Ministry of Defence

Joint Concept Note 1/18

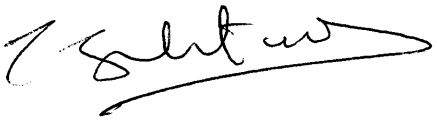
Human-Machine Teaming

Development, Concepts and Doctrine Centre

Joint Concept Note 1/18

Human-Machine Teaming

Joint Concept Note (JCN) 1/18, dated May 2018,
is promulgated as directed by the Chiefs of Staff



Director Concepts and Doctrine

Authorisation

The Development, Concepts and Doctrine Centre (DCDC) is responsible for publishing strategic trends, joint concepts and doctrine. If you wish to quote our publications as reference material in other work, you should confirm with our editors whether the particular publication and amendment state remains authoritative. We welcome your comments on factual accuracy or amendment proposals. Please send them to:

DCDC, Ministry of Defence Shrivenham, Swindon, Wiltshire, SN6 8RF

E-mail: DCDC-DocEds@mod.gov.uk Telephone: 01793 31 4216/4217/4220

Copyright

Distribution

Distributing Joint Concept Note (JCN) 1/18 is managed by the Forms and Publications Section, LCSLS Headquarters and Operations Centre, C16 Site, Ploughley Road, Arncott, Bicester, OX25 1LP. All of our other publications, including a regularly updated DCDC Publications Disk, can also be demanded from the LCSLS Operations Centre.

LLCSLS Help Desk: 01869 256197 Military Network: 94240 2197

Our publications are available to view and download on the Defence Intranet (RLI) at: http://defenceintranet.diif.r.mil.uk/Organisations/Orgs/JFC/Organisations/Orgs/DCDC

This publication is also available on the Internet at: www.gov.uk/mod/dcdc

Foreword

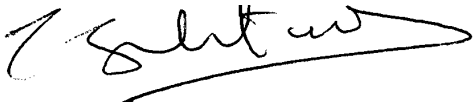
Throughout history, new technologies have been a driver of military adaptation and advantage. Whether moving from sail to steam, horses to tanks, or the introduction and exploitation of the aeroplane or radio, the results have often been transformative. When it has been transformative, strategy, tactics and technology have often evolved symbiotically; invariably when people figure out how best to exploit the full potential of the emerging combination of technologies.

Robotics and artificial intelligence (AI) offer the potential for another inflexion point in delivering military transformation and advantage. However, machines do not yet perform as well as a human brain. As John Boyd noted so wisely, 'Machines don't fight wars. People do, and they use their minds.' So realising this potential will depend on understanding the relative strengths of humans and machines, and how they best function in combination to outperform an opponent. Developing the right blend of human-machine teams – the effective integration of humans and machines into our war fighting systems – is the key; and we should not forget that we are in a race with our adversaries to unlock this advantage. The clock is ticking, as new technology capabilities accelerate.

Human-machine teaming considerations are not just a design factor for individual capability programmes; they will have systemic and institutional impacts on Defence. So there will need to be a clear public debate on how much, ultimately, we are prepared to 'trust' machines.

This joint concept note should be read by everyone who needs to understand how AI, robotics and data can change the future character of conflict, for us and our adversaries. In particular, it aims to inform and guide the development of policy, strategy, force and capability development so as to find the optimum balance of human-machine teams in Defence's future force structure.

Chief Scientific Advisor



Director, Development, Concepts and Doctrine Centre

"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 to best use it."

Paul Scharre,
Robotics on the Battlefield Part 1: Range, Persistence and Daring

Preface

Purpose

1. At the core of future military advantage will be the effective integration of humans and machines into war fighting systems that outperform our opponents. Joint Concept Note (JCN) 1/18, *Human-Machine Teaming* articulates the challenges and opportunities that robotic and artificial intelligence (AI) technologies offer, and identifies how we achieve advantage through human-machine teams. Its purpose is to guide coherent future force development and help frame Defence strategy and policy on automation and autonomy.

Aim

2. Remote and automated systems (RAS) continue to advance significantly, becoming more exploitable across all domains in multiple ways.[1] The aim of this JCN is to offer a long term, holistic view of these developments rather than predicting specific military applications. Tactics and technology evolve symbiotically and this JCN considers potential changes to the ways, as well as the means, with which we will fight.

Context

3. This document is subordinate to, and expands on, findings from JCN 1/17, *Future Force Concept* and is coherent with JCN 2/17, *Future of Command and Control*. It is set within the context provided by *Global Strategic Trends – Out to 2045* and *Future Operating Environment 2035*. It has been informed by: operational lessons and experimentation conducted by the North Atlantic Treaty Organization (NATO); international partners; Joint Forces Command; the Royal Navy; the British Army and Royal Air Force. It also draws from work by the Innovation and Research InSight Unit and the Defence Science and Technology Laboratory as well as a broad academic and industry network.

1. The term remote and automated systems (RAS) is frequently used with either 'automated' or 'autonomous' while remaining a collective term to describe unmanned aircraft systems, unmanned ground systems, unmanned surface vehicles and unmanned underwater vehicles.

Scope

4. The array of potential forms taken by RAS, and consequently how they interact in human-machine teams, is extremely varied. In size and complexity they could range from a future AI and robotically-enabled aircraft carrier retrofit, to a single, disposable nano-unmanned aerial vehicle.[2] We typically think of RAS as physical robotic systems in the battlespace, however, applying AI particularly for command and control functions and cyber operations will be increasingly common and important.

5. JCN 1/18 focuses on human-machine teaming and how we employ RAS out to 2035; however, its scope is also conditions based. It limits itself to the era of narrow AI that we have entered and that will continue for some time. Artificial general intelligence – a machine that can do all the things a human brain can do equally well – is beyond the scope of this JCN. When, or if, we approach the era of artificial general intelligence, the assumptions and deductions in this JCN will no longer be valid.[3]

6. The developing nature of the technologies in this field has created an array of terms and terminology which are often used interchangeably or differently by various commentators. Drawing distinct boundaries between those terms can often prove difficult, if not impossible. This may prove challenging for issues such as the proposition to ban lethal autonomous weapon systems – or even agree a common definition for these – which was being discussed in the United Nations at the time of this publication's release. Details on the considerations of autonomy and automation can be found at Annex A. For clarity, the Ministry of Defence's position, reiterated in 2017, is that 'we do not operate, and do not plan to develop, any lethal autonomous systems'.[4]

Audience

7. JCN 1/18 seeks to inform a wide audience. It should be read and understood by: those involved in policy and strategy formulation; science and technology personnel; personnel involved in concepts and force development; capability and acquisition staff; and operational commanders and their staffs.

2. A nano-unmanned aerial vehicle has a maximum take-off weight of 200g, typically fitting in the palm of an open hand. However, they can be potentially much smaller. Biomimetic unmanned aerial vehicles that mimic insects, and live insects fitted for electronic remote control, already exist.
3. Artificial narrow intelligence is sometimes also referred to as weak artificial intelligence (AI). It is AI optimised for specific, narrow, tasks; although with modification such tools can be applied to a variety of tasks. Strong AI, or artificial general intelligence, by contrast, is a machine with an ability to apply intelligence without modification to any problem, not just specific types of problem.
4. Joint Doctrine Publication (JDP) 0-30.2, *Unmanned Aircraft Systems*.

Structure

8. This document is separated into four chapters, deduction and insights and an annex.

a. Chapter 1 – Context. This chapter examines economic and technological trends and the likely impacts of AI and robotic systems on Defence.

b. Chapter 2 – The evolution of remote and automated systems. This chapter considers the potential evolutionary paths that robotic and AI systems in human-machine teams will take in conflict, including: headquarters and decision-making; cyber and information operations; and remote and automated platforms.

c. Chapter 3 – Impacts on conflict. This chapter considers the effects of robotic and AI development on conflict across the observe, orient, decide and act (OODA) loop.

d. Chapter 4 – Human-machine teaming. This chapter is the heart of the concept and examines why optimised human-machine teams will be essential to developing military advantage. It considers likely strengths and weaknesses, trust and confidence issues and how to optimise human-machine teaming.

e. Deductions and insights. A summary of the principal deductions and insights to guide coherent force development, strategy and policy.

f. Annex A – Understanding assessments of autonomy. Annex A briefly considers the meaning of the term autonomy in contemporary discussions about robotics and automated systems.

Contents

Chapter 1

Context

> Some technologies are so powerful as to be irresistible.
>
> Greg Allen

Section 1 – The impact of robotics and artificial intelligence

1.1. Robotics and artificial intelligence (AI) have the potential to be transformative military technologies on a par with radio, aircraft, computers and nuclear weapons. Because of the ubiquitous nature of the dual-use technologies of AI and robotics, the impacts on conflict are a matter of when, not if.[1] The effects of these technologies on economics, conflict and society are likely to be increasingly profound and, in the long term, offer new opportunities for strategic overmatch and operational advantage. The United States of America (US), China and Russia have all declared strategies to achieve offset advantage through robotics and AI.

