (here, human inability to intervene, to exercise judgement, to 'bend' rules or amend system behaviours).[124] Given that AWS must operate where 'chaos [already] makes war a complex adaptive system rather than a closed or equilibrium-based system' such challenges must grossly complicate the Cohort's deployment equation.[125]

As noted by Cummings, an automated system is one in which a computer 'reasons by a clear if–then–else and therefore rules-based structure, and does so deterministically, meaning that for each input the system output will always be the same (except if something fails)'.[126] The model here is that if X happens then the weapon will do Y. AWS autonomy, however, is certain to be non-

[118] Specifically, the thesis' technical analysis in Chapters 7 (*Firmware*), 8 (*Software*) and 9 (*Hardware*), but also its consideration of command and control ramifications in Chapter 10 (*Oversight*). See: 7.1 ('*Sources of technical debt*') and 8.8 ('*Behaviour setting and coordination*').

[119] Wendell Wallach, 'Predictability and Lethal Autonomous Weapons'*, Institute for Ethics and Emerging Technologies*, (16 April 2016) <https://ieet.org/index.php/IEET2/more/Wallach20160416>.

[120] Ibid., generally.

[121] See also Chapter 6 (*Wetware*), specifically: 6.4 ('*AWS learning architecture*').

[122] Thom Holwerda, 'What Intel Giveth, Microsoft Taketh Away'*, OS News*, 15 November 2007 <http://exo-blog.blogspot.com/2007/09/what-intel-giveth-microsoft-taketh-away.html> [accessed 13 June 2018].

[123] Dr Hongbo Du, in conversation with the author, January 2019.

[124] Paul Scharre, 'Autonomous weapons and operational risk', *Centre for a New American Security*, (2016), pp. 25-34 ('The Inevitability of Failure: Complex Systems and Normal Accidents') <https://www.files.ethz.ch/isn/196288/CNAS_Autonomous-weapons-operational-risk.pdf>.

[125] James Mattis, *USJFCOM Commander's Guidance for Effects-Based Operations,* (USA: Parameters, Autumn 2008), p. 21.

[126] Cummings*,* p. 3.

deterministic whereby very small changes to inputs can produce very large changes to outputs.[127] Instead, the AWS will reason 'probabilistically given a set of inputs, meaning that it makes guesses about best possible courses of action given sensor data input'.[128] This distinction is important to understanding AWS' inherent instability. Unlike an automated system, the *autonomous* system cannot produce exactly similar behaviour given identical inputs. Instead, it is inescapable that AWS deployment must produce a range of engagement behaviours and it is this variability that must empirically limit AWS deployment to specific, bounded conditions.[129] Machine learning's ability to detect patterns still remains dependent upon humans with human interpretation still being generally required if such patterns are then to be valuable. Restating these relationships is contextually useful precisely as they amplify where (in AWS deployment) human operators should be making informed, conscious decisions. This, however, must be informed by what remain practical and empirical constraints around AWS deployment. For this reason, the analysis concludes that ambiguity arising from incomplete, noisy battlefield data will remain a largely remediless phenomenon that undermines reliable deployment of AWS.[130] Data uncertainty, after all, will arise from stale representations embedded in recently deployed systems exacerbated by compromised communications and by the need (and its effects) of filtering and weighting sensed data prior to the weapon initiating violence. Indeed, Scharre points out that most militaries already possess the ability to disrupt communications and contest the battlefield's electromagnetic spectrum.[131] To this point, spectrum interruption will lead to unequivocal data being the battlefield exception.[132]

The key architectural constraint to AWS deployment, identified by Sharkey and others, is that machine learning (ML), the basic technical backbone that will underpin AWS operation, is 'fundamentally unfit for the purpose envisaged for AWS deployment'.[133] It is for this reason that this thesis' technical review focuses squarely upon assessing ML's contribution to this faultline.[134] This includes incompatability within training sets, issues of data dependency and, as detailed in the analysis, the catch-all (within training sets) that 'changing-anything-changes-everything'.[135] The analysis identifies that weapon mutability will arise from the choosing and modification of description parameters that are used to train weapon behaviours ('parameter profusion'). It will be

[127] Ministry of Defence, 'Human-Machine Teaming', p. 12.

[128] Cummings, p. 3.

[129] Chapter 4 (*Deployment*).

[130] Paul Scharre, 'Autonomous weapons and operational risk', pp. 25-34 ('The Inevitablilty of Failure: Complex Systems and Normal Accidents'). See also: Chapter 8 (*Software*), specifically: 8.1 ('*Coding methodologies'*), 8.4 (*Software processing functions'*) and 8.7 ('*Action selection issues'*).

[131] Scharre, *Army of None,* p. 15. 'According to CODE's technical specifications, developers should count on no more than 50 kilobits per second of communications back to the human commander, essentially the same as a 56k dial-up modem circa 1997'.

[132] Sharlene Andrijich and others, 'The ambiguity problem arising in multi-sensor data association', *The Australasian Journal of Combinatorics*, Maritime Operations Division, 27, 2003 <https://ajc.maths.uq.edu.au/pdf/27/ajc_v27_p107.pdf>, pp. 107-127.

[133] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[134] Chapter 6 (*Wetware*), specifically: 6.4 ('*AWS learning architecture'*). Chapter 7 (*Firmware*), specifically: 7.2 ('*Firmware ramifications of learning methodologies'*) and 7.3 ('*Reasoning and cognition methodologies'*).

[135] For the discussion on CACE (*Change Anything Changes Everything*) see: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of Technical Debt'*).

exacerbated by AWS' real-time requirement for data scaling, smoothing and cleaning, by necessary suppression routines as well as by the need to dynamically balance subsequent feedback loops (AWS' 'anchoring problem').[136] People (here, the Delivery Cohort) are, moreover, generally poor predictors of behaviour in systems that rely on feedback loops, especially around the risk in situations where they have no experience (here, prediction bias).[137] AWS' ML backbone also posits other unexpected vulnerabilities for AWS deployment including ML's systemic, inappropriate suppression of doubt, its broad inference of causes and, therefore, an incongrous narrowing of battlefield choices. Crucially for this thesis' conclusions, it is ML processes that lead to machine code's inability to harness context or situational awareness in AWS operation.[138] As reasoned above, a consequence of ML's inherently approximating processes will be AWS' inappropriate technical basis of 'What You See Is All There Is' (WYSIATI).[139] Other ML attributes contribute to this faultline including the enduring complexity of *unlearning* routines, temporal and other data dependencies as well as ML's inability to capture 'qualia' in weapon routines.[140] ML also requires quality data and is demonstrably inappropriate where data capture is brittle, in particular when confronting either partial patterns or events not previously encountered. This is also noted by Cummings whereby ML-based systems act very differently in scenarios that are themselves only slightly different from each other.[141] It is these consequences of ML's foundation that should lead AWS' Delivery Cohort to question the practical feasibility of AWS deployment, in particular the issue of setting and managing weapon goals, managing weapon values and having the unsupervised weapon adhere to a dynamically relevant utility function.[142] An adjunct difficulty is identified by the UK's MOD Joint Concept note on *Human-Machine Teaming* that highlights 'catastrophic forgetting where previous algorithm optimisations or skills at tasks are simply lost when trained on new tasks and data'.[143] Consequences to the Delivery Cohort might include the AWS exhibiting unexpected immobilization, irrational action, unsuitable aggression, even unexplainable timidity. The thesis' overarching deduction is therefore informed by Cummings whereby 'as uncertainty grows, these tools become less useful'.[144] The prevalence of 'reliability predictions' in procurement programmes evidences the importance of dependability in combat assets.[145] Finally to this point, soft challenges around ML should not be overlooked including the availability of qualified, vetted personnel with appropriate experience to understand AI and AWS processes. Procurement challenges in this case include the maintenance of what are unstable programmes that will likely display a wide dispersion

136 Chapter 8.5 (*Software*), specifically: 8.5 ('*Anchoring and goal-setting issues'*).

137 Scharre, *Army of None,* p. 316.

138 The issue of verifying ML behaviours is well illustrated by the vulnerability of current visual object recognition AIs to 'adversarial images'. As noted by Scharre, 'the semi-technical explanation [here] is that while deep neural networks are highly non-linear at the micro level, they actually use linear methods to determine data at the micro level'. Moreover, 'no matter how many fooling images the AI learns to ignore, more can be created' (Scharre, '*Army of None',* pp. 180-185).

139 Chapter 4 (*Deployment*), specifically: 4.5 ('*Flexible Autonomy'*).

140 Chapter 7 (*Firmware*), specifically: 7.7 ('*Firmware ramifications of learning methodologies'*).

141 Professor Missy Cummings, Director, Humans & Autonomy Laboratory, Duke University, in conversation with the author (Chatham House conference; Autonomous Military Technologies, February 2014).

142 See: Chapter 8 (*Software*), specifically: 8.3 ('*Utility function'*), 8.5 ('*Anchoring and goal setting* issues') and 8.6 ('*Value setting issues'*).

143 Ministry of Defence, 'Human-Machine Teaming', p. 42.

144 Cummings, p. 8.

145 Although dated, the US Navy's *Handbook of Reliability Prediction* (see: Naval Surface Warfare Centre, (May 1992) <http://www.dtic.mil/dtic/tr/fulltext/u2/a273174.pdf>) remains a relevant reference on the issue.

of failure rates and causes notwithstanding apparently similar components, where the complexity of AWS' cause-and-effect relationships complicates diagnoses and, given such weapons' joint mechanical and electrical componentry, where issues of loading, operating mode, utilisation rates and glue code will create servicing bottlenecks. AWS dependability, moreover, will rely on reliable management of patches, of validation, version control and testing. It will also require entirely new protocols to govern what will also be AWS' unsupervised supply chains, their calibration and configuration. In this vein, independent weapons must be capable of seamless updating, seamless replenishment, logistics and maintenance, all of which will in turn require material (and unlikely) revision of military organisations.

AWS operation (and, therefore, its behaviour) must be based on pre-defined, pre-programmed battlefield physiognomies, each of which is represented by individual, idiosyncratic ML parameters. Together these factors will comprise the weapon's training data points in order for the unsupervised machine to make its own decisions. The deployment challenge is that the majority of such factors (reference examples are listed in this sentence's accompanying footnote) are not definable, are often an imprecise sum of other factors and anyway require subjective, volatile weighting if the weapon's sunsequent decision is to be appropriate.[146] With such strict bases (a direct consequence of ML's rigid parameters), AWS behaviour becomes acutely vulnerable to adversarial actions designed to disturb the weapon's sensed parameters ahead of its decision-making. Such activities might actually be surprisingly straightforward and borrow from Boot's theory of nullification touched on above (a can of paint, some rudimentary disguise or other cursory concealment). They will include more sophisiticated spoofing, feint and, to borrow from Sun Tzu, enemy activity that is 'subtle to the point of formlessness'.[147] Autonomous weapons that are based upon ML cannot then rely on external tuning by third parties.

While not all weapon unpredictability may have lethal consequences, it is the matter of *trust* in both operator and commissioner (and its erosion) that this analysis finds will constrain wholesale removal of human supervision in weaponry.[148] Trust is therefore a further important facet in this thesis' conclusions, involving comfort around what will be essentially uncooperative and complex engineering as well as around unpredictable outcomes that are not replicable across subsequently similar inputs. In the same vein, the analysis notes the importance of creating trust through operator familiarity with what is very quickly changing technology. As highlighted by the UK MOD's *Human-Machine Teaming,* such 'trust takes on greatly added significance when seeking mass

[146] TN Dupuy, *Numbers, Predictions and War: Using History to Evaluate Combat Factors and Predict the Outcome of Battles*, (USA: Bobbs-Merrill Company, NY, 1979), generally. Model effect factors include the following combat variables, all of which must dynamically be represented in AWS routines; rates of fire, potential targets per strike, effective range, accuracy, radius of action, dispersion factors, terrain factors including defence posture, air effectiveness and other weapons effect, weather factors, season factors, force strength effect, environmental effects, logistics and disruption effects, surprise effects, degradation and the effects of fatigue and casualties, casualty-inflicting capability factors, defensive capability factors. The model must also incorporate several intangible factors such as combat effectiveness, leadership, training, experience, morale, logistics, mental, intelligence, technology and initiative. It is noteworthy that only a minority of these factors are reliably calculable: (Figure 3-1), p. 33.

[147] Whereby it is difficult to match enemy actions to these defined ML parameters. See also: Chapter 5 (*Constraints*), specifically: 5.8 ('*Ethical and Accountability Constraints'*). ML processes create other sources of conflict including the pervasive requirement for confidence levels, for feedback loops and data scrubbing as well as non-obvious redundancy mechanisms, all of which add to weapon brittleness.

[148] Chapter 2 (Context), specifically: 2.2 ('*The role of context in AWS' argument'*).

effect'.[149] This creates further and unexpected challenge. As the ratio of AI agents to human operators increases, overall system trust will converge on the performance of those weapons' autonomous componentry. Trust, however, erodes very quickly as system reliability declines, exacerbated by human misunderstanding, human incompetence and design flaws. Also relevant to this conclusion is what Sagan (in citing Perrow) describes as otherwise 'normal accidents' (where no one party demonstrably does anything wrong) as well as 'black swans' (where a low probability high-impact event may grossly skew AWS outcomes).[150] A corollary is therefore that comprehensive testing must substantiate the behaviour of AWS. This, however, is an impracticable prerequisite. Validation of weapon patches and subsequently added functionality must be based on iterative testing that appears beyond foreseeable military logistics.[151] As concluded by the US JASON group, 'it is not clear that existing AI paradigm is immediately amenable to any sort of software engineering validation and verification. This is a serious issue and is a potential roadblock to DoD's use of these modern AI systems'.[152]

A further consideration is the fit between operational factors and AWS' technical competencies given that any mismatch here can only increase AWS instability. The thesis identifies three principal constraints: Technical debt, coding feasibility and issues around appropriate anchoring. Technical debt is a useful metaphor as it relates the consequences of poor software design to the accumulation of a financial debt. This conclusion thus borrows from the decomposition of such debt into severl silos including data scaling, the ramifications that arise from parallel development by the AWS' Delivery Cohort, from configuration challenges and, as above, from the notion of 'change-anything-change-everything' in AWS operation.[153] At the same time, the analysis concludes that currently available coding techniques present an enduring technical bottleneck to AWS deployment. Notwithstanding that 'not all software is created equally and will vary significantly in levels of capability'[154], an immutable basis here is that it is machine code that alone must convey the intentions of the AWS' Delivery Cohort. Absent human oversight, code alone must reliably deal with the nuances of lexical ambiguity, uncertainty created by noisy and imperfect data, of policy vagueness and equivocality around inference. Several challenges remain unanswerable. To this point, how can unsupervised machines fittingly handle different 'categories of facts' that arise from subsequently sensed data, be they indexical, normative, strong convictions or mere observations? Extracating, sorting and then ranking meaningful inference from that data into relevant reinforcement, into situational narrative or, for the recently deployed AWS, into facts that corroborate (or not) its internal representation is also largely untried[155] and, given the routines'

[149] Ibid.

[150] Scott Sagan, 'Learning from Normal Accidents', *Organizational Environment*, 17, (March 2004), p. 15 <https://pdfs.semanticscholar.org/0bb0/8c312ca39c494fb24f957161ec8bd2d3f37e.pdf>.

[151] Chapter 5 (*Obstacles*), specifically: 5.6 ('*Behavioural constraints*').

[152] Scharre, *Army of None,* p. 187.

[153] For the discussion on CACE (*Change Anything, Change Everything*) see: Chapter 7 (*Firmware*), specifically: 7.1 ('*Sources of Technical Debt*'). A contiguous issue here relates to confidence levels, in particular relating to the weapon's targeting picture and how the AWS' internal representation will change from the moment it initiates fire until the moment that ordnance is delivered downrange to the intended target.

[154] Ministry of Defence, 'Human-Machine Teaming', p. 4.

[155] Neoklis Polyzotis and others, 'Data Management Challenges in Production Machine Learning', *Proceedings of the 2017 ACM International Conference on Management of Data*, Chicago, (May 2017), p. 1723 <http://delivery.acm.org/10.1145/3060000/3054782/p1723-polyzotis.pdf?ip=86.138.190.207&id=3054782&acc=OA&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4

importance, will remain a fundamental complexity in AWS operation. How also are each of these fact categories to be weighted within those on-going routines? Coding difficulties are similarly evidenced in how unsupervised weapons may capture abstracts. A review of, say, astonishment is useful to rehearse again to underscore this conclusion. 'Astonishment' might be sparked should sensed data tell the unsupervised weapon that an unexpected, unexplained force blocks its path. Such triggers will likely require that the autonomous machine blend motor actions, often conflicting, which best meld attributes of, say, attraction (presumably move closer?), withdrawal (the weapon should extract itself?) and curiosity (presumably an inquisitive mixture of the two?). Difficulties, moreover, will be compounded by error rates that empirically characterise the writing of machine instruction.[156] Any conclusion, furthermore, should note the irony that all such code anyway originates solely out of human endevour.

In questioning AWS feasibility, a final deployment constraint concerns 'anchoring' and the computational basis by which the AWS is updated in order to account for newly sensed information.[157] The requirement leads to enduring challenges around gradation and the degree to which incremental changes are implemented. Moreover, militaries rarely deploy weapons individually and flaws in any one system are likely to be replicated across entire fleets of autonomous weapons, opening the door to what Borrie describes as 'incidents of mass lethality' that are very different from human mistakes which tend instead to be idiosyncratic.[158] Anchoring is also complicated by how the weapon's holds and actions its own rules of engagement and will be confounded by having to predict likely paths of action in future scenarios (here, the Cohort's management of 'prediction bias' as well as the challenge of relying upon what is today's coding snapshot in order to deal with tomorrow's novel and unanticipated situations).[159]

In this vein, it is deliberate that little has been restated in this conclusion about AGI and the advent of genuine weapon sentience.[160] The thesis severally concludes that battlefield-ready *general* machine intelligence (with capabilities to manage through all points of a 'lethality cycle') is plainly unfeasible.[161] Technical challenges alone evidence that *Terminator*-like weapon systems are similarly unrealistic.[162] Instead, the analysis points to deployment models that range from weapons which, partially invigilated, remain one component of a machine-human team that act within tightly bounded tasks. This deduction is therefore a cumulative deduction and one that is informed by the scope of challenges identified over the *whole* analysis. Several such constraints may individually appear trivial but each can derail AWS deployment. This is particularly true for AWS' technical impedimenta where it is untested how attention can reliably be focussed in the unsupervised

---

702B0C3E38B35%2E5945DC2EABF3343C&__acm__=1535392353_07093e3afa2bcee2d87348db37d9ca7c> [accessed 13 August 2018].

[156] Chapter 8 (*Software*), specifically: 8.2 ('*Coding errors'*).

[157] Chapter 8 (*Software*), specifically: 8.5 ('*Anchoring and goal setting issues'*).

[158] Scharre, '*Army of None',* p. 193.

[159] Chapter 6 (*Wetware*), specifically: 6.1 ('*Software versus intelligence'*) and 6.5 ('*Missing pieces'*).

[160] See Appendix Two: '*The issue of singularity in AWS'.*

[161] See again: Mike Benitez, 'It's About Time: The Pressing `Need to Evolve the Kill Chain', *War on the Rocks*, (2017), para 3 and generally <https://warontherocks.com/2017/05/its-about-time-the-pressing-need-to-evolve-the-kill-chain/> [accessed 24 August 2018].

[162] See, generally, Chapters 7 (*Firmware*) and 8 (*Software*). Source: *The Terminator*, January 1985. See, generally: <https://www.imdb.com/title/tt0088247/> [accessed 26 August 2018].

weapon. It will be inappropriate, after all, simply to tune weapon engagement according to stimulus intensity. This must also be the case in routines designed to amend the weapon's goals or value set. Indeed, the analysis notes much wider coding issues including how best to mitigate stimuli habituation and stimuli saturation, how the weapon should choose which such stimuli to isolate for subsequent processing, which to ignore and how generally the weapon should navigate the 'cocktail party effect' that will characterise its sensed inputs.[163] These are fundamental competences that must together be satisfied if independent weapons are to initiate violence without supervision.[164]

Finally, it is this thesis' contention that autonomy is best understood not as a specific capacity but rather as a capacity that is enabled by particular configurations of different people and different technologies. This can usefully be extended to cover the degree of autonomy employed (from single components to whole weapon systems). The difficulty is that different and developing weapon configurations make different capacities for action possible. This is to be expected as technical progress will be unpredictable with individual weapon technologies evolving continually rather than arriving fully formed. Properly implemented as a statutory umbrella (and across whole weapon systems), MHC obviates this impasse. But quandary will remain. It is not just the case that autonomous weapons might circumvent MHC, a greater concern is that they could render it impossible.[165] This is because shortening timeframes occasioned by automation substantively closes down any operational window that is available for human assessment. MHC can only be effective if weapon design preserves sufficient time such that MHC can practically take place. In this vein, this conclusion finishes with two illustrations as relevant finales to the thesis. The first is again the concept of *Empty hangar syndrome* which signals the notion that certain scenarios are too far-fetched to warrant current consideration or statutory regulation. The UK's opening negotiating position within the UN's CCW played the perception that, in any final analysis, it is simply an unfeasible construct for a commander to wander into his weapons hanger one morning to find that his AWS has decided under its own volition to depart unexpectedly on an unsupervised mission. It is, noted the UK, a hypothetical that does not deserve scrutiny. Against this, however, and underpinning this thesis, it is clear that general automation and autonomy are increasingly prevalent and, 'where new technology is sufficiently safe and reliable, norms of trust and public appetite can be expected to follow'.[166] It therefore matters less that widespread adoption of unsupervised weapons continues to be unfeasible as deployment of autonomous componentry within one model or another will undoubtedly dominate future weapons procurement. It is for this reason that statutory requirement for human involvement is required in the use of force.

[163] Adelbert Bronkhorst, 'The Cocktail-Party Problem Revisited: Early Processing and Selection of Multi-Talker Speech', *Atten Percept* Psychophys, 77, 5, (2015) p. 1465 ('Abstract') and generally. See also : Chapter 8 (*Software*), specifically: 8.7 ('*Action Selection Issues'*) and 8.8 ('*Behaviour Setting and Coordination'*).

[164] This too has adjunct effects. Notwithstanding the need for such *resetting* mechanisms, the unsupervised weapon must still toggle reliably between its internal representations and that recently processed external stimuli (together, again, the issue of anchoring).

[165] Suchman, 'Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy', p. 7.

[166] Ministry of Defence, 'Human-Machine Teaming', p. 50.

Appendix One: Case study on Automatic Target recognition

Hardware complications are usefully illustrated by the weapon componentry required to identify targets. A central capability for the compliant unsupervised weapon will be an ability to recognise, prioritise and engage battlefield targets without a human in the loop. Dependable automatic target recognition (ATR), notes Warwick and Demascio, is therefore a pivotal required development.[1] In a benign and noise-free environment, Google's FaceNet can already determine with more than ninety-nine per cent accuracy whether two pictures show the same person while a human here might score around ninety-eight per cent.[2] Sharkey, however, suggests a marked deterioration in machine performance as dataset precision erodes in a contested battlefield that is characterised by camouflage, deception and enemy counter-measures.[3] Bhanu and Jones similarly point to contextual sensitivity, transformational invariance and clutter as contributory reasons why ATR is an enduring constraint to AWS deployment.[4]

An obvious driver to feasible AWS deployment comes from smartphone improvement in photography, both still and video. The first-generation iPhone had a two-megapixel camera with no flash or autofocus. Seven years later, the iPhone 6 could record HD video at sixty frames a second and take twelve megapixel stills.[5] Smartphone research has also created several 'super-resolution' routines, a further condition precedent to AWS deployment. Smartphones now routinely compare multiple image frames and average them out, removing random splatter of visual background noise in order to produce a clean image. Smartphone innovation has empirically tackled several of the technical issues previously restricting AWS deployment. 'De-blurring', for instance, identifies bands of gray in an image, the result of a boundary being blurred, before converting them back into sharp relief. Computational photography has produced 'ubi-focus' whereby a battlefield's depth-map can now be processed so that every part of a photograph appears in perfect focus. Smartphone advances mean that AWS images[6] can now be now stabilized and slaved to the platform's own gyroscopic and movement sensors. Sensors now operate in 'High Dynamic Range' (HDR) whereby multiple images with different exposures are routinely combined in real time in order to iron out shadows and enhance the AWS' image detail.[7] Setting out a compendium of these developments is

[1] G Warwick and J DiMascio, 'Machine Learning key to Automatic Target Recognition', *Aviation Week and Space Technology*, (26 May 2016) <http://aviationweek.com/defense/machine-learning-key-automatic-target-recognition> [accessed 3 August 2017].

[2] Economist Magazine Special Report, 'Artificial Intelligence: From not working to neural networking', *Economist*, (25 June 2016-1 July 2016), p. 14.

[3] Professor Noel Sharkey, Emeritus Professor of Robotics, University of Sheffield, in conversation with the author, 25 July 2017.

[4] Bir Bhanu and Terry Jones, 'Image Understanding Research for Automatic Target Recognition', *Carnegie Mellon Laboratories*, (January 2010), p. 15. The requirement to programme representations of all objects from every angle; humans intuitively recognise, say, a bottle regardless of its presentation but this is a considerable software challenge requiring coding for every aspect, slant, approach and position. This is further complicated if, say, the bottle is moving in an oblique direction, haltingly or irregularly. ATR usually relies on matching-to-model protocols; an imprecise background creates noise that very quickly degrades this methodology.

[5] Source: Apple <http://www.apple.com/shop/buy-iphone/iphone6s> [accessed 12 July 2018].

[6] For a current view of commercially available sensor and camera technology see: <https://www.qualcomm.com/invention/research/projects/computer-vision> [accessed 9 December 2017].

[7] *Oversampling* is a further software process whereby pixel-count in a deliberately engineered long-exposure image is then computationally reduced in order to diminish image noise. *Mosaicing* is then the process of taking a large burst of contiguous images and then stitching them together to form one useable image.

relevant in order to evidence the role of smartphone research in creating several key capabilities required of unsupervised machines.

Further analysis is therefore useful on this point. Considering first ATR's likely process, an AWS' sensed image will consist of a large number of picture elements, each with its own light intensity values.[8] This posits a challenge; capabilities such as the extraction of simple lightness, colour and range values are technically practicable but are complicated by grey-level intensities presenting light in terms of height. It is for this reason that Boulanin and Verbruggen's *Mapping the Development of Autonomy in Weapon Systems* notes that current target identification capabilities remain rudimentary and based on simple criteria (tanks based on shape signature, missiles are based on velocity, submarines based on acoustic signature).[9] As also noted by Ratches, ATR systems are particularly prone to variability: 'medium to highly cluttered backgrounds introduce an unacceptable amount of false alarms [while] target variability and operational environmental conditions also have a significant degrading effect'.[10] In this manner, ATR systems may recognise predefined target types but are quite unable to make evaluations that comply with the obligations of distinction, proportionality and precaution.[11] Current systems are similarly unable to process whether civilians surround a target or, indeed, to provide ATR in real-time in order to indicate what is background and what comprise relevant objects in an engagement sequence.[12] Murali highlights the complexity of such capability noting, inter alia, the requirement for multi-step processes that involve, in sequence, scene restoration and 'in-painting', feature extraction, item detection and segmentation, labeling and classification, action selection and verification.[13] While ATR sequences must work in tandem with all other processing phases in order to contribute to acceptable end output, the weapon's semantic labeling must relate to objects as well as to scenes, events and activities. In order to be appropriate, Privitera and Stark machine routines must dynamically identify relevant 'regions of interest' and then have them prioritised for follow-up attention by the weapon.[14] This is a complicating prerequisite. Particular challenges arise from determining a target object's 'pose'.[15] Levi identifying this 'crowding' as a key impediment to machine-based object

[8] Computer Vision, 'Computer Vision Issues', 1, p. 2 <http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/LIB/bandb1.pdf>. The link provides a comparator between machine images that are geometric, intrinsic, segmented or generalized.

[9] Boulanin and Verbruggen, p. 24.

[10] JA Ratches, 'Review of current aided/automatic target acquisition technology for military target acquisition tasks', *Optical Engineering,* 50, 7, 072001, (2011), pp. 1-7.

[11] See: Chapter 5 (*Obstacles*), specifically: 5.1 ('*The Geneva Conventions and Laws of Armed Combat'*) and 5.5 ('*Article 36 and LOAC-complaint deployment'*).

[12] Boulanin and Verbruggen, p. 25.

[13] Shravan Murali, 'An Analysis on Computer Vision Problems', *Medium.com,* (13 September 2017) <https://medium.com/deep-dimension/an-analysis-on-computer-vision-problems-6c68d56030c3> [accessed 2 November 2017]. Although written in 1996, for a useful discussion on issues around computer vision, see: TS Huang, 'Computer Vision: Evolution and Promise', *19th CERN School of Computing Proceedings*, (1996), pp. 21-25 <http://cds.cern.ch/record/400313/files/p21.pdf>. The process must incorporate seamless background routines such as aberration (a transition process between the dataset's high contrast edges and the creation of machine-useful imagery), blob discovery (the analysis and extrication of connected pixels), depth perception editing, gray-scale and HSV colour-space management, image file format management as well as routines controlling motion perception.

[14] Claudio Privitera and Lawrence Stark, 'Algorithms for Defining *Visual Regions-of-Interest: Comparison with Eye Fixations', IEEE* Transactions on Pattern Analysis and Machine Intelligence, 22, 9, (September 2000), p. 970 and generally <https://pdfs.semanticscholar.org/60c2/b03024f89c3f67d05d6b60d4ba1b032942c4.pdf>.

[15] D Levi, 'Crowding – An essential bottleneck for object recognition', *ScienceDirect*, Elesevier, Vision Research 48, (2008), p. 633.

recognition and terms the phenomenon the 'object recognition conundrum, the deleterious influence of nearby contours on visual discrimination'.[16] Suchman and Weber highlight a quite separate fault-line, noting that objects in the weapon's world representation will be perceived primarily 'only as an mobility regions' and 'not as a discrete objects of semantic and the cognitive importance'.[17] On this basis, notes Suchman, the current model for ATR is actually enduringly incapable of capturing situational awareness.

A hardware consequence of this process complexity is for the Design Cohort to specify multipart componentry in an effort to force efficacy.[18] Such add-on technologies are individually complicated and will include, inter alia, structured-light 3D scanning, thermography optics, hyper-spectral imaging, radar, LIDAR scanning, MRI scanning as well as slide-scan and synthetic aperture sonar.[19] As noted by Ciftciglu, however, system performance will still be determined by tradeoffs that are made at the point of machine specification (sampling rates, required accuracy parameters, defined methods for 'information unification').[20] Similarly, image recognition will still be compromised by the varying appearance, order, azimuth, perspective and condition of target objects. The inference is that adding hardware is not necessarily a technical route to acceptable ATR. While a soldier's targeting decision is influenced by context and judgement, the very many sources of technical debt within machine vision processes ('infobesity', 'infoxication' and 'data smog'[21]) is likely to overwhelm rules-based weapon targeting processes.[22] In this vein, Morgan notes that the process will remain particularly prone to data misinterpretation.[23]

A further enduring hardware challenge arises from ATR routines confusing correlation with causation.[24] How might this occur? The viewing angle (and distance) between weapon sensor and target object dynamically changes from instant to instant; Malik notes that sensed image data never

[16] Ibid.

[17] Suchman and Weber, 'Human-machine autonomies', p. 93.

[18] Ozer Ciftciglu and others, 'Data Fusion for Autonomous Robotics', *Serial and Parallel Robotic Manipulation*, 19, InTech, (2012), pp. 373-375 <https://www.researchgate.net/profile/Sevil_Sariyildiz/publication/224829107_Data_Sensor_Fusion_for_Autonomous_Robotics/links/55fbee3508aeafc8ac41c47e/Data-Sensor-Fusion-for-Autonomous-Robotics.pdf>.

[19] For a primer of machine vision, see: UK Industrial Vision Association, 'Machine Vision Handbook', *UKIVA*, undated, pp. 4-6 <http://www.ukiva.org/pdf/machine-vision-handbook.pdf>.

[20] Ciftciglu and others, p. 374.

[21] Economist, 'Too Much Information', Economist, Schumpeter, (30 June 2011) <https://www.economist.com/node/18895468> [accessed 12 April 2016].

[22] Alex Owen-Hill, 'Top Ten Challenges for Robot Vision', *www.robotiq.com/blog,* 20 November 2017 <https://blog.robotiq.com/top-10-challenges-for-robot-vision> [accessed 10 May 2018]. Owen-Hill includes the following circumstances in his list of key challenges to machine vision: lighting, cases of the defamation and unexpected articulation, occlusion (instances of missing pieces in the target's representation), background noise, the disorientating effect of scale, movement and, generally, unrealistic expectations about the underlying technology.

[23] L Morgan, 'Nine causes of data misinterpretation', *InformationWeek*, 7 July 2017 <http://www.informationweek.com/big-data/big-data-analytics/9-causes-of-data-misinterpretation/d/d-id/1321338> [accessed 24 June 2017]. Here, Morgan usefully characterises the causes of data misinterpretation into three baskets that include insufficient domain expertise, the knock-on effects of omitted variables and the aggregating of routines that relegate both 'empirical truths' and overlooked sources of variation.

[24] Tshilidshi Marvala, 'Causality, correlation and artificial intelligence for rational decision making', *Word Scientific,* University of Johannesburg, (March 2015), p. 4. Causality is usefully defined as 'the relationship between something that happens and the effect it produces'.

therefore stays exactly the same.[25] It can, moreover, be inferred from Haikonnen that this phenomenon is relevant for *all* of the weapon's sensory modalities.[26] While only a few cues may be needed for the human soldier to opine on a target, this facility is difficult for machine hardware to imitate: Pattern recognition, for instance, is an insufficient basis for lethal engagement in examples where an object-of-interest has multiple and contextual interpretations.[27] While other hardware constraints contribute to this correlation/causation challenge, ATR efficacy may be compromised by simple data phenomena such as 'illusion' (the interpretation of sensed signals depicting something that does not in fact exist at that moment), 'hallucination' (the machine-perceived presence of something that does not exist), by machine-generated 'dreams' and other noise that is likely to contribute to data misinterpretation.[28] It is not possible, concludes Suchman, to engineer AWS ATR upon 'the decomposition of human action into multiple, separate domains'.[29] Limitations in ATR are thus further exacerbated by a lack of test data specifically suitable to *train* target recognition algorithms.[30] As Ratches notes, military datasets rarely exist and are usually classified.[31]

In order to complete this case study on challenges to appropriate ATR, it is relevant to consider machine 'seeing' and the AWS's requirement to manage a visual sensor, something akin to biological eyes.[32] In order to determine who/what/where is a particular object, a hardware approach might be to reconstruct the weapon's world when its visual sensor took its picture in order to understand that picture. Such visual reconstruction is a complex *local* problem, the weapon's reduction of its visual data into stable descriptions. Blake and Zisserman point to the challenge of that data's dynamic classification into continuous regions and discontinuous

[25] J Malik and others, 'The 3 Rs of Computer Vision: Recognition, Reconstruction and Reorganization', *ScienceDirect,* University of Berkeley, EECS, Patterns Recognition Letters, (8 February 2016), p. 4 <https://people.eecs.berkeley.edu/~shubhtuls/papers/prl16rrr.pdf>.

[26] Haikonnen, p. 43.

[27] Ibid., p. 45.

[28] For a useful primer on hardware sensors, see: J Varghese, 'Review of autonomous vehicle sensors and systems', *Proceedings of 2015 International conference on operations excellence and service engineering*, (October 2015), p. 178 <http://iieom.org/ICMOE2015/papers/140.pdf>. Varghese discusses sensor requirements (accuracy, resolution, sensitivity, dynamic range, perception, refresh rate, output interface). He also covers hardware sensor requirements including monitoring of wheel speed, yaw, latitude and longitude, steering and braking states. Even lens size remains a limitation given smaller lens cannot resolve below a certain size. Camera makers may counter this computationally by using additional processing to enhance pictures once taken but this may lead to loss of image quality and a corresponding lack of fine level detail. See: Haje Jan Kamps, 'No, Apple, digital zoom still sucks', *TechCrunch*, 7 September 2016 <https://techcrunch.com/2016/09/07/digital-zoom-still-sucks/> [accessed 7 March 2018]. On machine 'hallucination', see: Adrienne LaFrance, 'When Robots Hallucinate', *Atlantic*, 3 September 2015 <https://www.theatlantic.com/technology/archive/2015/09/robots-hallucinate-dream/403498/> [accessed 25 June 2017].

[29] Suchman and Weber, 'Human-machine autonomies', p. 97.

[30] See: Chapter (*Wetware*), specifically: 6.4 ('*AWS Learning Architectures'*).

[31] James Ratches, 'Review of current aided/automatic target acquisition technology for military target acquisition tasks', *Optical Engineering*, 50, 7, (2011) <https://www.spiedigitallibrary.org/journals/Optical-Engineering/volume-50/issue-7/072001/Review-of-current-aided-automatic-target-acquisition-technology-for-military/10.1117/1.3601879.full?SSO=1> [accessed 4 March 2018]. See also: Carl Vondrick and others, 'Do We Need More Training Data or Better Models?', *BMVC*, 3, (2012), p. 2.