1.2. Maximising our ability to overmatch opponents will require leaders at all levels to be open to new ideas and encourage learning and experimentation. Success will be determined by outperforming potential opponents in an enduring cycle of development and countermove. Doing this will not only help deter possible opponents but, when deterrence fails, it will give us potential overmatch through: increased situational awareness; lighter physical and cognitive loads; sustainment with increased anticipation and efficiency; increased force protection; and, ultimately, superior manoeuvre options in and across all domains. The greatest advantages the confluence of artificial intelligence (AI) and robotics development will allow are:

- the ability to scale physical mass and battlefield points of presence increasingly independent of the numbers and locations of human combatants;

1. Dual-use technologies are those that can be used for both commercial and military purposes.

- extending the reach and persistence of our intelligence, surveillance and reconnaissance (ISR) and weapon systems; and

- information advantage[2] for understanding, decision-making, tempo of activity and assessment.

The importance of artificial intelligence

1.3. Fundamental to understanding the changes to conflict that are underway, is to understand the dependency remote and automated systems (RAS) have on AI. The rapid pace of development in AI since 2012 represents the maturation of a field that has existed in concept for over 50 years. However, the convergence of vast data sets, powerful hardware and advanced algorithms has made intelligent computing and increasingly capable robotics a reality.

Artificial intelligence will be a critical determining factor in remote and automated systems capability

2. For the purpose of this publication information advantage is defined as: the competitive advantage gained through the continuous, adaptive, decisive and resilient employment of information and information systems. Joint Concept Note 1/18, *Human-Machine Teaming*.

In short, autonomous systems are inherently, and irreducibly, artificially intelligent robots.

CNA Analysis and Solutions,
AI, Robots and Swarms[3]

1.4. One of the key limitations of using RAS is the balance between access to the electromagnetic spectrum (EMS) bandwidth required for remote operation, and the level of independent operation allowed by the automated system's capabilities. The more capable a platform's level of automation, the less remote control it needs. This creates a lower bandwidth demand and, if it does not need a permanent active remote control, the system has greater resilience of operation in contested or congested EMS.

1.5. Three elements are pivotal to creating AI systems, these are shown in Figure 1.1. They are:

- computing capability – hardware;
- developing advanced algorithms – software; and
- access to the sufficient quantities and quality of data – both the data to train the system and data to be exploited.

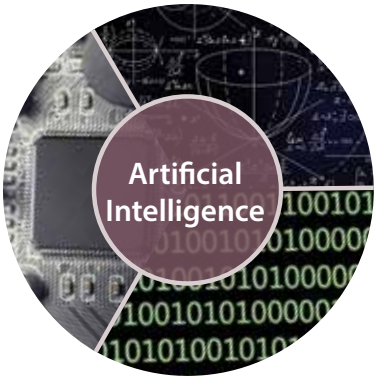


Figure 1.1 – The three elements of artificial intelligence

3. More information is available at https://www.cna.org/CNA_files/PDF/DRM-2017-U-014796-Final.pdf

1.6. Underpinning the three elements pivotal to creating AI systems are two critical indirect elements. These are:

- people with the qualifications, experience and proficiencies to understand and generate the algorithms required; and
- investment that allows access to the computing capability and data required.

Section 2 – Civil sector dominance

1.7. Civil commercial investment in AI and robotic technologies and the recruitment of subject matter experts dwarfs that of any state. Multiple Silicon Valley and Chinese companies spend more annually on AI and robotics research and development than the entire US government on research and development for all mathematics, robotics and computer science combined.[4] The impact is a shift in the relative rates of innovation from defence to commercial firms with the best systems already, and remaining, in the civilian sector. Military access to the best technologies will become a challenge, except in national crisis situations.[5] States and major technology firms will become increasingly averse to sharing the best AI systems, just as deep-level cryptographic algorithms are already highly valued and protected today.

Section 3 – Economic drivers for capability divergence

1.8. Technology trends are increasingly divergent, driven by economics. Those with the economic means to acquire the best hardware and programmes will have access to computing services and sensing capabilities far beyond those at market entry levels. Not all AI is created equal; it will vary in levels of capability. Actors able to acquire the best AI will have a significant advantage over those with less capable machine learning systems. In the most extreme case of

4. Allen, G. and Chan, T., *Artificial Intelligence and National Security*. China's declared state spending plan comes closest to comparable commercial investment levels.
5. Some Western commercial entities have publicly declared policies stating they will not contract with defence or security agencies which may compound the challenges facing the UK Ministry of Defence (MOD). This is in stark contrast to other states which have enshrined access rights to expertise, technology and data in their national legislation.

divergence, the outcome could even be a level of permanent advantage beyond that afforded by any prior technological revolution.

> AI's role as [an] innovation supercharger [could] deliver a strategic, and perhaps permanent, economic and military advantage to a country that develops a significant lead in exploiting AI applications. Because of this recursive-improvement property, and because AI applications also facilitate the automation of labor, it is possible to imagine a breakaway economic and innovation growth… which then guarantees it will be the first to discover the next generation of innovations, and so on.
>
> Greg Allen and Taniel Chan,
> *Artificial Intelligence and National Security*

1.9. Individuals or groups with access to advanced AI may accrue immense wealth while the pressure of automation simultaneously squeezes middle and lower incomes with potential socio-economic and military implications. National gross domestic product (GDP) may become a poor measure of the robotic and AI capability each state can employ. Subject matter expert human capital is more likely to become a critical factor. For example, Russia has a relatively low GDP, but a strong track record in maths and programming skills which are crucial to AI development. Likewise, a small but advanced city state, such as Singapore, could rapidly develop a military potency far beyond its population size through its industrial base in robotics and computer hardware, and access to programming expertise.

1.10. Technical capabilities like precision, automated navigation, remote operation and image recognition will become cheaper through exploiting commercially available systems in products like phones, quadcopters and self-driving vehicles. So, although the wealthiest actors will have exclusive access to the most capable systems, the cost of what have previously been considered expensive precision warfare capabilities will fall and become more widely attainable. As a result, minor actors will increasingly punch above their weight.

Section 4 – Skills shortage

1.11. The major strategic issue for all actors – nations or technology giants – is a chronic skills shortage. There is a significant shortage of skilled graduates, software engineers and computer technology staff with the necessary skills to develop the full breadth of possible AI enabled technologies. Early investment in education to generate subject matter experts may represent the critical long-term source of economic and military advantage for a nation.[6]

1.12. Competition for talent and investment is global, and it is therefore important to understand access in that context. Technology giants want to lock-in intellectual property as fast as possible and are finding the best way to do this is to lock-in the researchers and top talent in the field. Major technology firms have used mergers and acquisitions to secure personnel and simultaneously eliminate competition from capable start-ups. This process of buying the company to secure the personnel is conducted for sums that average $5 - 10 million per subject matter expert secured.[7]

Access to skilled personnel will be essential for achieving technological advantage

6. The UK Government has an extant strategy to support UK skills. It is important that Defence is engaged with this strategy to ensure the MOD's requirements are supported. This strategy is available at https://www.gov.uk/government/publications/uk-digital-strategy
7. 'Artificial Intelligence: the next digital frontier'. Available at http://mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-artificial -intelligence-can-deliver-real-value-to-companies

1.13. Defence will struggle to compete in such a recruiting market. We will need to be innovative to secure access to subject matter expertise. Defence will also need to nurture sufficient in-house knowledge and understanding to generate intelligent customer capabilities. This will be essential to: understand where AI should be exploited; translate operational requirements and constraints to non-military AI experts; and support the generation of effective military-industrial teams for innovation, problem solving and force development.[8]

Section 5 – UK expertise, industry and data access

1.14. Exploiting AI requires access to data. In addition, large, high-quality data sets can be a pre-requisite to train AI systems. Efficient use of AI systems has the potential for substantial cost reductions in Defence processes. However, a key challenge will be that many Ministry of Defence (MOD) data assets are fragmented or locked-in proprietary application programming interface solutions controlled by supplier companies.[9] Accessing the necessary data sets to train and test the solutions is a challenge for any commercial team seeking to deliver MOD-specific AI capabilities.

1.15. The UK has a strong pedigree in the theory and algorithmic side of AI development. There is also a strong academic and research base in the UK for RAS, including self-driving vehicle technologies and the software needed for designing swarming robotic systems.[10] Access to expertise is vital and, in addition to improving its own skill base, the MOD could maintain a register of security cleared UK nationals with AI and robotics skills.

1.16. Platforms and even subsystems, will be increasingly likely to have deeply embedded AI software and services. Obtaining guarantees and establishing assurance over the behaviour of such systems will be very difficult. Buying constituent elements of platforms from foreign suppliers will be increasingly risky. The UK needs to form a sufficient industrial and research base in robotic

8. For more information on the UK's human capital, see the MOD Innovation and Research Insight Unit's *Autonomy: Stocktake of the UK External Research Base.* This is only available to UK military personnel on the Defence Intranet at http://defenceintranet.diif.r.mil.uk/Organisations/Orgs/HOCS/Organisations/Orgs/CSA/DSTStrat/Pages/InnovationandResearchInSightUnit(IRIS).aspx
9. An application programming interface (API) is a set of routines, protocols, and tools for building software applications. An API specifies how software sub-components should interact. A good API makes it easier to develop a computer programme by providing all the building blocks, which are then put together by the programmer.
10. The remote and automated systems (RAS) industrial manufacturing base in the UK is weaker than the research base, partly due to our poor exploitation of research base innovations in the past.

and AI technologies onshore, or be capable of assuring components procured by clearly established AI, robotics and cyber security standards.

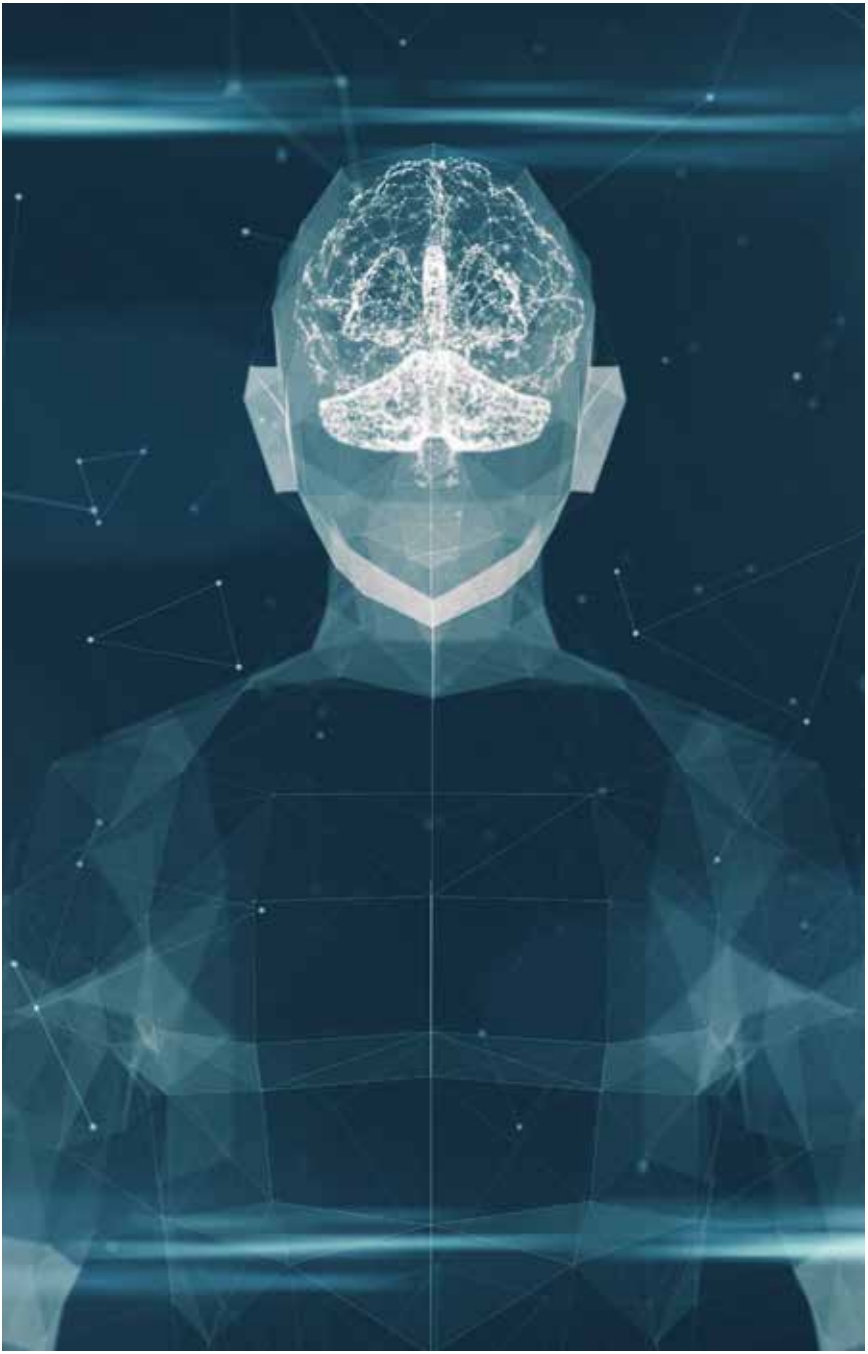
1.17. Asian states and the US currently host the majority of silicon chip and information technology manufacturers, vital for AI and robotics development. However, looking further ahead, the potential industrial impact of novel computer architectures and non-standard chip types may become important for AI systems. Low power AI chip technologies – where the UK has a good research base – will be vital for RAS. Defence-specific research into energy efficient AI and alternatives to deep learning algorithms could exploit UK based expertise and develop a source of national, military-technological advantage. For those with the underpinning technologies AI is going to: run faster; at lower costs; with lower power demands; and offer broader options to integrate AI into future military platforms.[11] This may be as valuable an advantage as the ability to fabricate high grade steel during the Victorian age.

Low power artificial intelligence chip technology will be vital for remote and automated systems

11. 'Deep learning startups in China: report from the leading edge'. Available at http://www.cogniteventures.com/2017/07/07/deep-learning-startups-in-china-report-from-the-leading-edge/

Key deductions and insights

- Robotics and artificial intelligence (AI) have the potential to be transformative and deliver long-term advantage.
- AI and robotics development will allow the ability to scale physical mass; extend reach and persistence; and enable better exploitation of information for advantage.
- In a restricted bandwidth environment greater automation capability offers the potential for advantage and resilience of operation.
- AI requires hardware, software and access to data; critically these are underpinned by skills and investment. Development will be led by the private sector.
- Actors able to acquire the best AI will have a significant advantage over those with less capable systems.
- Access to precision, automated navigation, remote operation and image recognition will become cheap and accessible.
- The major strategic issue for all actors – nations or technology giants – is a chronic skills shortage. Early investment in education to generate subject matter experts may represent the critical long-term source of economic and military advantage for a nation.
- The UK needs to form a sufficient industrial and research base in robotic and AI technologies onshore, or be capable of assuring components procured by clearly established AI, robotics and cyber security standards.



Chapter 2

The evolution of remote and automated systems

> Artificial intelligence is the future, not only for Russia, but for all humankind. It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become the ruler of the world.
>
> Vladimir Putin[12]

Section 1 – Phases of evolution

2.1. The principal constraints on using artificial intelligence (AI) and robotic capabilities are size, weight, power, cost, computing capability, algorithmic development, data access and communication bandwidth limitations. These limitations are being overcome – in particular, the size, weight and power as well as the cost of AI chips.[13] This will in turn decrease the bandwidth demand for control and increase the endurance of platforms. The economic drivers for such technological advances are the same as those that inexorably drove smartphone evolution; today's smartphones deliver what was only a decade ago the performance of a supercomputer.[14]

2.2. One further key constraint is associated with the need for industry to be able to verify, validate and certify defence products. Typically, defence systems are highly deterministic in terms of their behaviours; if X happens, they do Y. However, many advanced AI techniques can be both non-deterministic[15] and opaque in terms of the inability to explain why the system made a particular decision. This is an important consideration that will have a substantial impact

12. More information is available at https://www.wired.com/story/for-superpowers-artificial-intelligence-fuels-new-global-arms-race/
13. Examples include IBMs True North chip and the Neural Engine chip in the iPhoneX. Available at http://www.research.ibm.com/articles/brain-chip.shtml as well as https://www.wired.com/story/apples-neural-engine-infuses-the-iphone-with-ai-smarts/
14. Allen, G. and Chan, T., *Artificial Intelligence and National Security.*
15. Non-deterministic systems are characterised as those where very small changes to inputs can produce very large changes to outputs. Non-deterministic systems are associated with unpredictability.

upon all safety critical uses of AI, including on where and how defence can use AI until assurance mechanisms are developed.

2.3. In a similar fashion to trends in computer development, the capability growth of remote and automated systems (RAS) is likely to be exponential rather than linear. While development may appear low in earlier years, huge advantage will be available to those able to exploit these foundational developments in later years. The likely exponential nature of robotic and AI evolution makes identification of a timescale for development impossible to predict with any confidence, but does suggest three overlapping phases.

a. **Augment.** Initially, RAS should offer low-level augmentation to existing capabilities. Current force structures and operating concepts will require amendment rather than a complete revision. We should expect an acceleration in the move from manned to unmanned approaches for certain functions, particularly with: intelligence, surveillance and reconnaissance (ISR); cyber operations; and data processing tasks. RAS platforms will initially be mainly remotely operated with semi-autonomous supporting functionality and have limited technical integration with other systems. Intelligence and decision-support tools will be bespoke to narrow tasks, such as facial recognition software, rather than acting as holistic intelligent support systems.

b. **Parallel.** As RAS become more advanced, we should anticipate them operating alongside legacy systems as peer capabilities. Decision-support tools, automated logistic monitoring and remote ISR, as well as loitering munitions will offer step changes in military capability alongside contemporary platforms improved by retrofitting AI and robotic technologies. The key attributes are likely to be those listed below.