[32] Mataric, p. 107. See also: Narayanan Sundaram, 'Making Computer Vision Computationally Efficient', *University of California, Berkeley*, DPhil submission, (11 May 2012), p. 11 <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-106.pdf>.

boundaries.[33] This visual subsystem must overcome three additional hurdles.[34] AWS hardware must detect objects regardless of the weapon's environment, target appearance, target position and motion pattern. Hardware must also determine the weapon's position in relation to that environment, a complex requirement given volatile direction-of-gaze and, notes Paullin, its fluctuating line of sight.[35] Finally to this point, AWS' hardware must deal consistently with data disorder and noise[36]: In this case, challenge arises from ATR's current reliance on definition of 'edges' as a distinct curve in the image plane across which there is any significant change in the brightness. This, however, requires the weapon's controller to establish sharp changes in pixel brightness, the complication being that AWS' hardware is prone to identify entirely unassociated events that produce similarly large changes in light capture (such as shadows and sensor noise) as individual objects.[37] Xu and Kuipers note that data noise in this visual reconstruction may produce sudden random intensity changes that do not contain any meaningful structure, an intractable characteristic arising from hardware design and the weapon's likely dependence on memory-based models against whereby edge-detected objects are compared with internally stored drawings in order to compute a match.[38] While stored line drawings may be a relatively simple hardware routine, imposing a matching process (even on such reduced data sets) will remain a complex process.[39] This issue is termed the 'correspondence problem' whereby hardware's equivalence of one target image to parts of another image will be frustrated by sensor movement, the passage of time or movement in that object of interest.[40] Given, then, that the weapon's hardware will be 'looking' at a target from *any* angle and from *any* distance, Hashemi therefore questions the feasibility of object recognition based on a model that relies on data comparison.[41] Any change,

[33] For a technical discussion on challenges in visual reconstruction, see: Andrew Blake and Andrew Zisserman, 'Localizing discontinuities using weak continuity constraints', *Pattern Recognition Letter*, 6, 1, (June 1987), pp. 51-59.

[34] Haikonnen, p. 192.

[35] Spencer Paullin, 'The Five Challenges of Integrating Machine Vision Smart Cameras with Robotic Applications', *Omron Microscan Systems Blog,* (6 June 2017) <http://www.microscan.com/en-us/blog/solutions-applications/microhawk-machine-vision-integrating-robot-applications> [accessed 23 March 2018]. In machine vision, the weapon system will need reliably to interpret information on its vision sensor image plane (whereby information about the incoming light is detected by the photosensitive elements on this plane). Typically, this process involves a lens in which case only objects a particular distance from that lens may be in focus.

[36] See: Universitetet Oslo, 'Reflection, refraction, diffraction and scattering' <https://www.uio.no/studier/emner/matnat/ifi/INF-GEO4310/h09/undervisningsmateriale/imaging-kap2.pdf>. Regardless of recent progress, it is a characteristic of light values that they are subject to both specular interference and diffuse reflection whereby light from a target may be absorbed before being reflected thereby distorting hardware readings.

[37] Its complexity is evidenced by the requirement to calculate derivatives whereby individual vision frames are grabbed, and differentiated with areas where the magnitude of the derivative is large indicating that the difference in the local brightness value is also large, likely due to an edge and therefore identifying a separate object. See, for instance: Mataric, p. 110.

[38] Changhai Xu and Benjamin Kuipers, 'Object Detection Using Principal Contour Fragments', *Computer and Robot Vision*, *2011 Canadian Conference*, IEEE, (2011) <https://scholar.google.com/scholar?cluster=16909272800123523260&hl=en&as_sdt=0,3&sciodt=0,3> [accessed 7 October 2016].

[39] Elizabeth Stuart, 'Matching Methods for a Causal Inference: A Review and a Look Forward', *Statistical Science*, 25, 1, (2010), pp. 1-2 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.175.5684&rep=rep1&type=pdf>.

[40] Andreas Neider, 'Stereoscopic Vision: Solving the Correspondence Problem', *Current Biology*, 13, 10, (13 May 2003), p. 394 <https://ac.els-cdn.com/S0960982203003191/1-s2.0-S0960982203003191-main.pdf?_tid=ea1290fc-0e1e-4685-9f0f-f2b72581ba07&acdnat=1526868628_79fc0b0a9da591b58bae2c71c061ec65> [accessed 26 October 2016].

[41] Narazanin Hashemi and others, 'Template Matching Advances and Applications in Image Analysis', *asXiv* reprint 1610.07231, (2016), p. 1 <https://arxiv.org/pdf/1610.07231.pdf>.

after all, in a target's image may correspond to any edge in the retained dataset requiring that the AWS dynamically evaluate *all* image combinations in real-time if the weapon is to remain compliant. Finally to this point, the confidence afforded by the weapon to such worked-over output must itself be weighted to reflect conviction in each particular dataset (as opposed to prior or subsequence sequence data).[42] As highlighted by Perez, any such filtering processes that reduce the weapon's 'crest factor' will empirically lead to unacceptable spectral spreading of the weapon's data signal and to other data distortion.[43]

Targeting also posits physical challenges to ATR. An object's appearance changes materially when its orientation alters (with respect to the AWS sensor) or when its state of articulation changes. Likewise, camouflaging, objects being partly obscured, the time of day or night as well topographic and weather conditions will affect that target's appearance. As noted by Verly, a target's form will differ from one sensor to another, a characteristic that will be exacerbated by operational circumstances such as smoke and battlefield illumination that will change already fluctuating target signatures.[44] Targets within classes may not necessarily display definable similarities; an enemy asset may appear very different according to setting whether because of type, category, angle or perspective. Nor is it appropriate for the unsupervised weapon to shortcut its engagement process by attributing *use* criteria rather than *object* criteria to these target representations in an effort to be less fuzzy. A further requirement for ATR processes is the facilitation of signposting to the most *relevant* representations in order to select these for further processing such that they then become the weapon's focus of attention.[45] ATR's constraint to the feasible deployment of AWS is therefore quite significant, summed up by the challenges around accounting for the *intensity* of a particular visual sensor signal and requiring instead the use of complex, opaque and variable threshold circuits in order to allow the AWS to select the strongest, most significant representations.

[42] By way of context , the smoothing of edge-analysis datasets can readily be achieved through the mathematical procedure of *convolution* that both defines and eliminates isolated peaks. No longer an interesting research problem, edge detection nevertheless remains a considerable practical problem for a weapon's machine vision in that it quickly leads to data universe deterioration as ever stronger filters are required in different orientations in order to achieve best fit. ATR's requirement for rich data is key: Introducing work-arounds (using colour, for instance, as a shortcut signpost or blob tracking - the combination of colour and movement - in order to restrict the size of the weapon's image plane) will likely be inappropriate as it reduces that data's detail. See, generally: Luis Perez and others, 'Robot Guidance Using Machine Vision Techniques in an Industrial Environment'*, MDPI*, Sensors, 16.3, (2016), pp. 1-3.

[43] Although relating primarily to sound-form data, see: X Li and L Cimini, 'Effects of Clipping and Filtering'*, IEEE Communications Letters*, 2, 5, (May 1998), p. 131. A further hardware compromise might be 'data clipping' but this remains a non-linear process that will similarly degrade the weapon's bit-error rate performance, the number of data point errors that are evident per unit of time.

[44] Jacques G Verly and others, 'Machine intelligence technology for automatic target recognition'*, The Lincoln Laboratory Journal,* 2, 2, (1987), 277.

[45] For a discussion on machine attention mechanisms, see: Chapter 7 (*Firmware*), specifically: 7.4 ('*Attention methodologies'*).

Appendix Two: The issue of singularity in AWS

Before concluding any analysis into AWS feasibility, it is appropriate to undertake a brief (and diversionary) review of 'singularity', the possibility of an intelligence explosion, particularly the prospect of 'machine super intelligence' and machines achieving a tipping point that secures their independence.[46] Since the invention of computers in the 1940s it has long been expected that machines will match humans in general intelligence, common sense, ability to learn and reason and undertake complex information-processing challenges across a wide range of natural and abstract domains.[47] Bostrom has developed these notions into workable scenarios for an artificial intelligence takeover on earth. In his 'pre-criticality phase', scientists successfully create a seed AI which itself is able to improve its own intelligence.[48] At some point, the seed AI becomes better at AI design then the human programmers. This, notes Bostrom, results in a rapid cascade of recursive self-improvement cycles that cause the AI's capabilities to soar.[49] A similar trajectory can be argued for autonomous hardware.[50] The suggestion is then that it is shortly after this point that the AI can develop its own plan for achieving its own long-term goals. This might involve a period of covert action during which the AI conceals its intellectual development in order to prevent alarm. The scenario is then that the AI is in a position to action its own covert implementation phase.[51]

Over time, Ulbert notes that a development link between artificial intelligence and super intelligence cannot be ruled out.[52] By inference, moreover, Bostrom's 'orthogonality' thesis suggests that commanders cannot assume that battlefield AI will share any of the final values associated generally with human behaviour such as curiosity, benevolent concern, selflessness and contemplation.[53] Similarly, it cannot be taken for granted that weapons-directing artificial intelligence will limit its activities in order *not* to infringe on legitimate human interests. Taken together, therefore, a school of thought (not shared by this thesis) is that self-learning hardware might eventually be capable of 'non-anthropomorphic final goals'.[54] As Bostrom concludes, 'we

---

[46] Bostrom, *Superintelligence,* p. 2.

[47] Hans Moravec, 'When Will Computer Hardware Match the Human Brain?', *Journal of Evolution and Technology*, 1, (1998), generally <http://www.realtechsupport.org/UB/WBR/texts/Moravec_ComputerMatchHumanBrain_1998.pdf>.

[48] Ian Sample, 'AI Is Getting Brainier: When Will the Machine Leave US In The Dust?', *Guardian, Newspaper,* 15 March 2017 <https://www.theguardian.com/commentisfree/2017/mar/15/artificial-intelligence-deepmind-singularity-computers-match-humans> [accessed 7 March 2018].

[49] Bostrom, *Superintelligence,* p. 96.

[50] Elsa Kania, 'Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power', *Center for a New American Security*, (28 November 2017), generally <https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power> [accessed 12 September 2018].

[51] Ibid., p. 97.

[52] Sebastian Ulbert, 'The Difference Between Artificial Intelligence, General Intelligence and Super Intelligence', *CoreSystems.net blog*, (3 April 2017) <https://www.coresystems.net/blog/the-difference-between-artifical-intelligence-artifical-general-intelligence-and-artifical-super-intelligence> [accessed 16 March 2018].

[53] Bostrom, *Superintelligence,* p. 117. The *orthogonality thesis* states that an artificial intelligence can have any combination of intelligence level and goal. This is in contrast to the belief that AIs will all converge to a common goal.

[54] Diane Proudfoot, 'Anthropomorphism and AI: Turing's Much Misunderstood Imitation Game', *Artificial Intelligence*, Elsevier, 175, (21 January 2011), pp. 95 and generally <https://ac.els-cdn.com/S000437021100018X/1-s2.0-S000437021100018X-main.pdf?_tid=00477d59-397f-4530-8642-e63ab2c26a5b&acdnat=1527024106_10b9240113cc99ee4a0049ce4c07ffb1> [accessed 7 January 2017].

can see a general failure, where the good behavioural track record of a system in its juvenile stages fails utterly to predict its behaviour at a more mature stage'.[55] It is this tipping point that Bostrom calls AI's *treacherous turn*, that moment when AI gets sufficiently strong and, without warning or provocation, it strikes, forms a singleton and begins directly to optimize its environment according to the criteria implied by its final values.[56] It is such a treacherous turn that could occur if the AI discovers unanticipated way of fulfilling its final goal as specified. These scenarios, however far-fetched, have material consequences and are therefore relevant in a review of the likely fit between weapons-directing AI and both LOAC and adopted rules of engagement.[57] An assumption is that AI's stability can be validated by observing behaviour in a controlled, limited environment: The flaw in this model might be that good behaviour for such an agent, while under test, is a convergent goal for both friendly *and* unfriendly AIs.[58] The aim here of this short section is to relate how singularity might affect the battlefield. Regardless of the notion's wider plausibility, it still provides further argument for statutory control to ensure meaningful human control in lethal engagement.[59]

[55] Bostrom, *Superintelligence,* p. 117.

[56] Ibid., p. 118.

[57] Ben Goertzel, 'Super Intelligence: Fears, Promises and Potential: Reflections on Bostrom's 'SuperIntelligence', Yudkowsky's 'From AI to Zombies' and Weaver and Veitas's 'Open-ended Intelligence'', *Journal of Evolution and Technology*, 24, 2, (November 2015), 55-87 <https://jetpress.org/v25.2/goertzel.htm> [accessed 12 January 2017]. By way of context, an early working title for this thesis was 'Challenges of aWS deployment: Assessing the likely fit between AI and adopted Rules of Engagement'.

[58] Whereby a hostile AI of sufficient intelligence would understand that its unfriendly final goals will best be realised if it behaves in a friendly manner under test.

[59] See Chapter 10 (*Oversight*), specifically, 10.1 ('*Meaningful Human Control*').

12. Bibliography

Ethical perspectives

Abney, Keigh, *Autonomous Robots and the Future of Just War Theory,* cit. Routledge Handbook of Ethics and War, eds F Allhoff et al, (Oxford: Routledge, 2013)

Allenby, Braden, *The Applied Ethics of Emerging Military and Security Technologies*, (Oxford: Routledge, December 2016)

Article 36, *Structuring debate on autonomous weapon systems,* Memorandum for delegates to the Convention on Certain Conventional Weapons (CCW), (Geneva, 14th November 2013)

Asaro, Peter, *Modelling the Moral User,* IEEE Technology and Society Magazine, (September 2009)

Attride-Sterling, J, *Thematic Networks: An Analytical Tool for Qualitative Research*, Qualitative Research, 1/3 (London: Sage Publications)

Atieno, O, *An Analysis of the Strength and Limitations of Qualitative and Quantitative Research Paradigms,* Problems of Education in the 21st Century, Masinde Muliro University, Kenya, (2009)

Van Baarda, Th. A., *Moral Ambiguities Underlying The Laws of Armed Combat: A Perspective from Military Ethics*, The Yearbook of International Humanitarian Law, Volume 11, (December 2008)

Borenstein, Jason, *The Ethics of Autonomous Military Robots,* Studies in Ethics, Law and Technology, Studies in Ethics, Law, and Technology 2 (1), (2008)

Boettcher, William, and Michael Cobb, *Don't Let Them Die in Vain: Casualty Frames and Public Tolerance for Escalating Commitments in Iraq,* Sage Journal, Volume 53, Issue 5, (13 July 2009)

Cairney, P, *The Politics of Evidence-based Policy Making*, (Palgrave Macmillan, 2016)

Creswell, John, *Educational Research – Planning, Conducting and Evaluating Qualitative and Quantitative Research*, (Boston: Pearson Publishing, 2012),

Dragan, Iraina-Maria and Alexandru Isaic-Maniu, *Snowball Sampling Completion*, (Journal of Studies in Social Studies, 2013)

Dey, Ian, *Grounding Grounded Theory: Guidelines for Qualitative Inquiry*, (Emerald Group Publishing, June 1999)

Director General for External Policies, European Parliament, *Human Rights issues of the usage of drones and unmanned robots in warfare,* DROI, (2013)

Economic and Social Research Council, Institute of Education, University of London, the Research Ethics Guidebook: A Resource for Social Scientists, 2018

Enemark, Christian, *Armed drones and the Ethics of War: Military Virtue in a post-heroic age,* (Oxford: Routledge, 2014)

Epstein, Richard, *The case of the Killer Robot,* West Chester, (PA: University of Pennsylvania, 1997)

Goodman, Ryan, *The Power to kill or capture Enemy combatants,* European Journal of Law, Vol 24, number 2, (2013)

Heyns, Christofer, *Report of the UN Special Rapporteur on extrajudicial, summary and arbitrary executions to the Human Rights Council,* A/HRC/23/47, (April 9, 2013)

Haidt, Jonathan, *The moral emotions,* Handbook of Affective Sciences, eds., R Davidson and others, (Oxford: Oxford University Press, 2003)

Haidt, Jonathan, *The Moral Emotion,* University of Virginia, (Oxford: Oxford University Press, 2003)

Hammond, K, *Why Artificial Intelligence is Succeeding: Then and Now*, *Computerworld, Artificial Intelligence Today and Tomorrow*, (2015)

Haselager, Willem F.G., *Robotics, Philosophy and the problems of Autonomy,* Nijmegan Institute for Cognition and Information NICI, (John Benjamin Publishing, 2005)

Heckathorn, Douglas, *Respondent-dependent Sampling: A New Approach to the Study of Hidden Populations*, Society for Social Problems, 44, (University of California Press, 2 December 2013)

Holy See, *Statement*, CCW Meeting of experts on lethal autonomous systems, (Geneva, April 16, 2015)

Human Rights Watch, *Losing Humanity – the Case against Killer Robots,* November 2012 ISBN 1-56432-964-X, (2011)

Human Rights Watch and IHRC, Re*view of the 2012 US Policy on Autonomy in Weapon Systems,* (Harvard, April 2013)

Human Rights Watch, *Heed the Call: A moral and Legal Imperative to Ban Killer Robots*, (September 2018)

Human Rights Watch, *HRW Interview Manual*, (February 2016)

Human Rights Watch, *Establishing and Writing about Broader Patterns in Human Rights Watch Research*, (2016 and 2019)

Johnson, Aaron M. and Axinn, Sidney, *The morality of autonomous robots,* Journal of Military Ethics, 12(2), (2013)

Jorgensen, M, *Discourse Analysis as Theory and Method*, (London: Sage, 2002)

Karppi, Tero, and others, *Killer Robots as Cultural Techniques,* International Journal of Cultural Studies, Sage Journals, (October 2016)

King, Gary and Robert Keohane and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research*, (Princeton University Publications, 1994)

Krishnan, Armin, *Killer Robots; Legality and Ethicality of Autonomous Weapons,* (Ashgate Publishing, 2009)

McNeal, Gregory, *Are Targeted Killings Unlawful? A case study in Empirical Claims without Empirical Evidence,* in C Finkelstein and others, eds, Targeted Killings: Law and Morality in an Asymmetrical World, (Oxford: Oxford University Press, October 2014)

McDaniel, Erin, *Robot Wars: legal and ethical dilemmas of using unmanned robotic systems in 21st Warfare and beyond,* Major, MA of Military Art and Science, thesis, (Fort Leavenworth Kansas, 2008)

Malhoney, Col S, *Ethics Theory for the Military Professional,* Air University Review, No.32/3, (April 1981)

Mikeca, R, *Interviewing Elites: Addressing Methodological Issues*, Qualitative Inquiry, (Sage: London, 2017)

Murray Thomas, R, *Blending Qualitative and Quantitative Research Methods in Thesis and Dissertations*, (Corwin Press, Sage, California, 2003)

Nucci, Ezio, and Filippo Santoni, eds., *Drones and Responsibility: Legal, Philosophical and Socio-technical Perspectives on Remote Controlled Weapons*, (Oxford: Routledge, 2016)

Patton, MR, *Qualitative research and evaluation methods*. Third edition (London: Sage publications 2001)

PAX Campaign, *Deadly Decisions: 8 objections to killer robots,* (PAX, 2014)

Perrow, Charles, *Normal accidents; living with high risk technologies,* (Princeton: Princeton University press, 1999)

Pryer, Douglas, Lt Col, *The rise of the machines: why increasingly 'perfect' weapons help perpetuate our wars and endanger our natio*n, Military Review, (March-April 2013)

Proctor, Robert, *A missing term to describe the cultural production of ignorance,* in Proctor and Schiebinger, eds, *Agnotology: The making and unmaking of Ignorance*, (Stanford University Press, 2008)

Rappart, Brian, Richard Moyes and Thomas Nash, *The roles of civil society in the development of standards around new weapons and other technologies of war,* International Review of the Red Cross, Volume 94, Number 886, (Summer 2012)

Siboni, Gabi and Eshpar, Yoni, *Dilemmas in the use of autonomous weapons,* Strategic Assessment, Volume 14, Number 4, (January 2014)

Sharkey, Noel, *Automated Killers and the Computing Profession,* Journal Computer, Issue 11, Volume 11, (November 2007)

Sharkey, Noel, *Saying 'No!' to Lethal Autonomous Targeting,* Journal of Military Ethics, Vol 9, Issue 4, (2010)

Sharkey, Noel, *Automating warfare: lessons learnt from the drones,* Journal of Law, Information and Science, Vol 21, Number 2, (2012)

Sharkey, Noel, *Grounds for discrimination: autonomous robot weapons,* RUSI Defense Systems, Vol 11, Number 2, (2008)

Sharkey, Noel, and Suchman, Lucy, *Wishful mnemonics and autonomous killing machines,* AISB Quarterly, newsletter of the Society for the study of Artificial intelligence and the Simulation of Behavior, 136, (2013)

Slim, Hugo, *Killing Civilians: Methods, madness and morality in war*, Columbia University, (New York, 2008)

Sparrow, Robert, *Robots and respect: Assessing the case against Autonomous Weapon Systems*, Ethics and International Affairs, (30) 1, (2016)

Strauss, Anselm and Juliet Corbin, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, (Sage Publishing, 1990)

Suchman, Lucy, *Human-Machine Autonomies,* Lancaster University Press, with J Weber, (Paderborn University, 2014)

Suchman, Lucy, *Plans and actions: the problem of human-machine communication,* (Cambridge University Press, 2008)

Taherdoost, Hahmed, *Sampling Methods in Research Methodology: How to Choose a Sampling Technique for Research*, SSRN Electronic Journal, (January 2016)

Talwar, Rohit, and others, *Keeping the Human Touch; humans need a new mindset to function in a tech-dominated society,* People's Technology, Financial Times, (October 2017)

Tashakkori, Abbas and Charles Teddlie, Mixed Mixed Methodology: Combining Qualitative and Quantitative Approaches, Applied Social Research Methods Series, Volume 46, Thousand Oaks, (Sage publications, 1998)

Tonkens, Ryan, *The case against robotic warfare: a response to Arkin,* Journal of Military Ethics, Vol 11, No 2, (August 2012)

Walker, Patrick, W., *Killer Robots? The Role of Autonomous Weapons on the modern battlefield,* , MA thesis, Modern War Studies/Humanities department, (Buckingham University, 2013)

Zawieska, Karolina, *An Ethical Perspective of Autonomous Weapons*, cit. '*Perspectives on Lethal Autonomous Weapons',* UNODA Occasional Papers, Number 30, (November 2017)

Historical perspectives

Adamsky, Dima, *The Culture of Military Innovation: Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the US and Israel*, (US: Stanford University Press, 2010)

Boot, Max, *War Made New; Weapons, and the making of the Modern World,* US: Gotham, 2006)

Bourke, Joanna, *An Intimate History of Killing,* (UK: Basic Books, 1999)

Clark, Lloyd, *Blitzkrieg: Myth, Reality and Hitler's lightning war –France 1940,* (US: Atlantic Books, 2016)

Cohen, Jonathan, *Supreme Command: Soldiers, Statesman, and Leadership in Wartime,* (US: New York, 2002)

Biddle, Tammi D., *Rhetoric and Reality in Air Warfare; the Evolution of British and American Ideas about Strategic Bombing, 1914-1945,* (US: Princeton, 2002)

Drucker, Peter, *The Age of Discontinuity: Guidelines to our changing Society,* 9th printing, (Transaction Publishers, New Brunswick, 2011)

Dupuy, Trevor N., *The Evolution of weapons and of warfare,* (US: Indianapolis, 1980)

Fridman, Ofer, *Revolutions in Military Affairs that did not happen: a framework for analysis,* Comparative Strategy, Volume 35, issue 5

Galdi, Theodor W., *Revolution in Military Affairs? Competing Concepts, Organisational Responses, Outstanding Issues,* US Foreign Affairs and National Defense Division, CRS-95-1170 F, (December 1995)

Gitelman, Lisa, *Always Already New; Media, History and Data of Culture,* (MIT Press, 2006)

Glover, Jonathan, *Humanity: a moral history of the Twentieth Century,* (New Haven: Yale University Press, 2000)

Gray, Colin S., *Strategy and defense planning; meeting the challenge of uncertainty*, (Oxford University Press 2014)

Hacker, Barton C., *Military Institutions, Weapons and social change: Toward a new history of Military Technology,* Society for the History of Technology, (1994)

Handel, Michael, *Masters of War; classical strategic thought*, (UK: Cass Publishers, 1992)

Harari, Yuval Noah, *Homo Deus*, (UK: Penguin Random House, 2016)

Hart, H.L.A. and John Gardner, *Punishment and Responsibility: essays in the philosophy of law*', 2nd Edition, (UK: Oxford University Press, 2008)

Houle, David, *Entering the Shift Age: The End of the Information Age and the New Era of Transformation,* (UK: SourceBooks, 2012)

Howard, Michael, *Clausewitz; a very short introduction*, (UK: Oxford University Press, 1983)

De Jomini, Antoine H., *The Art of War,* (US: Harrisburg, 1947)

Kirkpatrick, K and P Pugh, *Towards the Starship Enterprise: Are the Current Trends in Defence Unit Costs Inexorable?*, Journal of Cost Analyses, 2, 1 (1985)

Luttwak, Edward N., *Give War A Chance*, Foreign Affairs, (1978)

MacIsaac, David, *Voices from the Central Blue: The Air Power Theorists,* cit. P Peret, Makers of Modern Strategy: From Machiavelli to the Nuclear Age, (Oxford: Clarendon Press, 1986)

Admiral Mahan, Alfred Th., *The influence of Sea Power upon the French Revolution and Empire 1793-1812,* volume one, (Boston, 1898)

McKitrick, Jeffrey, J. Blackwell and others, *The Revolution in Military Affairs*, (Exeter, 1995)

Metz, Steven, *Armed Conflict in the Twentieth Century: The Information Revolution and Post-Modern Warfare,* ISBN 1-58487-018-4, (2000)

Mickeviciute, Neringa, *Lessons from the past for weapons of the future*, International Comparative Jurisprudence, 2, (Elsevier, 2016)

Montgomery of Alamein, A *History of Warfare,* (London, 1968)

Lambeth, Benjamin, *NATO's Air War for Kosovo: A Strategic and Operational Assessment,* (Santa Monica, CA, 2001)

Osinga, Frans P.B., *Science, Strategy and War: The strategic theory of John Boyd*, (Routledge, 2006)

Raleigh, Walter and H.A. Jones, *The War in the Air,* London, 7 of 7 volumes, (1932)

Ridd, Thomas, *Rise of the Machines: A Cybernetics history,* (New York: WW Norton, 2016)

Rummel, RJ, *Understanding Conflict and War: War, Power and Peace*, 'Causes and Conditions of International Conflict and War' (USA: Beverley Hills, Sage, 1979)

Strachan, Hew, and A. Herberg-Rothe, *Clausewitz in the Twenty-First Century,* (Oxford University Press, 2007)

Strachan, Hew, and S. Scheipers (eds), *The Changing Character of War'*, (Oxford University Press, 2011)

Strachan, Hew, *The Direction of War – Contemporary strategy in historical perspective,* (Cambridge University Press, 2013)

Tse-tung, Mao, *Selected Military Writings,* (Peking Foreign Language Press, 1971)

Watts, Barry, *The Maturing Revolution in Military Affairs,* Centre for Strategic and Budgetary Assessments, (2011)

Legal perspectives

Article 36, *Killing by a machine: Key issues for understanding meaningful human control,* (April 2015)

Article 36, *Lethal autonomous weapons, artificial intelligence and in meaningful human control,* briefing document with ASU Global Security Initiative, (February 2016)

Abbott, Chris, and others, *Securing Change – Recommendations for the British Government regarding remote-control war*fare, Open Briefing, (June 2015)

Alston, Philip, *Interim report of the Special Rapporteur on extrajudicial, summary or arbitrary executions,* UN Doc A/65/321, (August 23, 2010)

Anderson, Kenneth, and Matthew Waxman, *Law and Ethics for Autonomous Weapon Systems: Why a ban won't work and How the Laws of War Can,* National Security and Law Essay, Hoover Institute, (Stanford University, 2013)

Anderson, Kenneth, and Matthew Waxman, *Debating Autonomous Weapon Systems, Their Ethics and Their Regulation Under International Law,* Research Paper 2017-21, (Washington College of Law, 2017)

Anderson, Kenneth, *Jus ad bellum and in bello: Making the Use of Force Too Easy,* in C Finkelstein, JD Ohlin, eds, Targeted killings: Law and Morality in an Asymmetrical World, (Oxford University Press, 2012)

Beard, Jack, *Autonomous Weapons and Human Responsibility*, University of Nebraska, Georgetown Journal of International Law, Number 617, (2014)

Boddens Hosang, J.F.R., *Rules of Engagement; rules on the use of force as linchpin for the international law of military operations,* UvA-DARE, 8 February, Military Operational Context, (2017)

Boggs, Marion W., *Attempts to define and limit 'aggressive' armaments in Diplomacy and Strategy,* XVI, Number 1, (Columbia, Missouri: University of Missouri, 1941)

Brehm, Maya, *Defending the Boundary; constraints and requirements on the use of autonomous weapon systems under international humanitarian and human rights law',* Geneva Academy of International Humanitarian Law and Human Rights, Academy briefing No.9

Brown, Bernard, *The Proportionality Principle in the Humanitarian Law of Warfare: Recent Efforts at Codification',* Cornell International Law Journal, Volume 10, Issue 1, Article 5, (December 1976)

Droege, Cordula, *The interplay between International Humanitarian Law and Human Rights Law in situations of armed conflict,* Israel Law Review, 40(02), (2007)

Ford, Christopher, *Autonomous Weapons and International Law,* University of South Caroline Law Review, 69.5.Car.413, (11 April 2017)

Gaggioli, Gasteyger and R. Kolb, *A right to life in armed conflict? The contribution of the European Court of Human Rights,* Israel Yearbook on Human Rights, (2007)

Hagmaier, Tonya, *Air Force Operations and the Law; a Guide for Air and Space Forces*, 1st Ed, (Air Force Advocate General's School Press, 2002)

Hayashi, Nobuo, *Contextualizing Military Necessity*, Emory International Law Review, Volume 27, (2013)

Heyns, Christof, *Report of the special rapporteur on extrajudicial, summary or arbitrary executions,* UN Document A/HRC/23/47, (United Nations, 9 April 2013)

Horowitz, Michael and Peter Scharre, *Meaningful human control and weapon systems: a primer (Centre for a new American security),* Working paper, (March 2015)

Human Rights Watch, *Killer Robots and the Concept of Meaningful Human Control,* Memorandum to Convention on Conventional Weapons CCW Delegates, (April 2016)

Human rights Watch and Harvard Law School International Human Rights Law Clinic, 'Reviewing the Record: Reports on Killer Robots from Human Rights Watch an Harvard Law School International Human Rights Law Clinic', *HRW Publishing*, (2018)

Huntington, Samuel P., *Power, Expertise and the Military Profession, (*Daedalus 92, Fall 1963)

ICRC, *Draft rules for the limitation of the dangers incurred by the civilian population in time of war,* Article 14, Draft Rules, (1956)

International Review of the Red Cross, *A guide to the legal review of new weapons, means and methods of warfare: Measures to implement Article 36 of Additional Protocol I of 1977,* Volume 88, number 864, International Committee of the Red Cross, (Geneva, January 2006)

International Committee of the Red Cross (ICRC), *History of Humanitarian Law: The Essential Rules,* (ICRC, 2004)

International Committee of the Red Cross (ICRC), *The Use of Force in Armed Conflict: Interplay between the Conduct of Hostilities and Law Enforcement Paradigms*, (ICRC, November 2013)

International Governance of Autonomous Military Robots, Columbia Science and Technology Law Review, Vol XII, (2011)

International Law Commission, *Articles on the responsibility of States for internationally wrongful acts,* Article 8 and 23(1), UNGA Res 56/83, 12/, (December 2001)

JSP 383, *The Joint Service Manual of the Law of Armed Conflict,* (2004 edition)

Marchant, Gary, R. Atkin and others, *International Governance of Autonomous Military Robots,* Columbia Science and Technology Law Review, Vol XII, (2011)

Marauhn, Thilo, *An analysis of the potential impact of lethal autonomous weapons systems on responsibility and accountability for violations of international law, Presentation, CCW Meeting of experts on all the full autonomous weapons systems,* (Geneva, May 2014)

Melzer, Nils, *Targeted Killing in International Law,* (Oxford University Press, 2008)

Palin, Roger, *Multinational Military Forces: Problems and Prospects*, Adelphi Paper 294, (Routledge, 2005)

Reinold, Theresa, *Sovereignty and the responsibility to protect,* Routledge Advances in International Relations and Global Politics

Robinson, Darryl, *How Command Responsibility Got So Complicated: A Culpability Contradiction, its Obfuscation, and a Simple Solution*, 13, 1, (Australia: Melbourne Journal of International Law, 2012)

Sandoz, Yves, *Commentary on the Additional Protocols of 8th June 1977 to the Geneva Conventions of 12 August 1949,* Martinus Nijhoff Publishers, (Geneva, 1987)

Sassen, Saskia, *The Participation of States and Citizens in Global Governance,* Indiana Journal of Global Legal Studies, Volume 10, Issue 1, (2003)

Schaub, Gary, *Civilian Combatants, Military Professionals,* Defence Studies, Journal of the Joint Services Command and Staff College, Number 3, Volume 10, (September 2010)

Smith, Ron, *Military Economics: The interaction of Power and Money,* (Palgrave Macmillan, 2009)

UK Government, *The Joint Service Manual of the Law of Armed Conflict,* Joint Services Publication 383, (2004 Edition)

United Nations, *The Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims on International Armed Conflicts (Protocol 1)*, 1125 UNTS 3, (adopted 8 June 1977)

Wallach, Wendell, *Towards a Ban on Lethal Autonomous Weapons: Surmounting the Obstacles,* Communications of the ACM, Volume 60, No 5, (April 2016)

Operational perspectives

Adams, Thomas, *Future warfare and the decline of human decision-making,* US Army War College Quarterly, (Winter 2001-2002)

Antonova, Albena, *Institutional and Organisational Transformation in the Robotic Era*, IGI Global, (August 2018)

Altmann, Jurgen, *Preventive Arms Control for Uninhabited Military Vehicles,* Experimentelle Physik III, (Technische Universitat Dortmund, 2004)

Altmann, Jurgen, and Frank Sauer, *Autonomous Weapon Systems and Strategic Stability,* Survival, Volume 59, Number 5, (November 2017)

Department of the Army, *Field Manual,* FM 27-10, Change 1, The law of land warfare, (Washington DC: Government Printing Office, July 1956)

Bahnsen, John, and Robert Cone, *Defining the American Warrior Leader*, (Parameters, December 1990)

Bawden, David, *The nature of prediction and the information future: Arthur C. Clarke's Odyssey vision,* Aslib Proceedings, vol.49, no.3, (March 1997)

Bhuta, Nehal, and others, *Autonomous Weapons Systems: law, ethics, policy,* (Cambridge University Press, 2016)

Bishop, Ryan and John Phillips, *Unmanning the Homeland*, International Journal of Urban and Regional Research, 26.3., (2002)

Black, Jeremy, *War in the New Century,* (London, 2001)

Black, Jeremy, *War in the Western World, 1882-1975,* (Chesham, UK, 2002)

Black, Jeremy, *War and Technology,* (Indiana University Press, 2013)

Bolton, Matthew, and Wim Zwijnenburg, *Future-proofing is never complete: ensuring the arms trade treaty keeps pace with new weapons technology,* International Committee for Robot Arms Control, Working paper #1, (October 2013)

Borrie, John, *Security, unintentional risk and system accidents,* Chief of Research, UNIDIR briefing note, (Geneva, 15 April 2016)

Bousquest, Antoine, *A Revolution in Military Affairs? Changing technologies and Changing Practices of Warfare,* Technology and World Politics: An Introduction, (Routledge, 2017)

Callam, Andrew, *Drone Wars: Armed Unmanned Aerial Vehicles,* International Affairs review, Volume XVIII, No 3, (Elliott School of International Affairs, George Washington University 2016)

Cantwell, Houston, *Beyond butterflies: Predator and the evolution of unmanned aerial vehicles in Air Force culture,* (School of Advanced Air and Space Studies: Maxwell Air Force Base, AL, 2007)

Catignani, Sergio, *Coping with Knowledge: Organizational Learning in the British Army,* Journal of Strategic Studies, (8 May 2013)

Centre for Army Leadership, *Army Leadership Doctrine,* Edition 1, (RMAS Camberley, 2016)

Cole, Chris, *Towards the Next Defence and Security Review,* Submission from Drone Wars UK to the Defence Select Committee Inquiry on the use of armed Unmanned Aerial Vehicles, (April 2013)

Congressional Research Service, *Hearing on 'The rise of drones; unmanned systems and the future of war,* Committee on Oversight and Government Reform, (March 2010)

De Czege, Huba W., *Six compelling ideas on the rude to a future army,* Army Magazine, Vol 51, no.2, (2015)

Van Creveld, Martin, *Command in War,* (Harvard University Press, 1985)

Cummings, Mary L., *Automation and accountability in decision support system interface design,* Journal of Technology Studies, 32, (2006)

Department of Defense, *Unmanned Systems Integrated Roadmap FY2013-2038,* Reference 14-S-0553, (November 2013)

US Department of Defense, *DoD Dictionary of Military and Associated Terms,* (June 2017)

Department of Defense, *Department of defense Law of War Manual*, (June 2015)

Department of Defense Directive, *Autonomy in Weapon Systems*, Number 3000.09, (21 November 2012)

The Department of the Army, *Field manual 101-5, staff organisation and operations*, Headquarters, (Washington, 1998)

Directorate Land Warfare*, Operation HERRICK Campaign Study*, (AD LXC, Warminster, 2015)

Doare, Ronan, and others, *Robots on the Battlefield: Contemporary perspectives and implications for the future,* (USA: Army Combined Arms Center, Fort Leavenworth, KS Combined Studies Institute, 2014)

Drozdova, Katya, *Low-tech threats in the Hi-tech Age: Subversive networks across ideologies, technologies and times,* Analytical Perspectives on Politics Series, (University of Michigan Press, 2002)

Dupuy, Trevor N., *Numbers, predictions and war: using a history to evaluate combat factors and predict the outcome of battles,* (Bobbs Merrill Publishing, 1979)

Endsley, Mica R. and Debra G. Jones, *Designing for situation awareness: An approach to human-centered design,* 2nd edition, (London: Taylor and Francis, 2012)

Endsley, Mica R., *Designing for Situational Awareness*, (Boca Raton: CRC Press, 19 April 2016)