- **Coverage.** RAS technologies will offer greater coverage of the battlespace. Increased range and endurance combined with low unit costs, will offer opportunities for affordable mass and increased points of presence.

- **Information volume.** RAS will execute **persistent stare** missions, exponentially increasing the volume of ISR data collected. This increase will necessitate automation of information flows, data synthesis and the use of decision-support technologies.

- **Integration.** The integration of functions and data transfer across multiple RAS types will improve as technical integration protocols become established. These protocols must include cyber security and the ability to assure communication systems, algorithms and data.

- **Command support and sustainment.** Communications, cyber and electromagnetic activities will also be enabled by cognitive electromagnetic spectrum (EMS) management systems and unmanned nodes.[16] Sustainment will be improved, in the first instance, by improved stock and platform monitoring and anticipation; but also by automated logistic delivery.

c. **Supersede.** In this phase we should anticipate decreasing unit costs and maturing capabilities that make some (but not all) current capabilities obsolete. Extremely expensive low population platforms or facilities, particularly those with demanding logistic tails, low mobility, or large electromagnetic signatures, risk becoming liabilities that can be neutralised, or destroyed, for a fraction of their cost. Obsolescence may be by direct overmatch, for example, unmanned air combat platforms operating faster than human reaction speeds, with beyond human endurance and tactical anticipation.[17] Or, it could be through indirect overmatch, for example, heavy armoured platforms whose high fuel demands cannot be met through the attrition that swarming ISR and loiter munitions cause to fuel tanker fleets.

Decreasing costs and maturing remote and automated systems will overwhelm some existing capabilities making them obsolete

16. Cognitive electromagnetic spectrum (EMS) management systems, such as cognitive radio or cognitive radar, monitor the EMS, including the behaviour of other agents using it, and alter their own actions to optimise their ability to: communicate; detect, jam or attack others; and remain undetected by opponents.
17. Artificial intelligence (AI) is already capable of dominating simulated aerial combat against United States of America (US) fighter pilots. It did so using no more than the processing power available in a tiny, affordable computer (Raspberry Pi) that sells for as little as $35. See 'New artificial intelligence beats tactical experts in combat simulation' available at http://magazine.uc.edu/editors_picks/recent_features/alpha.html

Section 2 – Lessons from the past – augment, parallel and supersede

2.4. The future may, in many respects, be similar to the inter-war years. In World War 1 battleships were the decisive weapon at sea, so during the inter-war period, battleships received the majority of naval investments. Hull displacement almost tripled, main batteries grew in power and doubled their range, secondary batteries improved, radar was installed, speed was increased by 50 percent, cruising range more than doubled, and armour was thickened.[18] None of these advances changed the fundamental capabilities of the battleship – they simply provided improvement on existing strengths. This is typical of mature technology – even massive investment leads to only incremental improvement.

2.5. By contrast, aviation was in its infancy in 1914; aircraft were slow, had limited range and were lightly armed. Aircraft merely supplemented, and were subsidiary to, ships at sea and armies on land; aircraft were primarily systems to conduct ISR and enable over-the-horizon fire control. Despite this lower priority, by 1941 carrier aviation dominated naval warfare with most of the advances in aircraft design and production initially developed for civilian use. Aircraft production, a highly competitive business, led to rapid technological advances.[19]

A Sopwith Camel takes off from the World War 1 aircraft carrier, HMS Pegasus

18. Breyer, S., *Battleships and Battlecruisers, 1905 - 1970.*
19. Hammes, T.X., *Technologies Converge and Power Diffuses: The Evolution of Small, Smart and Cheap Weapons.*

2.6. A relatively modest investment in new technologies resulted in massive increases in military capability. As a result, aircraft – the cheap and many – first augmented battleship main batteries conducting ISR and targeting for its firepower, then went on to operate in parallel. They subsequently went on to overwhelm and destroy the few and expensive battleships as aircraft carriers superseded them.[20] The key to the aircraft's evolutionary leap was its extreme usefulness and low cost for improvement. The battleship once dominated maritime surface warfare, but the aircraft and its ability to operate in the third dimension offered surveillance, attack, movement of stores, passengers and casualties, and control of the air.

2.7. The broad utility of AI and robotics already outstrips that of the many mature technologies, which are often many orders of magnitude more expensive to incrementally improve. Investment in emerging technologies will only accelerate their capabilities and pervasive use.

Section 3 – Evolution in headquarters and decision-making

2.8. The use of automation offers opportunities to better exploit information to improve understanding, decision-making and tempo. It will also enable smaller headquarters and more agile command and control.[21] Current UK command systems remain based on significant numbers of staff in static locations with large installed information technology systems. Current configurations are rigid, vulnerable to attack and expensive to reconfigure or redeploy. The move from paper-based to electronic-based workflows has added information awareness and data volume, but at the expense of reduced mobility or structural flexibility. In addition, future intelligence, surveillance, target acquisition and reconnaissance systems will generate much larger volumes of real-time data which will be impossible to process without automated support.[22] Data fusion, automated analysis support and visualisation technologies will be essential to achieving manageable cognitive loads, not just for commanders and staff, but also for platform operators, dismounted combatants and support staff.

20. *Ibid.*
21. For more details on agile command and control see Joint Concept Note (JCN) 2/17, *Future of Command and Control*, Chapter 4.
22. A single MQ-9 Reaper sortie already generates between 20 and 40 laptops worth of data.

2.9. Effort will be required to automate the information collection, processing and management cycle. Assuming the underlying information technology systems are migrated to a modern containerised and plug-and-play configuration, then the creation of a more modular and intelligent data service becomes viable. This requires using existing machine learning and visual analytic platforms, but not advanced AI. Bespoke AI may be required for specific applications, such as automating the analysis of visual and audio data flows. This process would make best use of core headquarters staff and reduce the need for augmentees to deploy a headquarters. As further improving technology enables continuing force development, the optimum number of staff for a headquarters will reduce.

2.10. Longer term research efforts should be focused on the use of intelligent software agents that manage all aspects of information processing. Ultimately, this could eliminate technological constraints that confine us to our current monolithic headquarters approaches. The whole system could be built on a federated, disaggregated and self-organising peer-to-peer command, control, communications, computers and intelligence (C4I) network – effectively a combat cloud. Such a system should be able to draw on reachback access to cloud-based servers, but be capable of resilient operation provided by command and control applications across a variety of in-theatre platforms. From an operator's perspective such a system will handle user requests for information and data passage as an intelligent assistant service.

2.11. As designs evolve, the software agents will be able to pre-filter, fuse and classify all data flows, eliminate paralysing information overload, and accelerate the observe, orient, decide and act (OODA) loop of decision-makers. Such AI enabled command and control systems could also be proactive in prompting decision-makers to examine emerging issues and anticipate demands.

2.12. We can begin to picture how this may feel from a user perspective now; our relationship with machines is changing. Technology is moving away from the banks of single purpose switches, keyboards and screens that required low level, simple, control inputs to dumb machines. As the machines get smarter, the interface can, and must, become intuitive and natural for humans. The gateway into the machine world is likely to become an AI avatar and an interactive three-dimensional representation of the environment showing only useful pre-filtered elements of the massively increased levels of incoming information.

This avoids drowning the user in data and supports increased decision-making quality and tempo.

2.13. Future headquarters could become a largely virtual service, with high levels of resilience, adaptability and lower operational costs. End users, including commanders, could access the service from their mobile platforms. The AI agents could also be used to support planning and coordinate actions, including modelling threats and anticipating opponent scenario responses. Such a dynamic headquarters model would also be easily mirrored in training, supporting a more immersive and representative learning experience delivered prior to conflict.

2.14. This evolution will generate new risks and dependencies for command and control, particularly an increasing dependence on cyber and electromagnetic defences for resilience. It will also demand that staff are practised in reversionary modes of operating. Effective integration of systems across such a C4I system will demand security and assurance of communications systems, AI algorithms, and data. For example, underpinning datasets or incoming information can be poisoned in ways that are invisible to the human eye, causing unintended, potentially dangerous, outcomes.[23]

Headquarters will increasingly become virtual in their nature

23. See Chapter 3, paragraphs 3.4 – 3.6.

Section 4 – Evolution in cyber and information operations

2.15. The application of AI and automation to cyber systems is the most immediate arena for evolution and advantage. The cyber domain's intrinsically codified nature, the volume of data, and the ability to connect the most powerful hardware and algorithms with few constraints of EMS bandwidth, power access, or limits on speed and repeatability of actions creates an environment where AI can rapidly evolve and optimise to their assigned tasks.

2.16. The most challenging type of cyber attack that organisations deal with is the advanced persistent threat (APT).[24] With an APT, the attacker is actively hunting for weaknesses in the defender's security, constantly primed and waiting for the defender to make a mistake. Currently this is a labour-intensive activity and requires highly skilled personnel. With the growing capabilities in machine learning and AI, 'hunting for weaknesses' will be automated to a degree that is not currently possible, and, critically, it will occur faster than human-controlled defences can respond. This will demand AI-enabled adaptive defensive capabilities.

2.17. Not only will AI increase the variety and tempo of cyber attacks; it will also decrease the cost and increase the variety of actors able to undertake this activity. As the requirement for skilled specialists involved in the attack diminishes, the limitation will become access to the AI algorithms needed to conduct such an attack. In other words, any actor with the financial resources to buy, or steal, an AI APT system could gain access to tremendous offensive cyber capability; even if that actor is relatively ignorant of Internet security technology. Given that the cost of replicating software can be nearly zero, that may hardly present any constraint at all; this is likely to be a live issue by 2020 or soon thereafter. For example, the state-of-the-art AI is being trained in tactical reasoning by playing computer strategy games.[25] AI's like this could then be

24. Musa, S., 'Advanced Persistent Threat,' (2014). 'Advanced Persistent Threat (APT) is a set of stealthy and continuous hacking processes often orchestrated by human targeting of a specific entity. APT consists of three major components: advanced; persistent; and threat. **Advanced** signifies sophisticated techniques to exploit vulnerabilities in systems. **Persistent** indicates that an external command and control is continuously monitoring and extracting data from the target. **Threat** indicates the intent to attack as vulnerabilities are identified.'
25. One of the examples of the best cutting edge technology being developed in this way would be DeepMind's state-of-the-art system which is being trained to play StarCraft II as an AI research environment. More information is available at https://deepmind.com/blog/deepmind-and-blizzard-open-starcraft-ii-ai-research-environment/

readily adapted to drive APT cyber attack tactics, where the AI is competing against human or non-adaptive automated cyber defenders.

2.18. There appears to be no obvious stable outcome in terms of state *versus* non-state power, or offensive *versus* defensive cyber advantage. Advantage will depend on:

- the balance of research and development investments by all actors – civil and military;
- commercial espionage; and
- the speed with which actors can exploit emerging technologies.

2.19. These factors will help drive a thriving black market for stolen AI systems. As the best AI will be expensive, digital theft will pay.[26] We can expect to see sophisticated cyber attacks against companies like Google's DeepMind, IBM or Facebook if they are perceived to have the best AI code. Defending such civil/commercial assets may become an issue of national security. Furthermore, this will tend to shorten windows of technological advantage for states and companies alike, and potentially even give individual actors access to cutting edge technologies, offering advantage, for narrow windows of time.

A thriving black market will develop for stolen AI systems

26. A contemporary example is available at https://techcrunch.com/2017/06/28/uber-unaware-of-trade-secret-theft/

2.20. We must consider that the evolving cyber domain will be a complex ecosystem containing billions of competing AI agents. In the civil sector alone, before any combatant AI systems engage, there will be intelligent agents competing over: cyber security; finance; media influence; virtual currency mining; advertising; social media influence; pornography; and every other form of web-based interaction. Furthermore, the Internet of things is dissolving boundaries between the online and physical world.[27] Any deployed cyber system will be exposed to, and become part of, this wider ecosystem; an ecosystem that will also be increasingly indivisible from civil critical national infrastructure.

2.21. Automated systems can make quick decisions, much faster than humans can monitor and restrain them without the aid of machines. As online agents become more common, the probability for unexpected interactions that rapidly spiral out of control will increase. One example was the United States of America (US) stock market Flash Crash, in May 2010. A small trader's spoofing algorithm caused banks automated trading systems to enter an online loop that crashed the stock market, temporarily devaluing it by trillions of dollars, all in under 36 minutes. Chaos and friction will remain key elements of the nature of war in the virtual domain.

The probability of unexpected interactions will increase as online agents become more common

27. The Internet of things is the network of physical devices, vehicles, home appliances and other items embedded with electronics, software, sensors, actuators and network connectivity which enable these objects to collect and exchange data.

2.22. We should not think about future APT approaches simply as extensions of current thinking, which assumes we are discussing destructive computer viruses. APT cyber activities will be for surveillance, espionage, sabotage (viruses), deception, social engineering, psychological operations and compound cocktails of those activities. The emerging capacity of AI to create photo-real fake images, video and audio will have a major impact on the ability to influence a population.[28] They may also delay and degrade intelligence products, or damage confidence in their veracity. AI-enhanced forgery of audio and video has improved in quality and decreased in cost. When untrained amateurs, or automated social engineering web robots (bots) can produce fake videos at a higher quality than today's Hollywood computer-generated imagery, forgeries are likely to constitute a large proportion of online content.[29] Such forgeries will challenge trust in, and between, institutions.

2.23. Combined with cyber attacks and social media bot networks, AI-enhanced forged media could apply an overwhelming tempo and volume of online material that influences perceptions and even threatens social, political or economic stability. For example, consider the impact of the Syrian hacker who took control of the Associated Press' Twitter account announcing: 'BREAKING: Two Explosions in the White House and Barack Obama is injured'; in the two minutes following the tweet, the US stock market lost nearly $136 billion in value.[30]

Section 5 – Evolution in remote and automated platforms

2.24. The confluence of AI and robotics development will allow us to scale physical mass and battlefield points of presence increasingly independently of numbers and locations of human combatants. This is similar to the way the Internet has enabled access to information and projection of influence at scale and across the globe by individuals in the virtual domain. Cheap and relatively simple systems are already altering the economics of warfare; an area where the North Atlantic Treaty Organization (NATO) has enjoyed a technological-economic

28. Shunsuke, S., *et al.* 'Photorealistic Facial Texture Inference Using Deep Neural Networks'. Available at https://arxiv.org/pdf/1612.00523.pdf
29. Allen, G., 'Artificial Intelligence will make forging anything entirely too easy.' Available at https://www.wired.com/story/ai-will-make-forging-anything-entirely-too-easy/
30. *The Washington Post*, 'Syrian hackers claim AP hack that tipped stock market by $136 billion. Is it terrorism?', 23 April 2013.

advantage since the 1980's. In March 2017, US Army General David Perkins revealed a US ally had used a $3 million Patriot missile against a quadcopter that cost $200 from Amazon.[31] Shortly thereafter, it emerged that Houthi rebels, in Yemen, had employed low-cost drones to disable Patriot missile systems in Saudi Arabia. As General Perkins joked, "I'm not sure that's a good economic exchange ratio."[32] Future options, such as pilot tunnelling, where defensive systems are overwhelmed by employing massed cheap systems, are increasingly viable.[33] Understanding what this means for the way we fight and force development will be significant.

2.25. Novel combinations of human-machine teaming will offer a range of new capabilities.[34] They will present opportunities to augment human teams and manned platforms and even create massed effect, such as swarms. Networked mass – large numbers of interconnected sensors and soldiers, vehicles, ships and aircraft – contribute to resilient ISR networks, understanding and enable manoeuvre. Cheap, smart systems can provide resilience by absorbing casualties on a scale that will not be viable, or desirable, using a solely manned force; they will also be used to overwhelm an opponent's defences.

2.26. Optimising command and control of such systems will be essential. Manoeuvre will consist not only of the intelligent employment of advanced remote and automated capabilities, but also the rapid redesign and fielding of such systems. Cheap, bespoke systems are likely to offer opportunities in mass which are unaffordable with platforms of extremely high quality and cost. However, to realise advantage from cheap bespoke-to-task mass, demands shorter equipment life cycles, and far more agility in our procurement and logistic systems.

31. More information available at https://www.youtube.com/watch?v=6v7nfB5bV3E&feature=youtu.be&t=14m54s
32. *DefenseNews*, 'Report: Houthi rebels flying Iranian-made 'Kamikaze drones' into surveillance radars'. Available at http://www.defensenews.com/articles/report-houthi-rebels-flying-iranian-made-kamikaze-drones-into-surveillance-radars
33. Hammes, T.X., 'The Future of Warfare: Small, Many, Smart VS. Few & Exquisite?' *War on the Rocks*. Available at https://warontherocks.com/2014/07/the-future-of-warfare-small-many-smart-vs-few-exquisite/
34. Capabilities such as: smart minefields; small, cheap unmanned aircraft systems; unmanned ground systems; unmanned surface vehicles and unmanned underwater vehicles. Cheap, smart munitions such as successors to the US Switchblade system and telexistence systems that allow remote interaction in dangerous (explosive ordnance device), dirty (chemical, biological, radiological and nuclear) or dull (occasional oversight of repetitive and low risk tasks) environments.

Remote and automated systems will change the tactics and economics of warfare

2.27. Advances in these technology fields are not the preserve of the wealthiest nations. Unmanned aircraft systems (UAS) have been fielded by combatants on all sides of the conflicts in Syria, Iraq and Ukraine, both state and non-state. The decreasing costs, increasing capabilities and proliferation of RAS may change the economics and character of conflict.[35] Furthermore, the ability of actors to employ RAS while avoiding attribution will become increasingly possible, especially where adapted commercial RAS are used. This offers particular opportunities to those seeking to exploit complex conflict situations, foster disorder or escalate conflicts.