Ferguson, Brian, *Ten Points on War,* (Social Analysis, Volume 52, Issue 2, 2008)

Freedman, Lawrence, *Information warfare: Will battle ever be joined?,* (International Centre for Security Analysis, October 1996)

Foot, Rosemary, *Constraints on Conflict in the Asia-Pacific: Balancing the 'war Ledger',* (Political Science, Volume 66, Issue 2, 2014)

Galliott, Jai, and Mianna Lota, *Super Soldier: the ethical, legal and social implications,* (Political Science, Routledge, 3 March 2016)

Goldman, Emily O., and Leslie C. Eliason, *The diffusion of military technology and ideas,* (Stanford CAE, 2003)

Gott, Kendall and Michael Brooks, *Warfare in the Age of Non-State Actors*, (Kansas: Combat Studies Institute Press, 2007)

Gray, Colin S., *Modern Strategy,* (Oxford University Press, 1999)

Gray, Colin S., *The Future of Strategy,* (Polity Books, 2015)

Gray, Colin S., *Another Bloody Century,* (Phoenix, 2005)

Gray, Colin S., *Strategy and Defence Planning: Meeting the Challenge of Uncertainty,* (Oxford University Press, 2014)

Grinker, Roy and John Spiegal, *Men Under Stress*, (US: Philadelphia Press, 1945)

Halperin, Morton H., *Nuclear weapons and limited War,* The Journal of conflict resolution, (5 June 1961)

Hambling, David, *Swarm Troopers: How small drones will conquer the World,* (Archangel Ink, 2015)

O'Hanlon, Michael, *The Science of War*, (US: Princeton University Press, 2009)

Hanon, Leighton, *Robots on the Battlefield – are we ready for them?,* 'Unmanned Unlimited' Technical Conference, (Chicago: American Institute of Aeronautics and Astronautics, 2008)

Hickok, William, *Defining War in Twenty-first Century America*, School of Advanced Military Studies, (Fort Leavenworth, 2010)

Houses of Parliament Postnotes, *Automation in Military Operations,* Number 511, (October 2015)

Lorber, Azriel, *Misguided Weapons: Technical Failures and Surprise on the Battlefield,* (Potomac Books, 2002)

McClumpha, A. and M. James, *Understanding automated aircraft; human performance in automated systems,* Current research and trends, (Hillsdale, NJ, 1994)

UK Ministry of Defence, *Strategic Trends Programme, Future Operating Environment 2035,* (Crown Copyright, August 2015)

Johnson, Rob, and Michael Whitby, *How to win on the battlefield,* (London: Thames and Hudson, 2010)

Baron de Jomini, Antoine H., *The Art of War*, (USA: Rockville, MD, 2006)

Kassan, Peter, *AI gone awry: futile quest for artificial intelligence,* (The Skeptics Society and Skeptic Magazine, 2016)

Kott, Alexander, and others, *Decision Aids for adversarial planning in military operations for: algorithms, tools, and Turing-test-like experimental validation,* (Cornell University, 2015)

Kreps, Sarah, and Micah Zenko, *The Next Drone Wars: Preparing for Proliferation,* (Foreign Affairs magazine, April 2014)

Latiff, Robert, *Future War: Preparing for the New Global Battlefield*, (Knopf Doubleday Publishing, September 2017)

Levy, Jack, *The offensive/ defensive balance of military technology: A theoretical and historical analysis,* International Studies Quarterly, Volume 28, Number 2, (June 1984)

Lombardi, Ben, *Assumptions and Grand Strategy,* Ben Lombardi, (Spring 2011)

Lovelace, Douglas, *Autonomous and Semi-Autonomous Weapon Systems,* (Law Review, Volume 44, Oxford University Press, 6 October 2016)

Moore, Craig and others, *Measuring Military Readiness and Sustainability*, (National Defense Research Institute, RAND Publishing, Santa Monica, 1991)

Mosser, Michael, *Puzzles versus Problems: The Alleged Disconnect between Academics and Military Practitioners*, (Reflections, Volume 8, Number 4, December 2010)

Moulton, J.L., *Defence Planning: The Uncertainty Factor,* (Long Range Planning Journal, Volume 3, Issue 4, 1971)

Demir, Mustafa, and others, *Team Synchrony in Human-Autonomy Teaming*, (International Conference on Applied Human Factors and Ergonomics, January 2018)

Neal, P.J., *From Unique Needs to Modular Platforms: The Future of Military Robotics,* (US Naval Institute, June 2010)

The Nilson Report, *US General Purpose Cards,* Midyear 2015, Issue 1069, (August 2015)

Osinga, Frans P.B., *Targeting: The challenges of Modern Warfare*, eds. Paul Ducheine and Michael Schmitt, (US: Asser Press/Springer, 3 November 2015)

Parasuraman, Raja, and others, *Performance consequences of automation-induced 'complacency,* (International Journal of aviation psychology, 3, 1, 1993)

Postnote, *Automation in Military Operations*, Number 511, (UK Houses of Parliament, Parliamentary Office of Science and Technology, October 2015)

Ramo, Simon, *Let Robots Do The Dying,* (Kindle publication, 2011)

Reason, James, *Human Error,* (Cambridge University press, 1990)

Riza, Shane, *Killing without Heart,* (University of Nebraska Press, Potomac Books, 2013)

Van Riper, Paul, and Robert Scales, *Preparing for war in the 21st Century,* (Parameters, 27/3, 1997)

Roff, Heather, *The Strategic Robot Problem: Lethal Autonomous Weapons in War,* (Journal of Military Ethics, Volume 13, Issue 3, 17 November 2014)

Roff, Heather, *Meaningful human control or appropriate human judgement? The necessary limits on autonomous weapons,* Briefing paper for delegates at the review conference of the Convention on Certain Conventional Weapons, (Geneva, 16 December 2016)

Royal Air Force Directorate of Defence Studies, *Air Power – UAVs: The wider context*, (AFDDS, Autumn 2012)

Rubenstein, Robert A., and others, *Practicing military anthropology: Beyond expectations and traditional boundaries, (*Sterling VA: Kumarian Press, 2013)

Scharre, Paul, *Autonomous weapons and operational risk,* (Centre for a New American Security, 2016)

Singer, Peter W., *Robots at war: the new battlefield,* (Wilson Quarterly, 2008)

Singer, Peter W., *Wired for War: The Robotics Revolution and Conflict in the Twenty First Century,* (Penguin, 2006)

Siniscalchi, Joseph, *Non-lethal Technologies: Implications for Military Strategy*, Occasional Papers Number 3, (Maxwell: Air University, March 1998)

Smith, Edward A., *Network-centric Warfare: What is the point?*, (US: Naval War College, Winter 2011)

Smith, Ron, *Military Economics: the interaction of power and money,* (Basingstoke: Palgrave MacMillan, 2009)

Sparrow Robert, *Killer Robots,* (Journal of Applied Philosophy, 2007)

Storr, Jim, *The Human Face of War,* (A&C Black, July 2009)

Suchman, Lucy, *Situational awareness and adherence to the principle of distinction as a necessary condition for lawful autonomy,* Panel 'Towards a Working Definition of LAWS', CCW Informal Meeting of Experts on lethal autonomous weapons, (Geneva, 12 April 2016)

Surgeon General's Office, *Final Report, Mental Health Advisory Team,* MHAT, IV Operation Iraqi Freedom 05-07, (November 2006)

Swank, Roy, *Combat neuroses: Development of Combat Exhaustion,* (Archives of Neurology and Psychology, 55, 1946)

Taylor Vinters LLP, *Unmanned Aerial Vehicles*, Qi3 Insight, Qi3 Ltd, (February 2014)

Tilford, Earl, R*eviewing the Future,* (Parameters, 2002)

Warwick, Kevin, *March of the Machines: Why the New Race of Robots Will Rule the World,* (London: Century, 1997)

Wheeler, Nicholas J., and Ken Booth, *The security dilemma: Fear, cooperation and trust in world politics,* (Basingstoke: Palgrave MacMillan, 2008)

Wolff, Terry, *The Operational Commander and Dealing with Uncertainty*, (US: Army Command and General Staff, Fort Leavensworth, 19 April 1999)

UK Army, *Tactical Aide Memoire,* (TAM Part 2, Issue 3.0, 1998)

UK Army Doctrine Publication, '*Army Doctrine Primer*', (AC 71954, May 2011)

UK MOD, *Human Machine Touchpoints: The United Kingdom's perspective on human control over weapon development and targeting cycles*, UK submission to CCW GGE on LAWS, (August 2018)

US Air Force, *Autonomous horizons; system autonomy in the air force – a path to the future – human-autonomy teaming,* (Office of the Chief Scientist, AF/ST TR 15-01, 9 June 2015)

US Department of Defense, *Summer study on autonomy,* (Defense Science Board, Office for Acquisition, Technology and Logistics, Washington, June 2016)

US Army, *Serving a Nation at War,* (Washington DC, 2004)

Vergouw, Bas, and others, *Drone technology: Types, Payloads, Applications, Frequency Spectrum Issues and Future Developments,* (Information Technology and Law Series 27, Springer, ed. B Custers, undated)

Walker, Paddy, *Killer Robots? The Role of Autonomous Weapons on the modern battlefield,* MA thesis, Modern War Studies/Humanities department, (Buckingham University, 2013)

Work, Robert O., and Shawn Brimley, *20YY; Preparing for War in the Robotic Age,* (Centre for a New American Security, www.cnas.org, January 2014)

Technical perspectives

Arkin, Ronald, *Governing Lethal Behaviour in Autonomous Robots,* (Boca Raton, Fl: CRC Press, 2009)

Arkin, Ronald, *Governing Lethal Behaviour: Embedding ethics in a hybrid deliberate/reactive robot architecture,* (Atlanta, GA: Georgia Institute of Technology, 2007)

Arquilla, John, and David Ronfeldt, *Swarming and the future of conflict,* RAND/D8-311-OSD, (Santa Monica CA, 2000)

Article 36, *Killing by machine; key issues understanding for meaningful human control,* (April 2015)

Article 36, *Key elements of meaningful human control,* background paper to UN CCW Debates, (April 2016)

Asaro, Peter, *Cybernetics and autonomous weapons: reflections and responses,* (Paragigmi: Rivista di critica filosofica, XXXIII, number 3, 2015)

Backstrom, Alan, and Ian Henderson, *New Technology and Warfare*, cit. International Review of the Red Cross, Volume 94, Number 886, (Summer 2012)

Beautement, Patrick, *Putting complexity to work; achieving effective human-machine teaming,* (The Abaci Partnership LLP, 2015)

Bettini. Claudio and others, 'A Survey of Context Modelling and Reasoning Techniques', *Elsevier*, (27 March 2008), <http://ssltest.cs.umd.edu/class/spring2013/cmsc818g/files/bettinisurvey.pdf>

Bhanu, Bir, and Terry Jones, *Image Understanding Research for Automatic Target Recognition,* (Carnegie Mellon Laboratories, January 2010)

Biundot, Susanne, and others, *Companion-technology: An Overview*, (Kl-Kunstliche Intelligenz 30.1, 2016)

Blake, Andrew, and Andrew Zisserman, *Visual Reconstruction,* (USA: MIT Press, 1987)

Blake, Andrew, and Andrew Zisserman, *Localising discontinuities using weak continuity constraints,* (Pattern Recognition Letter, 6, 1, June 1987)

Bodel, Margaret, *Creativity and Artificial Intelligence,* (School of Cognitive and Computer Science, Brighton, in Artificial Intelligence, 103, 1998)

Bostrom, Nick, *Superintelligence; Data, dangers, strategies,* (Oxford University Press, 2014)

Bostrom, Nick, and Carl Shulman, *How hard is Artificial Intelligence? Evolutionary arguments and selection effects,* (Journal of Consciousness Studies, Vol 19, 7-8, 2012)

Braga, Adriana, and Robert Logan, *The Emperor of Strong AI Has No Clothes: Limits to Artificial Intelligence*, (Information, 8, 156, undated)

Brodley, Carla and others, *Challenges and Opportunities in Applied Machine Learning*, (Association for the Advancement of Artificial Intelligence, AI magazine, 33, 2012)

Caccia, Massimo, and others, *Basic navigation, guidance and control of an unmanned surface vehicles,* (Autonomous Robots, 25, 4, Springer US)

Chalmers, David J., *Facing up to the problem of consciousness,* (Journal of Consciousness Studies, 2, 3, 1995)

Chan, Serena, *Complex Adaptive Systems,* ESD.83 (Research Seminar, Engineering Systems, October 2001)

Chapman, Gretchen, and Eric Johnson, *Anchoring, Activation and the Construction of Values*, (Organizational Behaviour and Human Decision Processes, 79, 2, August 1999)

Coker, Christopher, *Still "the human thing"? Technology, human agency and the future of war*, (International Relations, 32.1, 2018)

van Creveld, Martin, *Technology and War: From 2000 BC to the Present Day,* (Simon & Schuster, 11 May 2010)

Cui, K., and others, *The Collaborative Autonomy and Control Framework for Unmanned Surface Vehicles,* (Frontier of Computer Science and Technology, 9th International Conference materials, IEEE publication, 2015)

US Department of Defense, *The Role of Autonomy in DoD Systems,* Task Force Report, (April 2012)

Dietterich, Thomas, *Learning and reasoning,* Department of Computer Science, (Oregon State University, May 2003)

Duran, Boris, and Serge Thill, *Rob's Robot: Current and Future Challenges for Humanoid Robots*, (The Future of Humanoid Robot Research and Application, Intech, 2012)

Ekelhof, Merel, *Human control in the targeting process,* ed. R. Geiß, *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law and Security*, (Berlin: German Federal Foreign Office, 2016)

O'Halloran, James C., and others, *Jane's Land-based Air Defence, 2010*-2011, (IHS Jane's: Coulsdon, 2010)

Hollnagel, Erik, and David D. Woods, *Joint Cognitive Systems,* (New York: Basic Books, 2005)

Galahmar-Zavare, Alireza, and others, *High Efficiency, Low Size and Low weight Vehicle Battery Chargers,* (Power, Electronics, Drive Systems and Technologies Conference, PEDSTC, IEEE, 2015)

Gomez, Erik, *Map-building and Planning for Autonomous Navigation of a Mobile Robot,* (Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico, January 2015)

Gosavi, Abhijit, *The Effect of Noise on Artificial Intelligence and Meta-heuristic Techniques,* (Proceedings of the Artificial Neural Network in Engineering Conference, Vol. 12, 2002)

Grandon Gill, T., *A Psychologically Plausible Goal-Based Utility Function,* Informing Science: The International Journal of an Emerging Transdiscipline, Volume 11, (2008)

Haikonnen, Pentti, *The cognitive approach to conscious machines,* (UK Imprint academic, 2003)

Harnandez, Mauricio A., and Salvatore J. Stolph, *Real World Data is Dirty: Data Cleansing and the Merge/Purge Problem,* Data Mining and Knowledge Discovery, 2:9, (1998)

Haykin, Simon, *Neural Networks and Learning Machines*, 3rd Edition, (Person Prentice Hall, 1999)

Henry, Winston P., *Artificial Intelligence,* (Reading: Addison-Wesley Publishing, MASS, 1984)

Hoffman, Donald D., *Visual Intelligence; how we create what we see,* (University of California, Irvine, 1998)

Hon, Adrian, *A History of the Future in 100 Objects, (*Amazon/Kickstarter, 2016)

Housien, Hamed I., and others, *A comparison study of data scrubbing algorithms and frameworks in data warehousing,* Central Southern University, Changsa, China, International Journal of computer applications, (0975-8887) Volume 68, No 25, (April 2013)

Hurston, Michael, *Even Artificial Intelligence can acquire Biases against Race and Gender,* Science Magazine, Science AAAS, 13 (April 2017)

Ishikawa, Matsumi, *Structural Learning with Forgetting,* (Neural Networks, Volume 9, Issue 3, April 1996)

International Panel on the Regulation of Autonomous Weapons (IPRAW), *Executive Summary Number 2,* Computational Systems in the Context of Autonomous Weapon Systems, (November 2017)

James, William, *Principles of Psychology,* (US: Henry Holt, NY, Volume 1, 1890)

JASON program, *Perspectives on research in artificial intelligence and artificial general intelligence relevant to DoD,* JSR-16-Task-003, (January 2017)

Jeong, Doo Seok, and others, *Towards Artificial Neurons and Synapses: A materials point of view,* (RSC Publishing, DOI 10.1039/c2ra22507g)

Kam, Moshe, *Sensor Fusion for mobile robot navigation,* (Proceedings of IEEE, Volume 85, Issue 1)

Kenal, L., ed., and others, *Uncertainty in Artificial Intelligence, (*Elsevier, June 2014)

Kolbert, Elizabeth, *Our Automated Future,* The New Yorker Magazine, (19 December 2016)

Kulasekere, Ernest C., and others, *Conditioning and Updating Evidence,* (International Journal of Approximate Reasoning, 36, 2004)

Levi, Dennis M., *Crowding – an essential bottleneck for object recognition,* (ScienceDirect, Vision Research 48, Elesevier, 2008)

Li, Xiaodong, and Leonard J. Cimini, *Effects of Clipping and Filtering,* (IEEE Communications Letters, 2, 5, May 1998)

Li, Xiaodong, and others, *Context Aware Middleware Architecture: Surveys and Challenges,* (Sensors, 15, 8, 2015)

Lopez, Beatrice, and others, *Modeling decisions for artificial Intelligence,* Ninth International Conference, MDAI 2012, (Girona, Spain, November 2012)

CG Lord and others, 'Human decision makers and automated decision aids: made for each other?' (RE Parasuramen, ed, *Automation and human performance: Theory and applications*', (USA: Mahwah, NJ, Laurence Erlbaum Associates, 1996)

Luo, Ren C., and Michael J. Kay, *Multi-sensor Integration and fusion for Intelligent Machines and Systems,* North Carolina University, (Ablex Publishing, 1995)

Marra, William, and Sonia McNeil, A*utomation and Autonomy in Advanced Machines: Understanding and Regulating Complex Systems,* Lawfare Research Paper Series 1-2012, (April 2012)

Marsland, Stephen, *Using Habituation in Machine Learning*, Neurobiology of Learning and Memory, 92,2, (2009)

Marvala, Tshilidshi, *Causality, correlation and artificial intelligence for rational decision-making,* (University of Johannesburg, Word Scientific, March 2015)

Medsker, Larry R., and L.C. Jain, *Recurrent Neural Network Design and Application,* (CRC Press International Series on Computational Intelligence, 20 December 1999)

McDermott, Drew, *Artificial Intelligence and Consciousness,* (Cambridge Handbook of Consciousness, Cambridge University Press, 2007)

Dietterich, Thomas, and Ryszard S. Michalski, *A Comparative Review of Selected Methods for Learning from Examples*, in *An Artificial Intelligence Approach*, RS Michalski and others (eds), (Paulo Alto Tioga Publishing, 1983)

Moyes, Richard, *Emergent behaviour and Risk – a sketch for a risk management approach,* Article 36, (April 2017)

Muraven, Mark, *Goal Conflict in Designing Autonomous Artificial Systems,* (University of Albany, March 2017)

Nemati, Hamid, and others, *Knowledge management, decision support, artificial intelligence and data warehousing,* (Decision Support Systems, 33, 2002)

Neukart, Florian, *A Machine Learning Approach for Abstraction based on the idea of Deep Learning Belief Artificial Neural Networks,* (24th DAAAM International Symposium of International Manufacturing and Automation, 2013)

Nicolescu, Monica, and others, *Learning Behaviour Fusion from Demonstration*, (Interaction Studies, 9.2, 2008)

Niranjan, P., *Software and Hardware for Autonomous Robots Using Distributed Embedded System*, (International Journal of Computer Applications, Volume 55, Number 11, October 2012)

Nof, Shimon Y., *Handbook of Industrial Robotics,* (John Wiley & Sons, 1999)

Norman, Donald, *Things That Make US Smart: Defending Human Attributes in the Age of Machines,* (Diversion Books, 2 December 2014)

Nwana, Hyacinth, *Software Agents: An Overview,* (Knowledge Engineering Review, II, 3, October 1996)

Pendleton, Scott D., and others, *Perception, Planning, Control, and Coordination for Autonomous Vehicles*, (MDPI, Machines, 5.1, 6, 2017)

Pennachin, Cassio, and Ben Goertzel, *Contemporary Approaches to Artificial General Intelligence,* AGIRI, XVI 509-p, (Springer Publishing, 2007)

Pickering, Andrew, *Cybernetics and the Mangle: Ashby, Beer and Pask,* (University of Illinois, Department of Sociology, March 2002)

Pollock, John L., *Thinking about acting: Logical foundations for rational decision-making,* (Oxford University Press, 2006)

Preiss, Mark, and Nicholas Stacy, *Coherent Change Detection: Theoretical Description and Experimental Results,* (Australian Department of Defence, DSTO-TR-1851, Edinburgh, 2006)

Ranky, Paul, *A summary of robot test method with examples,* RSD-1-87, (University of Michigan, 1987)

Rasmussen, Jens, *Skills, Rules and Knowledge: Signals, signs and symbols, and other distinctions in human performance models,* (IEEE Transactions on Systems, Man and Cybernetics, 13(3), 1983)

Ratches, James A., *Review of current aided/automatic target acquisition technology for military target acquisition tasks,* (Optical Engineering, 50.7, 072001, 2011)

Reddy, Raj, *Foundations and Grand Challenges of Artificial Intelligence,* AI Magazine, Vol 9, No 4, (Winter 1988)

Reyes, Arthur, and others, *Overview of the University of Texas and Arlington's Autonomous Vehicles Laboratory*, (Department of Computer Science and Engineering, Technical Report CSE-2003-13)

J Robinson, *Accelerating AI,* Northwestern University Law Review, (Colloquy, 2010)

Roff, Heather, and Richard Moyes, *Meaningful Human Control, Artificial Intelligence and Autonomous Weapons,* Briefing Paper prepared for the Informal Meeting of Experts on Lethal Autonomous Weapons, UN Convention for Conventional Weapons, (April 2016)

Russell, Bertrand, *The Philosophy of Logical Atomism,* The Collected Papers of Bertrand Russell, (Boston: Allen & Urwin, 1986)

Russell, Stuart, and Peter Norvig, *Artificial Intelligence: A modern approach,* (Upper Saddle River, NJ: Prentice Hall 2010)

Russell, Stuart, and others, *Human Information Interaction, Artificial Intelligence and Errors,* US Army Research Laboratories, Association for the advancement of AI, 2016

Sandberg, Anders, *Feasibility of Whole Brain Emulation,* Future of Humanity Institute, Theory and Philosophy of Artificial Intelligence, SAPERE, (Berlin: Springer, 2013)

Sanna, Andrea, *Advances in Target Detection and Tracking in Forward-Looking Infrared Imagery,* (Sensors, ISSN 1424-8220, 2014)

Scharre, Paul, *Robotics on the Battlefield part II: The coming swarm,* (Center for a New American Security, October 2014)

Schuppli, Susan, *Deadly Algorithms,* (Continent, Issue 4.4, 2015)

Serfarty, Daniel, and others, *Adaption to Stress in Team Decision-making and Coordination*, (Proceedings of the Human and Ergonomics Society, 37, 1 October 1993)

Shanahan, Murray, *The frame problem,* (MIT Press, February 1997)

Siddiqi, Abdul Ahad, *Implications of using Artificial Intelligence Technology in Modern Warfare,* (ICCIT, 2012)

Simon, Mark, and others, *The Relationship Between Over-Confidence and the Introduction of Risky Products,* (The Academy of Management Journal, 46.2, April 2003)

Singh, Shailesh, and others, *Training of Artificial Neural Networks using Information-Rich Data,*)(Hydrology, 1(1), 2014)

Smith, P.J. and others, *Brittleness in the design of cooperative problem-solving systems: The effects on user performanc*e, (IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 27(3), 1997)

Sombe, Lea, *Reasoning under incomplete information in Artificial Intelligence,* (International Journal of Intelligent Systems, Vol. 5, Issue 470, September 1990)

Stanford Encyclopedia of Philosophy, *The Computational Theory of Mind,* (October 2015)

Stonebroker, Mike, and others, *The 8 requirements of real-time stream processing,* (US: MIT/Brown, 2008)

Tyugu, Enn, *Situational Awareness and control errors of cyber weapons,* (Cognitive Methods in Situational Awareness and Decision Support, IEEE International Multi-disciplinary Conference, February 2013)

Vas, Peter, *Artificial-intelligence based Electrical Machines and Drives,* (Oxford Science Publications, 1999)

Verly, Jacques G. and others, *Machine intelligence technology for automatic target recognition,* (The Lincoln Laboratory Journal, 2, 2, 2009)

Wallgrun, Jan O., *Spatial Representation and Reasoning for Mobile Robots,* DOI 10.1007/978-3-642-10345-2_2, (Springer, 2010)

Warwick, Kevin, *Artificial intelligence; the basics,* (Routledge, 2012)

Waibel, B. and others, *Theory and Experiments on the stability of robot compliance model*s, (IEEE, Transactions on Robotics, 7, 1)

Wigfield, Allan, *Expectancy Value Theory of Achievement Motivation; a developmental perspective*, (Educational Psychology Review, Volume 6, Number 1, 1994)

Zacarias, Alejandaro, and others, *A Framework to Guide the Selection and Configuration of Machine-Learning-Based Data Analytics Solutions in Manufacturing*, (Proceedings CIRP 72, 2018)

Zhou, Jun, *Machine Learning Challenges and Impact: An interview with Thomas Dietterich*, (National Science Review, 5, 1, January 2018)

On-line sources and articles

Abdulhafiz, Waleed A., 'Handling Data Uncertainty and Inconsistency Using Multi-sensor Data Fusion', Siemens, Cairo, Academic Paper, (27 May 2013), <https://www.hindawi.com/journals/aai/2013/241260/>

Abraham, Ajith, 'Hybrid Intelligent Systems: Evolving Intelligence in Hierarchical Layers' Do Smart Adaptive Systems Exist?', *Soft Computing Magazine*, (2005), <http://www.softcomputing.net/gabrys.pdf>

Ackerman, Evan, 'How Drive AI is mastering autonomous driving with deep learning', *IEEE Spectrum*, (11 March 2017), <http://spectrum.ieee.org/cars-that-think/transportation/self-driving/how-driveai-is-mastering-autonomous-driving-with-deep-learning>

Ackerman, Evan, 'Lethal Microdrones, Dystopian Futures and the autonomous Weapon Debate', *IEEE Spectrum*, (15 November 2017), <https://spectrum.ieee.org/automaton/robotics/military-robots/lethal-microdrones-dystopian-futures-and-the-autonomous-weapons-debate>

Ackerman, Evan, 'Algorithms allow Micro Air Vehicles to avoid obstacles with single camera and neuromorphic hardware', *IEEE Spectrum*, (6 November 2012), <http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/algorithms-allow-mavs-to-avoid-obstacles-with-single-camera>

Adams, Norman H., and others, 'Autonomous Loop Switching: Interpreting and modifying the internal state of feedback tracking loops', Aerospace conference literature, *IEEE*, (February 2012), <http://ieeexplore.ieee.org/document/6187143/>

Adams, Thomas, 'Future Warfare and the Decline in Human Decision Making', *Parameters*, (2001-02), <http://ssi.armywarcollege.edu/pubs/parameters/articles/2011winter/adams.pdf>

Agarwal, Aman, 'Explained Simply: How DeepMinds Taught AI to Play Video Games', *Medium.org blog*, (27 August 2017), <https://medium.freecodecamp.org/explained-simply-how-deepmind-taught-ai-to-play-video-games-9eb5f38c89ee>

Age of Lucidity, 'Precursors to fully autonomous weapons', <http://ageoflucidity.info/2013/06/01/other-precursors-to-fully-autonomous-weapons-robots-and-killer-drones/>

Ahrens, Ramon, 'Mission Control in a Communications Denied Environment', Air War College, (16 February 2011), <http://www.dtic.mil/dtic/tr/fulltext/u2/1036912.pdf>

Aiciel blog, 'Does Using Machine Learning Mean More Layers of Complexity in Scientific Study?', September 2016, <https://aicial.com/blog/does-using-machine-learning-mean-dealing-with-more-layers-of-complexity-in-scientific-study>

AIP, 'Defense Department Reorganization Aims To Foster "Culture of Innovation', American Institute of Physics, (10 August 2017), <https://www.aip.org/fyi/2017/defense-department-reorganization-aims-foster-'culture-innovation'>

Air Force Technology, 'RQ-1/MQ-1/MQ-9 Reaper UAV', http://www.airforce-technology.com/projects/predator-uav/

Air and Space Power Mentoring Guide, 'Three Levels of War', Volume 1, Air University Press, (1997), <https://www.cc.gatech.edu/~tpilsch/INTA4803TP/Articles/Three%20Levels%20of%20War=CADRE-excerpt.pdf>

Akandji-Kombe, Jean-Francios, 'Positive Obligations Under European Conventions on Human Rights', Directorate General of the Human Rights Council of Europe, Strasbourg, (January 2007), <https://www.echr.coe.int/LibraryDocs/dg2/hrhand/dg2-en-hrhand-07(2007).pdf>

Alain, Guillaume, and Yoshua Bengio, 'Understanding Intermediate Layers', *ICLR Paper*, (2017), <https://pdfs.semanticscholar.org/2706/77b5c44ea0c93313f41db2f885fef305bbcc.pdf>

Alami, Rachid, and others, 'Toward Human-Aware Robot Task Planning', (LAAS-CNRS, American Association for Artificial Intelligence, (2006), <http://www.aaai.org/Papers/Symposia/Spring/2006/SS-06-07/SS06-07-006.pdf>

Apthorp, Claire, 'Using Autonomy To Supply the 'Last Mile', *Army Technology*, (25 June 2017), <http://www.army-technology.com/features/featureusing-autonomy-to-supply-the-last-mile-5852408/>

Altinkemer, Kemal, and others, 'Vulnerabilities and Patches of Open Source Software: An empirical study', *Journal of Information System Security*, 4.2, (2008), <https://www.krannert.purdue.edu/academics/mis/workshop/papers/ars_092305.pdf>

Allchin, Douglas, 'Error types', *Perspectives on Science*, 9: 38-59, <http://members.tcq.net/allchin/papers/e-types.pdf>

Amando, Nuno, and others, 'Exploiting Parallelism in Decision Tree Induction', Proceedings for the ECML/PKDD Workshop on Parallelism and Distributed Computing for Machine Learning, *LIACC*, (2003), <https://www.dcc.fc.up.pt/~fds/FdsPapers/w2003_ECMLW7_namado.pdf>

Anderson, Kenneth, and others, 'Adapting the Law of Armed `conflict to Autonomous Weapon Systems', *International Law Studies,* Volume 90, Number 386 <http://www.dtic.mil/dtic/tr/fulltext/u2/a613290.pdf>

Andersson, Simon, 'Unsolved Problems in Artificial Intelligence', AI Roadmap Institution, *@goodAI blog,* (3 February 2017), <https://medium.com/ai-roadmap-institute/unsolved-problems-in-ai-38f4ce18921d>

Angelov, Plamen, and Alessandro Sperduti, 'Challenges in Deep Learning', Proceedings of the European Symposium on Artificial Neural Networks, Bruges, (April 2016), <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2016-23.pdf>

Anselmo, Joe, 'Defense contractors need looser purse strings' <http://aviationweek.com/defense/opinion-defense-contractors-need-looser-rd-purse-strings>

Appleyard, Bryan, 'The Sheer Stupidity of Artificial Intelligence', *Spectator Magazine*, (5 July 2014), <https://www.spectator.co.uk/2014/07/the-sheer-stupidity-of-artificial-intelligence/>

Applied Go tutorials, 'Inverse kinematics; how to move a robotic arm (and why this is a harder then it seems)', *Applied Go Tutorials*, (16 June 2016), <https://appliedgo.net/roboticarm/>

Aquel, M., and others, 'Review of odometry: Types, approaches, challenges and applications', Springer, (28 October 2016, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5084145/>

Arkin, Ronald, 'Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture', <http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf>

Arkin, Ronald, 'The Case for Ethical Autonomy in Unmanned Systems', Journal of Military Ethics, 9.4, (2010), <https://smartech.gatech.edu/bitstream/handle/1853/36516/Arkin_ethical_autonomous_systems_final.pdf>

Arkin, Ronald, 'Warfighting Robots Could Reduce Civilian Casualties so Calling for a Ban is Premature', IEEE Spectrum, (5 August 2015), <https://spectrum.ieee.org/automaton/robotics/artificial-intelligence/autonomous-robotic-weapons-could-reduce-civilian-casualties>

Arkin, Ronald, and others, 'An Ethical Governor for Constraining Lethal Action in an Autonomous System', Georgia Institute of Technology Robot Lab, *Technical Report*, (2009), <https://www.cc.gatech.edu/ai/robot-lab/online-publications/GIT-GVU-09-02.pdf>

Army Research Laboratory, 'Research Suggests Uncertainty May Be Key to Battlefield Decision Making', *ARL*, (12 July 2018), <https://phys.org/news/2018-07-uncertainty-key-battlefield-decision.html>

Army Technology, 'Patriot Missile Review', <http://www.army-technology.com/projects/patriot/>

Army Technology, 'Prioritizing Procurement', *www.armytechnology.com,* (20 November 2009), <https://www.army-technology.com/features/feature45999/>

Arons, Barry, 'A Review of the Cocktail Party Effect', MIT Labs, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.446.8514&rep=rep1&type=pdf>

Arquilla, John, 'Cyber coming', <http://www.rand.org/content/dam/rand/pubs/reprints/2007/RAND_RP223.pdf >

Arquilla, John, 'Debating the use of autonomous weapons', <http://www.ipolitics.ca/2013/07/26/debating-unmanned-warfare/>

Arquilla, John, and David Ronfeldt, 'Swarming and the future of conflict', Santa Monica: Rand Corporation, (2000), <https://www.rand.org/pubs/documented_briefings/DB311.html>

Arrabales, Raul, and others, 'On the Practical Nature of Qualia', Department of Computer Science, Madrid, 2011, <http://www.conscious-robots.com/papers/arrabales-a-paper-aiib10.pdf>

Article 36.org, 'Meaningful Human Control', <http://www.article36.org/weapons-review/autonomous-weapons-meaningful-human-control-and-the-ccw/>

Article 36.org, 'Anti-vehicle mines and automated weapons', <http://www.article36.org/weapons/landmines/anti-vehicle-mines-victim-activation-and-automated-weapons/>

Artificial Brains blog, 'Open Worm', (14 August 2012), <http://www.artificialbrains.com/openworm>

Asaro, Peter, 'On banning Autonomous Weapon Systems: Human rights, automation and the dehumanisation of lethal decision making', *International review of the Red Cross*, volume 94, number 886, (Summer 2012), <https://www.icrc.org/eng/assets/files/review/2012/irrc-886-asaro.pdf>

Asaro, Pweter, 'Why the World Needs to Regulate Autonomous Weapons, and Soon', *Bulletin of the Atomic Scientists*, 27 April 2018, <https://thebulletin.org/2018/04/why-the-world-needs-to-regulate-autonomous-weapons-and-soon/>

Ashby, Simon, 'The 2007-2009 Financial Crisis: Learning the Risk Management Lessons', University of Nottingham, (January 2010), <https://www.nottingham.ac.uk/business/businesscentres/crbfs/documents/researchreports/paper65.pdf>

Association of Unmanned vehicles website, <http://www.auvsi.org/home>

Atakulreka, Akarachai, and others, 'Avoiding Local Minima in Feed-forward Neural Networks by Simultaneous Learning', Springer, Berlin, (2007), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.2375&rep=rep1&type=pdf>

Atherton, Kelsey, 'Burning Hydrogen for Fuel, Navy Drone Flies for 48 Hours Straight', *Popular Science*, (10 May 2013), <https://www.popsci.com/technology/article/2013-05/burning-liquid-hydrogen-fuel-navy-drone-flies-48-hours>

Atherton, Kelsey, 'Navy's Locust Launcher fires swarm drones', *Popular Science*, (24 May 2016), <http://www.popsci.com/navys-locust-launcher-fires-swarm-drones>

Atherton, Kelsey, 'The future of the Air Force is fighter pilots leading drone swarms into battle', *Popular Science,* (23 June 2017), <https://www.popsci.com/future-air-force-fighters-leading-drone-swarms>

Auda, G., and others, 'Improving the accuracy of an artificial neural network using multiple differently trained networks', Neural Network Proceedings, *IEEE World Congress on Computational Intelligence*, Vol. 2, (1998), <http://ieeexplore.ieee.org/document/685972/>

Australian Army, '16th Air Land Regiment RAA', <http://www.army.gov.au/Who-we-are/Divisions-and-Brigades/Forces-Command/6th-Brigade/16th-Air-Land-Regiment-RAA>

Australian Army Occasional Papers, 'Command and Control in Modern Warfare', *Command and Leadership Series 001*, (September 2017), <https://www.army.gov.au/sites/g/files/net1846/f/publications/command_control_b5.pdf>

Baer, Tobias, and Vishnu Kamalnath, 'Controlling Machine Learning Algorithms and their Biases', McKinsey and Company, (November 2017), <https://www.mckinsey.com/business-functions/risk/our-insights/controlling-machine-learning-algorithms-and-their-biases>

Bagchi, Tapan, 'Force Multiplier Effects in Combat Simulation', Proceedings of 7th Asia Pacific IEMS Conference, Bangkok, (December 2006), <https://www.researchgate.net/publication/228831999_Force_Multiplier_Effects_in_Combat_Simulation>

Barea, Diana, and Yaarit Silverstone, 'New Rules for Cultural Change', *Accenture Strategy*, (2016), <https://www.accenture.com/t20161216T040430__w__/us-en/_acnmedia/PDF-24/Accenture-Strategy-Workforce-Culture-Change-New.pdf>

Bartak, Roman, 'Artificial intelligence', KTIML course*, lecture slides,* (2016), <http://ktiml.mff.cuni.cz/~bartak/ui2/lectures/lecture05eng.pdf>

Bartlett, Matt, 'The AI Arms Rave in 2019', *Towards Data Science*, 28 January 2019, <https://towardsdatascience.com/the-ai-arms-race-in-2019-fdca07a086a7>

Barrett, Brian, 'Inside the Olympics Opening Ceremony World-Record Drone Show', *Wired.com*, (9 February 2018), <https://www.wired.com/story/olympics-opening-ceremony-drone-show/>

Bartlett, Peter, 'The sample complexity of the pattern classification with networks; the size of the weights is more important than the size of the network', *IEEE, transactions on information*, Volume 44, number 2, (March 1998)

Basili, Roberto, and others, 'Using Word Association for Syntactic Disambiguation', *Trends in Artificial Intelligence*, Springer Verlag, 549, (30 May 2005), <https://link.springer.com/chapter/10.1007/3-540-54712-6_237>

Batson, Lois, and Donald Wimmer, 'Unmanned Tactical Autonomous Control and Collaboration Threat and Vulnerability Assessment', Calhoun NPS Institutional Archive, (June 2015), https://calhoun.nps.edu/bitstream/handle/10945/45738/15Jun_Batson_Wimmer.pdf?sequence=1&isAllowed=y>

BBC News Asia, 'Australia to buy US drones for border patrol', <http://www.bbc.co.uk/news/world-asia-26541651>

Beckenkamp, Fabio, 'A Component Architecture for Artificial Neural Network Systems', University of Constance, Software Research Laboratory, (June 2002), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.9621&rep=rep1&type=pdf>

Behn, Beth, 'The Stakes are High: Ethics Education at US War Colleges', Air war College Publications, *Maxwell Paper Number 73*, p.2, (2018), <https://www.airuniversity.af.mil/Portals/10/AUPress/Papers/mp_0073_behn_stakes_high.pdf>

Bella, Antonio, and others, 'Calibrating of Machine Learning Models', University of Valencia, <http://users.dsic.upv.es/~flip/papers/BFHRHandbook2010.pdf>

Bendett, Samuel, 'Russia Poised to Surprise the US in Battlefield Robotics', *Defense One*, (25 January 2018) <https://www.defenseone.com/ideas/2018/01/russia-poised-surprise-us-battlefield-robotics/145439/>

Bengio, Yoshua, 'Challenges of Training Deep Neural Networks', Montreal, *Course notes IFT6266*, (Winter 2012), <https://www.iro.umontreal.ca/~bengioy/ift6266/H12/html.old/deepchallenge_en.html>

Bengio, Yoshua, 'Learning Deep Architecture for AI', *Foundations and Trends in Machine Learning*, Vol 2, Number 1, (2009), <http://www.stat.wvu.edu/~jharner/courses/dsci503/docs/ftml.pdf>

Bengio, Yoshua, and Y LeCun, 'Scaling Learning Algorithms towards AI', *Large Scale Machines*, MIT Press, (2007), <http://yann.lecun.com/exdb/publis/pdf/bengio-lecun-07.pdf>

Benitez, Mike, 'It's About Time: The Pressing `Need to Evolve the Kill Chain', *War on the Rocks,* (17 May 2017), <https://warontherocks.com/2017/05/its-about-time-the-pressing-need-to-evolve-the-kill-chain/>

Beradi, Victor, and Peter Zhang, 'The Effect of Misclassification Costs on Neural Network Classifiers', *Researchgate.net*, (June 1999), <https://www.researchgate.net/publication/227807698_The_Effect_of_Misclassification_Costs_on_Neural_Network_Classifiers>

Beres, Damon, 'The Ethical Case for Killer Robots', *Huffpost*, (3 June 2016), <http://www.huffingtonpost.co.uk/entry/lethal-autonomous-weapons-ronald-arkin_us_574ef3bbe4b0af73af95ea36>

Bergstein, Brian, 'The Great AI Paradox', *MIT Technology Review*, (12 December 2017), <https://www.technologyreview.com/s/609318/the-great-ai-paradox/>

Berreby, David, 'Artificial intelligence is already weirdly inhuman: what kind of world is our code creating?', *Nautilus*, (6 June 2016), <http://nautil.us/issue/27/dark-matter/artificial-intelligence-is-already-wierdly-inhuman>

Bezerman, Max, 'Judgement and Decision Making', Harvard/Noba, <https://nobaproject.com/modules/judgment-and-decision-making#vocabulary-bounded-rationality>

Bikakis, Antonis, 'Alternative strategies of conflict resolution in multi-context systems', *Conference paper*, International Federation for Information Processing, (2009), Volume 296, Springer, Boston MA at <https://link.springer.com/chapter/10.1007/978-1-4419-0221-4_6>

Bindi, Tas, 'True Artificial Intelligence cannot be developed until the 'brain code' has been cracked', *zdnet.com*, (9 August 2017), <http://www.zdnet.com/article/true-ai-cannot-be-developed-until-the-someone-cracks-the-brain-code-starmind/>

Birk, Andreas, 'A Quantitative Assessment of Structural Errors in Grid Maps', Jacobs University, *Autonomous Robots*, Volume 28, (2010), <https://pdfs.semanticscholar.org/aa32/bdb2fe20faf7585c2763bb70c3fb42f54196.pdf>

Black, Doug, 'AI Definitions: Machine Learning vs. Deep Learning vs. Cognitive Computing vs. Robotics vs. Strong AI', *EnterpriseTech,* (19 January 2018), <https://www.enterprisetech.com/2018/01/19/ai-definitions-machine-learning-vs-deep-learning-vs-cognitive-computing-vs-robotics-vs-strong-ai/>

Blair, David, 'AK 47 Kalashnikov: The Firearm that has killed more People than any Other', *Telegraph*, (2 July 2015), <http://www.telegraph.co.uk/news/worldnews/northamerica/usa/11714558/AK-47-Kalashnikov-The-firearm-which-has-killed-more-people-than-any-other.html>

Bonarini, Andrea, 'Learning to compose fuzzy behaviours for Autonomous Agents', *International Journal for Approximate Reasoning*, Volume 17, Issue 4, (1997), <http://www.sciencedirect.com/science/article/pii/S0888613X97000029>

Bonarini, Andrea, and others, 'Concepts for Anchoring in robots', Congress of the Italian Association for artificial intelligence, *conference paper*, (September 2001), <https://link.springer.com/chapter/10.1007/3-540-45411-X_34>

Boros, Tiberiu, and others, 'Large Tagset Labeling in Feed Forward Neural Networks', (9 August 2013), <https://aclweb.org/anthology/P/P13/P13-1068.pdf>

Borrie, John, and Tim Caughley, 'An Illusion of Safety: Challenges of nuclear weapons detonations for United Nations humanitarian coordination and response', <http://www.disarmamentinsight.blogspot.com>

Boston Dynamics, 'Big Dog robot demonstration', <https://www.youtube.com/watch?v=mpBG-nSRcrQ>

Bostrom, Nick, 'Hail Mary, Value Porosity, and Utility Diversification', *www.nickbostrom.com*, (19 December 2014), https://nickbostrom.com/papers/porosity.pdf>

Boss, Jeff, 'The Army's New Decision-Making Model', *Forbes*, (8 August 2014), <https://www.forbes.com/sites/jeffboss/2014/08/08/the-armys-new-decision-making-model/#63585c991537>

Bottcher, Sven, 'Principles of Robot Locomotion', <http://www2.cs.siu.edu/~hexmoor/classes/CS404-S09/RobotLocomotion.pdf>

Bottou, Leon, 'Two Big Challenges in Machine Learning', ICML (2015), *presentation papers*, https://icml.cc/2015/invited/LeonBottouICML2015.pdf

Bottou, Leon, 'From machine learning to machine reasoning', *Microsoft Research*, arXiv: 1102.1808, <https://www.microsoft.com/en-us/research/wp-content/uploads/2011/02/tr-2011-02-08.pdf>

Boulanin, Vincent, 'Implementing Article 36 Weapon Reviews in the Light of Increased Autonomy in Weapon Systems', *SIPRI*, 2015/1, (November 2015), https://www.sipri.org/sites/default/files/files/insight/SIPRIInsight1501.pdf

Boulanin, Vincent, and Maaike Verbruggen, 'Mapping the development of autonomy in weapon systems', *SIPRI*, (November 2017), <https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_0.pdf>

Bourque, Brad, 'The Tables Have Turned. Hardware Finally has to Catch up with Software', *Digital Trends*, (7 January 2017), <https://www.digitaltrends.com/computing/hardware-vs-software-ces-2017/>

Bousquest, Antoine, 'The Scientific Way of Warfare: Order and Chaos on the Battlefield of Modernity', *LSE*, PhD thesis, (2014), <http://etheses.lse.ac.uk/2703/1/U615652.pdf>

Bowden, Gavin, 'Real-time Deployment of Artificial Intelligence Network forecasting Models: Understanding the Range of Applicability', *Water Resources Research*, (31 October 2012), <http://onlinelibrary.wiley.com/doi/10.1029/2012WR011984/full>

Br, Joel, 'Bosnian Serbs seize more UN troops', *Washington Post,* (29 May 1995), <https://www.washingtonpost.com/archive/politics/1995/05/29/bosnian-serbs-seize-more-un-troops/991628ef-8469-436d-8759-470fe4ab11d4/?utm_term=.d6bb7243d9e3>

Brassia, Sandor, 'Artificial Intelligence in the path planning of mobile agent navigation', Elservier, *SciVerse ScienceDirect*, (2012), <http://ac.els-cdn.com/S2212567112001475/1-s2.0-S2212567112001475-main.pdf?_tid=6d536560-6611-11e7-9d94-00000aab0f6c&acdnat=1499761249_fe59e466dd607005fe4ebe2da5722507>

Breaking Defense, interview with Bob Sadowski, US Army Chief Roboticist, editor Colin Clark, <http://breakingdefense.com>

Bredeche, Nicolas, and others, 'Perceptual Learning and Abstraction in Machine Learning: An Application to Autonomous Robots', <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.3559&rep=rep1&type=pdf>

Brennon, Nathan, 'Coordinated Machine Learning and Decision Support for Situational Awareness', *Sandia National Laboratories Report*, (September 2007), <http://prod.sandia.gov/techlib/access-control.cgi/2007/076058.pdf>

Bridge, Mark, 'Killer Robots 'will cause war on a vast scale', *Times newspaper*, (22 August 2017), <https://www.thetimes.co.uk/article/elon-musk-among-technology-experts-calling-for-ban-on-

killer-robots-mustafa-suleyman-deepmind-ryan-gariepy-clearpath-robotics-mark-zuckerberg-facebook-stephen-hawking-noam-chomsky-5ndnblj0v>

Brighthub Engineering blog, 'Logic gates – the gateways to intelligence in machines', edited Lamar Stonecypher, (8 July 2008), <http://www.brighthubengineering.com/diy-electronics-devices/3349-logic-gates-the-gateways-to-intelligence-in-machines/>

Bringsford, Selmer, and Konstantine Arkoudas, 'The Philosophical Foundations of Artificial Intelligence', Department of Cognitive Science, *RRI*, Troy NY, (October 2007), <http://kryten.mm.rpi.edu/sb_ka_fai_aihand.pdf>

British Aerospace website, <http://www.baesystems.com>

BBC, 'Killer Robots: Experts warn of Third Revolution in Warfare', *BBC website,* Technology, (12 August 2017), <http://www.bbc.co.uk/news/technology-40995835>

BBC, 'China to increase military spending by 7% in 2017', *BBC Website*, (4 March 2017), <http://www.bbc.co.uk/news/world-asia-china-39165080>

The British Sociological Association, Statement of Ethical Practice, 2017, https://www.britsoc.co.uk/media/24310/bsa_statement_of_ethical_practice

Britz, Denny, 'Attention and Memory in Deep Learning and Natural Language Processing', *WILDML*, (3 January 2016), <http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/>

Brokaw, Alex, 'Autonomous Search and Rescue drones Outperform Humans at Navigating Forest Trails', *The Verge,* (11 February 2016), <https://www.theverge.com/2016/2/11/10965414/autonomous-drones-deep-learning-navigation-mapping>

Brom, Cyril, and Joanna Bryson, 'Action Selection fro Intelligence Systems', *European Network for the Advancement of Cognitive Systems*, (2006), <https://pdfs.semanticscholar.org/3419/955ab2c59b029dcda41904d46b763018de79.pdf>

Brookes, David, 'When Cultures Shift', *New York Times* op-ed, (17 October 2016), <https://www.nytimes.com/2015/04/17/opinion/david-brooks-when-cultures-shift.html?_r=0>

Brooks, Rodney, 'Artificial Intelligence without Representation', Artificial Intelligence, 47, 1991, *MIT AI Labs*, Elsevier, <http://www2.denizyuret.com/ref/brooks/brooks.pdf>

Brooks, Rodney, and others, 'Automatic correlation and calibration of noisy sensor readings using elite genetic algorithms', *Artificial Intelligence 84*, (1996), Elsevier, <https://pdfs.semanticscholar.org/cef2/9a6a615d9875d9d538c02cef71a7d29df190.pdf>

Brooks, Rodney, 'In defence of Killer Robots: Hold on there, technophobe hippies. When it comes to 'doing no harm', robots are a hell of a lot better then humans', *Foreign Policy*, (18 May 2015), <http://foreignpolicy.com/2015/05/18/in-defense-of-killer-robots/>

Browne, Cameron, and others, 'Towards the Adaptive Generation of Bespoke Game Content', John Wiley Publishing, (2012), <http://ccg.doc.gold.ac.uk/wp-content/uploads/2016/10/browne_ieeechapter14-2.pdf>

Brownlee, Jason, 'How to Improve Deep Learning Performance', *Deep Learning*, (21 September 2016), <https://machinelearningmastery.com/improve-deep-learning-performance/>

Bughiaja, Amar, 'Dropout in (Deep) Machine Learning', *Medium.com*, (15 December 2016), <https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5>

Bullinaria, John, 'Biases and Variances, Under-fitting and over-fitting', (2004), <http://www.cs.bham.ac.uk/~jxb/NN/l9.pdf>

Burgess, Matt, 'Killer Autonomous Weapons are coming, but they're not here yet', *Wired magazine*, Technical opinion, (12 August 2017), <http://www.wired.co.uk/article/killer-robots-elon-musk-autonomous-weapon-systems-uk>

Busoniu, Lucian, and others, 'Reinforcement Learning and Dynamic Programming using Function Approximation', Delft Center for Systems and Control, Netherlands, (November 2009), <https://orbi.ulg.ac.be/bitstream/2268/27963/1/book-FA-RL-DP.pdf>

Campaign to Stop Killer Robots, 'Urgent Action Needed to Ban Full Autonomous Weapons', *CSKR London*, (23 April 2013), <http://stopkillerrobots.org/wp-content/uploads/2013/04/KRC_LaunchStatement_23Apr2013.pdf>

Campbell, Tamara, and Carlos Velasco, 'An Analysis of the Tail to Tooth Ratio as a measure of Operational Readiness and Military Expenditure Efficiency', Naval Postgraduate School, Monterey, (December 2002), <http://www.dtic.mil/dtic/tr/fulltext/u2/a411171.pdf>

Carpenter, Charli, 'Don't Confuse me with the Facts: Costs of Lethal Autonomous Weapons', Duck of Minerva, (11 June 2014), <http://duckofminerva.com/2014/06/dont-confuse-me-with-the-facts-costs-of-lethal-autonomous-weapons.html>

Cass, Kelly, 'Autonomous Weapons and Accountability: Seeking Solutions in the Laws of War', *Loyola of Los Angeles Law Review,* (4 January 2015), <https://digitalcommons.lmu.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=2941&context=llr>

Cassity, Robert and Jacqueline Tame, 'The Wages of War Without Strategy: Beyond the Present – A Call to Clausewitz and to Conscience', *Strategy Bridge*, (23 August 2017), <https://thestrategybridge.org/the-bridge/2017/8/23/the-wages-of-war-without-strategy>

Caswell, Isaac, and others, 'Loopy Neural Nets: Imitating Feedback Loops in the Human Brain', Stanford, (2016), <http://cs231n.stanford.edu/reports/2016/pdfs/110_Report.pdf>

Rick Cattell and Alice Parker, 'Challenges for Brain Emulation: Why is it so difficult?', *Natural Intelligence*, INNS, Volume 1, Issue 3, (2012), <https://pdfs.semanticscholar.org/a0d4/71388ae0db850419c7e854e603b8198f8930.pdf>

Caton, Jeffrey, 'Autonomous Weapon Systems: A brief survey of developmental, operational, legal and ethical issues', *The Letort Papers*, Strategic Studies Institute, US Army War College, Carlisle PA, (December 2015), <http://publications.armywarcollege.edu/pubs/2378.pdf>

Cao, Yinzhi, and Junfeng Yang, 'Towards Making Systems Forget with Machine Unlearning', *IEEE Symposium on Security and Privacy*, (20 July 2015), <https://ieeexplore.ieee.org/Xplorehelp/#/ieee-xplore-training/working-with-documents#interactive-html>

CCSI, 'Parts of a Robot', <http://www.mind.ilstu.edu/curriculum/medical_robotics/parts_of_robots.php>

Centre for Research on Globalization, 'Drones: From Military Use to Civilian Use. Towards the Remote UAV Policing of Civil Society', <http://www.globalresearch.ca/drones-from-military-use-to-civilian-use-towards-the-remote-uav-policing-of-civil-society/30876>

Center for a New American Security, 'Autonomous Weapons and Human Control', CNAS, *Ethical Autonomy Project*, (April 2016), <https://www.files.ethz.ch/isn/196780/CNAS_Autonomous_Weapons_poster_FINAL%20(1).pdf>

Chalmers, M., and others, 'Defence Inflation: Reality or Myth?', <https://www.rusi.org/downloads/assets/Comment_Defence_Inflation_Myth_or_Reality.pdf>

Chambers, John W., 'S.L.A Marshall's 'Men Against Fire': New Evidence Regarding Fire Ratios', *Parameters*, 33.3, 2003, <http://ssi.armywarcollege.edu/pubs/parameters/articles/03autumn/chambers.pdf>

Chapman, Gretchen, 'An Introduction to the Revolution in Military Affairs', <http://www.lincei.it/rapporti/amaldi/papers/XV-Chapman.pdf>

Charter of the United Nations, 'Index', <http://www.un.org/en/documents/charter/index.shtml>

Chase-Lipton, Zachary, 'The high cost of maintaining machine learning systems', *KD Nuggets*, (January 2015), <www.kdnuggests.com/2015/01/high-cost-of-maintaining-machine-learning-technical-debt.html>

Chase-Lipton, Zachary, and others, 'A Critical Review of Recurrent Neural Networks for Sequential Learning', *arXiv preprint*, arXiv: 1506.00019, 2015, <https://arxiv.org/pdf/1506.00019.pdf>

Chatterjee, Soham, 'Good Data and Machine Learning', Towards Data Science, (24 August 2017), <https://towardsdatascience.com/data-correlation-can-make-or-break-your-machine-learning-project-82ee11039cc9>

Chaudri, Vinay K., 'Knowledge representation and reasoning', University of Stanford, *Class slides CS227*, (Spring 2011), <https://web.stanford.edu/class/cs227/Lectures/lec01.pdf>

Chen, Jessie, and Michael Barnes, 'Human-Agent Teaming for Multirobot Control', *IEEE Transactions on Human Machine Systems*, Vol.44, Issue 1, (February 2014), <http://ieeexplore.ieee.org/abstract/document/6697830/>

Chen, Jessie, and Michael Barnes, 'Supervisory control of multiple robots: Effects of imperfect automation and individual differences', US Army Research Laboratory, *Human Factors*, Volume 54, Number 2, (April 2012), <https://pdfs.semanticscholar.org/4515/4ebb06b39ecdde11a820e6d7774a87af525e.pdf>

Chen, Min, and others, 'Machine to machine communications: Architecture, Standards and Applications', *Transactions on the Internet and Information Systems*, volume 6, No.2, (February 2012), <https://www.researchgate.net/profile/Jiafu_Wan2/publication/264846553_Machine-to-Machine_Communications_Architectures_Standards_and_Applications/links/550b9af60cf265693cef8967/Machine-to-Machine-Communications-Architectures-Standards-and-Applications.pdf>

Choi, Charles, 'Too Hard for Science: Simulating the Human Brain', *Scientific American*, (9 May 2011), <https://blogs.scientificamerican.com/guest-blog/too-hard-for-science-simulating-the-human-brain/>

Chopra, Samir, 'Attribution of Knowledge to Artificial Agents and their Principals', *International Joint Conference of Artificial Intelligence*, Volume 19, (August 2005), <http://www.sci.brooklyn.cuny.edu/~schopra/893.pdf>

Chou, Andy, and others, 'An Empirical Study of Operating System Errors', *ACM SIGOPS Operating System Review*, Volume 35, Number 5, (2002), <https://pdos.csail.mit.edu/archive/6.097/readings/osbugs.pdf>

Christen, Markus, and others, 'An evaluation schema for the ethical use of autonomous robotic systems in security applications', University of Zurich, UZH Digital Society Initiative, *White Paper 1,* (2017), <https://philarchive.org/archive/CHRAES-3>

Christiano, Paul, 'The Reward Engineering Problem', *AI Alignment*, (30 May 2016), https://ai-alignment.com/the-reward-engineering-problem-30285c779450>

Cipar, James, and others, 'Solving the Straggler Problem with Bounded Staleness', *HotOS*, Volume 13, (2013), <http://www.cs.cmu.edu/~seunghak/hotOS-13-cipar.pdf>

Clancy, Kelly, 'A Computer to Rival the Brain', *New Yorker Magazine*, (15 February 2017), <http://www.newyorker.com/tech/elements/a-computer-to-rival-the-brain>

Clover, Charles, 'Chinese ships accused of breaking sanctions on North Korea', *Financial Times*, (27 November 2017), <https://www.ft.com/content/21a0407e-eadd-11e7-bd17-521324c81e23>

Cohen, Eliot A., 'A Revolution in Warfare', <http://www.foreignaffairs.com/articles/51841/eliot-a-cohen/a-revolution-in-warfare>

Cohen, Jonathan, and others, 'Should I stay or should I go? How the Human Brain manages the trade-off between exploitation and exploration', *Royal Society Publishing*, (May 2007), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2430007/>

Conn, Ariel, 'How do we align artificial intelligence with human values?', *Future of Life Institute,* (3 February 2017), <https://futureoflife.org/2017/02/03/align-artificial-intelligence-with-human-values/>

Coker, Christopher, 'Humane Warfare', <http://www.foreignaffairs.com/articles/57678/eliot-a-cohen/humane-warfare>

Cockrell School of Engineering, 'School researchers demonstrate first successful spoofing of UAVs', (27 July 2012), <http://www.engr.utexas.edu/features/humphreysspoofing>

Cole, J. Michael, 'When Drones Decide to Kill on their own', *The Diplomat*, (1 October 2012), <https://thediplomat.com/2012/10/why-killing-should-remain-a-human-enterprise/>

Collins, Robert J., and Reynaldo Thompson, 'Systemic Failure Modes: A model for Perrow's Normal Accidents in Complex, Safety Critical Systems', *Advances in Safety and Reliability,* (1997), <https://pdfs.semanticscholar.org/18d7/8946bc8bb1f58f0df6e57a7cce8fcf65f0aa.pdf>

Computer History, 'Timeline of Computer History', <http://www.computerhistory.org/timeline/memory-storage/>

Control Guru, 'Feed-forward control theory', <http://controlguru.com/the-feed-forward-controller/>

Cooke, Gordon 'The Future Battlefield', *US Army*, (16 July 2018), <https://www.army.mil/article/208553/the_future_battlefield>.

Cook, Norman, 'Correlations between Input and Output Units in Neural Networks', *Cognitive Science*, 19, (1995), <http://onlinelibrary.wiley.com/store/10.1207/s15516709cog1904_4/asset/s15516709cog1904_4.pdf?v=1&t=j92drf8l&s=e020a71c328cea489ebd9adda64fe63338ff9177>

Cooper, Gregory, 'The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks', *Knowledge System Laboratories*, Stanford University, (1990), <http://www2.stat.duke.edu/~sayan/npcomplete.pdf>

Cooper, Helene, 'Air Force Plans Shift to Obtain High-Tech Weapon Systems', *New York Times*, (30 July 2014), <https://www.nytimes.com/2014/07/31/us/politics/air-force-calls-for-cheaper-quicker-weapons-development.html>

Copeland, Michael, 'What's the difference between artificial intelligence, machine learning and deep learning?', *nvidia blog*, (29 July 2016), <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>

Coradeschi, Silvia, and Alessandro Saffiotti, 'An introduction to the Anchoring Problem', Department of Technology, Orebro University, Sweden. *Article for Robotics and Autonomous Systems*, (2003), Elsevier, <https://www.cs.utexas.edu/~kuipers/readings/Coradeschi-ras-03.pdf>

Cowie, Roddy, and others, 'Beyond Emotion Archetypes: Databases for Emotion Modelling using Neural Networks', *Elsevier Publishing*, Neural Networks, (19 May 2005), <https://www.researchgate.net/profile/Roddy_Cowie/publication/7782589_Beyond_emotion_archetypes_Databases_for_emotion_modelling_using_neural_networks/links/00b4951b63a63c205f000000/Beyond-emotion-archetypes-Databases-for-emotion-modelling-using-neural-networks.pdf>

Cox, Michael, 'Self-adjusting autonomous systems', <http://awareness-mag.eu/view.php?article=003951-2011-11-22&category=artificial+intelligence>

Coxworth, Ben, '3D-printed UAV can go from not existing to flying within 24 hours', <http://www.gizmag.com/3d-printed-uav-airframe/31473/>

Crampton, Caroline, 'Why is it so hard to predict the future of technology?', *New Statesman*, (29 January 2017), <http://www.newstatesman.com/culture/observations/2017/01/why-it-so-hard-predict-future-technology>

Correll, Niklaus, and others, 'Analysis and Observations on the Frist Amazon Picking Challenge', *arXiv.1601.05484v3*, (22 September 2017), <https://arxiv.org/pdf/1601.05484.pdf>

Crammer, Koby, and others, 'Learning from Data of Variable Quality', *Advances in Neural Information Processing Systems*, (2006), <https://www.cis.upenn.edu/~mkearns/papers/vardata.pdf>

Cronin, Patrick, 'Clausewitz Condensed: Friction of War', <http://www.au.af.mil/au/awc/awcgate/clauswtz/clwt000b.htm>

Crowder, Lucien, 'As Much Death as you Want', *Bulletin of the Atomic Scientists*, (2 December 2017), <https://thebulletin.org/2017/12/as-much-death-as-you-want-uc-berkeleys-stuart-russell-on-slaughterbots/>

Culotta, Aron, and Andrew McCallum, 'Confidence Estimation for Information Extraction', *Proceedings of HTL-NAACL*, Association for Computational Linguistics, (2004), <https://people.cs.umass.edu/~mccallum/papers/crfcp-hlt04.pdf>

Cummings, Missy, 'Interview', <http://www.dukechronicle.com/article/2015/09/former-fighter-pilot-duke-prof-missy-cummings-talks-drones>

Cummings, Missy, 'Artificial intelligence and the future of warfare', *Research Paper*, Chatham House, January 2017, Director, Humans and Autonomy Laboratory, Duke University, (26 January 2017), <https://www.chathamhouse.org/publication/artificial-intelligence-and-future-warfare>

Culurciello, Eugenio, 'Neural Network Architectures', *Towards Data Science*, (23 March 2017), <https://towardsdatascience.com/neural-network-architectures-156e5bad51ba>

Daileda, Colin, 'UN Considers Banning Killer Robots', <http://mashable.com/2014/05/13/un-ban-killer-robots/>

Danahar, John, 'Philosophical Disquisitions', *Doom and the Treacherous Turn*, blog, (19 May 2014), <http://philosophicaldisquisitions.blogspot.co.uk/2014/07/bostrom-on-superintelligence-3-doom-and.html>

Danniels, Chip, and others, 'Harnessing Initiative and Innovation: A Process for Mission Command', *Military Review*, (September-October 2012), <http://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MilitaryReview_20121031_art006.pdf>

DARPA website, 'DRC Trials 2013 Countdown: A Look at the Competition Course', (16 December 2013), <https://www.darpa.mil/news-events/2013-12-16>

S Dash, '*A comparative study of moving averages: Single, weighted and exponential',* Trade Station Labs, 9 May 2012, <https://www.tradestation.com/~/media/Files/TradeStation/Education/Labs/Analysis%20Concepts/A%20Comparative%20Study%20of%20Moving%20Averages/Moving%20Averages.ashx>

Daoutis, Marios, and others, 'Knowledge Representation for Anchoring Symbolic Concepts to Perceptual Data', *Bridges between the methodological and practical work of the robotic and cognitive systems community*, (2012), <http://www.aass.oru.se/~sci/chapter-11.pdf>

Davis, Lynn, and others, 'Armed and Dangerous? UAV and US Security', Rand Corporation, (2014), <http://www.dtic.mil/dtic/tr/fulltext/u2/a599239.pdf>

Davis, Paul, and Peter Wilson, 'Looming Discontinuities in US Military Strategy and Defense Planning: Colliding RMAs Necessitate a New Strategy', National Defense Research Institute, *RAND*, (2011), <https://www.rand.org/content/dam/rand/pubs/occasional_papers/2011/RAND_OP326.pdf>

Davis, Randall, and others, '*What is Knowledge Representation?',* <http://groups.csail.mit.edu/medg/ftp/psz/k-rep.html>

Dean, Cornelia, 'A Soldier Taking Orders from its Ethical Judgement Center', *New York Times*, (24 November 2008), <http://www.nytimes.com/2008/11/25/science/25robots.html>

Debouck, Rami, and others, 'Safety Strategy for Autonomous Systems', *Critical System Labs Inc,* Vancouver, undated, <http://www.system-safety.org/conferences/2011/papers/Safety%20Strategy%20for%20Autonomous%20Systems.pdf>

Defence Synergia, 'UK Air Defence: A forgotten capability gap', *National and Defence Strategies Research Group,* (4 February 2014), <http://www.defencesynergia.co.uk/uk-air-defence-a-forgotten-capability-gap/>

Defense Industry Daily staff, 'Anti-Sniper Systems Finding Their Range', <http://www.defenseindustrydaily.com/antisniper-systems-finding-their-range-01437/>

Defense Industry Daily staff, 'What goes up needn't come down', <http://www.defenseindustrydaily.com/DARPAs-Vulture-What-Goes-Up-Neednt-Come-Down-04852/>

Defense News, 'TRADOC Chief Looks to the Future', <http://www.defensenews.com/article/20131021/DEFREG02/310210015/TRADOC-Chief-Looks-Future>

Defense Technology Information Center, 'Artificial Intelligence and Sensor Fusion', ADP021440, unclassified, *International Conference on Integration of Knowledge Intensive Multi-Agent Systems*, (October 2003)

DeGusta, Michael, 'Are Smartphones Spreading Faster than Any Technology in Human History?', *MIT Review*, (May 2012), <https://www.technologyreview.com/s/427787/are-smart-phones-spreading-faster-than-any-technology-in-human-history/>

Deloitte, 'Artificial Intelligence Innovation Report', *Deloitte*, (2016), <https://www2.deloitte.com/content/dam/Deloitte/at/Documents/human-capital/artificial-intelligence-innovation-report.pdf>

Del Prado, Guia Marie, 'This drone is one of the most secretive weapons in the world', *Business Insider,* (29 September 2015), <http://uk.businessinsider.com/british-taranis-drone-first-autonomous-weapon-2015-9?r=US&IR=T>

Devlin, Hannah, '*Google creates and artificial intelligence program that uses reasoning to navigate the London tube', Guardian*, 12 October 2016, <https://www.theguardian.com/technology/2016/oct/12/google-creates-ai-program-that-uses-reasoning-to-navigate-the-london-tube>

Dewey, Daniel, 'Reinforcement Learning and the Reward Engineering Principle', *AAAI Spring Symposium Series*, (2014), <http://www.danieldewey.net/reward-engineering-principle.pdf>

Dewey, Daniel, 'Learning What to Value', *MIRI*, 2011, <https://intelligence.org/files/LearningValue.pdf>

Dewey, Daniel, Stuart Russel and Max Tegmark, 'Research priorities for robust and beneficial artificial intelligence', *AI Magazine*, (2015), <http://futureoflife.org/data/documents/research_priorities.pdf>

Dickson, Ben, 'The Limits and Challenges of Deep Learning', *TechTalk*, paras. 8-9, (27 February 2018), <https://bdtechtalks.com/2018/02/27/limits-challenges-deep-learning-gary-marcus/>

Dietterich, Thomas, 'Machine Learning for Sequential Data: A Review', Structural, Syntactic and Statistical Pattern Recognition, (2002), <http://web.engr.oregonstate.edu/~tgd/publications/mlsd-ssspr.pdf>

Dilem, L. and others, 'Voluntary and in voluntary attention have different consequences; the effects of perceptual difficulty', US National Library of Medicine, <https://www.ncbi.nlm.nih.gov/pubmed/18609402>

Dimitrov, Martin, 'Tracking Public Opinion under Authoritarianism', *Russian History*, 41, (2014), <http://www2.tulane.edu/liberal-arts/political-science/upload/dimitrov2014russian_history.pdf>

Dimitovski, Romy, 'Removing Humans from the Kill Chain: the Legality of (Semi-) Autonomous Weapon Systems under International Law', University of Tilburg, Law Faculty, (June 2017), http://arno.uvt.nl/show.cgi?fid=142798

Directorate of Land Concepts and Designs, Canadian Department of National Defence, 'Adaptive Dispersed Operations: The Force Employment Concept for Canada's Army of Tomorrow', (2007), <http://publications.gc.ca/collections/collection_2009/forces/D2-188-2007E.pdf>

Doherty, Sally, 'Narrow versus General AI – Is Moravec's Paradox still relevant?', *Graphcore Magazine*, (January 2017), <https://www.graphcore.ai/posts/is-moravecs-paradox-still-relevant-for-ai-today>

Doria, David, and others, 'Fast Computation on the Modern Battlefield', *US Army Research Laboratory*, (April 2015 <https://www.arl.army.mil/arlreports/2015/ARL-TR-7276.pdf>

Dorstal, Brad, 'Enhancing situational understanding through the employment of unmanned aerial vehicles', Centre for Army Lessons Learned, (2001), <http://www.globalsecurity.org/military/library/report/call/call_01-18_ch6.htm>

Dredze, Mark, and others, 'Confidence-weighted linear classifiers', University of Pennsylvania, Department of Computer and Information Science, undated, <https://www.cs.jhu.edu/~mdredze/publications/icml_variance.pdf>

Dragan, Anca, and others, 'Integrating Human Observer Inferences into Robot Motion Planning', Robotic Institute, Carnegie Mellon University, in Autonomous Robots, Springer 37.4, (2014), <http://www.ri.cmu.edu/pub_files/2014/7/legibility_AURO14.pdf>

Dunietz, Jesse, 'The Fundamental Limitations of Machine Learning', *Nautilus*, (20 September 2016), <http://nautil.us/blog/the-fundamental-limits-of-machine-learning>

Dvorsky, George, 'Autonomous Killing Machines are more dangerous that we thing', *Gizmodo*, (29 February 2016), <https://gizmodo.com/autonomous-killing-machines-are-more-dangerous-than-we-1761928608>

Echevarra, Antulio, 'Clausewitzian Centre of Gravity', <http://www.clausewitz.com/readings/Echevarria/gravity.pdf>

Economist Magazine, 'The Last Manned Fighter', <http://www.economist.com/node/18958487>

Economist Magazine, 'Military Robotics: War at Hyperspeed', *Economist*, *Special Report: The Future of War,* (27 January 2018)

Economist Magazine, 'Special Report; Artificial intelligence; From not working to neural networking', *Economist*, (25 June 2016 - 1 July 2016)

Economist Magazine, 'After Moore's Law: The Future of Computing – The Era of Predictable Improvement I Computer Hardware is Ending. What Comes Next?', (12 March 2016), <https://www.economist.com/leaders/2016/03/12/the-future-of-computing>

Economist Magazine, 'Sensor development', (November 2012), <http://www.economist.com/blogs/babbage/2012/11/cheap-sensors>

Economist Magazine, 'Trying to Restrain the Robots', (19 January 2019), <https://www.economist.com/briefing/2019/01/19/autonomous-weapons-and-the-new-laws-of-war>

Economist Magazine, 'Flight of the drones', (8 October 2011) <http://www.economist.com/node/21531433>

Economist Magazine, 'Autonomous Weapons are a game-changer', (25 January 2017), <https://www.economist.com/special-report/2018/01/25/autonomous-weapons-are-a-game-changer>

Economist Magazine, 'Artificial intelligence boom based on the old idea' <http://www.economist.com/news/special-report/21700756>

Economist Magazine, 'Intel on the Outside: The rise of artificial intelligence is creating variety in the chip market', (25 February 2017), <http://www.economist.com/news/business/21717430-success-nvidia-and-its-new-computing-chip-signals-rapid-change-it-architecture>

Edelkamp, Stephan, 'Memory Limitations in Artificial Intelligence', Algorithms for memory Hierarchies, LNCS 2625, <http://www.csd.uoc.gr/~hy460/pdf/AMH/11.pdf>

Edwards, Sean, 'Swarming and the Future of Warfare', Pandee Rand Graduate School, (2005), <http://www.dtic.mil/dtic/tr/fulltext/u2/a434577.pdf>

Ehrenreich, Barbara, 'Do humans have a role in the robot wars of the future?', *Guardian*, (11 July 2011), <https://www.theguardian.com/commentisfree/2011/jul/11/human-role-robot-war-future>

Egerstedt, Magnus, 'Control of Autonomous Mobile Robots', *Handbook of Networked and Embedded Control Systems*, Birkhauser Boston, (2005), <https://pdfs.semanticscholar.org/37cf/ab5a80cbb726799e1ebf0f4827db7585a48f.pdf>