35. Russia, China and the US all have major modernisation programmes based on this assumption. The US Department of Defense third offset strategy, which makes extensive reference to man-machine teaming and weapon systems 'autonomy' is the most well-known to UK military audiences, however its core constituents mirror development programmes announced by Russia and China that pre-date the US strategy.

Section 6 – Countering remote and automated systems

2.28. It is essential that we develop the means to disrupt and defeat the range of robotic threats we face. Just as RAS is not one single entity or technology, there is no single approach or technology for countering it. Since the technologies and uses of RAS are evolving, identifying how such systems can be defeated is necessarily reactive and evolving. However, there are areas for exploration and potential advantage. The greatest sources of advantage are likely to lie in counter-RAS systems that can attack common vulnerabilities efficiently and economically – rather than those that can only deal with a few or are far more expensive to employ than the threats they neutralise. There are several means to counter remote and automated systems.

a. Cyber operations. Advanced computing and AI capabilities are fundamental enablers to many RAS systems, and particularly the more capable ones. Systems that can attack or compromise software dynamically – through the Internet or using the EMS as a delivery medium – either during conflict, or in advance of conflict[36] offer means to counter or even subvert RAS systems.

b. Electronic warfare. In addition to cyber effects on software, electronic warfare – including electromagnetic pulse weapons – can attack hardware, and control or reporting signals between RAS systems and their controllers, through jamming or deception techniques. As AI capability grows, the dependence of RAS on continuous control signals is likely to decrease.

c. Adversarial artificial intelligence approaches. Capable AI systems are already able to subvert other target AI systems without access to their internal workings or details of their programming.[37] At a simple programming level this has been identified in the subversion of Internet bots by other, more capable bots, to influence massed online content. This is not just an online activity, AI systems have been subverted by influencing physical visual cues that are unnoticeable or appear innocuous to human inspection.[38]

36. For example, the infection of components prior to manufacture.
37. 'Researchers fooled a Google AI into thinking a rifle was a helicopter.' Available at https://www.wired.com/story/researcher-fooled-a-google-ai-into-thinking-a-rifle-was-a-helicopter/
38. See Chapter 3, paragraphs 3.3 – 3.5 for more information.

d. **Direct attack.** Conventional systems, like counter-rocket, artillery and mortar exist but may struggle as RAS numbers and swarming technologies improve. Countering swarms is an area for consideration in its own right and a significant challenge for contemporary weapon systems, requiring high rates of fire, rapid targeting, discrimination, reliability and large magazine capacities. Directed energy weapons (DEW) may offer better options in future if technical challenges can be overcome. The most widely recognised DEW are laser weapons; however, the term includes radio frequency directed energy applications and beams can create several effects. DEW have particularly pertinent qualities relevant to RAS, including low cost of shot, extremely deep magazine capacities and negligible time of flight, but they also have power generation, heat dissipation and dwell time challenges.

The USS Ponce is able to make use of an installed directed energy weapon system

Key deductions and insights

- Artificial intelligence (AI) will transform war fighting. Pursuing it will be non-negotiable. Full exploitation of the potential of AI will be constrained by what can be assured.
- The likely exponential nature of robotic and AI evolution suggests there will be three overlapping phases: augment; parallel; and supersede. Early adopters have the potential to reap significant advantage.
- The use of automation offers opportunities to better exploit information to improve understanding, decision-making and tempo. The larger volumes of data generated in the future will be impossible to process without automated support.
- Greater automation will generate new risks and dependencies for command and control, particularly an increasing dependence on cyber and electromagnetic defences for resilience.
- The application of AI to cyber systems is the most immediate area for evolution and advantage. AI-enabled cyber attacks will demand AI enabled cyber defences.
- Gaining access to the best AI offers potentially significant advantage. Defending such AI assets – including those owned by the commercial sector – may become an issue of national security.
- Remote and automated systems may alter the economics and character of conflict.
- Realising mass from cheap systems demands shorter equipment life cycles and greater agility in procurement and logistic systems.
- It is essential to develop the means to disrupt and defeat the range of robotic threats.

Notes:


Military

Chapter 3

Impacts on conflict

> I am certainly questioning my original premise of a fundamental nature of war that does not change… You have got to question that now. I just don't have the answer.
>
> James Mattis[39]

3.1. This chapter considers the effects of robotic and artificial intelligence (AI) development on conflict across the observe, orient, decide and act (OODA) loop. The OODA loop is summarised in Figure 3.1.

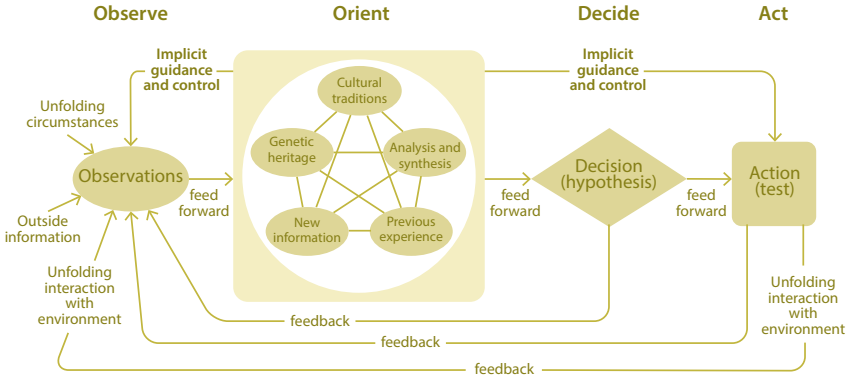


Figure 3.1 – The OODA loop

Section 1 – Observe

3.2. Robotic and AI systems are likely to revolutionise and dominate observation. The proliferation of sensors and machine learning systems outperforming humans at recognition and pattern detection is likely to increase. Systems are being developed to enable change and anomaly detection that are

39. More information is available at https://phys.org/news/2018-02-artificial-intelligence-poses-nature-war.html

platform agnostic, allowing systems to not only recognise specified targets, but also detect the unusual or out of place.[40]

3.3. Lessons from contemporary conflicts in the Middle East and Ukraine where the use of unmanned aircraft systems (UAS), with small radar cross sections and low heat emissions, highlight a trend for the future. Previous assumptions over air supremacy guaranteeing a benevolent sky for any side are increasingly obsolete. Even where enemy aircraft have been neutralised, being observed (and hence targeted) by remote and automated systems (RAS) or remotely hacked civilian sensors (phones and cameras amongst others) must be continuously treated as a risk. RAS that perch or harvest power from solar or wind energy are being developed, significantly increasing the endurance potential of intelligence, surveillance and reconnaissance (ISR) systems.[41]

3.4. Assuming that everything on future battlefields will be observed at all times is wrong. AI can get it wrong, it can be fooled and it can have biases.[42] Attributing infallibility to either AI or more conventional programming is similarly erroneous. Algorithmic decisions are not automatically equitable just by virtue of being the products of complex processes, and the procedural consistency of algorithms is not equivalent to objectivity. Data, including images or audio, can be poisoned, sometimes in ways too subtle to be detectable to a human, in order to fool a target AI. An example is shown in Figure 3.2 below.[43]

Figure 3.2 – Deception of an artificial intelligence by an adversary

40. Boeing's Corvus is one such system. The Royal Navy also achieved impressive results in identifying abnormal actor behaviours in Exercise INFORMATION WARRIOR 17.
41. Examples include, Airbus's Zephyr unmanned aircraft system (UAS), or the University of Sherbrooke's fixed wing perching UAS. These examples are available at https://www.gov.uk/government/news/mod-buys-third-record-breaking-uav and https://spectrum.ieee.org/automaton/robotics/drones/reliable-perching-makes-fixedwing-uavs-much-more-useful
42. These biases are not necessarily the same as human biases, but could be; if the AI is trained from data that represents society then societal biases mirrored in that data may be learned.
43. More information is available at https://blog.openai.com/adversarial-example-research/

3.5. The first image in Figure 3.2 is a normal image. The second is a magnification of changes to the colours of the pixels, indistinguishable to the human eye, of that panda designed by an adversarial AI. The third image is the resultant picture, the target AI is fooled into identifying the image as a gibbon.

3.6. Real world experimental threat examples have included stickers applied to road signs causing self-driving cars to read them as entirely different signs.[44] Similarly, the artist Adam Harvey created a fractal like pattern for clothing that convinces facial recognition cameras that thousands or millions of faces are present. And it is not only a visual problem – signals inaudible to the human ear can trick voice-controlled assistants like Amazon's Alexa into taking unwanted actions, for example, visiting a website and downloading malware. Any input type can be targeted if it can be accessed and its AI algorithm or training data identified. AI is vulnerable, in part, because it lacks actual intelligence;[45] deception is alive and well in the age of AI.