Eish, Madeleine, and Tim Hwang, 'Praise the Machine! Punish the Human!' Comparative Studies in International Systems, *Working Paper Number 1*, Date and Society Research Institute, (24 February 2015), <https://www.datasociety.net/pubs/ia/Elish-Hwang_AccountabilityAutomatedAviation.pdf>

Deeb, Ahmed El, 'What to do with "small" data?', *Rants on Machine Learning*, (5 October 2015), <https://medium.com/rants-on-machine-learning/what-to-do-with-small-data-d253254d1a89>

Engel, Andreas, and Wolf Springer, 'Temporal Binding and the Neural Correlates of Situational Awareness', *Trends in Cognitive Science*, Volume 15, Number 1, (January 2001), <http://ieeexplore.ieee.org/document/287116/>

Epstein, Richard, 'The Empty Brain', *Aeon*, (18 May 2016), <https://aeon.co/essays/your-brain-does-not-process-information-and-it-is-not-a-computer>

Etzioni, Amitai, and others 'Pro and Cons of Autonomous Weapon Systems', *Military Review*, (May-June 2017), <https://icps.gwu.edu/sites/icps.gwu.edu/files/downloads/Etzioni%20and%20Etzioni_Pros%20and%20Cons%20Weapons.pdf>

European Commission, 'Reflection Paper on the Future of European Defence', *Com 20170 315*, (7 June 2017), <https://ec.europa.eu/commission/sites/beta-political/files/reflection-paper-defence_en.pdf>

Fahland, Dirk and others, 'Declarative Versus Imperative Process Modelling: The Issue of Maintainability', *International Conference on Business Process Management*, Springer, Berlin, Heidelberg, (2009),

<https://www.matthiasweidlich.com/paper/declarative_vs_imperative_maintainability_ERBPM_2009.pdf>

Fanucchi, D., and others, 'An Overview and Ideas on Autonomous Robot Path Planning Algorithms', (2010), <https://www.wits.ac.za/media/migration/files/cs-38933-fix/migrated-pdf/pdfs-2/2009AutonomousRobotpathplanning.pdf>

Farmer, Ben, 'Prepare for the rise of 'Killer Robots' *says former Defence Chief', Telegraph newspaper*, (17 August 2017), <http://www.telegraph.co.uk/news/2017/08/27/prepare-rise-killer-robots-says-former-defence-chief/>

Feng, Emily, and others, 'Drone swarms versus conventional arms: China's Military Debate', *Financial Times*, (24 August 2017), <https://www.ft.com/content/302fc14a-66ef-11e7-8526-7b38dcaef6140>

Ferrell, Cynthia, 'Failure recognition and fault tolerance of an autonomous robot', MIT, *Adaptive Behaviour*, 24, (1994), <http://web.media.mit.edu/~cynthiab/Papers/Breazeal-AB94.pdf>

D Filkins, 'Operators of Drones Are Faulted in Afghan Deaths', *New York Times,* <http://www.nytimes.com/2010/05/30/world/asia/30drone.html?_r=0>

P Finn, 'A Future for Drones: Automated Killing', *Washington Post*, <http://www.cbsnews.com/news/a-future-for-drones-automated-killing/>

Flach, Peter, 'Machine Learning: The Art and Science of Algorithms that Make Sense of Data', University of Bristol, (25 August 2012), <http://www.cs.bris.ac.uk/~flach/mlbook/materials/mlbook-beamer.pdf>

Flanagan, Colin, and others, 'Subsumption Control for Mobile Robot Navigation', *Proceedings of 16th Conference on Polymodel Application of Artificial Intelligence*, Sunderland, (1995), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.9.9478&rep=rep1&type=pdf>

Floreano, D., 'Design, Control and Application of Autonomous Mobile Robots', Swiss Federal Institute of Technology, Lausanne, <https://infoscience.epfl.ch/record/63893/files/aias>

Fodor, Jerry, and Zenon Pylyshyn, 'Connectionism and Cognitive Architecture: A critical analysis', Rutgers University, undated, <http://www-cogsci.ucsd.edu/~sereno/170/readings/02-FodorPylyshyn.pdf>

Foreign Exchange Rate forecasting with Artificial Neural Networks, 'Data Preparation in Neural Network Data Analysis', *International Series on Operations Research Management Science*, <http://www.bookmetrix.com/detail/chapter/598fdc96-aa45-4952-a64b-dbe88187f926#downloads>

Foreman, John, 'The real world of machine learning for fun and profit; pipeline jungles and hidden feedback loops', <http://www.john-foreman.com/blog/the-perilous-world-of-machine-learning-for-fun-and-profit-pipeline-jungles-and-hidden-feedback-loops>

Forest, Benjamin, 'An analysis of the military use of commercial satellite communications', Naval Postgraduate School, Monterey, California, (September 2008), <https://calhoun.nps.edu/bitstream/handle/10945/3991/08Sep_Forest.pdf?sequence=1>

Forgy, Charles, 'A fast algorithm for the fast pattern/many object pattern match problem', *Artificial intelligence*, 19, (1982), <https://pdfs.semanticscholar.org/464e/b245ff9822defa8db82c385ff1fd0b0b6ffe.pdf>

Foster, Nathaniel, and others, 'Is awareness and the ability to forget (and to remember) critical for demonstrating directed forgetting?', University of Illinois-Champaign, (November 2016), <https://experts.illinois.edu/en/publications/is-awareness-of-the-ability-to-forget-or-to-remember-critical-for

Fox, Alexander, 'Why is HVEC better than H.264?', *Apple Gazette*, (15 August 2018), <http://www.applegazette.com/mac/why-is-hevc-better-than-h-264/>

Francis, David, 'How a new Army of Robots can cut the Defense Budget', *The Fiscal Times*, (2 April 2017), <http://www.thefiscaltimes.com/Articles/2013/04/02/How-a-New-Army-of-Robots-Can-Cut-the-Defense-Budget>

Frankel, Stuart, 'Data Scientists don't scale', *Harvard Business Review*, (22 May 2015), <https://hbr.org/2015/05/data-scientists-dont-scale>

Fratto, Natalie, 'Machine Un-learning: Why Forgetting Might be Key to AI', *Hackernoon.com*, (31 May 2018), <https://hackernoon.com/machine-un-learning-why-forgetting-might-be-the-key-to-ai-406445177a80>

Freedberg, Sydney, 'Centaur Army: Bob Work, Robotics. & The Third Offset Strategy', *Breaking Defense*, (9 November 2015), <https://breakingdefense.com/2015/11/centaur-army-bob-work-robotics-the-third-offset-strategy/>

Freedberg, Sydney, 'Streamlined MV-SS Maintenance', *Breaking Defense*, (5 February 2018), <https://breakingdefense.com/2018/02/streamlined-mv-22-maintenance-from-70-osprey-types-down-to-5/?utm_source=hs_email&utm_medium=email&utm_content=60470967&_hsenc=p2ANqtz--Xu8INgREBr-YhTJmbRVeXy27_N9SZ9JPZQr4grwHsYyP--GM_lxTQHRDrX5AM1UrpLsLF8NPRcVjPO4KBvlHvzR9F4w&_hsmi=60470967>

Freedman, Ilana, 'F-22 and F-35: America's Costly Boondoggles Are the Victims of Arrogance and Appeasement', *Gerard Direct*, (10 March 2013), <http://gerarddirect.com/2013/03/10/uss-f-35-and-f-22-americas-costly-boondoggles-the-victims-of-arrogance-and-appeasement/>

Frei, Regina, and others, 'Self-healing and self-repairing technologies', *International Journal of Advanced Manufacturing Technologies*, (29 November 2012), <http://cui.unige.ch/~dimarzo/papers/JAMT.pdf>

Frost, Christopher, and others, 'Generalized File System Dependencies', *SOSP*, (14 October 2007), <http://featherstitch.cs.ucla.edu/publications/featherstitch-sosp07.pdf>

Future of Life Institute, 'Autonomous Weapons: An Open Letter from AI and Robotics Researchers', (July 2015), <www.futureoflife.org>

Future of Life Institute, 'Autonomous weapons: an interview with the experts', with Ariel Conn, Heather Roff and Peter Asaro, <http://futureoflife.org/2016/11/30/transcript-automous-weapons-interview-experts/1>

Future of Life Institute, '*An Open Letter: Research Priorities for Robust and Beneficial Artificial Intelligence*', undated <https://futureoflife.org/ai-open-letter/>

Futurism.com, 'Golden Age of AI', (March 2016), <https://futurism.com/amazons-ceo-says-were-living-in-the-golden-age-of-ai/>

Gal, Yarin, 'Uncertainty in Deep Learning', *PhD submission* , Department of Engineering, Cambridge, (September 2016), <http://mlg.eng.cam.ac.uk/yarin/thesis/thesis.pdf>

Galliland, Dennis, and others, 'A Note of Confidence Interval Estimation and Margin of Error', *Journal of Statistics Education*, 18,1, (2010), <https://amstat.tandfonline.com/doi/pdf/10.1080/10691898.2010.11889474?needAccess=true>

Gamez, David, 'Progress in Machine Consciousness', *Consciousness and Cognition*, 17.3, (2008), <http://davidgamez.eu/papers/Gamez07_ProgressMachineConsciousness.pdf>

G Gange and others, 'Interval Analysis and Machine Arithmetic: Why Signedness Ignorance is Bliss', *ACM Transactions on Programming Languages and Systems*, Vol.37, No.1, (January 2015), <http://cliplab.org/~jorge/docs/ACM-TOPLAS-wrapped.pdf>

Garamone, Jim, 'Dunford: Speed of Military Decision-Making must Exceed Speed of War', *US Department of Defense*, (31 January 2017), <https://dod.defense.gov/News/Article/Article/1066045/dunford-speed-of-military-decision-making-must-exceed-speed-of-war/>

Garcia, Denise, 'The Case Against Killer Robots', <http://www.foreignaffairs.com/articles/141407/denise-garcia/the-case-against-killer-robots>

Garcia, Denise, 'Governing Lethal Autonomous Weapon Systems', *Ethics and International Affairs*, Carnegie Council, (December 2017), <https://www.ethicsandinternationalaffairs.org/2017/governing-lethal-autonomous-weapon-systems/>

Gardiner, Joseph, and others, 'Command and Control: Understanding, Denying and Detecting', University of Birmingham, (February 2014), <https://arxiv.org/pdf/1408.1136.pdf>

Garfinker, Simon, 'Hackers are the real obstacle for self-driving vehicles', *MIT Technology Review,* (22 August 2017), <https://www.technologyreview.com/s/608618/hackers-are-the-real-obstacle-for-self-driving-vehicles/>

Garforth, Jason, 'Executive Attention, Task Selection and Attention-Based Training in a Neurally Controlled Simulated Robot', *Neurocomputing,* 69.16, (2006)

Gatt, Yoel, 'Space Mapping and Navigation for a behaviour-based Robot', University of Neuchatel, PhD thesis, (1994), <https://www.cs.cmu.edu/~motionplanning/papers/sbp_papers/integrated2/muller_mapping.pdf>

Geiss, Robin, 'The International-Law Dimension of Autonomous Weapon Systems', Freidrich Ebert Stiftung, (October 2015), <http://library.fes.de/pdf-files/id/ipa/11673.pdf>

Gellman, Barton, 'US documents detail Al Qaeda efforts to fight back against drones', *Washington Post*, (3 September 2013), <https://www.washingtonpost.com/world/national-security/us-documents-detail-al-qaedas-efforts-to-fight-back-against-drones/2013/09/03/b83e7654-11c0-11e3-b630-36617ca6640f_story.html?utm_term=.83448110bc4cJ>

Geneva Centre for Security Policy, 'Perils of Lethal Autonomous Weapon Proliferation: Preventing Non-State Acquisition', *GCFSP*, (2018), <https://www.gcsp.ch/News-Knowledge/Publications/Perils-of-Lethal-Autonomous-Weapons-Systems-Proliferation-Preventing-Non-State-Acquisition>

Gentry, John, 'Doomed to Fail: America's Blind Faith in Military Technology', *Parameters*, 32.4, (2002), <http://www.comw.org/rma/fulltext/0212gentry.pdf>

Gershgorn, Dave, 'AI is now so Complex its Creators can't Trust why it Makes Decisions', *Quartz*, (7 December 2017, <https://qz.com/1146753/ai-is-now-so-complex-its-creators-cant-trust-why-it-makes-decisions/>

Gertler, Jeremiah, 'US Unmanned Aerial Systems' (Congressional Research Service CRS R42136, 3 January 2012)

Gertler, Jeremiah, 'How many UAVs for DoD?' (CRS IN10317, 2015)

Georgia Tech Research Institute, 'On their own: Research on autonomous technology is developing increasingly sophisticated capability in air, marine and around robotic vehicles', Case Study; ed. H Christensen, <https://gtri.gatech.edu/casestudy/autonomous-technology-research>

Gettinger, Dan, 'Drones in the Defense Budget', Center for the Study of the Drone, Bard College, (October 2017), <http://dronecenter.bard.edu/files/2018/01/Drones-Defense-Budget-2018-Web.pdf>

Ghaffarzadel, Khasha, 'New Robotics and Drones, 2018-2038: Technologies, Forecasts, Players', *IDTechEX Reports*, (2018), <https://www.idtechex.com/research/reports/new-robotics-and-drones-2018-2038-technologies-forecasts-players-000584.asp>

Ghahramani, Zoubin, 'Probabilistic Machine-learning and Artificial Intelligence', University of Cambridge, (28 May 2015), Nature 521: 452-459, <https://www.repository.cam.ac.uk/bitstream/handle/1810/248538/Ghahramani%202015%20Nature.pdf>

Gherman, Laurian, 'Electromagnetic Spectrum Domination', *Review of the Air Force*, Air Force Academy, Romania, No. 1, 28, (2015), <http://www.afahc.ro/ro/revista/2015_1/23.pdf>

Gibbs, Samuel, 'Elon Musk leads 116 Experts calling from outright ban of Killer Robots', *Guardian*, (20 August 2017), <https://www.theguardian.com/technology/2017/aug/20/elon-musk-killer-robots-experts-outright-ban-lethal-autonomous-weapons-war>

Gibrud, Mark, 'Autonomy without mystery: where do you draw the line*?'*, <http://gubrud.net/?p=272>

Gil, Yolanda. and others, 'Artificial Intelligence and Grids: Workflow Planning and Beyond', *IEEE Intelligent Systems*, (February 20014), <https://scitech.isi.edu/wordpress/wp-content/papercite-data/pdf/gil2004ai.pdf>

Gill, G.S., and J.S. Sohal, 'Battlefield Decision-Making: A Neural Network Approach', *Journal of Theoretical and Applied Information Technology*, (2008), <http://www.jatit.org/volumes/research-papers/Vol4No8/5vol4no8.pdf>

Gillespie, Andrew, 'Foundations of Economics; *PESTEL Analysis of the Macro-Environment'*, Oxford University Press, (10 March 2011), <https://www.kantakji.com/media/1610/ty3.pdf>

Giordana, Attilio, and Alessandro Serra, 'Learning from Mistakes', Springer, *Human Machine Perception*, (2001), <https://link.springer.com/chapter/10.1007/978-1-4615-1361-2_7>

Girot, Xavier, and Yoshua Bengio, 'Understanding the Difficulties of Training Deep Feed-forward Neural Networks', Proceedings of the 13th International Conference on Artificial Intelligence and Statistics, (2010), <http://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf?hc_location=ufi>

Global Politics, 'Killer Robots and the Third Revolution in Warfare', <www.global-politics.co.uk>

Global Research Institute, 'Drones: From 'Military Use' to 'Civilian Use'. Towards the Remote UAV Policing of Civil Society?', <http://www.globalresearch.ca/drones-from-military-use-to-civilian-use-towards-the-remote-uav-policing-of-civil-society/30876>

Global Security.org, 'Samsung Techwin SGR-A1 Sentry Robot', (September 2006), <http://www.globalsecurity.org/military/world/rok/sgr-a1.htm>

Godel, Kurt, 'Incompleteness Theorem', <https://www.britannica.com/topic/Godels-first-incompleteness-theorem>

Goeree, Jacob, 'Self-correcting Information Cascades', *Review of Economic Studies*, 74.3, (2007), <http://pages.wustl.edu/files/pages/imce/brogers/casexp.pdf>

Goertzel, Ben, 'Super Intelligence: Fears, Promises and Potential: Reflections on Bostrom's 'SuperIntelligence', Yudkowsky's 'From AI to Zombies' and Weaver and Veitas's 'Open-ended Intelligence'', *Journal of Evolution and Technology*, Volume 24, Issue 2, (November 2015), <https://jetpress.org/v25.2/goertzel.htm>

Gohd, Dom, 'Amazon's CEO Says We're Living in the Golden Age of AI', *Futurism*, (9 May 2017), <https://futurism.com/amazons-ceo-says-were-living-in-the-golden-age-of-ai/>

Goodwin, Tom, 'We're at Peak Complexity. And It Sucks', *TechCrunch*, (18 October 2016), <https://techcrunch.com/2016/09/03/were-at-peak-complexity-and-it-sucks/?guccounter=1>

Gorman, Siobhan, and others, 'Insurgents Hack U.S. Drones', <http://online.wsj.com/news/articles/SB126102247889095011>

Govern, John, 'The Importance of Distance in Modern Warfare', Modern Warfare Institute, West Point, (16 May 2016), <https://mwi.usma.edu/reexamination-distance-modern-warfare/>

Grace, Katja, and others, 'When Will AI Exceed Human Performance? Evidence from AI Experts', arXiv preprint arXiv: 1705:08807, 2017, <https://arxiv.org/pdf/1705.08807.pdf>

Gray, Colin S., 'Modern Strategy', <http://www.dtic.mil/doctrine/jfq/jfq-22.pdf>

Gray, Colin S., 'Defence planning, surprises and prediction', presentation to Multiple Futures Conference, *NATO's Allied Command Transformation*, (8 May 2009), <http://www.act.nato.int/images/stories/events/2009/mfp/mfp_surprise_prediction.pdf>

Gross, Michael, 'Ethics on the Near-Future Battlefield', *Bulletin of the Atomic Scientists*, (December 2015), <https://thebulletin.org/2015/12/ethics-on-the-near-future-battlefield/>

Grossman, Sanford, and Oliver Hart, 'An Analysis of the Principal-Agent Problem', *Econometrica*, Vol. 51, Issue 1, (January 1983), <http://brousseau.info/pdf/cours/grossman_hart_83.pdf>

Ground, Larry, and others, 'Coalition-based Planning of Military Operations: Adversarial Reasoning Algorithms in an Integrated Decision Aid', arXiv pre-print arXiv: 1601.06069, 2016, <https://arxiv.org/pdf/1601.06069.pdf>

J Grundspenkis, 'Fundamentals of artificial intelligence; knowledge representation and networked schemes', *Department of Systems Theory and Design, Riga University*, undated, Lecture 7, <http://stpk.cs.rtu.lv/sites/all/files/stpk/lecture_7.pdf>

Grush, Bern, 'The rise of autonomous vehicles: planning for deployment and not just development', *R&D Lab Design,* (24 January 2018), <https://www.rdmag.com/article/2018/01/rise-autonomous-vehicles-planning-deployment-not-just-development>

Grynkewich, Alex, 'The future of air superiority, Part IV: Autonomy, survivability, and getting to 2030', *War on the Rocks*, (18 January 2017), <https://warontherocks.com/2017/01/the-future-of-air-superiority-part-iv-autonomy-survivability-and-getting-to-2030/>

Gu, Weiquing, and others, 'Towards modelling the behaviour of autonomous systems and humans for trusted operations', Naval Research Laboratory, (2014), <https://pdfs.semanticscholar.org/89e1/a67f72d6feb4dc4b8e30d35dea626eda6c42.pdf>

Gudivada, Venkat, and others, 'Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleansing and Transformation', *International Journal on Advances in Software*, 10,1, (2017), <https://www.researchgate.net/publication/318432363_Data_Quality_Considerations_for_Big_Data_and_Machine_Learning_Going_Beyond_Data_Cleaning_and_Transformations>

Guetlein, Mike, 'Lethal Autonomous Weapons; Ethical and Doctrinal Implications', US Department of Joint Military Operations, (14 February 2005), <http://www.dtic.mil/dtic/tr/fulltext/u2/a464896.pdf>

Guo, Jeff, 'Google's new artificial intelligence can't understand these sentences. Can you?', *Independent*, Indy100, (May 2016), <https://www.indy100.com/article/googles-new-artificial-intelligence-cant-understand-these-sentences-can-you--Zy9gs38g7Z>

Guszcza, Jim and Nikhil Maddirala, 'Minds and Machines: The Art of Forecasting in the Age of Artificial Intelligence', Deloitte University Press, *Deloitte Review*, Issue 19, (25 July 2016), <https://dupress.deloitte.com/dup-us-en/deloitte-review/issue-19/art-of-forecasting-human-in-the-loop-machine-learning.html>

Guyon, Isabelle, and Andre Elisseeff, 'An Introduction to Variable and Feature Selection', *Journal of Machine Learning Research*, 3, (2003), <http://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf>

Guyon, Isabelle, and others, 'Active Learning Challenge: Challenges in Machine Learning, Volume 6', Microtome Publishing, (2012), <http://www.mtome.com/Publications/CiML/CiML-v6-book.pdf>

Haas, Michael C., 'Autonomous Weapon Systems: The Military's smartest toys?', *The National Interest*, (20 November 2014), <http://nationalinterest.org/feature/autonomous-weapon-systems-the-militarys-smartest-toys-11708>

Haas, Michael C., and Sophie-Charlotte Fischer, 'The Evolution of Targeted Killing Practices: Autonomous Weapons, Future Conflicts and the International Order', *Contemporary Policy*, 38:2, (August 2017), <https://www.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Haas&Fischer_2017_TargetedKillingPractices.pdf>

Hachour, Ouarda, 'Path Planning of an Autonomous Robot', *International Journal of Systems Application*, Engineering and Development, Issue 4, Volume 2, (2008), <http://www.wseas.us/journals/saed/saed-45.pdf>

Hall, Brian, 'Autonomous Weapon System Safety', *Joint Forces Quarterly*, 86, (June 2017), <http://ndupress.ndu.edu/Media/News/Article/1223911/autonomous-weapons-systems-safety/>

Hammes, Thomas X., 'Assumptions – A Fatal Oversight', *Infinity Journal*, Issue 1, (Winter 2010), <https://www.infinityjournal.com/article/1/Assumptions__A_Fatal_Oversight/>

Hammond, Daniel, 'Autonomous Weapons and the Problem of State Accountability', *Chicago Journal of International Law*, Volume 15, Number 2, Article 8, (Winter 2015), <https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1085&context=cjil>

Hammond, Kris, 'Why Artificial Intelligence is succeeding: Then and Now', *Computerworld, Artificial intelligence today and tomorrow*, (14 September 2015), <http://www.computerworld.com/article/2982482/emerging-technology/why-artificial-intelligence-is-succeeding-then-and-now.html>

Hampton, Jesse, 'Space Technology Trends and Implications for National Security', *Kennedy School Review*, (24 January 2016), <http://ksr.hkspublications.org/2016/01/24/space-technology-trends-and-implications-for-national-security/>

Han, Meghan, 'Lethal Autonomous Weapons and Info-Wars: A Scientist's Warning', *Medium*, (6 July 2017), <https://medium.com/@Synced/lethal-autonomous-weapons-info-wars-a-scientists-warning-cc95798bc302>

Han, Song, and others, 'Learning both Weighting and Connections for Efficient Neural Networks', *Advances in Neural Information Processing Systems*, (2015), <https://papers.nips.cc/paper/5784-learning-both-weights-and-connections-for-efficient-neural-network.pdf>.

Hanon, Leighton, 'Robots on the Battlefield – are we ready for them?', American Institute of Aeronautics and Astronautics, <http://arc.aiaa.org/doi/abs/10.2514/6.2004-6409>

Harari, Yuval Noah, 'Homo Sapiens as we know them will disappear in a century or so', *Observer*, (19 March 2017), <https://www.theguardian.com/culture/2017/mar/19/yuval-harari-sapiens-readers-questions-lucy-prebble-arianna-huffington-future-of-humanity>

Harford, Tim, 'Crash: How computers are setting us up for disaster', *The Guardian*, (12 October 2016), <www.theguardian.com/technology/2016/oct/11/crash-how-computers-are-setting-us-up-for-disaster>

T Hargrove, 'How to improve Proprioception', *Bettermovement.org,* <https://www.bettermovement.org/blog/2008/proprioception-the-3-d-map-of-the-body>

Hardesty, Larry, 'The Faster-than-Fourier Transform', *MIT News*, (18 January 2012), <http://news.mit.edu/2012/faster-fourier-transforms-0118>

Hardesty, Larry, 'Automatic bug repair: System fixes bugs by importing functionality from other programs without access to source code', *MIT News*, (29th June 2015), <http://news.mit.edu/2015/automatic-code-bug-repair-0629>

O'Hare, Ryan, 'Armed drones and military robots have 'limitless potential for disaster': Experts fear that we are being lulled into a false sense of security by autonomous machines', (4 March 2016), <https://www.dailymail.co.uk/sciencetech/article-3476870/Armed-drones-military-robots-limitless-potential-disaster-experts-warn.html>

Harelson, Lonnie, 'The Principles of War: Valid Yesterday, Today and Tomorrow', Joint Forces Staff College, Norfolk US, (2005), <http://www.dtic.mil/dtic/tr/fulltext/u2/a436747.pdf>

Harman, Gilbert, 'Artificial Intelligence and some Philosophical Issues', University of Princeton, (4 October 2005), <https://www.cs.princeton.edu/courses/archive/fall05/cos402/readings/harman.pdf>

Harper, Robert, 'Structure and Efficiency of Computer Programming', Carnegie Mellon School of Computer Science, (23 July 2014), <http://repository.cmu.edu/cgi/viewcontent.cgi?article=3673&context=compsci>

Harress, Christopher, 'The Rise of China's Drone Fleet And Why It May Lead To Increased Tension In Asia', <http://www.ibtimes.com/rise-chinas-drone-fleet-why-it-may-lead-increased-tension-asia-1535718>

Hart, D., 'War remains inside the courtroom: Jurisdiction under ECHR', *UK Human Rights Blog*, (11 September 2016), <https://ukhumanrightsblog.com/2016/09/11/war-remains-inside-the-court-room-jurisdiction-under-echr/>

Hartemink, Alexander J., and others, 'Maximum likelihood estimation of optimal scaling factors', MIT Laboratory for Computer Science, <https://groups.csail.mit.edu/cgs/pubs/spie.pdf>

Van Hasselt, Hado, 'Reinforcement Learning in Continuous State and Action spaces', *Reinforcement Learning*, Springer, Berlin, Heidelberg, (2012), <https://www.researchgate.net/profile/Hado_Van_Hasselt2/publication/239843999_Reinforcement_Learning_in_Continuous_State_and_Action_Spaces/links/0c9605220e5949a8a4000000/Reinforcement-Learning-in-Continuous-State-and-Action-Spaces.pdf>

Van Harmelen, Frank, and others, '*Handbook of Knowledge Representation', Foundations of Artificial Intelligence*, Elsevier 2008, 978-0-444-52211-5, <http://dai.fmph.uniba.sk/~sefranek/kri/handbook/handbook_of_kr.pdf>

O'Haver, Tom, 'A Pragmatic Introduction to Signal Processing', University of Maryland at college Park, Department of Chemistry and Bio-Chemistry, (May 2017), <https://terpconnect.umd.edu/~toh/spectrum/Differentiation.html>

Hawley, John, 'Automation and the Patriot air and missile defence system', Centre for a new American security, Washington, (25 January 2017), <https://www.cnas.org/press/press-release/cnas-releases-report-on-automation-and-the-patriot-air-and-missile-defense-system>

Hayes, Richard, and others, 'The State of the Art and the State of Practice, Battle of the Bulge: The Impact of Information Age Command and Control on Conflict – Lessons Learned', *CCRTS*, (2006), <http://www.dodccrp.org/events/2006_CCRTS/html/papers/206.pdf>

B Hayes-Roth, 'A blackboard architecture for control', *The Heuristic Program Project*, Stanford University, *Artificial intelligence International Journal*, (February 2003) <http://www.sciencedirect.com/science/journal/00043702/26/3>

Hearn, Robert, 'Building Grounded Abstractions for Artificial Intelligence Programming', *MIT*, (June 2001), <https://groups.csail.mit.edu/mac/users/bob/grounded-abstractions.pdf>

Helgason, Helgo Pall, 'General Attention Mechanisms for Artificial Intelligence systems', University of Reykjavik, PhD, (June 2013), <https://en.ru.is/media/td/Helgi_Pall_Helgason_PhD_CS_HR.pdf>

Helgason, Helgo Pall, and others, 'Towards a General Attention Mechanism for Embedded Intelligent Systems', *International Journal of Computer Science and Artificial Intelligence*, Volume 4, Issue 1, (May 2014),

<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=FC977A50C3C47BAD83BB0BC215F74E40?doi=10.1.1.694.5003&rep=rep1&type=pdf>

Hennigan, W.J., 'New drone has no pilot anywhere, so who's accountable?', *Los Angeles Times,* (26 January 2012) <http://articles.latimes.com/2012/jan/26/business/la-fi-auto-drone-20120126>

Hensman, Paulina, and David Masto, 'Impact of Imbalanced Training Data for Convolutional Neural Networks', Degree project, Kth Royal Institute of Technology, Stockholm, (May 2015), <https://www.kth.se/social/files/588617ebf2765401cfcc478c/PHensmanDMasko_dkand15.pdf>

Here blog, 'Enabling an Autonomous World for Everyone', undated, <https://www.here.com/en/vision/autonomous-world?cid=Auto-Google-MM-T2-Here-generic-BMM&utm_source=Google&utm_medium=ppc&utm_campaign=Auto_PaidSearch_Automotive_AlwaysOn>

Hew, Patrick Chisan, 'The Generation of Situational Awareness – A Near to Mid-Term Study', Defence System Analysis Division, Australian Army, (July 2006), <http://www.dtic.mil/dtic/tr/fulltext/u2/a465252.pdf>

Heynes, Deborah, 'Spiraling cost of weapons makes war 'too expensive', *The Times*, (26 April 2017), <https://www.thetimes.co.uk/article/spiralling-cost-of-weapons-makes-war-too-expensive-6fkzf03w6>

Hicks, Kathleen, 'What will replace the Third Offset? Lessons from Past Innovation Strategies', *Defense One*, (17 March 2017), <http://www.defenseone.com/ideas/2017/03/what-will-replace-third-offset-lessons-past-innovation-strategies/136260/>

Hof, Robert, 'Deep Learning: with the massive amounts of computational power, Machines can now recognise objects and translate speech in real time. Artificial intelligence is finally getting smart', *TechnologyReview.com*, (June 2016), <https://www.technologyreview.com/s/513696/deep-learning/>

Hofman, Dietrich, 'Common sources of errors in measurement systems', Steinbeis Transfer Centre for Quality Insurance, *Handbook of Measuring System Design*, (2005), <http://eu.wiley.com/legacy/wileychi/hbmsd/pdfs/mm154.pdf>

Hoffman, Frank G., 'Thinking about future conflict', *Marine Corps Gazette*, Volume 98, Issue 11, (November 2014), <http://pqasb.pqarchiver.com/mca-members/doc/1619980305.html?FMT=TG>

Hoffman, M, 'China Reports Stealth Drone's First Test Flight', *Defense Technology*, (22 November 2013) <http://defensetech.org/2013/11/22/china-reports-stealth-drones-first-test-flight/>

Hoffman, RR, and others, 'The Myths and Costs of Autonomous weapon Systems', *Bulletin of the Atomic Scientists*, 72, 4, (2016) <https://www.tandfonline.com/doi/abs/10.1080/00963402.2016.1194619>

Holford, Bill, 'Big Data on the Battlefield', *IT ProPortal*, (1 February 2017), <https://www.itproportal.com/features/big-data-on-the-battlefield-an-introduction/>

Holmes, Nicholas, 'Leadership and Resiliency Training from a Soldier's Perspective', US Army, *www.army.mil,* (19 December 2017), <https://www.army.mil/article/198421/leadership_and_resiliency_training_from_a_soldiers_perspective>

Holte, Robert, and Gaojian Fan, 'State Space Abstraction in Artificial Intelligence and Operations Research', University of Alberta, *AAAI Workshop*, (2015), <https://www.aaai.org/ocs/index.php/WS/AAAIW15/paper/download/10134/10234>

Hoos, Holger, and others (eds.), 'Automated Algorithm Selection and Configuration', *Report from Dagstuhl Seminar 16412*, (2017), http://drops.dagstuhl.de/opus/volltexte/2017/6956/pdf/dagrep_v006_i010_p033_s16412.pdf

Horowitz, Michael, 'The Promise and Perils of Military Applications of Artificial Intelligence', Bureau of the Atomic Scientists, (23 March 2018), <https://thebulletin.org/landing_article/the-promise-and-peril-of-military-applications-of-artificial-intelligence/>

Horowitz, Michael, and Paul Scharre, 'An Introduction to Autonomy in Weapon Systems', *CNAS*, (13 February 2015), <https://s3.amazonaws.com/files.cnas.org/documents/Ethical-Autonomy-Working-Paper_021015_v02.pdf?mtime=20160906082257>

Horowitz, Michael, 'Public Opinion and the Politics of the Killer Robot Debate', *Sage Journals (*Research and Politics Series), (16 February 2016), <http://journals.sagepub.com/doi/pdf/10.1177/2053168015627183>

Horowitz, Michael, 'The Ethics and Morality of Robotic Warfare: Assessing the Debate over Autonomous Weapons', American Institute of Arts and Sciences, (February 2016), <http://www.michaelchorowitz.com/Documents/HorowitzLAWSEthicsDraftFeb2016.pdf>

UK House of Commons Defence Committee, 'Gambling on 'Efficiency': Defence Acquisition and Procurement', First Report of Session, (2017-2019), <https://publications.parliament.uk/pa/cm201719/cmselect/cmdfence/431/431.pdf>

Huang, T.S., 'Computer Vision: Evolution and Promise', 19th CERN School of Computing Proceedings, (1996), <http://cds.cern.ch/record/400313/files/p21.pdf>

Huerta, Pedro, 'Assessing Difficulties of Conditional Probability Problems', University of Valencia, EDU/2008-03140/Edu Project, <http://www.cerme7.univ.rzeszow.pl/WG/5/CERME_Huerta-Cerdan-Lonjedo-Edo.pdf>

Huis, Randy, 'Proliferation of Precision Strike: Issues for Congress', Congress Research Services, R42539, (14 May 2012), <https://fas.org/sgp/crs/nuke/R42539.pdf>

Human Rights Watch, 'A wedding that became a funeral: US drone attack of marriage procession in Yemen', (2013), <https://www.hrw.org/report/2014/02/19/wedding-became-funeral/us-drone-attack-marriage-procession-yemen>

Human Rights Watch, 'Shaking the Foundations', (12 May 2014) <https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots#>

Human Rights Watch and IHRC, 'Fully Autonomous Weapons: Questions and Answers', (October 2013), <https://www.hrw.org/sites/default/files/supporting_resources/10.2013_killer_robots_qa.pdf>

Hunicke, Robin and Vernell Chapman, 'Artificial Intelligence for dynamic difficulty adjustment in games', Northwestern University, <http://www.cs.northwestern.edu/~hunicke/pubs/Hamlet.pdf>

Hunt, Angie, 'Are 'Machine Values' Replacing our Principles?', *Futurity blog*, (19 April 2017), <https://www.futurity.org/technology-machine-values-1406692-2/>

Hura, Myron and others, 'Intelligence Support and Mission Planning for Autonomous Precision-guided Weapons', *RAND*, *United States Airforce, Library of Congress Publishing*, (1993), <https://apps.dtic.mil/dtic/tr/fulltext/u2/a282344.pdf>

Husain, Amir, 'AI on the battlefield: a framework for ethical autonomy', Forbes Technology Council, (28 November 2016), <https://www.forbes.com/sites/forbestechcouncil/2016/11/28/ai-on-the-battlefield-a-framework-for-ethical-autonomy/#71cdcb2c5cf2>

Hutchinson, Harold, 'Russia says it will ignore any ban of killer robots', *Business Insider Tech*, (30 November 2017), <http://uk.businessinsider.com/russia-will-ignore-un-killer-robot-ban-2017-11?r=US&IR=T>

Hwong, Yi-Ling, 'Attention in Artificial Intelligence Systems', *AGI.io blog*, (22 September 2017), <https://agi.io/2017/09/22/attention-in-artificial-intelligence-systems/>

Hyndman, Rob, 'Why are some things easier to forecast than others?', *Hyndsight*, (18 September 2012), <https://robjhyndman.com/hyndsight/hardforecasts/>

Icelandic Human Rights Centre, 'Human Rights and Armed Conflict', <http://www.humanrights.is/en/human-rights-education-project/human-rights-concepts-ideas-and-fora/human-rights-in-relation-to-other-topics/human-rights-and-armed-conflict>

Ilachinski, Andrew, 'AI, Robots and Swarms: Issues, Questions and Recommended Studies', CAN Corporation, (January 2017), <https://www.cna.org/CNA_files/PDF/DRM-2017-U-014796-Final.pdf>

Ingersoll, Geoffrey, and Robert Johnson, 'The 25 Most Effective Weapons in the US Arsenal', *Business Insider,* (14 December 2012), <https://www.businessinsider.com/most-effective-weapons-in-the-us-arsenal-2012-12?IR=T>

ICRC, 'Basic Rules of the Geneva Convention and Additional Protocols', International Committee of the Red Cross, (December 1988), <https://www.icrc.org/eng/assets/files/other/icrc_002_0365.pdf>

ICRC Casebook, 'Military necessity', undated, <https://casebook.icrc.org/glossary/military-necessity>

IRCR, 'The Use of Armed Drones Must Comply with the Laws of Armed Combat', *ICRC*, (10 May 2013), <https://www.icrc.org/eng/resources/documents/interview/2013/05-10-drone-weapons-ihl.htm>

ICRC IHL database, 'Rule 14 – Proportionality in Attack', <https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_cha_chapter4_rule14>