3.7. Emission control, stealth and decoys will be increasingly important in limiting the adversary's situational awareness. RAS, using speed and widely distributed platforms, may successfully create many false positives simultaneously by creating clutter and decoy signals.

Section 2 – Orient

3.8. The vast quantities of data gathered from ISR and open source systems is too much to be handled by humans in a timely and effective manner. The use of joint action to influence actors is dependent on exploiting this information despite a burgeoning data deluge.[46] The increases in data collection, promised by funded and future capabilities in the equipment programme, are not currently matched by single information environment integration, or by automated analysis to support decision-making. Militaries risk failing to capitalise on improving ISR capabilities where this is not addressed.

3.9. Assumptions that automation will inevitably make understanding and therefore decision-making easier, or effortlessly increase spans of command,

44. More information is available at https://arxiv.org/pdf/1602.02697.pdf
45. The ability of systems to test for deception will begin with voting systems that cross-check different means of assessment for consistency, but will only begin to approach critical introspection of assessments beyond the era of narrow artificial intelligence.
46. More information about joint action can be found in Joint Concept Note (JCN) 1/17, *Future Force Concept* and Joint Doctrine Publication (JDP) 3-00, *Campaign Execution*, 3rd Edition, Change 1.

are simplistic. Automated systems typically have harder limits and less ability to function in situations outside their design parameters. They will therefore tend to either fail catastrophically, or recognise that they are reaching their limits and demand human attention at points of high stress, potentially handing over the problem to an insufficiently engaged human with no opportunity for them to understand the issue in time to then act to avert disaster. Despite the capability of modern autopilots, airline pilots remain essential for those few occasions when the autopilot can no longer cope.

3.10. Challenges also exist on the human side of the interaction. Skills that go unpractised (including because of automation) wither.[47] Furthermore human attention is neither constant nor consistent. Simply monitoring systems holds people's attention poorly. It is often very difficult for a previously unengaged person to be able to ramp up their mental alertness at a point of crisis, or orient themselves sufficiently quickly to the key variables and context in time to act.

> How do you establish vigilance at the proper time? 23 hours and 59 minutes of boredom followed by one minute of panic.
>
> Major General Michael Vane[48]

3.11. Humans interacting with machines tend to be far more mentally engaged when:

- **searching** for an already understood and defined object; or
- **exploring** for things of interest, for example, boundaries and anomalies, or undefined targets of the 'I'll know it when I see it,' variety.

Therefore, while details will vary, good human-machine teaming will require intuitive human-machine interfaces and be optimized for 'searching and exploring' tasks for their operators. Our systems will also need a means of monitoring the cognitive workload of the human commanders and operators such that information is represented to optimise human attention, and even take on work where operators become overloaded.

47. A textbook case of skill fade and a failure to understand the situation can be found in the case study of the crash of Air France flight 447 in 2009. More information is available at https://www.vanityfair.com/news/business/2014/10/air-france-flight-447-crash
48. Hawley, J., 'Patriot wars', Centre for New American Security, available at https://www.cnas.org/publications/reports/patriot-wars

Section 3 – Decide

3.12. Increased connectivity has previously led to a tendency for senior decision-makers to monitor and intercede in low-level tactical action in real time. Improvements in RAS technologies and human-machine teaming are likely to reverse this trend. The requirement for speed at the tactical level will benefit the side able to optimise their human-machine decision-making. The team will seek to exploit the detection, recognition, optimisation and efficiency advantages of AI in the OODA loop where the risk appetite of the human commander in the human-machine team judges its advantages to outweigh risk and it is lawful to do so. This is especially the case once hostilities are initiated, and more so where conflict pits AI capable of rapid tactical acts against one another. Context will be critical in establishing where the bounds of automation are optimised; for example, discrete, non-complex areas of battlespace only containing combatants are likely to be dominated by RAS and AI actions.

3.13. Tactical RAS actions, cycling decisions rapidly against a similarly equipped adversary, will still demand human oversight. Firstly, to monitor for emergent effects in the interactions between RAS that depart from the goals of its human managers, in that moment. Secondly, to intervene where circumstances go beyond the capability of the RAS and to exploit human mental strengths and mitigate machine weakness. Just as military professionals today must understand the capabilities and limits of weapons and platforms, in the future they must understand RAS. Advantage will not automatically lie with the force that has the newest or most expensive algorithm, but more likely with the most effective human-machine team.

Human oversight is required to monitor and respond to unexpected emergent behaviours of highly automated systems

3.14. The impact of RAS in future conflicts is often discussed only from a technological perspective, but war is also psychological. Surprise, or shock, has often dislocated and defeated material advantages in fighting power. Surprise can be attained by deception, either passively through concealment, or actively, by using false signalling such as feints or decoys. Surprise can also be attained through speed of manoeuvre, and here RAS offer a distinct tempo of action advantage. Combined with precision and the distributed firepower of massed RAS, the advantages in acting first may be considerable. However, an automated system encountering an unexpected adversary move will not possess initiative, but nor will it be susceptible to the dramatic cognitive effects of shock, including paralysis of decision-making. Rival AIs will engage in high-speed battles of pattern detection, deception and spoofing.

Section 4 – Action

3.15. Historically, a qualitative edge in speed of decision, action and precision at a critical point has often overcome advantages of mass – even where platforms have been relatively equal in performance.[49] The approach of powerful actors has traditionally involved concentrating forces in time and space, creating favourable local force ratios to defeat less organised enemy forces.[50] In turn, weaker actors typically seek to offset opponents' strengths through deception, dispersal, fortification of positions and use of terrain.[51]

3.16. Concentrated combat power, whether in defence or offence, is difficult to coordinate and conceal. Massed forces are often cumbersome to manoeuvre and vulnerable to attrition. Dispersed networks, conversely, create problems of command and control and sustainment. Dispersed formations generate less concentrated firepower and are susceptible to defeat in detail. Advances in RAS are likely to challenge these dynamics. Command and control problems are reduced once high-level instructions become viable to RAS and they can report back in low-bandwidth, metadata terms rather than continuous streams of data.[52] Furthermore, AI enabled RAS can share information locally in combat clouds or swarms, thereby learning and improving performance, without necessarily

49. A classic example would be General 'Stonewall' Jackson's wide flanking offensive at the battle of Chancellorsville 30th April – 6th May 1863 during the American Civil War.
50. Such as the success of the 7th Panzer Division drive through France in 1940.
51. An historic example would be the Hezbollah's tactics in the 2006 Lebanon conflict.
52. For example, the human initiated orders, "Go to this grid and conduct surveillance in this specified area of interest, reporting enemy military units, remain silent otherwise, less a platform health status update every three days."

referring back to a hierarchical command and control system. This tactical learning, combined with better detection, recognition and precision increases lethality and intensifies the imperative to identify, understand and target quicker than an opponent.[53]

3.17. Small, high-quality distributed AI networks will have the potential to defeat mass and concentration by those using older, or inferior, AI. This ought to favour the defence, since attackers will be reluctant to concentrate for fear of attrition, and they will have only fleeting targets to concentrate against if they do attack. This could generate potential for the re-emergence of advantage for the tactically defensive;[54] but also the operationally offensive actor. Acting before an opponent enables an actor to establish a network and place sensors without interference or observation. The network can observe patterns of life and survey the electromagnetic spectrum. In doing so it gains better data and contextual information against which to spot anomalies. The operationally slower actor then risks detection as soon as they begin to manoeuvre, given AI's acuity in pattern recognition. Meanwhile, the operationally aggressive actor – even if surprised by the timing of an opponent's tactical attack – can then move quickly into an active defence posture or a counter-attack. Because RAS accentuates speed of decision and situational awareness, through pattern recognition, tactical AI systems undercut the traditional advantages of shock. The interdependencies between technologies, tactics and strategy are likely to be complex; understanding them demands force development experimentation in laboratories, wargaming and live exercises.

People, ideas, and hardware – In that order!

Colonel John R. Boyd

53. The imperative to rapidly achieve identification and targeting solutions does not imply immediately striking at the point of identification; often the key will be to balance rapidly understanding an opponent's network without overexposing our own forces before a position of decisive advantage can be achieved.
54. Technologies that emerged during the era of the American Civil War – barb wire, mines, automatic weapons – offered advantages to the *tactically defensive* that would endure through to the latter years of World War 1.

Key deductions and insights

- Robotic and artificial intelligence (AI) systems will revolutionise and dominate observation.
- AI is vulnerable, in part, because it lacks actual intelligence. AI tends not to degrade gracefully at the edge of its capabilities, it tends to fail catastrophically.
- The vast quantities of data gathered by intelligence, surveillance and reconnaissance and open source systems will be too large to be handled by humans in a timely and effective manner.
- Advantage will not automatically lie with the force with the newest or most expensive algorithm, but with the most effective human-machine teams.
- Effective human-machine teaming will require intuitive and optimised human-machine interfaces.
- AI-enabled tactical learning, combined with better detection, recognition and precision will increase lethality.
- The interdependencies between technologies, tactics and strategy are likely to be complex. Our understanding must be improved through research, experimentation and training.