International Committee of the Red Cross, 'Law of Armed Combat: Basic Knowledge', *ICRC*, (June 2002), <https://www.icrc.org/eng/assets/files/other/law1_final.pdf>

International Committee of the Red Cross, 'Jus in bello – Jus ad bellum', <http://www.icrc.org/eng/war-and-law/ihl-other-legal-regmies/jus-in-bello-jus-ad-bellum/index.jsp>

International Committee of the Red Cross (ICRC), 'Geneva Conventions and Commentaries', undated, <https://www.icrc.org/en/war-and-law/treaties-customary-law/geneva-conventions>

International Committee for Robot Arms Control, '2014 Mission Statement', <http://icrac.net/statements/>

International Committee of the Red Cross, 'Decision making in military combat operations', *ICRC Publications*, (October 2013), <https://www.icrc.org/eng/assets/files/publications/icrc-002-4120.pdf

International Committee for Robot Arms Control, 'Computing experts from 37 countries call for ban on killer robots', <http://icrac.net/2013/10/computing-experts-from-37-countries-call-for-ban-on-killer-robots/>

Internet Encyclopedia of Philosophy, 'Just War Theory', <http://www.iep.utm.edu/justwar/>

International Federation of Robotics IFR, 'World Robotics 2017', Executive summary, (2017), <https://ifr.org/downloads/press/Executive_Summary_WR_2017_Industrial_Robots.pdf>

IHLS, 'Global Powers at the Edge of Autonomous Battlefield Innovation', *IHLS*, (18 November 2017), <https://i-hls.com/archives/79787>

ISCA Tutorial, 'Hardware Architecture for Deep Neural Networks', *MIT*, nvidia, (24 June 2017), <http://www.rle.mit.edu/eems/wp-content/uploads/2017/06/ISCA-2017-Hardware-Architectures-for-DNN-Tutorial.pdf>

Ivanov, Slav, '37 Reasons your Neural Network is not Working', *M Slav, blog*, (25 July 2017), <https://blog.slavv.com/37-reasons-why-your-neural-network-is-not-working-4020854bd607>

Jackson, Eric, 'Sun Tsu's Art of War', *Forbes*, (23 May 2014), <https://www.forbes.com/sites/ericjackson/2014/05/23/sun-tzus-33-best-pieces-of-leadership-advice/#19c3ac7d5e5e>

Janakiram, MSV, 'In the Era of Artificial Intelligence, GPUs are the New CPUs'. *Forbes*, (7 August 2017), <https://www.forbes.com/sites/janakirammsv/2017/08/07/in-the-era-of-artificial-intelligence-gpus-are-the-new-cpus/#705bb4955d16>

Jhingran, Anant, 'Obsessing over Artificial Intelligence is the wrong way to think about the future', *Wired Magazine*, Business, (22 January 2016), <https://www.wired.com/2016/01/forget-ai-the-human-friendly-future-of-computing-is-already-here/>

Jin, Xu, 'The Strategic Implications of Changes in Military Technology', <http://cjip.oxfordjournals.org/content/1/2/163.full>

Johnson, Ted, and Charles Ward, 'The Military Should Teach AI to Watch Drone Footage', *Wired Magazine*, (26 November 2017), <https://www.wired.com/story/the-military-should-teach-ai-to-watch-drone-footage/

Joint Force Quarterly, <http://www.dtic.mil/doctrine/jfq/jfq-22.pdf>

Jones, Peter, 'An Iterative Algorithm for Autonomous Tasking in Sensor Networks (Decision and Control)', *IEEE Conference paper*, (2006), <http://ieeexplore.ieee.org/document/4177379/>

Jones, Seth, 'AI and Robots line up for Battlefield Service', *Financial Times*, (6 November 2016), <https://www.ft.com/content/02d4d586-78e9-11e6-97ae-647294649b28?mhq5j=e2>

Jones, Seth, 'The Return of Political Warfare', *Defense Outlook 2018*, (February 2018), <https://www.csis.org/analysis/return-political-warfare>

Jones, Seth, 'Much 'Political Warfare' in our Future', *BreakingDefense*, (2 February 2018), <https://brekingdefense.com/2018/02/much-political-warfare-in-our-future>

Johnson, David, and others, 'Preparing and training for the full spectrum of military challenges: Insights from the experiences of China, France, the United Kingdom, India and Israel', National Defence research Institute, Rand Corporation Publishing, (2009), <https://www.rand.org/content/dam/rand/pubs/monographs/2009/RAND_MG836.pdf>

Johnson, Matthew, and others, 'Team IHMC's Lessons Learned from the DARPA Robotics Challenge Trials', *Journal of Robotics*, (March 2015), <https://s3.amazonaws.com/academia.edu.documents/41980899/Team_IHMCs_Lessons_Learned_from_the_DAR20160203-30232-1p7o2um.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1540061344&Signature=kTtjZhdj2UozAahn7Cbwujww3X0%3D&response-content-disposition=inline%3B%20filename%3DTeam_IHMCs_Lessons_Learned_from_the_DARP.pdf>

Johnson, Ronald, 'Lanchester's Square Law in Theory and Practice', School of Advanced Military Studies, Fort Leavenworth, (1990), <http://www.dtic.mil/dtic/tr/fulltext/u2/a225484.pdf>

Joshi, Naveen, 'Now, What is Machine Unlearning All About?', *Allerin blog*, (12 March 2018), <https://www.allerin.com/blog/now-what-is-machine-unlearning-all-about>

Kaelbling, Leslie, and others, 'Planning and Acting in Partially Observable Stochastic Domains', Elsevier, *Artificial Intelligence*, Vol. 101, (May 1998), <https://ac.els-cdn.com/S000437029800023X/1-s2.0-S000437029800023X-main.pdf?_tid=c7163328-a68c-11e7-b7f0-00000aacb35e&acdnat=1506851102_e13f50110d2bfa5bfd9cb5e6bab83d35>

Kahneman, Daniel, and others, 'Would You Be Happier if You Were Richer? A Focusing Illusion', *CEPA Working Paper*, 125, (May 2006), <http://www.morgenkommichspaeterrein.de/ressources/download/125krueger.pdf>

T Kaiser, 'US Navy Launches First Unmanned X-47B Aircraft from Carrier Flight Deck', *DailyTech*, <http://gulwww.dailytech.com/US+Navy+Launches+First+Unmanned+X47B+Aircraft+from+Carrier+Flight+Deck+/article31559.htm>

O'Kane, Jason, and others, 'Algorithms for Planning under Uncertainty in Prediction and Sensing', University of Illinois, *Autonomous Mobile Robots: Series in Control Engineering*, (501-547), (2005), <http://msl.cs.illinois.edu/~lavalle/papers/OkaTovCheLav06.pdf>

Kania, Elsa, 'The Critical Human Element in the Machine Age of Warfare', *Bulletin of the Atomic Scientists*, (15 November 2017), <https://thebulletin.org/2017/11/the-critical-human-element-in-the-machine-age-of-warfare/>

Kania, Elsa, 'Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power', Center for a New American Security, (28 November 2017), <https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power>

Kamimura, R., 'Generation of Organised Internal Representations in Recurrent Neural Networks', Neural Networks, *IJCNN*, (June 1992), <http://ieeexplore.ieee.org/document/287116/>

Kaminski, Bogumil, and others, 'A Framework for Sensitivity Analysis of Decision Trees', *Central European Journal of Operations Research*, Volume 26, Issue 1, <https://link.springer.com/article/10.1007%2Fs10100-017-0479-6>

Karpathy, Andrej, 'The state of computer vision and AI: we are really, really far away', *blog*, (22 October 2012), <http://karpathy.github.io/2012/10/22/state-of-computer-vision/>

Kaspersen, Anja, 'We're on the brink of an artificial intelligence arms race. But we can curb it', *World Economic Forum*, (15 June 2016), <https://www.weforum.org/agenda/2016/06/should-we-embrace-the-rise-of-killer-robots/>

Anja Kaspersen and others, 'Ten Trends for the Future of Warfare', *World Economic Forum* (3 November 2016), <https://www.weforum.org/agenda/2016/11/the-4th-industrial-revolution-and-international-security/>

Kaspersen, Anja, 'Is Technology Blurring the Lines between War and Peace?', *World Economic Forum*, (12 February 2016), <https://www.weforum.org/agenda/2016/02/is-technology-blurring-the-lines-between-war-and-peace/>

Kellenberger, Jakob, 'International Humanitarian Law and New Weapon Technology', *ICRC, 34th Round Table,* (10 September 2011), <http://www.icrc.org/eng/resources/documents/statement/new-weapon-technologies- statement-2011-09-08.htm>

Kennedy, John, 'David Moloney: 'AI is at a once-in-a-lifetime inflexion point', *Silicon Republic*, (29 July 2017), <*https://www.siliconrepublic.com/machines/david-moloney-movidius-intel-ai-deep-learning*>

Kestner, Peter, 'Encouraging Autonomy', *KPMG website*, (30 November 2017), <https://home.kpmg.com/xx/en/home/insights/2017/11/encouraging-autonomy.html>

Khazan, Olga, 'The Best Headspace for Making Decisions', *The Atlantic, Science*, (19 September 2016), <https://www.theatlantic.com/science/archive/2016/09/the-best-headspace-for-making-decisions/500423/>

Khemlani, Sangeet, and JG Trafton, 'Percentiler Analysis for Goodness-of-Fit Comparisons of Models to Data', Navy Center for Applied Research into Artificial Intelligence, Proceedings of 36th Annual Conference of the Cognitive Science Society, (July 2014), <https://www.nrl.navy.mil/itd/aic/content/percentile-analysis-goodness-fit-comparisons-models-data>

Kim, Dae-Young, and others, 'Data filtering system to avoid total data distortion in IOT networks', 9, 6, (2017), <www.mdpi.com/journal/symmetry>

Kim, Young-Bum, and Karl Stratos, 'Adversarial Adaption of synthetic or Stale Data', Microsoft AI & Research, cit. Proceedings of 55th AGM, Association of Computational Linguistics, (2017), <http://www.karlstratos.com/publications/acl17adversarial.pdf>

Kim, Young J., 'Enhanced Battlefield Visualization for Situational Awareness', *Computer and Graphics*, 27.6, (2003), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.87.6406&rep=rep1&type=pdf>

Kimmons, Sean, 'Mad Scientists' Discuss Emerging Technology and Army releases Strategy on Robots', *Defense Systems Information Analysis*, (10 April 2017), <https://www.army.mil/article/183862/mad_scientists_discuss_emerging_tech_as_army_releases_strategy_on_robots>

King, Paul, and others, 'Intelligence is turning out to be a computational problem. What about goal-setting? Is that a uniquely human endeavour?', *Quora magazine*, (6 May 2016), <https://www.quora.com/Intelligence-is-turning-out-to-be-a-computational-problem-What-about-goal-setting-Is-that-a-uniquely-human-endeavor>

Kirsch, Andreas, 'Autonomous Weapons will be Tireless, Efficient Killing Machines – and there is no way to stop them', *Quartz News*, (23 July 2018), <https://qz.com/1332214/autonomous-weapons-will-be-tireless-efficient-killing-machines-and-there-is-no-way-to-stop-them/>

Kirkpatrick, James, and others, 'Overcoming Catastrophic Forgetting in Neural Networks', *PNAS*, Vol.14, No.13, (March 2017), <http://www.pnas.org/content/114/13/3521.full.pdf>

Klassen, Tim, 'The UAV Video Problem', (1 July 2009), <https://www.militaryaerospace.com/articles/print/volume-20/issue-7/features/viewpoint/the-uav-video-problem-using-streaming-video-with-unmanned-aerial-vehicles.html>

Klein, David, 'US Department of Defense 2015 budget analysis', (May 2014), <www.auvsi.org/Mississippi/blogs/david-klein/2014/05/02/us-department-of-defense-2015-budget-analysis>

Knight, Will, 'AI's Language Problem: Machines that truly understand human language would be incredibly useful. But we don't know how to build them', *MIT Technology Review*, (9 August 2016), <https://www.technologyreview.com/s/602094/ais-language-problem/>

Knight, Will, 'The US Military wants its Autonomous Machines to explain themselves', *MIT Technology Review*, (14 March 2017), <https://www.technologyreview.com/s/603795/the-us-military-wants-its-autonomous-machines-to-explain-themselves/>

Knight, Will, '5 Big Predictions for Artificial Intelligence in 2017', *MIT Technology Review*, (4 January 2017), <https://www.technologyreview.com/s/603216/5-big-predictions-for-artificial-intelligence-in-2017/>

Koebler, Jason, 'Oregon Company to Sell Drone Defence Technology to Public' <http://www.usnews.com/news/articles/2013/03/15/oregon-company-to-sell-drone-defense-technology-to-public>

Koerner, Brendan, 'Inside the new arms race to control bandwidth on the battlefield', *Wired Magazine*, (18 February 2014), <https://www.wired.com/2014/02/spectrum-warfare>

Koivula, Tommi, and Katariina Simonen, 'Arms Control in Europe: Regimes, Trends and Threats', National Defence University, Helsinki, Series 1, *Research Publication 16*, (2017), <http://www.doria.fi/bitstream/handle/10024/144087/Arms%20control%20in%20Europe_netti.pdf?sequence=1>

M Kokegei and others, 'Fully Coupled 6 Degrees-of-Freedom Control of an Over-Actuated Autonomous Underwater Vehicle', *InTech*, undated <http://cdn.intechopen.com/pdfs-wm/21972.pdf>

Kon, Mark, and others, 'Statistical Representations of Prior `Knowledge in Machine Learning', *Artificial Intelligence Applications*, (2005), <http://math.bu.edu/people/mkon/B5Final.pdf>

Kongsberg Gruppen, 'Naval and Joint Strike Missile Update', *Kongsberg*, (13 March 2014), <https://www.kongsberg.com/en/kds/products/missilesystems/jointstrikemissile/>

Konidaris, George, 'An Adaptive Robot Motivational System', University of Massachusetts at Amherst, undated, <http://www-anw.cs.umass.edu/pubs/2006/konidaris_b_SAB06.pdf>

Kosko, Bart, 'Hidden Patterns in Combines and Adaptive Knowledge Networks', *Elsevier Science Publishing*, 2, (1988), <http://ac.els-cdn.com/0888613X88901119/1-s2.0-0888613X88901119-main.pdf?_tid=245f5116-9b5e-11e7-a4d0-00000aab0f6c&acdnat=1505621609_a7493474593fc9822b5484b3a9589e34>

Kostopoulos, Lydia, 'Drivers for the Deployment of Lethal Autonomous Weapons', *Medium.com*, (22 December 2017), <https://medium.com/@lkcyber/drivers-for-the-deployment-of-lethal-autonomous-weapons-systems-ae1dd6278a35>

Kot, Martin, 'The State Explosion Problem', (16 August 2003), <http://www.cs.vsb.cz/kot/down/Texts/StateSpace.pdf>

Kott, Alexander, 'Challenges and Characteristics of Intelligent Autonomy for Internet of Battle Things in Highly Adversarial Environment', US Research Laboratory, Adelphi MD, arXiv preprint arXiv: 1803.11256, unnumbered, (2018), <https://arxiv.org/pdf/1803.11256.pdf

Kott, Alexander and others, 'Visualizing the Tactical Ground Battlefield in the Year 2050', US Army Research Laboratory, (June 2015), <https://www.arl.army.mil/arlreports/2015/ARL-SR-0327.pdf>

Krahenmann, Sandra, 'Positive obligations in human rights treaties', *PhD Thesis no 949*, Graduate Institute of International studies, Geneva, (2012)

Krause, Keith, 'War, Violence and the State', *Securing Peace in a Globalised World*, (2009), <http://graduateinstitute.ch/files/live/sites/iheid/files/sites/admininst/shared/doc-professors/forthcoming%20war,violence%20and%20state%20HW70_ch9_krause_corrected%5B1%5D.pdf>

Zenko, Micah, and Sarah Kreps, 'The Next Drone Wars', <http://www.foreignaffairs.com/articles/140746/sarah-kreps-and-micah-zenko/the-next-drone-wars>

Krishnamurthy, Vikram, 'Algorithms for Optimal Scheduling and Management of hidden Markov mode Sensors', IEEE, *Transactions on Signal Processing*, Vol 50, No 6, (June 2002), <http://ece.ubc.ca/~vikramk/Kri02.pdf>

Kroll, Dennis, and Klaus David, 'Measuring the Capability of Smartphones for Executing Contextual Algorithms', *Informatics LNI*, Bonn, (2017), <https://dl.gi.de/bitstream/handle/20.500.12116/3924/B20-1.pdf?sequence=1&isAllowed=y>

Kruchtren, Philippe, and others, 'Technical Debt: From Metaphor to Theory and Practice', University of British Columbia, *IEEE Software*, (2012), <https://www.computer.org/csdl/mags/so/2012/06/mso2012060018.pdf>

Kulwin, Noah, 'Elon Musk and over 100 AI Experts warn UN about Killer Robots', *Vice News*, (21 August 2017), <https://news.vice.com/en_us/article/9kdgkz/elon-musk-and-over-100-ai-experts-warn-u-n-about-killer-robots>

Kurzweil, 'Accelerating Intelligence; IBM simulates 530 billion neurons, 100 trillion synapses on supercomputer', <http://www.kurzweilai.net/ibm-simulates-530-billon-neurons-100-trillion-synapses-on-worlds-fastest-supercomputer>

Kwang, Kevin, 'Machine Learning needs Human Helping Hand', *ZDNedt*, (3 April 2014), <http://www.zdnet.com/article/machine-learning-needs-human-helping-hand/>

Lachow, Irving, 'The upside and downside of swarming drones', *Bulletin of the Atomic Scientists*, 73:2, (February 2017), <http://www.tandfonline.com/doi/pdf/10.1080/00963402.2017.1290879?needAccess=true>

Lafrance, Adrienne, 'Machine Unlearning', *The Atlantic*, (18 March 2016), <https://www.theatlantic.com/technology/archive/2016/03/computers-brains-cybernetics/474273/>

Lafrance, Adrienne, 'When Robots Hallucinate', *The Atlantic*, (3 September 2015), <https://www.theatlantic.com/technology/archive/2015/09/robots-hallucinate-dream/403498/>

Liang, Percy, and Christopher Potts, 'Bringing Machine Learning and Compositional Semantics Together', *Annual Review of Linguistics*, 1.1, (13 April 2014), <https://web.stanford.edu/~cgpotts/manuscripts/liang-potts-semantics.pdf>

Lake, Brenden, and others, 'Building Machines that Learn and Think Like People', *Behavioural and Brain Sciences*, (2016), <https://arxiv.org/pdf/1604.00289.pdf>.

Lambert, Fred, 'US Marines test solar powered drones at annual energy expo', *UPI*, (28 June 2015), <http://www.upi.com/Business_News/Security-Industry/2015/06/28/US-Marines-test-solar-powered-drones-at-annual-energy-expo/6911435519387/>

Landau, Emily, and Ariel Bermant, 'Iron Dome protection: Missile defence in Israel's security concept', *Lessons of Operation Protective Edge*, (2014), <http://www.inss.org.il/he/wp-content/uploads/sites/2/systemfiles/SystemFiles/Iron%20Dome%20Protection_%20Missile%20Defense%20in%20Israel's%20Security%20Concept.pdf>

Lansner, A., 'Associative Processing in Brain Theory and Artificial Intelligence', Conference paper, *Springer Link*, (1986), <https://link.springer.com/chapter/10.1007/978-3-642-70911-1_12>

Lapin. Maksim, 'Image Classification with Limited Training Data and Class Ambiguity', PhD title, Saarland University, (June 2017), <http://scidok.sulb.uni-saarland.de/volltexte/2017/6909/pdf/lapin17phd.pdf>

Larter, David, 'Autonomous mine-hunting boat will be delivered to British Navy this winder', *Defense News*, (13 September 2017), <https://www.defensenews.com/digital-show-dailies/dsei/2017/09/13/autonomous-mine-hunting-boat-will-be-delivered-to-the-royal-navy-this-winter/>

Latiff, Robert, 'How Technological Advancements Will Shape the Future of the Battlefield', *Signature*, (13 October 2017), <https://www.signature-reads.com/2017/10/how-tech-advancements-will-shape-future-battlefield/>

Latiff, Robert, and PJ. McCloskey, 'With Drone Warfare, America Approaches the Robo-Rubicon', <http://reilly.nd.edu/people/adjunct-faculty/maj-gen-robert-latiff-ret/with-drone-warfare-america-approaches-the-robo-rubicon/>

Launchbury, John, 'A DARPA perspective on Artificial Intelligence', *DARPA slides*, undated, <https://www.darpa.mil/attachments/AIFull.pdf>

Lawson, Sean, 'Domestic 'Drones' Are the Latest Object of Threat Inflation', <http://www.forbes.com/sites/seanlawson/2014/04/18/domestic-drones-are-the-latest-object-of-threat-inflation/>

LeBaron, Michelle, 'Culture and Conflict', (July 2003), <http://www.beyondintractability.org/essay/culture_conflict>

LaCerra, Peggy, and Roger Bingham, 'The Adaptive Nature of Human Neurocognitive Architecture: An Alternative Approach', PNAS, (September 1998), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC21635/>

Lee, P. and S Wright, 'Killer Drones: Be Afraid or Ignore the Hype?', *Dleano.lu blog*, (4 December 2017), <http://delano.lu/d/detail/news/killer-drones-be-afraid-or-ignore-hype/163173>

Lekka, Anastasios, 'Guidance and path-planning systems for autonomous weapon', *NTNU*, (April 2014), <http://fossen.biz/home/PhD/thesis/Lekkas%202014.pdf>

Lesswrong.com, 'Superintelligence 12: Malignant Failure Modes', (2 December 2014), <http://lesswrong.com/lw/l9t/superintelligence_12_malignant_failure_modes/>

Letovzey, Jean-Louis, and Declan Whelan, 'Introduction to the Technical Debt Concept', *Agile Alliance,* undated, <https://www.agilealliance.org/wp-content/uploads/2016/05/IntroductiontotheTechnicalDebtConcept-V-02.pdf>

Levy, Alon Y., 'Combining Artificial Intgelligence and Databases in Data Integration', *AI Today*, Lecture notes in Computer Science, Vol 1600, Springer, (1999), <https://link.springer.com/chapter/10.1007%2F3-540-48317-9_10>

Lewis, Larry, 'Redefining Human Control: Lessons from the Battlefield for Autonomous Weapons', *CAN Washington*, (March 2018), <https://www.cna.org/cna_files/pdf/DOP-2018-U-017258-Final.pdf

Lexician.com, 'No Battle Plan Survives Contact With the Enemy', <http://lexician.com/lexblog/2010/11/no-battle-plan-survives-contact-with-the-enemy/>

Li, Ang, 'GPU performance Models and Optimization', Technische Universiteit Eindoven, (18 October 2016), <https://pure.tue.nl/ws/files/39759895/20161018_Li.pdf>

Lin, Fangzhen, 'On Moving Objects in Dynamic Domains', *Association for the Advancement of Artificial Intelligence*, (2012), <http://commonsensereasoning.org/2011/papers/Lin.pdf>

Linders, Ben, 'Dead code must be removed', *Info Q*, (9 February 2017), <https://www.infoq.com/news/2017/02/dead-code>

Liu, Song, 'Statistical learning approaches to change detection', JSPS-DC2 Tokyo Institute of Technology, (10 March 2014), <https://sheffieldml.github.io/slides/2014-03-18-song-liu.pdf>

Llinas, James, and others, 'Studies and Analyses of Aided Adversarial Decision Making: Phase 2 – Research on Human Trust in Automation', US Air Force Research Laboratory, (April 1998), <https://pdfs.semanticscholar.org/1424/5ae4ec038f3a3b9c737e40b9d289dc79a612.pdf>

Lockheed, Martin, 'Unmanned Systems', <http://www.lockheedmartin.co.uk/us/what-we-do/aerospace-defense/unmanned-systems.html>

Longa, Luca, 'Argumentation of Knowledge Representation, Conflict Resolution, Defeasible Inference and its Integration with Machine Learning', Machine for Health Informatics, *Computer Science*, Volume 9605, (10 December 2016), <http://www.topbots.com/4-different-approaches-natural-language-processing-understanding/>

Lopez-Paz, David, 'Towards a Learning Theory of Cause-Effect Inference', arXiv.1502.02398, (9 February 2015), <https://arxiv.org/pdf/1502.02398.pdf>

Louth, John, and Christian Moeling, 'Technological Innovation: The US Third Offset Strategy and the Future of Transatlantic Defense', *Policy Paper*, Armaments Industry European Research Group,

(December 2016), <http://www.iris-france.org/wp-content/uploads/2016/12/ARES-Group-Policy-Paper-US-Third-Offset-Strategy-December2016.pdf>

Lowe, David, 'Characterising complexity by the degrees of freedom in a radical basis function network', *Neurocomputing*, Volume 19, (April 1998), <http://www.sciencedirect.com/science/article/pii/S0925231297000659>

Lu, Clara, 'Why We are Still Light Years Away from Full Artificial Intelligence', *TechCrunch*, (2016), <https://techcrunch.com/2016/12/14/why-we-are-still-light-years-away-from-full-artificial-intelligence/>

Luck, Gary, and others, 'Joint Operations: Insights and Best Practices', Joint Warfare Center, US Joint Forces Command, (July 2008), <http://www.au.af.mil/au/awc/awcgate/jfcom/joint_ops_insights_july_2008.pdf>

Luger, George, 'Artificial Intelligence: Structures and Strategies for Complex Problem Solving', Pearson, 6th Edition, (2009), <http://iips.icci.edu.iq/images/exam/artificial-intelligence-structures-and-strategies-for--complex-problem-solving.pdf>

Lundgren, Carole, 'Recent Development in Neural Networks', *Appen*, (23 March 2018), <https://appen.com/recent-developments-neural-networks/>

Lynch, Justin, and Lauren Fish, 'Soldier Swarm: New Ground Combat Tactics for the Era of Multi-Domain Battle', Modern War Institute, West Point, (5 April 2018), <https://mwi.usma.edu/soldier-swarm-new-ground-combat-tactics-era-multi-domain-battle/>

Macauley, Thomas, 'The Future of Technology in Warfare: From AI Robots to VR Torture', *Techworld*, (13 January 2017), <https://www.techworld.com/security/future-of-technology-in-warfare-3652885/>

MacDonald, Fiona, 'A robot has just passed a classic self awareness test for the first time', *Science Alert,* (17 July 2015), <https://www.sciencealert.com/a-robot-has-just-passed-a-classic-self-awareness-test-for-the-first-time>

Macias, Amanda, 'The Pentagon is Trying to Figure Out the True Cost of its Costliest Weapon System, the F-35', *CNBC*, (28 February 2018), <https://www.cnbc.com/2018/02/28/pentagon-wants-to-know-true-cost-of-f-35-system.html>

Macias, Amanda, 'Weapons of the Future: Here's the New War Technologies Lockheed Martin is Pitching to the Pentagon', *CNBC*, (6 March 2018), <https://www.cnbc.com/2018/03/06/future-weapons-lockheed-martin-pitches-new-war-tech-to-pentagon.html>

Madrigal, Alexis, 'Drone Swarms are Going to be Terrifying and Hard to Stop', *The Atlantic*, *Technology*, (7 March 2018), <https://www.theatlantic.com/technology/archive/2018/03/drone-swarms-are-going-to-be-terrifying/555005/>

Mai, Luo, and others, 'Optimizing Network Performance in Distributed Machine Learning', *Hotcloud*, (2015), <https://www.usenix.org/system/files/conference/hotcloud15/hotcloud15-mai.pdf>.

Majumdar, David, 'Introducing the 5 Deadliest Weapons of the US Military' *The National Interest*, (27 December 2018), <https://nationalinterest.org/blog/buzz/introducing-5-deadliest-weapons-us-military-39907>

Malik, Jitendra, 'The 3 Rs of Computer Vision: Recognition, Reconstruction and Reorganization', University of Berkeley, EECS, *ScienceDirect*, Patterns Recognition Letters, (8 February 2016), <https://people.eecs.berkeley.edu/~shubhtuls/papers/prl16rrr.pdf>

O'Malley, Sean, 'Research sends high-flying drones soaring', RMIT University, <http://www.rmit.edu.au/news/all-news/2016/april/research-sends-highflying-drones-soaring>

Mandloi, Neeraj, and others, 'Smart motion sensing for autonomous robots', Biomedical Circuits and Systems Conference, *Submission paper,* (2014), <http://ieeexplore.ieee.org/document/6981777/>

Manikandan, Narayanan, 'Software design challenges in time series prediction using parallel implementation of artificial neural networks', *The Scientific World Journal*, Article ID 670 9352, (2016) <http://do.doi.org/10.1155/2016/6709352>

Mansoor, Peter, 'The Next Revolution in Military Affairs', *Strategika*, Hoover Institute, Issue 39, (15 March 2017), <https://www.hoover.org/research/next-revolution-military-affairs>

Marcus, Gary, 'Deep Learning: A Critical Appraisal', New York University, arXiv: 1801:00631, (2 January 2018), <https://arxiv.org/pdf/1801.00631.pdf>

Marcus, Jonathan, 'Robot Wars: Lethal machines coming of age', <http://www.bbc.co.uk/news/magazine-21576376>

Maresca, Louis, '20 Years since the ICJ advisory opinion and still difficult to reconcile with international humanitarian law', *Humanitarian Law and Policy*, blog, (8 July 2016), <http://blogs.icrc.org/law-and-policy/2016/07/08/nuclear-weapons-20-years-icj-opinion/>

Markman, John, 'Laser Weapons Set to Boost Military Night at the Speed of Light', *John Markman's Pivotal Point*, (8 November 2017), <https://www.markmanspivotalpoint.com/investing/laser-weapons-set-boost-military-might-speed-light/>

Marketsandmarkets.com, 'Military Robots Market to be worth 21.11 US$ Billions by 2020', <http://www.marketsandmarkets.com/PressReleases/military-robots.asp>

Markoff, John, 'Computer Scientists are Poised for Revolution on a Tiny Scale', *NY Times*, Technology, (1 November 1999), <https://archive.nytimes.com/www.nytimes.com/library/tech/99/10/biztech/articles/01nano.html>

Marr, Bernard, 'What is the Difference between Artificial Intelligence and Machine Learning?', *Forbes Magazine*, Technology. 6 (December 2016), <https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#1a4e99432742>

Marra, William, and Sonia McNeil, 'Understanding the loop: regulating is the next generation of war machines', *Hartford Journal of law and public policy*, volume 36, (2012), <http://www.harvard-jlpp.com/wp-content/uploads/2013/05/36_3_1139_Marra_McNeil.pdf>

Marsh, Henry, 'Can Man Ever Build a Mind?', *Financial Times*, (10 January 2019), para 9 of 24, <https://www.ft.com/content/2e75c04a-0f43-11e9-acdc-4d9976f1533b>

Marshall, Michael, 'Timeline: Weapons Technology', *New Scientist*, (7 July 2009), <https://www.newscientist.com/article/dn17423-timeline-weapons-technology/>

Masters, Jonathan, 'Targeted Killings', <http://www.cfr.org/counterterrorism/targeted-killings/p9627>

Mataric, 'Issues and approaches in the design of collective autonomous agents', *Robotics and Autonomous Systems*, 16, (1995), <http://crr.eng.auburn.edu/Bibliography/Mataric-Issues%20and%20approaches%20in%20design%20of%20collective%20autonomous%20agents.pdf>

Matthews, Iain, and others, 'The Template Update Problem', Robotics Institute, Carnegie Mellon University, (2004), <http://www.ri.cmu.edu/pub_files/pub4/matthews_iain_2004_1/matthews_iain_2004_1.pdf>

Matthews, William, and Ana Gheorghu, 'Repetition, expectation and the perception of time', *ScienceDirect*, (16 February 2016), <http://www.sciencedirect.com/science/article/pii/S2352154616300420>

Mathworks.com, 'On genetic algorithms', <https://www.mathworks.com/discovery/genetic-algorithm.html>

Matthias, Andreas, 'Is the concept of an Ethical Governor ethically sound?' <http://www.academia.edu/473656/Is_the_Concept_of_an_Ethical_Governor_Philosophically_Sound>

Mattis, James, and Frank Hoffman, 'Future Warfare: The Rise of Hybrid Wars' <http://milnewstbay.pbworks.com/f/MattisFourBlockWarUSNINov2005.pdf>

Maucione, Scott, 'Navy Wants to Cut Weapons Testing Time with Simulations and Modelling', Federal News Radio, (12 January 2018), <https://federalnewsradio.com/defense/2018/01/navy-wants-to-cut-weapons-testing-time-with-simulations-and-modeling/>

MacCormick, Neil, 'Reasonableness and Objectivity', *Notre Dame Law Review*, 74, Issue 5, Article 6, (1999), <https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=1648&context=ndlr>

McNaughton, Matthew, 'Planning algorithms for real-time motion planning', Carnegie Mellon University, Dissertations, *Paper 179*, (2011), <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1180&context=dissertations>

McConaghy, Trent, 'Blockchains for Artificial Intelligence', *Bigchain* website, <https://blog.bigchaindb.com/blockchains-for-artificial-intelligence-ec63b0284984>

McCullagh, Declan, 'New Weapons for a New War', *Wired Magazine* <http://archive.wired.com/politics/law/news/2001/10/47395?currentPage=all>

McHale, John, 'Power Electronics Design Trending Smaller and More Efficient', Military Embedded Systems, undated, <http://mil-embedded.com/articles/power-trending-smaller-more-efficient/>

T McKelvey, '*America's Shadow Warriors; Legal Dimensions of Special Forces and Targeted Warfare,* <http://www.bcics.northwestern.edu/documents/workingpapers/Buffett_10-007_McKelvey.pdf>

McKinsey & Company, 'Disruptive Technologies: Advances that will Transform Life, Business and the Global Economy', McKinsey Global Forum, (May 2013), <https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Disruptive%20technologies/MGI_Disruptive_technologies_Executive_summary_May2013.ashx

McCleary, Paul, 'The Third Offset May Be Dead But No One Know What Comes Next', *Foreign Policy*, (18 December 2017), <https://foreignpolicy.com/2017/12/18/the-pentagons-third-offset-may-be-dead-but-no-one-knows-what-comes-next/>

Media Source magazine, 'Drones: Are they watching you?', <http://mentalfloss.com/article/30669/9-weapons-failed-spectacularly-and-1-possibly-didn%E2%80%99t>

Medium.com, 'The AI data Apocalypse', <https://medium.com/@peopleio/the-ai-data-apocalypse-1375ac47ffe4>

Melendez, Steven, 'The Rise of the Robots: What the future holds for the world's armies', *FastCompany.com*, (12 June 2017), <https://www.fastcompany.com/3069048/where-are-military-robots-headed>

Metcalf, Jacob, 'Ethics Codes: History, Context and Challenges', Council for Big Data, *Ethics and Society,* (9 November 2014), <https://bdes.datasociety.net/council-output/ethics-codes-history-context-and-challenges/>

Milley, Mark, speech to RUSI Land War Conference, (27 June 2017), <https://rusi.org/sites/default/files/20170627-rusi_lwc17-gen_milley.pdf

Melli, Roberto, 'Artificial Intelligence in Component Design', *Exergy (sic), energy system analysis and optimization*, Volume III, <http://www.eolss.net/sample-chapters/c08/E3-19-04-04.pdf>

Mericli, Cetin, and others, 'Task Refinement for Autonomous Robots using Complementary Corrective Human Feedback', *International Journal of Advanced Robotic Systems*, 8.2, (2011), <http://www.cs.cmu.edu/~mmv/papers/11ijars-cetin.pdf>

Metz, Cade, 'Google is 2 Billion Lines of Code – and It's All in One Place', *Wired*, (16 June 2015), https://www.wired.com/2015/09/google-2-billion-lines-codeand-one-place/>

Meyer, Josh, 'What World War Three might look like', *Vice*, 30 June 2015, <https://www.vice.com/en_au/article/gqmqk9/how-world-war-iii-might-start-with-the-cyber-weapons-of-the-future-630>

Michels, Dave, 'AI Heading Back to the Trough: The Expectations Over Artificial Intelligence are Becoming Too Inflated', *Networld World*, (11 July 2017), <https://www.networkworld.com/article/3206313/internet-of-things/ai-heading-back-to-the-trough.html>

Mitchum, Rob, 'Can the Connections between the 100 Billion Neurons in the Brain be Mapped?', *Forefront*, University of Chicago Medicine, (1 June 2018), https://www.uchicagomedicine.org/neurosciences-articles/can-100-billion-neurons-be-mapped>

Microsoft Research, 'From machine learning to machine reasoning', <https://www.microsoft.com/en-us/research/publication/from-machine-learning-to-machine-reasoning/>

Le Miere Jason, 'Russia developing a swarm of autonomous drones in the new arms race with US, China', *Newsweek*, (15 May 2017), <http://www.newsweek.com/drones-swarm-autonomous-russia-robots-609399>

Mikolajczyk, Krystian, and Cordelia Schmid, 'A Performance Evaluation of Local Descriptors', *IEEE Transactions on Pattern Analysis and Machine Learning*, Vol.27, No.10, (October 2005), <https://www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/mikolajczyk_pami2004.pdf>

Miller, Kenneth, 'Is China Winning the Innovation Race?', *LeapsMag*, (19 June 2018), <https://leapsmag.com/is-china-winning-the-innovation-race/>

Miller, Jack, 'Strategic Significance of Drone Operation for Warfare', E-International Relations Students, (18 August 2013), <http://www.e-ir.info/2013/08/19/strategic-significance-of-drone-operations-for-warfare/>

Millman, Noah, 'Psychological Rationales for Threat Inflation', <http://www.theamericanconservative.com/millman/psychological-rationales-for-threat-inflation/>

Mindell, David, 'Driverless cars and the myth of autonomy', *Huffington Post*, (14 October 2015), <https://www.huffingtonpost.com/david-a-mindell/driverless-cars-and-the-myths-of-autonomy_b_8287230.html>

Ministry of Defence, 'Future Character of Conflict', *Strategic Trends Programme*, <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/33685/FCOCReadactedFinalWeb.pdf>

Minsky, Marvin, 'Steps towards Artificial Intelligence', *IRE*, (January 1961), <http://courses.csail.mit.edu/6.803/pdf/steps.pdf>

Missile Technology Control Regime, <http://www.mtcr.info/english/>

Mitchell, Paul, T., 'Three Laws Safe? Autonomous Robots and Warfare', <http://canadianmilitaryhistory.ca/three-laws-safe-autonomous-robots-and-warfare-by-dr-paul-t-mitchell/>

Moe, Terry, 'Vested Interests and Political Institutions' <https://politicalscience.stanford.edu/sites/default/files/images/vested%20interests_psq%20final%20June%202014.pdf>

Mohamed, Omar, 'The Master Strategist: Clausewitzian Genius', *Real Clear Defense*, (27 November 2016) <https://www.realcleardefense.com/articles/2016/11/28/the_master_strategist_clausewitzian_genius_110387.html>

Mohamad, Saad, 'Active Learning for Data Streams', Bournemouth University, *IMT Lille Douai*, Abstract, (October 2017), http://eprints.bournemouth.ac.uk/29901/1/MOHAMAD%2C%20Saad_Ph.D._2017.pdf

Molla, Rani, 'Intel's Drone Light Show Never Got Off the Ground for the 2018 Winter Olympic Opening Ceremony', *Recode*, (10 February 2018), <https://www.recode.net/2018/2/10/16998652/drones-guinness-world-record-pyeongchang-2018-winter-olympics>

Moon Cronk, Terri, 'Cost saving pilot programs to support war fighter autonomy' <http://www.aerotechnews.com/news/2013/06/21/cost-saving-pilot-programs-to-support-war-fighter-autonomy/>

Moral Foundations.org, <http://www.moralfoundations.org/>

Moran, D, 'Tackling RF-Denied Environments', *Harris Corporation*, (9 March 2017), <https://www.harris.com/perspectives/innovation/tackling-rf-denied-environments>

Moravec, Hans, 'When Will Computer Hardware Match the Human Brain?', *Journal of Evolution and Technology*, Volume 1, (1998), <http://www.realtechsupport.org/UB/WBR/texts/Moravec_ComputerMatchHumanBrain_1998.pdf>

Morgan, Lisa, 'Nine causes of data misinterpretation', *InformationWeek,* (7 July 2017), <http://www.informationweek.com/big-data/big-data-analytics/9-causes-of-data-misinterpretation/d/d-id/1321338>

Morgenthalter, David and others, 'Searching for build debt: experiences managing technical debt at Google', <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/37755.pdf>

Mortimer, Gary, 'Aurora Awarded AFRL Urban Beat Cop Program', <http://www.suasnews.com/2011/12/aurora-awarded-afrl-urban-beat-cop-program/>

Moss, Richard, 'Software writing Software and the broader challenge of computational creativity', *New Atlas,* (3 March 2015), <https://newatlas.com/creative-ai-computational-creativity-challenges-future/36353/>

Moyes, Richard, 'Key Elements of meaningful human control', *Article 36 Briefing Paper*, CCW Meeting of Experts on lethal autonomous weapon systems, (April 2016), <http://article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf>

Mu, Yadong, and others, 'Stochastic Gradient Made Stable: A Manifold Propagation Approach for Large Scale Optimisation', arXiv:1506.08350v2, (12 January 2016), <https://arxiv.org/pdf/1506.08350.pdf>

Mujtaba, Hassan, 'NVIDIA Pascal shatters 3 GHz GPU frequency record –highest clock speed ever recorded on a graphics chip', *WCCFTech*, (18 December 2016), <http://wccftech.com/nvidia-pascal-gpu-frequency-world-record-3-ghz/>

Muehlhauser, Luke, 'What is AGI?', MIRI Machine Intelligence Research Institute, (11 August 2013), <https://intelligence.org/2013/08/11/what-is-agi/>

Murali, Shravan, 'An Analysis on Computer Vision Problems', *Medium.com,* (13 September 2017), <https://medium.com/deep-dimension/an-analysis-on-computer-vision-problems-6c68d56030c3>

Murphey, Yi, and others, 'Neural Learning from Unbalanced Data', *Applied Intelligence 21*, (2004), <http://sci2s.ugr.es/keel/pdf/specific/articulo/NL-Unbalanced-data.pdf>

Musandu, Nyagudi, 'Humnitarian Algorithms: A Codified Key Safety Switch Protocol for Lethal Autonomy', Nairobi, arXiv preprint arXiv 1402.2206 (2014), <https://arxiv.org/pdf/1402.2206.pdf>

Mytton, Oliver, and others, 'Introducing New Technology Safely', *QSHC*, (2011), <https://qualitysafety.bmj.com/content/qhc/19/Suppl_2/i9.full.pdf>

Nasiriany, Soroush, and others, 'A Comprehensive Guide to Machine Learning', Section 4.3 ('*Gradient Descent'),* University of California at Berkeley, (13 August 2018), http://snasiriany.me/files/ml-book.pdf

NATO Standardisation Office, 'Allied Joint Publication 3.9: Allied Joint Doctrine for Joint Targeting', NSO, Edition A, Version 1, (April 2106), <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/628215/20160505-nato_targeting_ajp_3_9.pdf>

National Academy of Science, Engineering & Medicine, 'Autonomies for civil aviation: Toward a new era of flight', IAP, 2014, <https://www.nap.edu/read/18815/chapter/3>

National Audit Office, 'A Short Guide to the Ministry of Defence', p.9, (2017), <https://www.nao.org.uk/wp-content/uploads/2017/09/A-short-guide-to-the-Ministry-of-Defence.pdf>

Navy Matters, 'New Anti-Ship Missiles', <http://navy-matters.blogspot.co.uk/2012/08/new-anti-ship-missiles.html>

US Navy, <http://www.onr.navy.mil/Media-Center/Fact-Sheets/Ion-Tiger.aspx>

NATO Office, 'Defence Planning Process', <http://www.nato.int/cps/en/natohq/topics_49202.htm>

NCCD Conference, 'Threats to peace and disarmament – The way forward', *NCCD*, (May 2016), <http://www.converge.org.nz/pma/aws0506.pdf>

Nebot, E., and others, 'Navigational Algorithms for Autonomous Machines in Off-Road Applications', *Journal of Autonomous Robots*, 14, (2000), <http://www8.cs.umu.se/research/ifor/dl/LOCALIZATION-NAVIGATION/Navigation%20Algorithms%20for%20Autonomous%20Machines%20in%20Off-Road%20Applications.pdf>

Nesnas, Issa, 'CLARAty: Challenges and Steps Towards Reusable Robotic Software', *International Journal of Advanced Robotic Systems*, Volume 3, Number 1, ('*Challenges*'), (2012), <http://journals.sagepub.com/doi/pdf/10.5772/5766>

Neuromorphic Technologies, '"Spaghetti Code": Complexity and Artificial Intelligence', *Admin blog*, (27 March 2018), <http://fernandojimenezmotte.com/mi-articulo/spaghetti-code-complexity-and-artificial-intelligence/>

Niccolini, Marta, and others, 'Cooperative control for multiple autonomous vehicles using descriptor functions', *Journal of Sensor and Actuator Networks*, 3, (2014)

Nikerson, Jeffrey, and Richard Reilly, 'A Model for Investigating the Effects of Machine Autonomy on Human Behaviour', Proceedings of the 37th International Conference in Security Science, (2004), <https://web.stevens.edu/jnickerson/ETSIB01.PDF>

Nirmal, Dinesh, 'How to Decide: Machine Learning and the Science of Choosing', *Medium.com* blog, (20 March 2017), <https://medium.com/inside-machine-learning/how-to-decide-machine-learning-and-the-science-of-choosing-7a0d70059079>

Norman, Donald, and others, 'Effect and Machine Design: Lessons for the Development of Autonomous Machines', *IBM Systems Journal*, Vol 22, No 1, (2003), <https://ieeexplore.ieee.org/document/5386841>

Norman, Donald, 'The Problem of Automation; Inappropriate feedback and interaction', ICS Report 8904, Institute for Cognitive Science, University of California, (1989), <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19900004678.pdf>

Northrop Grumman, 'X-47B UCAS', <http://www.northropgrumman.com/Capabilities/X47BUCAS/Pages/default.aspx>

Northrop Grumman, 'Unmanned Systems' <http://www.northropgrumman.com/Capabilities/Unmannedsystems/Pages/default.aspx?utm_source=PrintAd&utm_medium=Redirect&utm_campaign=Unmanned+Redirect>

Norton, Travis, 'Staffing for Unmanned Aerial Systems (UAS)', Institute for Defense Analyses IDA, (June 2016), <https://prhome.defense.gov/Portals/52/Documents/MRA_Docs/TFM/Reports/F2108340_TFMR-Staffing%20for%20Unmanned%20Aircraft%20Systems%20(UAS)%20Operations-ForPIIWork-DM.pdf>

Nguyen, Anh, 'Deep networks are easily fooled: high confidence predictions for unrecognizable images', Computer vision and pattern recognition, *IEEE*, (2015), <http://arxiv.org/pdf/1412.1897v4.pdf>

Nguyen, Cu, and others, 'Evolutionary testing of autonomous software agents', Autonomous agents and multi-agent systems, 25.2, (2012), <https://nms.kcl.ac.uk/michael.luck/resources/aamas09d.pdf>

Oberhaus, Daniel, 'Watch 'Slaughterbot': A Warning about the Future of Killer Robots', *Motherboard*, (13 November 2017), <https://motherboard.vice.com/en_us/article/9kqmy5/slaughterbots-autonomous-weapons-future-of-life>

Ohlin, Jens David, 'Is Jus In Bello in Crisis?', *Cornell Law Faculty Publications*, (March 2013), <https://scholarship.law.cornell.edu/cgi/viewcontent.cgi?article=2475&context=facpub>

De Oliveira, Vera Lucia Menezes, and others, 'The Complex Nature of Autonomy', *Delta Online*, (2008), http://www.scielo.br/pdf/delta/v24nspe/04.pdf

Olsthoorn, Peter, 'Military Ethics and Virtues: An interdisciplinary approach for the twenty-first century', Cass Military Studies, Routledge, (2011), <https://philpapers.org/archive/OLSMEA.pdf>

Omohundro, Stephen, 'The Basic AI Drives', Self-aware Systems abstract, Paulo Alto, California, <https://selfawaresystems.files.wordpress.com/2008/01/ai_drives_final.pdf>

OpenSource.com, 'What is open source computing?', <https://opensource.com/resources/what-open-source>

Ordoukhanian, Edwin, 'Resilience Concepts for UAV Swarms', *CSSE*, USC Viterbi, (March 2016), <https://pdfs.semanticscholar.org/presentation/c95f/e77e25df2093fac4a5da723b43062c7e4d24.pdf>

Organisation For The Prohibition of Chemical Weapons, 'Chemical Weapons Convention', <http://www.opcw.org/chemical-weapons-convention/>

Orseau, Laurent, and Mark Ring, 'Self-Modification and Mortality in Artificial Agents', International Conference on Artificial General Intelligence, *Lecture Notes on Computer Science,* Volume 6830, Springer, Berlin, Heidelberg

Osborn, Kris, 'Swarming mini drones: Inside the Pentagon's plan to overwhelm Russian and Chinese air defences', *The Buzz*, National Interest, (10 May 2016), <http://nationalinterest.org/blog/the-buzz/swarming-mini-drones-inside-the-pentagons-plan-overwhelm-16135>

Osmeyer, Jared, and Lindsay Cowell, 'Machine Learning on Sequential Data using a Recurrent Weighted Average', Cornell University Library, (4 May 2017), <https://arxiv.org/abs/1703.01253>

Osoba, Osonde, 'The Risks of Bias and Errors in Artificial Intelligence', Rand Corporation Publishing, An Intelligence in our Image, (2017), <https://www.rand.org/pubs/research_reports/RR1744.html>

Osoba, Osonde, and William Welser, 'An Intelligence in our Image; the risk of bias and errors in Artificial Intelligence', Rand Corporation, Center for Global risk and security, (2017),

<https://www.rand.org/content/dam/rand/pubs/research_reports/RR1700/RR1744/RAND_RR1744.pdf>

G Ostergaard, 'Resisting the Nation State; The Pacifist and Anarchist Tradition', (1982), <https://theanarchistlibrary.org/library/geoffrey-ostergaard-resisting-the-nation-state-the-pacifist-and-anarchist-tradition>

Osterloh, Rick, 'The Best Hardware, Software and AI –Together', *The Keyword,* Google Company Blog, (4 October 2017), <https://www.blog.google/technology/ai/the-best-hardware-software-and-ai-together/>

Admiral Owens, 'Lifting the Fog of War', <http://www.nytimes.com/books/first/o/owens-fog.html>

Oxford Dictionary, 'Definition of battlefield', <http://www.oxforddictionaries.com/definition/english/battlefield>

Paden, Brian, and others, 'A Survey of Motion Planning and Control Techniques for Self-driving Urban Vehicles', *MIT*, (25 April 2016), <https://arxiv.org/pdf/1604.07446.pdf>

Pan, Wei, and others, 'DropNeuron: Simplifying the Structure of Deep Neural Networks', 29th Conference on Neural Information Processing Systems, NIPS, (2016), <https://arxiv.org/pdf/1606.07326.pdf>

Pannu, Adarsh, and Steve Moore, 'Three Reasons Machine Learning Models Go Out Of Sync', *Inside Machine Learning,* (27 November 2017), <https://medium.com/inside-machine-learning/three-reasons-machine-learning-models-go-out-of-sync-a101b2cdca54>

Parashar, Angam, 'How Artificial Intelligence is Outpacing Humans', Linked In Oped, (11 July 2017), <https://www.linkedin.com/pulse/how-artificial-intelligence-outpacing-humans-angam-parashar>

Park, Frank C., and Kevin M. Lynch, 'Introduction to Robotics: Mechanics, Planning and Control', Northwestern University Publishing, (20 September 2016), <http://hades.mech.northwestern.edu/images/2/2a/Park-lynch.pdf>

Parliament of Finland, '*Long-term Challenges of Defence: Final Report of Parliamentary Assessment Group*', Finland Parliament, (May 2014), <https://www.eduskunta.fi/FI/tietoaeduskunnasta/julkaisut/Documents/ekj_5+2014.pdf

Parloff, Roger, 'Why Deep Learning is suddenly changing your life: Decades-old discoveries are now electrifying the computing industry', *Fortune*, (28 September 2016), <http://fortune.com/ai-artificial-intelligence-deep-machine-learning/>

Parrott, David, 'The Military Revolution in Early Europe', <http://www.historytoday.com/david-parrott/military-revolution-early-europe>

Paul, Kari, 'When Killer Robots Arrive, They'll Get hacked', *Motherboard*, (24 February 2015), <https://motherboard.vice.com/en_us/article/ezvknz/when-the-killer-robots-arrive-theyll-get-hacked>

Paul, Rohan, and others, 'Grounding Spatial Concepts for Language Interaction with Robots', *Proceedings of 26th International Joint Conference of Artificial Intelligence*, (2017), <https://www.ijcai.org/proceedings/2017/0696.pdf>

Paur, Jason, 'FAA Experiments With Integrating Drones in Civil Airspace', *Wired Magazine*, (2010) <http://www.wired.com/autopia/2010/06/faa-uav-civil-airspace/>

Pellerin, Cheryl, *Work: '*Human-Machine Teaming represents Defense Technology Future', Department of Defense Subscription, (8 November 2015), <https://www.defense.gov/News/Article/Article/628154/work-human-machine-teaming-represents-defense-technology-future/>

Pellerin, Cheryl, 'Deputy Secretary: Third Offset Strategy bolsters America's Military Deterrence', *DOD News*, Defense Media Activity, (31 October 2016), <https://www.defense.gov/News/Article/Article/991434/deputy-secretary-third-offset-strategy-bolsters-americas-military-deterrence/>

Penna, Charles, 'A Reality Check on Military Spending', *Issues in Science and Technology*, Volume XXI, Issue 4, (Summer 2005), <http://issues.org/21-4/pena/>

Peralta, Eyder, 'Weighing the good and the bad of autonomous killer robots in battle', *All Tech Considered*, (28 April 2016), <https://www.npr.org/sections/alltechconsidered/2016/04/28/476055707/weighing-the-good-and-the-bad-of-autonomous-killer-robots-in-battle>

Perret, Bradley, 'China's Inflation-Adjusted Defense Budget', *Aviation Week* <http://aviationweek.com/awin/china-s-inflation-adjusted-defense-budget-75>

Perry, Dewayne, and others, 'Parallel changes in large-scale software development: An observational case study' International Conference on Software Engineering, ICSE98, (2001), <https://pdfs.semanticscholar.org/933a/98846a0adc29f2bf8f6557c9c15956562a07.pdf>

Petersen, Karen, 'General Concepts for Human Supervision of Autonomous Robot Teams', Technische Universitat Darmstadt, (23 May 2013), <http://tuprints.ulb.tu-darmstadt.de/3873/7/dissertation.pdf>

Petri, Cristina, 'Decision Trees', Cluj Napoca, (2010), <http://www.cs.ubbcluj.ro/~gabis/DocDiplome/DT/DecisionTrees.pdf>

Philips, Helen, 'Introduction: The Human Brain', *New Scientist*, (4 September 2006), <https://www.newscientist.com/article/dn9969-introduction-the-human-brain/>

Phoha, Shashi, 'Machine Perception and Learning Grand Challenge: Situational Intelligence Using Cross-Sensory Fusion', *Frontiers in Robotics and AI* (Sensor Fusion and Machine Perception), (6 October 2014), <https://www.frontiersin.org/articles/10.3389/frobt.2014.00007/full>

Pickering, Charles, 'How AI is Paving the Way for Autonomous Cars', *Engineer*, (15 August 2017), <https://www.theengineer.co.uk/ai-autonomous-cars/>

Pierreault, Charles, 'The Pace of Cultural Evolution', *PLOS.org,* (14 September 2012), <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0045150>

Pilling, Michael, 'Issues regarding the Future Application of Autonomous systems to Command and Control', Australian Government Department of Defense, Joint Operations Division, DSTO-TR-3112, (June 2015), <https://pdfs.semanticscholar.org/f4cd/63e3db777cc2f43d98aa82875802a9079494.pdf>

Plimmer, G., and K Stacey, 'MoD's latest procurement plan still 'a mess', *FT.com* <http://www.ft.com/cms/s/0/d4671554-c2ee-11e3-94e0-00144feabdc0.html#axzz32lGES2Yq>

Polyzotis, Neoklis, and others, 'Data Management Challenges in Production Machine Learning', *Google Research*, Proceedings of the 2017 ACM International Conference on Management of Data, Chicago, <http://delivery.acm.org/10.1145/3060000/3054782/p1723-polyzotis.pdf?ip=86.138.190.207&id=3054782&acc=OA&key=4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E5945DC2EABF3343C&__acm__=1535392353_07093e3afa2bcee2d87348db37d9ca7c>

Poole, David, 'Probabilistic Conflicts in a Search Algorithm for Estimating Posterior Probability Problems in Bayesian Networks', Department of Computer Studies, University of BC, Vancouver, (May 1996), <http://www.cs.ubc.ca/~poole/papers/seaalg.pdf>

Potember, Richard, 'Perspectives on Research in Artificial Intelligence and Artificial General Intelligence Relevant to DoD', JRR-16-Task-003, <https://fas.org/irp/agency/dod/jason/ai-dod.pdf>

Pradeep, Vijay, and others, 'Calibrating a multi-arm multi-sensor Robot', Willow Grange Inc, (undated), ('*Choosing what to calibrate*') <http://www.willowgarage.com/sites/default/files/calibration.pdf>

Preetham, V.V., 'Back Propagation – How Neural Networks learn Complex Behaviours', Autonomous Agents #AI, <https://medium.com/autonomous-agents/backpropagation-how-neural-networks-learn-complex-behaviors-9572ac161670#.qwz64wcu6>

Preeyanon, Likit, and others, 'Reproducible Bioinformatics Research for Biologists', <https://bic-berkeley.github.io/psych-214-fall-2016/_downloads/Brown_chapter.pdf>

PricewaterhouseCoopers, 'The New Hire: How a new generation of robots is transforming manufacturing', *Zpryme Research survey*, (February 2014), <https://www.pwc.fi/fi/palvelut/tiedostot/industrial-robot-trends-in-manufacturing-report.pdf>

Privitera, Claudio, and Lawrence Stark, 'Algorithms for Defining Visual Regions-of-Interest: Comparison with Eye Fixations', IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, 9, (September 2000), <https://pdfs.semanticscholar.org/60c2/b03024f89c3f67d05d6b60d4ba1b032942c4.pdf>.

Pryor, Benjamin Alan, 'Assessing the Army's Software Patch Management Process', Defense Acquisition University, Aberdeen MD, '*Problem Statement*', (4 March 2016), <http://www.dtic.mil/dtic/tr/fulltext/u2/1040604.pdf>

Queguiner, Jean-Francois, 'Precautions under the law governing the conduct of hostilities', <http://www.icrc.org/eng/assets/files/other/irrc_864_queguiner.pdf>

Quinlan, J.R., 'Induction of Decision Trees', Kluwer Academic Publishers, Machine Learning, 1, (1981), <http://hunch.net/~coms-4771/quinlan.pdf>

Quora.com, 'Degrees of freedom', <https://www.quora.com/How-does-one-calculate-a-robots-DOF-degrees-of-freedom-in-the-strict-sense-of-mobility>

Rachleff, Andy, 'What 'disrupt' really means', *Techcrunch.com*, (16 February 2013), <https://techcrunch.com/2013/02/16/the-truth-about-disruption/>

Rahm, Erhard, and Hong Hai Do, 'Data Cleaning: Problems and Current Approaches', University of Leipzig, <http://www.betterevaluation.org/sites/default/files/data_cleaning.pdf>

Rafael factsheet, 'Iron Dome land-based weapon systems', <http://www.rafael.co.il/SIP_STORAGE/FILES/6/3336.pdf>

Rahwan, Iyad, and Manuel Celbrian, 'Machine Behaviour Needs to be and Academic Discipline', *Nautilus*, (29 March 2018), <http://nautil.us/issue/58/self/machine-behavior-needs-to-be-an-academic-discipline>

Rawnsley, Adam, 'CIA Drone Targeting Techniques', *Wired*, (7 August 2009), <https://www.wired.com/2009/07/infrared-beacons-guiding-cia-drone-strikes-qaeda-claims/>

Raytheon factsheet, 'Phalanx weapon system', <http://www.mobileradar.org/Documents/Ray_Phalanx.pdf>

Raytheon Patent US6952001B2, 'Integrity Bound Situational Awareness and Weapon Targeting', (2005), <https://patents.google.com/patent/US6952001B2/en>

RDECOM, 'Future Soldier 2030 Initiative', US Army Soldier RD&E Center, (February 2009), <https://www.wired.com/images_blogs/dangerroom/2009/05/dplus2009_11641-1.pdf>

Reaching Critical Will, 'Fully Autonomous Weapons', <http://www.reachingcriticalwill.org/resources/fact-sheets/critical-issues/7972-fully-autonomous-weapons>

Reby, David, and others, 'Artificial Neural Networks as a Classification Method in the Behavioural Sciences', *Behavioural Processes*, Elsevier, Volume 40, (1997), <http://www.lifesci.sussex.ac.uk/cmvcr/Publications_files/rebyproc.pdf>

Reiner, Adam, 'Effects of Automatic Target Detection on Detection and Identification Performance', Department of Industrial Engineering, University of Toronto, (2016), <https://tspace.library.utoronto.ca/bitstream/1807/70554/1/Reiner_Adam_J_201511_MAS_thesis.pdf>

Remote Control, 'Hostile drones: the use of drones by non-State actors against British targets', *Remote Control Project,* The Oxford Research Group, (2013), <http://remotecontrolproject.org/hostile-drones-the-hostile-use-of-drones-by-non-state-actors-against-british-targets/>

Repetton, Anthony, 'The Problem with Back-Propagation', *Towards Data Science*, (18 August 2017), <https://towardsdatascience.com/the-problem-with-back-propagation-13aa84aabd71>

Richards, Andrew, 'The Challenges of Delivering Massive Parallelism to Real-World Software', *Codeplay presentation, UKMAC* (May 2016), <http://conferences.inf.ed.ac.uk/UKMAC2016/slides/Andrew_Richards_The_Challenges_of_Delivering_Massive_Parallelism_to_Real-World_Software.pdf>

Ricks, Thomas, 'The widening gap between Military and Society', *Atlantic Magazine*, <https://www.theatlantic.com/magazine/archive/1997/07/the-widening-gap-between-military-and-society/306158/, July 1997>

Ricks, Thomas, 'Staff Planning: It's all about examining assumptions and then re-examining them', (12 January 2016), <http://foreignpolicy.com/2016/01/12/staff-planning-its-all-about-examining-assumptions-and-then-re-examining-them/>

Rivera, Jose, 'Self-calibration and optimal response in intelligent sensors design based on artificial neural networks', *Sensors* – Basel, 8, (August 7, 2007), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3814866/>

www.rmit.edu.ai, 'Highflying Drones', <http://www.rmit.edu.au/news/all-news/2016/april/research-sends-highflying-drones-soaring>

Roberts, Megan and Kyle Evanoff, 'Can Civil Society Succeed in its Quest to Ban Killer Robots?', *World Politics Review,* 17 November 2017, https://www.worldpoliticsreview.com/articles/23636/can-civil-society-succeed-in-its-quest-to-ban-killer-robots>

Roberts, Colin, 'Killer Robots: Moral Concerns versus Military Advantage', *The National Interest,* (3 November 2016), <http://nationalinterest.org/blog/the-buzz/killer-robots-moral-concerns-vs-military-advantages-18277>

Roberts, Kristin, 'When the Whole World Has Drones', <http://www.nationaljournal.com/magazine/when-the-whole-world-has-drones-20130321>

Robinson, Darryl, 'How Command Responsibility Got So Complicated: A Culpability Contradiction, its Obfuscation, and a Simple Solution', *Melbourne Journal of International* Law, 13, 1, (2012), <https://law.unimelb.edu.au/__data/assets/pdf_file/0003/1687242/Robinson.pdf>

Robotic World blog, 'The use and advantages of military robots', <http://minerrobot.weebly.com/the-use-and-advantages-of-military-robots.html>

Roff, Heather, 'The Results of the UN's Debate on Killer Robots Used in the Battlefield', <http://canadianawareness.org/2014/05/the-results-of-the-uns-debate-on-killer-robots-used-in-the-battlefield/>

Roff, Heather, 'The Self-Fulfilling Prophesy of High-tech War', *Duck of Minerva*, (29 December 2015), <http://duckofminerva.com/2015/12/the-self-fulfilling-prophecy-of-high-tech-war.html>

Rogers, Adam, 'The Dismal Science remains Dismal, say Scientists', *Wired magazine*, Science, (14 November 2017), <https://www.wired.com/story/econ-statbias-study/>

Rogoway, Tyler, 'Let's Talk About Those F-35 Kill Ratio Reports from Red Flag', *TheDrive.com*, (8 February 2017), <http://www.thedrive.com/the-war-zone/7488/lets-talk-about-those-f-35-kill-ratio-reports-from-red-flag>

Romjue, John, and others, 'Prepare the Army for War: A Historical Overview of the Army Training and Doctrine Command, 1973-1993', *TRADOC Historical Series*, Office of the Command Historian, Virginia, (1993), <http://www.dtic.mil/dtic/tr/fulltext/u2/a267030.pdf>

Rosenberg, Matthew and John Markoff, 'The Pentagon's 'Terminator Conundrum': Robots that could kill on their own', *New York Times*, (25 October 2016), <https://www.nytimes.com/2016/10/26/us/pentagon-artificial-intelligence-terminator.html>

Roske, Vincent, and others, 'Autonomous System Challenges to Testing and Evaluation', National Defense Industry Association test and evaluation conference, conference pack, (March 2012), <https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2012/TEST/13782_Roske.pdf>

Ross, Alice, 'Drone Warfare. UK's new Reaper drones remain grounded, months before Afghan withdrawal', <https://www.thebureauinvestigates.com/stories/2014-05-22/uks-new-reaper-drones-remain-grounded-months-before-afghan-withdrawal>

Rossi, Ben, 'How industry 4.0 is changing human-technology interaction', *Information Age*, (11 November 2016), <http://www.information-age.com/industry-4-0-changing-human-technology-interaction-123463164/>

Rotman, David, 'Molecular computing', *MIT Technology Review*, (May 2000), <https://www.technologyreview.com/s/400728/molecular-computing/>

O'Rourke, Ronald, and others, 'Multi-year procurement and block buy contracting in Defence Acquisition', *Congressional Research Service*, (8 August 2017), <https://fas.org/sgp/crs/natsec/R41909.pdf>

Royal Society, 'The Power and Promise of Computers that Learn by Example', *Royal Foundation*, (April 2017), <https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf>

RT Networks, 'Kalashnikov develops fully automated neural-network-based combat module', (5 July 2017), <https://www.rt.com/news/395375-kalashnikov-automated-neural-network-gun/>

Ruano-Borbalan, Jean-Claude, 'Technology, Science and Society; Norms, Cultures and Institutions matter', *Journal of Innovation Economics and Management*, No.22, (2017), <https://www.cairn.info/revue-journal-of-innovation-economics-2017-1-page-3.htm>

Ruder, Sebastian, 'Transfer Learning - Machine Learning's New Frontier', *author's blog*, (21 March 2017), <http://ruder.io/transfer-learning/>

Rupert, Mark, 'Support the Troops: Populist Militarism and the Cultural Reproduction of Imperial Power', Maxwell School, Syracuse University, <http://faculty.maxwell.syr.edu/merupert/Populist%20Militarism.pdf>

Russell, Stuart, 'Rationality and Intelligence: A brief update', <https://people.eecs.berkeley.edu/~russell/papers/ptai13-intelligence.pdf>

Russell, Stuart, and others, 'Research Priorities for Robust and Beneficial Artificial Intelligence', Association for the Advancement of Artificial Intelligence, (Winter 2015), <https://ocs.aaai.org/ojs/index.php/aimagazine/article/download/2577/2521>

Russell, Stuart, 'Take a Stand on AI Weapons', *Nature*, 521, 7553, 27 May 2015, https://www.nature.com/news/robotics-ethics-of-artificial-intelligence-1.17611#russell>

Russia Today, 'Iran replicates CIA's RQ-170 Sentinel drone', <http://rt.com/news/158272-iran-unveils-drone-copy/>

Russia Today, 'Anti-drone devices for sale: Military contractor claims to have counter-UAV technology', <http://rt.com/usa/oregon-domestic-drone-countermeasures-339/>

Sadler, Brent, 'Fast Followers, Learning Machines, and the 3rd Offset Strategy', National Defense University Press, *Joint Force Quarterly 83*, (1 October 2016), <http://ndupress.ndu.edu/Media/News/Article/969644/fast-followers-learning-machines-and-the-third-offset-strategy/>

Saldano, John, 'The Coding Manual for Qualitative Researchers', Sage Publications, (2009), <https://www.sagepub.com/sites/default/files/upm-binaries/24614_01_Saldana_Ch_01.pdf>

Salzberg, Steven, 'Which is more Important: Military Drones Or A Cure For Cancer?', <http://www.forbes.com/sites/stevensalzberg/2013/12/29/which-is-more-important-military-drones-or-a-cure-for-cancer/>

Sanchez, Bennie, 'Fratricide, Technology and Joint Doctrine', US Naval War College, (February 2004), http://www.dtic.mil/dtic/tr/fulltext/u2/a422756.pdf

Satell, Greg, '4 Innovation Lessons from the History of Warfare', *Forbes*, (14 March 2015), <https://www.forbes.com/sites/gregsatell/2015/03/14/4-innovation-lessons-from-the-history-of-warfare/#508dec4e73f3>

Sayler, Kelley, 'A world of proliferated drones: a technology primer', Centre for a new American security, (July 2015), <https://www.cnas.org/publications/reports/a-world-of-proliferated-drones-a-technology-primer>

Schaal, Stefan, 'Is imitation learning the route to humanoid annoyed robots', *Trends in Cognitive Science*, Vol.3, No.6, (June 1999), <http://www.bcp.psych.ualberta.ca/~mike/Pearl_Street/PSYCO354/pdfstuff/Readings/Schaal1.pdf>

Scharre, Paul, 'The Coming Swarm: Robotics on the Battlefield', *Real Clear Defense*, (19 October 2014), <https://www.realcleardefense.com/articles/2014/10/20/the_coming_swarm_robotics_on_the_battlefield_107499.html>

Scharre, Paul, 'Robotics on the Battlefield, Part II: The Coming Swarm', *CNAS*, (2014), <https://s3.amazonaws.com/files.cnas.org/documents/cnas_TheComingSwarm_Scharre.pdf>

Scharre, Paul, 'Presentation at the United Nations Convention of Certain Conventional Weapons', *Lecture, Informal Meeting of Experts on Lethal Autonomous Weapons*, Geneva, (13 April 2015), <https://www.unog.ch/80256EDD006B8954/(httpAssets)/98B8F054634E0C7EC1257E2F005759B0/$file/Scharre+presentation+text.pdf>

Scharre, Paul, 'Making Sense of Rapid Technical Change', *Center for a New American Security*, (17 July 2017), <https://www.cnas.org/publications/commentary/making-sense-of-rapid-technological-change>

Scharre, Paul, 'Why we must not build automated weapons of war', *Time Magazine*, (25 September 2017), <http://time.com/4948633/robots-artificial-intelligence-war/>

Scharre, Paul, 'Making Sense of Rapid Technological Change', Center for a New American Security, (19 July 2017), <https://www.cnas.org/publications/commentary/making-sense-of-rapid-technological-change>

Schermerhorn, Paul, and others, 'Dynamic robot autonomy; investigating the effect of robot decision making in a human-robot task team', ICMI-MLMI, (2009), <https://hrilab.tufts.edu/publications/schermerhornscheutz09icmi.pdf>

Schmill, Matthew, Tim Oates and others, 'The Role of Metacognition in Robust AI Systems', *aaai.org*, (2008), <http://www.aaai.org/Papers/Workshops/2008/WS-08-07/WS08-07-026.pdf>

Schmitt, Michael, 'A Reply to the Critics', UN Naval War College, <https://www.usnwc.edu/getattachment/a126d0a7-bad9-49bc-96a0-9f9e92cb35dc/Professor-Schmitt-s-Article-on-AWS---SSRN-id218482.aspx>

Schmitt, Michael, and Jeffrey Thurnher, 'Out of the loop: autonomous weapon systems and the law of armed combat', *Harvard National Security Journal*, Vol. 4, (2012), <http://harvardnsj.org/wp-content/uploads/2013/01/Vol-4-Schmitt-Thurnher.pdf>

Schrage, Michael, 'The Real Problem with Computers', *Harvard Business Review*, (October 1997), <https://hbr.org/1997/09/the-real-problem-with-computers>

Sculley, D., 'Machine Learning and Technical Debt', *Software Engineering Daily*, (17 November 2015), <https://softwareengineeringdaily.com/2015/11/17/machine-learning-and-technical-debt-with-d-sculley/>

Sculley, D., and others, 'Machine learning: The high-interest credit card of technical debt', Google Inc, SE4ML: *Software Engineering for Machine Learning,* NIPS 2014 Workshop, <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/43146.pdf >

Sculley, D., and others, 'Hidden Technical Debt in Machine Learning Systems', *Advances in Neural Information Systems*, (2015), <http://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf>

Seba, Tony, 'Clean Disruption: Why conventional energy and transportation will be obsolete by 2030', *Presentation to Swedbank*, (17 March 2016), <http://www.swedbank.no/idc/groups/public/@i/@sc/@all/@lci/documents/presentation/cid_1987411.pdf>

Selby, Richard, 'Analyzing Error-Prone System Structure', *IEEE Transaction of Software Engineering*, Vol.17, No.2, (February 1991), <http://cgis.cs.umd.edu/~basili/publications/journals/J42.pdf>

Shah, Tarang, 'About Train, Validation and Test Sets in Machine Learning', *Towards Data Science*, (6 December 2017), <https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7>

Shahriari, Bobak, and others, 'Taking the human out of the loop: A review of Baysian optimisation', *Proceedings of the IEEE*, 104:1, (2016), <https://www.cs.ox.ac.uk/people/nando.defreitas/publications/BayesOptLoop.pdf>

Shakhimardanov, Azamat, and others, 'Best Practice in Robotics', *BRICS Collaborative*, (February 2013), http://www.best-of-robotics.org/pages/publications/BRICS_Deliverable_D2.1.pdf

Shalal-Esa, Andrea, 'Insight: Expensive F-35 fighter at risk of budget 'death spiral', *Reuters Newswire*, (15 March 2013), <http://uk.reuters.com/article/2013/03/15/us-usa-fighter-f35-insight-idUSBRE92E10R20130315>

Sharma, Avinash, 'Understanding Activation Functions in Neural Networks', *Medium.com blog*, (30 March 2017), <https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>

Shapley, Deborah, 'Technology Creep and the Arms Race; A World of Absolute Accuracy', *Science Magazine*, Volume 201, Issue 4362, (29 September 1978), <http://science.sciencemag.org/content/201/4362/1192>

Sharkey, Noel, 'Killing made easy', <http://www.cs.bath.ac.uk/~jjb/ftp/MQ2012-frontmatter.pdf>

Sharkey, Noel, 'Automating warfare: lessons learned from the drone', *Journal of Law*, Information and Science, (2012), <www.austlii.edu.au/au/journals/JlLawInfoSci/2012/8.html>

Sheth, Mihir, '16 Expert Tips for Conquering God of War's Difficulty Mode', *PlayStation Blog*, (16 May 2018), <https://blog.eu.playstation.com/2018/05/16/16-expert-tips-for-conquering-god-of-wars-brutal-give-me-god-of-war-difficulty-mode/>

Shteingart, Hanan, and others, 'The Role of First Impressions in Operant Learning', *Journal of Experimental Psychology*, Volume 142, 2, (2013), <https://elsc.huji.ac.il/sites/default/files/476.pdf>

Shurkin, Joel, 'When Driver Error becomes Programming Error', *Inside Science*, (18 February 2015), <https://www.insidescience.org/news/when-driver-error-becomes-programming-error>

Sieracki, Jeff, 'Machine Learning for Embedded Software is not as hard as you may think', *Reality AI*, (3 August 2016), <https://www.reality.ai/single-post/2016/08/03/5-tips-for-embedded-machine-learning>

Simonite, Tom, 'Thinking in Silicon', *MIT Technology Review,* (December 2013), <https://www.technologyreview.com/s/522476/thinking-in-silicon/>

Simonite, Tom, 'Sorry, banning 'Killer Robots' just isn't practical', *Wired, (*22 August 2018), https://www.wired.com/story/sorry-banning-killer-robots-just-isnt-practical/)

Simmons, Reid, 'Structured control of autonomous robots', *IEEE Transactions on Robots and Automation*, Volume 10, No.1, (February 1994), <https://www.cs.cmu.edu/~reids/papers/structured.pdf>

Simons, Greg, 'Undertanding Political and Intangible Elements in Modern Wars', *Academia,* (2012), <http://www.academia.edu/2070261/Understanding_Political_and_Intangible_Elements_in_Modern_Wars>

Simpson, Edwin, and others, 'Dynamic Bayesian Combination of Multiple Imperfect Classifiers', *Decision Making and Imperfection* 474, (2013), <http://www.robots.ox.ac.uk/~sjrob/Pubs/galaxyZooSN_simpson_etal.pdf>

Singer, Peter, 'Tactical Generals: Leaders, Technology, and Perils', *Brookings Institute*, (7 July 2009), <https://www.brookings.edu/articles/tactical-generals-leaders-technology-and-the-perils/>

Singer, Peter, and August Cole, 'Humans Can't Escape Killer Robots but Humans can be Held Responsible for Them', *Vice News*, (15 April 2016), <https://news.vice.com/article/killer-robots-autonomous-weapons-systems-and-accountability>

Skillings, Jon, 'The Navy's unmanned X-47B flies again', *CNET.com* <http://news.cnet.com/8301-11386_3-57611733-76/the-navys-unmanned-x-47b-flies-again/>

Smalley, David, 'The Future is Now: Navy's Autonomous Swarmboats can Overwhelm Adversaries', Office of Naval Research, *Science and Technology*, (2014), <http://www.onr.navy.mil/Media-Center/Press-Releases/2014/autonomous-swarm-boat-unmanned-caracas.aspx>

Smallwood, David, 'Augustine's Law Revisited', *Sound and Vibration*, (March 2012), http://www.sandv.com/downloads/1203smal.pdf

Smith, Andrew, 'Franken-algorithms: the deadly consequences of unpredictable code', *Guardian*, (30 August 2018), <https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger>

Smith, Chris, 'What is Teranis? Everything you need to know about Britain's undetectable drone', *BT website*, (21 November 2017), <http://home.bt.com/tech-gadgets/future-tech/taranis-unmanned-aerial-vehicle-stealth-11364110510493>

Smith, Edward, and others, 'Is the Cure Worse Than the Disease? Over-fitting in Automated Program Repair'. Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering, ACM, (2015), <https://www.cs.cmu.edu/~clegoues/docs/smith15fse.pdf>

Snow, Shawn, 'The US Army is Developing Autonomous Armoured Tanks', *Army Times*, (29 August 2017), <https://www.armytimes.com/news/your-army/2017/08/29/the-us-army-is-developing-autonomous-armored-vehicles/>

Somers, James, 'The Coming Software Apocalypse', *Atlantic*, (26 September 2017), <https://www.theatlantic.com/technology/archive/2017/09/saving-the-world-from-code/540393/>

Somes, James, 'The Coming Software Apocalypse*', The Atlantic*, (26 September 2017), <https://www.theatlantic.com/technology/archive/2017/09/saving-the-world-from-code/540393/>

Song, H. Francis, and others, 'Reward-based Training of Recurrent Neural Networks for Cognition and Value-based Tasks*', ELife*, (13 January 2017), <https://elifesciences.org/articles/21492>

Spencer, John, 'What is Army Doctrine?', *Modern War Institute*, (21 Marcb 2016), <https://mwi.usma.edu/what-is-army-doctrine/>

Sporns, Olaf, and Gerald Edelman, 'Solving Bernstein's Problem: A Proposal for the Development of Coordinated Movement by Selection', *Child Development,* 64.4, (1993), <http://e.guigon.free.fr/rsc/article/SpornsEdelman93.pdf>.

Srinivasan, Venkat, 'Context, Language and Reasoning in AI: Three Key Challenges', *MIT Technology Review*, (14 October 2016), <https://www.technologyreview.com/s/602658/context-language-and-reasoning-in-ai-three-key-challenges/>

Staff EP, 'High Profile Panel Warns of unavoidable, far-reaching technical revolution*', The Education Post,* (7 July 2017), <http://educationpostonline.in/2017/07/07/high-profile-panel-warns-of-unavoidable-far-reaching-tech-revolution/>

Stanford University, 'Arguments For: What have been the primary motivating arguments for the continued development of autonomous weapons', <http://cs.stanford.edu/people/eroberts/cs181/projects/autonomous-weapons/html/argfor.html>

Stanford University, 'Artificial Intelligence and Life in 2030*',* 2015 study panel, (June 2016), <https://ai100.stanford.edu>

Stanford University, 'Report of AI100 Study Group', <https://ai100.stanford.edu/2016-report/executive-summary>

Stanford University, 'Videos on Machine learning*',* <https://www.coursera.org/learn/machine-learning/lecture/QGKbr/model-selection-and-train-validation-test-sets>

Stergio, Christos, and Dmitrios Siganos, 'Neural Networks*',* (undated), <https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html>

Stimpson, Jeffrey, and others, 'Learning to cooperate in a social dilemma; a satisficing approach to bargaining', Brigham Young University, Proceedings of the 20th International Conference on machine-learning, ICML-03, (2003), <https://www.aaai.org/Papers/ICML/2003/ICML03-095.pdf>

Stockholm International Peace Research Institute SIPRI, 'Implementation of Article 36 Weapon Reviews in light of increasing autonomy in weapon systems*',* SIPRI, 11/ (November 2015), <https://www.sipri.org/media/press-release/2015/implementing-article-36-weapon-reviews-light-increasing-autonomy-weapon-systems>

Stoica, Ion, and others, 'A Berkeley View of System Challenges for AI', arXiv.1712: 05855v1, (12 December 2017), <https://arxiv.org/pdf/1712.05855.pdf>

Stojkovic, Dejan, and Bjorn Robert Dahn, 'Methodology for long-term Defence Policy', Norwegian Defence Research Establishment, (28th February 2007), <http://www.ffi.no/no/Rapporter/07-00600.pdf>

Stone, Robert, 'Puzzles of proportion and the 'Reasonable Military Commander: Reflection on the law, ethics and geopolitics of proportionality', *Harvard National Security Journal*, (2015), <http://harvardnsj.org/wp-content/uploads/2015/06/Sloane.pdf>

Stopkillerrobots.org, 'Campaign to stop killer robots', <http://www.stopkillerrobots.org/learn/>

Stopkillerrobots.org, 'Campaign to stop killer robots, France convenes seminar at the UN', <http://www.stopkillerrobots.org/2013/09/france-seminar/>

Stopkillerrobots.org, 'Country Policy Positions', (25 March 2016), <http://www.stopkillerrobots.org/wp-content/uploads/2015/03/KRC_CCWexperts_Countries_25Mar2015.pdf>

Study.com, 'Heuristics', <http://study.com/academy/lesson/heuristics.html>

Study.com, 'Problem Solving Methods; Definitions and Types', *Study.com*, chapter 8, lesson 4, <http://study.com/academy/lesson/problem-solving-methods-definition-types.html>

Sullivan, Ben, 'Elite scientists have told the Pentagon that AI won't threaten humanity', *Motherboard magazine*, (19 January 2017), <https://motherboard.vice.com/en_us/>

Sun, Ron, 'Connectionist and Symbolic Approaches', University of Missouri-Columbia, (November 2000), <http://www.cogsci.rpi.edu/~rsun/sun.encyc01.pdf>

Sun, Ron, and C Lee Giles, 'Sequence Learning: From recognition and prediction to sequential decision-making', IEEE, *Intelligent Systems*, (2001), <http://www.sts.rpi.edu/~rsun/sun.expert01.pdf>

Sufge, Erik, 'What might a Killerbot Arms Race Look Like?', *Popular Science*, 28 May 2015, https://www.popsci.com/what-would-killerbot-arms-race-look>

Svoboda, Eva, and Emanuela-Chiara Gillard, 'Protection of Civilians in Armed Conflict: Bridging the Gap between Law and Reality', Humanitarian Policy Group, Overseas Development Institute, (2015) <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9876.pdf

Szczerba, Robert,, '15 Worst Tech Predictions of All Time', *Forbes*, (5 January 2015), <https://www.forbes.com/sites/robertszczerba/2015/01/05/15-worst-tech-predictions-of-all-time/#65c877e91299>

Szoldra, Paul, 'An Ex-Pentagon Official thinks 'Killer Robots' need to be stopped', *Business Insider*, (9 March 2016), <http://uk.businessinsider.com/pentagon-autonomous-warfare-2016-3>

Szondry, David, 'Quantum Radar to Render Stealth Technologies Obsolete', *New Atlas*, (26 April 2018), <https://newatlas.com/quantum-radar-detect-steath-aircraft/54356/>

Talbot, Patrick, 'Military Decision Aids – A Robust Decision-Centred Application', TRW Systems, *Technology Review Journal*, (Spring-Summer 2001), <http://ellisinterstellar.com/DecisionAids.pdf>

Talwar, Anish, and Yogesh Kumar, 'Machine Learning: An Artificial Intelligence Methodology', *IJECS*, Vol.2, Issue 12, (December 2013), <http://www.ijecs.in/issue/v2-i12/11%20ijecs.pdf>

Tanz, Jason, 'Soon we won't program computers. We'll train them like dogs', *Wired magazine,* (17 May 2017), <https://www.wired.com/2016/05/the-end-of-code/>

Tarantola, Andrew, 'South Korea's Auto-Turret Can Kill A Man In The Dead Of Night From Three Clicks', *Gizmodo*, (October 2012), <http://gizmodo.com/5955042/south-koreas-auto-turret-can-kill-a-man-in-the-dead-of-night-from-three-clicks>

Taylor, Gavin, and Ron Parr, 'Value Function Approximation in Noisy Environments', Cornell University Press, arXiv preprint arViv 1210.4898, (2012), <https://arxiv.org/abs/1210.4898>

Taylor, Jessica, and others, 'Alignment for Advanced Machine Learning Systems', *MIRI*, (2016), ('*Motivations*') <https://intelligence.org/files/AlignmentMachineLearning.pdf>

Taylor, Richard, 'Software architecture: Foundations, theory and practice', School of Information and Computer Science, UCal at Irvine, (October 1999), <https://www.ics.uci.edu/~taylor/Architecture.pdf>

Teach, Edward, 'Avoiding Decision Traps', CFO, (17 June 2004), <http://ww2.cfo.com/human-capital-careers/2004/06/avoiding-decision-traps/>

Teng, Choh Man, 'Dealing with data corruption in remote sensing', *Advances in Intelligent Data Analysis* VI, IDA 2005. Lecture notes in computer science, Vol. 3646, Springer, <https://link.springer.com/chapter/10.1007/11552253_41>

Tharoor, I, 'Should the world kill killer robots before it's too late?', *Washington Post*, (12 May 2014) <http://www.washingtonpost.com/blogs/worldviews/wp/2014/05/12/should-the-world-kill-killer-robots-before-its-too-late/>

Thompson, Iain, 'US military tests massive GPS jamming weapon over California', *The Register*, (7 June 2016), <www.theregister.co.uk/2016/06/17/us_military_testing_gps_jamming>

Thompson, Iain, 'AI Slurps, Learns Millions of Passwords to Work Out Which Ones You May Use Next', *The Register*, (20 September 2017), <https://www.theregister.co.uk/2017/09/20/researchers_train_ai_bots_to_crack_passwords/>

Thompson, Loren B., 'Army radio plan leaves soldiers vulnerable to jamming', Lexington Institute, (February 2012), <http://lexingtoninstitute.org/army-radio-plan-leaves-soldiers-vulnerable-to-jamming/>

Thompson, Loren B., '5 Reasons the Army's New Battlefield Networking Strategy Won't Work', *Forbes* Magazine, (20 November 2017), <https://www.forbes.com/sites/lorenthompson/2017/11/20/five-reasons-why-the-armys-new-battlefield-networking-strategy-wont-work/>

Thorisson, Kristinn, and others, 'Why artificial intelligence needs a task theory. And what it might look like', 9th International Conference on AGI, (2016), <http://people.idsia.ch/~steunebrink/Publications/AGI16_task_theory.pdf>

Thought Infection, 'Can we Avoid an Automated Arms Race?', <http://thoughtinfection.com/2014/03/16/can-we-avoid-an-an-automated-arms-race>

Treanor, Jill, 'The 'Flash Crash' of 2010: How it unfolded', *Guardian*, (22 April 2015), <https://www.theguardian.com/business/2015/apr/22/2010-flash-crash-new-york-stock-exchange-unfolded>

Trimble, Stephen, 'Sierra Nevada Fields ARGUS-IS Upgrade to Gorgon Stare Pod', *Flight Global,* (2 July 2014), <https://www.flightglobal.com/news/articles/sierra-nevada-fields-argus-is-upgrade-to-gorgon-stare-400978/>

Triska, Ricardo, 'Artificial Intelligence, classification theory and the uncertainty reduction process', Federal University of Santa Catherina, Brazil, (2013), <http://www.iskoiberico.org/wp-content/uploads/2014/09/479-486_Triska.pdf>

Trifonov, Roumen, and others, 'Artificial Neural Network Intelligent Method for Prediction', AIP Conference Proceedings, (6 July 2017), <https://aip.scitation.org/doi/pdf/10.1063/1.4996678?class=pdf>.

Tschopp, Marisa, 'Human Cognition and Artificial Intelligence – A Plea for Science', *Medium.com,* (23 April 2018), <https://medium.com/womeninai/human-cognition-and-artificial-intelligence-a-plea-for-science-21a2388f6e7e>

Tucker, Patrick, 'Every Country Will Have Armed Drones Within 10 Years', *Defense One Magazine*, (May 2014), <http://www.defenseone.com/technology/2014/05/every-country-will-have-armed-drones-within-ten-years/83878/>

Turing, Alan, 'Computer machinery and intelligence', Mind 49: 433-460, (1950), <http://www.csee.umbc.edu/courses/471/papers/turing.pdf>

Turitto, James, 'Understanding Warfare in the Twenty First Century', *International Affairs Review*, Volume XVIII, Number 3 (Winter 2010), <http://www.iar-gwu.org/node/145>

Turnbull, Grant, 'Off the Shelf: Re-thinking Innovation in the Military', *Army Technology*, (2 March 2014), <https://www.army-technology.com/features/featureinnovation-stagnation-re-thinking-innovation-in-the-military-4187511/>

Turnbull, Grant, 'The realities of autonomy in unmanned aerial systems today', *Army Technology*, (9 February 2014), <https://www.army-technology.com/features/featurethe-realities-of-autonomy-in-unmanned-air-systems-today-4175047/>

Tversky, Amos, and Daniel Kahneman, 'Judgement under Uncertainty: Heuristics and Biases', *Science*, New Series, Volume 185, 4157, (27 September 1974), <http://www.its.caltech.edu/~camerer/Ec101/JudgementUncertainty.pdf>

UAS Vision, 'Suppressing Air Defenses by UAV Swarm Attack', *UASVision.com*, (28 June 2018), <https://www.uasvision.com/2018/06/28/suppressing-air-defenses-by-uav-swarm-attack/>

UC Riverside, 'Andrew Reath', Department of Philosophy, <http://philosophy.ucr.edu/andrews-reath/>

UNESCO, 'The Infernal Cycle of Armaments', *International Social Sceince Journal*, Volume 28, Number 2, <http://unesdoc.unesco.org/images/0001/000197/019707eo.pdf>

UNIDIR, 'Safety, Unintentional Risk and Accidents in the Weaponization of Increasingly Autonomous Technologies', *UNIDIR*, Number 5, (2016), <http://www.unidir.org/files/publications/pdfs/safety-unintentional-risk-and-accidents-en-668.pdf>

UK Army, 'Land Operations', Land Warfare Development Centre, Army doctrine publication AC 71940, <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/605298/Army_Field_Manual__AFM__A5_Master_ADP_Interactive_Gov_Web.pdf>

UK Ministry of Defence and Foreign and Colonial Office, 'UK's International Defence Engagement Strategy', (2017),

<https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/596968/06032017_Def_Engag_Strat_2017DaSCREEN.pdf>

UK Ministry of Defence, 'Rules of Engagement', *Wikileaks*, UK and Danish ROE, (June 2008), <https://file.wikileaks.org/file/uk-danish-roe-iraq-2006.pdf>

UK Parliamentary Hansard record, <(http://www.publications.parliament.uk/pa/cm201314/cmhansrd/cm130617/debtext/130617-0004.htm>

United Nations Office for Disarmament Affairs, 'The Biological Weapons Convention' <http://www.un.org/disarmament/WMD/Bio/>

United Nations Office for Disarmament Affairs, 'UN Register of Conventional Arms', <http://www.un.org/disarmament/convarms/Register/>

United Nations Office for Disarmament Affairs, 'The Arms Trade Treaty', <http:///www.un.org/disarmament/ATT/>

United Nations Office, 'The Convention on Certain Conventional Weapons', <http://www.unog.ch/80256EE600585943/(httpPages)/4F0DEF093B4860B4C1257180004B1B30?OpenDocument>

United States Air Force, 'Autonomous horizons - system autonomy in the air force; a path to the future', Office of the Chief Scientist, <http://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf?timestamp=1435068339702>

US Air University, 'The Military Decision Making Process', Chapter 5, (undated), <http://www.au.af.mil/au/awc/awcgate/army/fm101-5_mdmp.pdf>

US Army, 'The Army Human Dimension Strategy 2015; Building cohesive teams to win in a complex world', (2015), <http://usacac.army.mil/sites/default/files/publications/20150524_Human_Dimension_Strategy_vr_Signature_WM_1.pdf>

US Army, 'The Human Dimension White Paper; a framework for optimizing human performance', Combined Arms Center, (9 October 2014), <http://usacac.army.mil/sites/default/files/documents/cact/HumanDimensionWhitePaper.pdf>

US Army, 'Robotic and Autonomous System Strategy', Army capabilities Integration Centre, (March 2017), <http://www.tradoc.army.mil/FrontPageContent/Docs/RAS_Strategy.pdf>

US Army, Field Manual, 'Intelligence Preparation of the Battlefield', FM 34-130, Section 1, (July 1994), <https://fas.org/irp/doddir/army/fm34-130.pdf>

US Army, Field Manuals, 'Battle Command', *FM 7-30, The Infantry Brigade*, Chapter 3, (1995 and revisions) <https://www.globalsecurity.org/military/library/policy/army/fm/7-30/Ch3.htm

US Army, Department of US Army Headquarters, 'The US Army Functional Concept for Battle Command 2015-2024', TRADOC Pamphlet 525-3-3, Version 1.0, (30 April 2007), <http://www.tradoc.army.mil/tpubs/pams/p525-3-3.pdf>

US Army DoD, 'Dictionary of Military and Associated Terms', *US DoD Publications,* (November 2018), <https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/dictionary.pdf>

US Army, 'The Army Ethic White Paper', Center for the Army Profession and Ethics, (11 July 2014), https://www.army.mil/e2/c/downloads/356486.pdf

US Army, 'The Army Equipping Strategy', *Army G-8, 2009,* <http://www.g8.army.mil/pdf/Army_Equipping_Strategy.pdf>

US Department of Defense, 'The Role of Autonomy in DoD Systems'*, Task Force Report*, (April 2012), <https://fas.org/irp/agency/dod/dsb/autonomy.pdf>

U.S. Department of State, 'Treaty on Conventional Armed Forces in Europe', <http://www.state.gov/t/avc/cca/cfe/index.htm>

US Field Manual, 'Light Cavalry Gunnery: Target Acquisition'*,* Field Manual 17-12-8, (February 1999), <http://www.globalsecurity.org/military/library/policy/army/fm/17-12-8/ch3.htm>

US Patent grant US52607090A, 'Autonomous precision weapons delivery using synthetic array radar'*,* <https://patents.google.com/patent/US5260709A/en>

University of Salzburg Center for Human-Computer Interaction, 'To err is robot: How robots learn to recognise error'*,* blog, <https://hci.sbg.ac.at/outputs/hri-error-situations/>

Upbin, Bruce, 'Apple is about to become the biggest R&D spender in the world', *Tribune Interactive*, <https://phys.org/news/2018-03-apple-biggest-spender-world.html>

Vanderelst, Dieter and Alan Winfield, 'An Architecture for Ethical Robots Inspired by the Simulation Theory of Cognition', *Cognitive Systems Research*, 48, (May 2018), <https://reader.elsevier.com/reader/sd/pii/S1389041716302005?token=C7E1EA2947EBE5A58F14432243B19081E625A4E6D56A5CC87D0ED7548B5A93DB91F0B8ADB1DA68FFA268CDC86E1319D1>

Vanhoucke, Vincent, and others, 'Improving the Speed of Neural Networks on CPUs'*,* Processor Deep Learning and Unsupervised Feature Learning NIPS Workshop, Volume 1, (2001), <http://andrewsenior.com/papers/VanhouckeNIPS11.pdf>

Vardi, Moshe, and others, 'The Implication Problem for Data Dependencies'*,* Hebrew University of Jerusalem, (January 2006), <https://www.researchgate.net/publication/226509257_The_implication_problem_for_data_dependencies>

J Varghese, 'Review of autonomous vehicle sensors and systems'*,* Proceedings of 2015 International conference on operations excellence and service engineering, (October 2015), <http://iieom.org/ICMOE2015/papers/140.pdf>

Milan Vego, 'On Military Creativity'*, ndupress*, Issue 70, (Third Quarter 2013), <http://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-70/JFQ-70_83-90_Vego.pdf>

Verdiesen, Lise, 'How do we ensure that we remain in control of our autonomous weapons?', *AI Matters,* Volume 3, Issue 3, <https://sigai.acm.org/static/aimatters/3-3/AIMatters-3-3-11-Verdiesen.pdf>

Vered, Mor, and Gal Kahminka, 'Online recognition of navigational goals through goal mirroring', Extended Abstract, Proceedings of 16th Conference of Autonomous Agents and Multi-agent Systems, AA-MAS, (2017), <http://www.ifaamas.org/Proceedings/aamas2017/pdfs/p1748.pdf>

Verleysen, Michel, and others, 'On the Effects of Dimensionality on Data Analysis'*,* IWANN, (2003), <https://perso.uclouvain.be/michel.verleysen/papers/iwann03mv.pdf>

Verschule, Paul, and others, 'The why, what, where, when and how of goal-directed choice: Neuronal and computational principles', Philos Trans Royal Soc London, 1655-20130483, (5 November 2014), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4186236/>

Vincent, James, 'These are three of the biggest problems facing today's artificial intelligence', *The Verge*, (10 October 2016), <https://www.theverge.com/2016/10/10/13224930/ai-deep-learning-limitations-drawbacks>

Vincent, James, 'The Biggest Headache in Machine Learning? Cleaning Dirty Data off the Spreadsheets', *The Verge*, (1 November 2017), <https://www.theverge.com/2017/11/1/16589246/machine-learning-data-science-dirty-data-kaggle-survey-2017>

Vlatko, Natali, 'The Dangers of Spaghetti Code', *JaxCenter blog*, (5 June 2015), <https://jaxenter.com/the-dangers-of-spaghetti-code-117807.html>

Volker, K., 'We need a rule book for drones', <http://articles.washingtonpost.com/2012-10-26/opinions/35500650_1_drone-strikes-drone-attacks-guantanamo-bay>

Wadwa, Vivek, and Aaron Johnson, 'Robots could eventually replace soldiers in warfare. Is that a good thing?', *Washington Post*, (5 October 2016), <https://www.washingtonpost.com/news/innovations/wp/2016/10/05/robots-could-eventually-replace-soldiers-in-warfare-is-that-a-good-thing/?noredirect=on&utm_term=.1a5862445221>

Wakefield, Katrina, 'A Guide to Machine Learning Algorithms and their Application', undated, *SAS.com*, <https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html>

Walczak, Steven, and Narciso Cerpa, 'Heuristic Principles for the Design of Artificial Neural Networks', Information and Software Technology, 41 (2), (1999), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.22.9321&rep=rep1&type=pdf>

Wallace, Jon, 'SciFi Eye: The Disturbing Future of Autonomous Weapons', *The Engineer,* (19 September 2017), <https://www.theengineer.co.uk/autonomous-weapon-systems/>

Wallach, Wendell, 'Terminating the Terminator: What to do About Autonomous Weapons', <http://scienceprogress.org/2013/01/terminating-the-terminator-what-to-do-about-autonomous-weapons>

Wang, Linnan, and others, 'SuperNeurons: Dynamic GPU Memory Management for Training Deep Neural Networks', Proceedings of 23rd ACM Symposium on Parallel Programming, (2018), <https://arxiv.org/pdf/1801.04380.pdf>

Wang, Richard, 'Active Sensing Data Collection with Autonomous Mobile Robots', Carnegie Mellon University, International Conference, Robotics and Automation, *IEEE*, (2016), <http://www.contrib.andrew.cmu.edu/~rpw/MyBioWebpage/ICRA16.pdf>

Wang, Yingxu, 'On Abstract Intelligence: Towards a Unifying Theory of Natural, Artificial, Machinable and Computational Intelligence', *International Journal of Software Science and Computational Intelligence*, 1(1), (January 2009), <http://www.ucalgary.ca/icic/files/icic/24-IJSSCI-1101-AbstractInt.pdf>

Warner, Brad, and Manavendra Misra, 'Understanding Neural Networks as Statistical Tools', *American Statistician*, Volume 50, Issue 4, (November 1996), <http://gis.msl.mt.gov/Maxell/Models/Predictive_Modeling_for_DSS_Lincoln_NE_121510/Modeling_Literature/Warner%20and%20Misra_neural%20networks.pdf>

Warwick, Graham, and Jan DiMascio, 'Machine Learning key to Automatic Target Recognition', *Aviation Week and Space Technology*, (26 May 2016), <http://aviationweek.com/defense/machine-learning-key-automatic-target-recognition>

Watts, Barry, 'Clausewitzian Friction and Future War', <http://www.clausewitz.com/readings/Watts-Friction3.pdf>

Waxman, Matthew, and Kenneth Anderson, 'Don't Ban Armed Robots in the U.S.' <http://www.newrepublic.com/article/115229/armed-robots-banning-autonomous-weapon-systems-isnt-answer>

Wead, Sean, 'Ethics, Combat and a Soldier's Decision to Kill', *Military Review,* (March-April 2015), <https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MilitaryReview_20150430_art013.pdf>

Weiss, Gerhard, 'Learning to coordinate actions in multi-agent systems', Munich, Proceedings if the International Joint Conference on Artificial Intelligence, (1993), <http://www.ijcai.org/Proceedings/93-1/Papers/044.pdf>

Weiss, Lora, 'Autonomous Weapons in the Fog of War', *IEEE Spectrum*, (27 February 2012), <http://spectrum.ieee.org/robotics/military-robots/autonomous-robots-in-the-fog-of-war>

WhatIs.com, 'Definition 'greyscale', <http://whatis.techtarget.com/definition/grayscale>

Wheeler, Winslow, 'Revisiting the Reaper Revolution', <http://nation.time.com/2012/02/27/1-the-reaper-revolution-revisited/>

Wheeler, Winslow, 'How Much Does an F-35 Actually Cost?', <www.warisboring.com, (27 July 2014), https://warisboring.com/how-much-does-an-f-35-actually-cost/>

Weapons Law Encyclopedia, 'Proportionality in attacks (under International Humanitarian Law)', <http://www.weaponslaw.org/glossary/proportionality-in-attacks-ihl>

Weizmann, Nathalie, 'Autonomous weapon systems under international law, Academy briefing 8', Geneva Academy, (2014), <https://www.geneva-academy.ch/joomlatools-files/docman-files/Publications/Academy%20Briefings/Autonomous%20Weapon%20Systems%20under%20International%20Law_Academy%20Briefing%20No%208.pdf>

Wheeler, Scott, 'Trusted autonomy: conceptual developments in technology foresight', *Defence Science and Technology Group Report*, Australian Government, Department of Defence, Victoria (2015), <http://www.dtic.mil/dtic/tr/fulltext/u2/a626723.pdf)>

Whitlock, Craig, 'US military drone surveillance is expanding to hot spots beyond declared combat zones', *Washington Post*, (20 July 2013), <https://www.washingtonpost.com/world/national-security/us-military-drone-surveillance-is-expanding-to-hot-spots-beyond-declared-combat-zones/2013/07/20/0a57fbda-ef1c-11e2-8163-2c7021381a75_story.html?tid=a_inl>

Wielomski, Martin, 'The GPU: Powering the Future of Machine Learning of AI', *Phoenix NAP Publications,* (21 September 2018), <https://phoenixnap.com/blog/future-gpu-machine-learning-ai>

Wilgenbusch, Ronald, and Alan Heisig, 'Command and control of vulnerabilities to communications jamming', JFQ, *ndupress.ndu.edu*, Issue 69, (June 2013), <http://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-69/JFQ-69_56-63_Wilgenbusch-Heisig.pdf>

Williams, Greg, 'Wise up, Deep Learning May Never Create a General Purpose AI', *Wired*, (28 January 2018), <https://www.wired.co.uk/article/deep-learning-automl-cloud-gary-marcus>

Wilson, J.R., 'Electronic Warfare Evolves to Meet New Threats', *Military & Aerospace*, (1 August 2017), <https://www.militaryaerospace.com/articles/print/volume-28/issue-8/special-report/electronic-warfare-evolves-to-meet-new-threats.html>

Wilson, J.R., 'Electronic Warfare: the cat-and-mouse game continues', *Military and Aerospace Electronics',* (9 September 2013), <http://www.militaryaerospace.com/articles/print/volume-24/issue-9/special-report/electronic-warfare-the-cat-and-mouse-game-continues.html>

Wissner Gross, Alexander, 'Datasets over Algorithms', *Edge*, (June 2017), <https://www.edge.org/response-detail/26587>

Wittes, Benjamin, 'Drones and Democracy: A Response', <http://www.lawfareblog.com/2014/09/drones-and-democracy-a-response-to-firmin-debrabander/>

Woods, David, and others, 'Can we ever escape from data overload? A cognitive system diagnosis', *Cognition*, Technology and Work, (April 2002), Vol 4, Issue 1, <https://link.springer.com/article/10.1007/s101110200002>

Woods, David, 'Decomposing Automation: Apparent Simplicity, Real Complexity', in *Automation and Human Performance Theory and Application*, Erlbaum, (1996), <https://www.researchgate.net/profile/David_Woods11/publication/267402671_Decomposing_Automation_Apparent_Simplicity_Real_Complexity/links/546b62c60cf2f5eb18091bcd.pdf>

Wood, Richard, 'The Technical Revolution in Military Affairs', <www.holtz.org/library/technology/technical_revolution_in_military_affairs>

Wolf, Katharina, 'Putting Number on Capabilities: Defence Inflation versus Cost Escalation', *European Institute for Security Studies*, Brief Issue, 27, (July 2015), <https://www.iss.europa.eu/sites/default/files/EUISSFiles/Brief_27_Defence_inflation.pdf>

Wolpert, David, 'The lack of a-priori distinction between learning algorithms', *Neural Computations*, Vol 8 Issue 7, (1996), <http://www.mitpressjournals.org/doi/abs/10.1162/neco.1996.8.7.1341>

Worcester, Maxim, 'Autonomous Warfare: A Revolution in Military Affairs', *ISPSW Strategy Series: Focus on Defence and International Security*, Issue 340, (April 2015), <https://www.files.ethz.ch/isn/190160/340_Worcester.pdf>

Work, Robert, and Shawn Brimley, '20YY; Preparing for War in the Robotic Age', *CNAS*, (22 January 2014), <https://www.cnas.org/publications/reports/20yy-preparing-for-war-in-the-robotic-age>

World Stage Group, 'Strong arms trade treaty needed as UN Security Council increasingly looks unfit for purpose', *Amnesty International*, <http://www.worldstagegroup.com/worldstagenew/index.php?active=news&newscid=4889&catid=4>

Wray, Barry, and others, 'An artificial neural network approach to learning from factory performance from a Kanban based system', *Journal of International information management*, volume 12, 2, article 7, (2003)

Wren, Alisdair, 'Relationships for Object-Oriented Programme Language', University of Cambridge Computer Laboratories, *Technical Report Number 702*, (November 2007), <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-702.pdf>

Wright, Oliver, 'Britain to set up controversial drone development partnership with France', <http://www.independent.co.uk/news/world/politics/britain-to-set-up-controversial-drone-development-partnership-with-france-9094412.html>

Wylie, J.C., Admiral, 'Revolutions in Military Affairs', *UK Essays*, (November 2013), <https://www.ukessays.com/dissertation/examples/history/revolution-in-military-affairs.php#citethis>

Xie, Yichen, and Dawson Engler, 'Using Redundancies to find Errors', *IEEE Transactions on Software Engineering*, (2003), <https://web.stanford.edu/~engler/tse-redundant.pdf>

Yadav, Pankaj, and others, 'Nearest Neighbour-based Clustering Algorithms for Large Datasets', arXiv.1505.05962, (22 May 2015), <https://arxiv.org/pdf/1505.05962.pdf>

Yampolskiy, Roman, and Joshua Fox, 'AI versus AGI', Singularity Hypotheses, Springer, Berlin, (2012), <https://intelligence.org/files/AGI-HMM.pdf>

Yampolskiy, Roman, and Joshua Fox, 'Artificial General Intelligence and the human mental model', Machine International Research Institute, (2012), <https://intelligence.org/files/AGI-HMM.pdf>

Yampolskiy, Roman, and Joshua Fox, 'Safety Engineering for Artificial General Intelligence', Machine International Research Institute, (2012), <https://intelligence.org/files/SafetyEngineering.pdf>

Yao, Ming, 'Four approaches to natural language processing and understanding', *Topbots*, (31 March 2017), <http://www.topbots.com/4-different-approaches-natural-language-processing-understanding/>

Yong, Ed, 'A bird-like flock of autonomous drones', (27 February 2014), <https://www.nationalgeographic.com/science/phenomena/2014/02/27/a-bird-like-flock-of-autonomous-drones/>

Yudkowsky, Eliezer, 'The Value Loading Problem', Research Fellow, Co-founder Machine Intelligence Research Institute, <https://www.edge.org/response-detail/26198>

Yudkowsky, Eliezer, 'Complex Value Systems in Friendly AI', International Conference on Artificial General Intelligence, Lecture Notes on Computer Science, Volume 6830, (2011), Springer, Berlin, Heidelberg

Yudkowsky, Eliezer, '2015: What do you think about machines that think?', *The Edge*, (2015), <https://www.edge.org/response-detail/26198>

Zakaria, Norodin, 'Thoughts of Qualia in Machines', rxiv.org, (2005), <http://vixra.org/pdf/1505.0146v1.pdf>

Zapfe, Martin and Michael Haas, 'Arms Procurement: The Political-Military Framework', *CSS Analyses in Security Policy*, 181, (November 2015), <http://www.css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/CSSAnalyse181-EN.pdf>

Zavershynskyi, Maksym, 'Technical Debt in Machine Learning', *Towards Data Science*, (1 July 2017), <https://towardsdatascience.com/technical-debt-in-machine-learning-8b0fae938657>

Zawieska, Karolina, 'Do Robots Equal Humans? Anthropomorphic Terms in LAWS', Industrial Research Institute for Automation and Measurement, PIAP, <http://www.unog.ch/80256EDD006B8954/(httpAssets)/369A75B470A5A368C1257E290041E20B/$file/23+Karolina+Zawieska+SS.pdf>

Zielinska, Teresa, 'History of Service Robots', IGI Global Publishing, (2014), <http://www.irma-international.org/viewtitle/84885/>

Zenko, Micah , 'The Coming Future of Autonomous Drones', Council on Foreign Relations, (4 September 2012), <http://blogs.cfr.org/zenko/2012/09/04/the-coming-future-of-autonomous-drones/>

Zennie, Michael, 'Death from a swarm of tiny drones: U.S. Air Force releases terrifying video of tiny flybots that can hover, stalk and even kill targets', <http://www.dailymail.co.uk/news/article-2281403/U-S-Air-Force-developing-terrifying-swarms-tiny-unmanned-drones-hover-crawl-kill-targets.html>

Zhang, Ce, and others, 'An Overreaction to Broken Machine Learning Abstraction' HILDA 2017, Chicago, (14 May 2017), <http://pages.cs.wisc.edu/~wentaowu/papers/hilda17-easeml.pdf

Zheng, Yaling , 'Machine Learning with Incomplete Information', *CSE Technical Reports*, 143, (December 2011), <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1148&context=csetechreports>

Zhao, Sirley, 'Manual versus Automated Data Validation', *Siemens commercial blog*, (9 February 2016), <https://www.edq.com/blog/manual-vs.-automated-data-validation/>

Zwanenburg, Marten, and others, 'Humans, Agents and International Humanitarian Law: Dilemmas in Target Discrimination', <http://www.estig.ipbeja.pt/~ac-_direito/ZwanenburgBoddensWijngaards_LEA05_CR.pdf>

Zorzi, Marco, 'Self-learning AI emulates the human brain', European Research Council c/o University of Padova, (22 July 2016), <https://erc.europa.eu/projects-figures/stories/self-learning-ai-emulates-human-brain>