Notes:

Chapter 4

Human-machine teaming

> The real problem is not whether machines think but whether men do.
>
> B.F. Skinner

Section 1 – Why human-machine teaming is essential

4.1. At the core of future military advantage will be the effective integration of humans, artificial intelligence (AI) and robotics into warfighting systems – human-machine teams – that exploit the capabilities of people and technologies to outperform our opponents. The game of chess provides an excellent example of human-computer collaboration and a cautionary tale about over-extrapolating when computers outperform humans. In 1997, IBM's Deep Blue beat the chess grandmaster Garry Kasparov. Many observers regarded this, and the subsequent triumph of DeepMind's artificial intelligence (AI) at the game of Go, along with AI that consistently beats Top Gun instructors in air-to-air combat, as the beginning of the end for human cognitive dominance.[55] However, evidence suggests that the future is more complex than machine beats human. A useful example comes from chess in 2005; a competition was held allowing any combination of human and computer chess players to compete. The competition resulted in an unexpected victory that Kasparov later reflected on:

> "The surprise came at the conclusion of the event. The winner was revealed to be not a grandmaster with a state-of-the-art PC but a pair of amateur American chess players using three computers at the same time. Their skill at manipulating and 'coaching' their computers to look very deeply into positions effectively counteracted the superior chess understanding of their

55. Ernest, N., *et al*. 'Genetic Fuzzy based Artificial Intelligence for Unmanned Combat Aerial Vehicle Control in Simulated Air Combat Missions', *Journal of Defense Management*, available at https://www.omicsonline.org/open-access/genetic-fuzzy-based-artificial-intelligence-for-unmanned-combat-aerialvehicle-control-in-simulated-air-combat-missions-2167-0374-1000144.pdf

> grandmaster opponents and the greater computational power of other participants. Weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process… Human strategic guidance combined with the tactical acuity of a computer was overwhelming."[56]

4.2. United States (US) automated air defence post-incident lessons and Defence Science and Technology Laboratory work on variable autonomy shows that optimised human integration into combat systems is critical to the effectiveness of remote and automated systems (RAS) in guarding against unanticipated catastrophic error.[57] Catastrophic error is not a term used to exaggerate; as conventionally programmed automated systems become more complex, when they fail they do not gracefully degrade, they will collapse.

Section 2 – Humans and machine strengths and weaknesses

4.3. There is a tendency to assume that the difficulty of automating a task is proportional to the amount of human mental effort associated with that task, but that is a poor assumption. A useful rule of thumb when considering how well machines can be applied to a task is to understand how readily the activity can be codified. The clearer the rules, metrics and recognition features a task has, the higher the likelihood that a machine can be optimised to undertake the task. This is leading to surprising outcomes: roles traditionally considered to be challenging and that are often highly paid that involve data sorting or deterministic analysis like accounting, insurance estimation, legal documentation reviews and medical diagnostics, are proving to be automatable. Whereas waiting on tables or care assistance for the elderly – often much lower wage attracting roles – are proving difficult to automate. The last jobs to be automated in society will not simply be those of highly paid professionals. Actions that we as humans struggle to comprehend will be very difficult to codify and ultimately automate.

4.4. Significant staff efficiencies can be made if we adopt automation in data-centric and readily codified roles. Defence must consider how to automate whilst retaining understanding of the processes being automated. Furthermore,

56. Kasparov, G., 'The Chess Master and the Computer', *The New York Review of Books*, available at http://www.nybooks.com/articles/2010/02/11/the-chess-master-and-the-computer/
57. Hawley, J., 'Patriot Wars', *Centre for New American Security*, 25 January 2017. Available at https://www.cnas.org/publications/reports/patriot-wars

approaches to human-machine teaming that adopt the automate what you can, leave the humans to fill in the remainder view are likely to build systems that are cheap, but less resilient or effective. No network, organisation or system can be completely resilient; they experience constant change, operate under varying degrees of uncertainty and face evolving threats. The key to resilience in force and system design is therefore tied to adaptability and understanding what humans are best at and what machines are best at in the era of narrow AI.

4.5. Broadly, computer algorithms are good at sorting and searching through large amounts of structured data (for example, text and document processing, people and enterprise information, and genetics), doing deterministic analysis (for example, counting, classifying and game playing), and producing predictable mechanical interactions (for example, manufacturing, flying and driving). Computer algorithms are not as good at understanding complex unstructured data (for example, images, acoustics and environment structure or context), doing non-deterministic analysis (for example, road scene understanding or predicting human behaviour), and undertaking dexterous actions (for example, fine manipulation requiring touch and pressure feedback or handling deformable objects). Despite these being more challenging fields for machines, it must be understood that machines are increasingly outperforming humans at some of these challenging tasks, including image recognition. They do not suffer from concentration lapses, or fatigue, assuming access to a constant power supply.[58]

Automated systems are increasingly outperforming humans in activities that can be codified and have clearly defined goals

58. It is worth understanding the relative strengths of man and machine in the energy efficiency of data processing for tasks machines find difficult. Low power artificial intelligence (AI) chips are emerging but, as yet, producing an equivalent amount of processing power to the human brain for the energy efficiency of the human brain – which only uses 20 watts of power – remains beyond computing technology.

4.6. Essentially, computer algorithms are challenged by uncertainty and ambiguity in both data and decision-making. As a result, humans outperform machines at understanding context, and are likely to continue to do so for a long time. Machines are poor at exercising nuanced judgement on the complex or ambiguous contexts that then moderate decisions. Also, because machines are programmed or trained using established datasets relevant to a task or problem, encountering a new problem or something wildly divergent from established datasets tends to cause failure.[59] In contrast, the human ability to adapt to new situations is generally far superior, even imperfect responses are likely to be more functional. This is in part because humans use mental substitutions or approximations from familiar skills or tasks to approximate answers. AI technologies are typically able to conduct mental substitutions appropriate to new contexts only in specific narrow confines and can even suffer from catastrophic forgetting, where previous algorithm optimisations or skills at tasks are simply lost when trained on new tasks and data.[60]

4.7. These factors mean that the last roles likely to be automated will be where personnel conduct activities that demand contextual assessment and agile versatility in complex, cluttered and congested operating areas. This will apply across domains but, as an example to make the point, consider the dismounted combatant conducting an assault in an urban environment. While RAS will offer a lot of new forms of advantage in urban conflict in general, in the assault in close complex terrain humans dominate the ability to exercise continuous contextual judgement and readjustment – is it a child who has picked up a gun, or a combatant? Likewise, the ability to open doors, use varied tools, ropes, ladders, or move debris to manoeuvre indoors are simple to the point of instinctive for the human, but exceptionally difficult or impossible for the robot.

4.8. Force design and concepts of operation must also consider legal and societal factors of employment. This tends to revolve around the targeting debate, and while considerations about targeting are highly relevant, it is an oversimplification to assume this is the totality of the issue. The reality for military operations – which are broader than just war – will be more diverse, more complex and highly contextualised. For example, a US unmanned underwater

59. Generative adversarial networks can operate with sparse datasets. These approaches may eventually trump the machine learning approaches that require large datasets and can be used to explore new problems. Another example is DeepMind's AlphaZero a generic reinforcement learning algorithm which, only through playing itself, learned to play chess at the very highest level in only four hours.
60. More information is available at https://arxiv.org/pdf/1612.00796.pdf

vehicle was pulled from the ocean by the Chinese Navy, prior to holding and handing it back to the US five days later. The lack of certainty in international law on the status of such vessels is likely to have caused the Chinese to treat the vessel differently than they would had it been a manned warship.[61] Similarly, unmanned systems are unlikely to be considered a comparable commitment by populations, allies or adversaries to 'boots on the ground' in assessments of military commitment, political risk and demonstrations of national will. Balancing imperatives to deploy humans against the moral and legal imperatives to minimise risk to life and the potential advantages of employing more disposable RAS will be complex in some instances.

While remote and automated systems provide advantage in complex terrain, human versatility and judgement remain essential

61. More information is available at https://www.icrc.org/en/international-review/article/international-law-and-military-use-unmanned-maritime-systems

Section 3 – Human-machine teaming and force design

Mission command in an artificial intelligence age

4.9. Future force design must find the optimal mix of manned and unmanned platforms, and balance employment of human and machine cognition for various tasks. Because RAS will be a key means of generating mass, there will be a high ratio of AI driven systems – both physical and virtual – to people. There will be proportionally fewer points of human consciousness within the system. Optimising how we use human mental and physical capacity within such a force will become a key factor in out-manoeuvring and out-thinking opponents. It follows that AI must be used to free up human mental capacity in a flexible and adaptable way. At the heart of mission command is optimising independence of subordinate action to allow initiative and generate tempo, balanced against measures to create unity of effort and managing risk. Risk is assessed within context, and will remain a human responsibility. Dynamically managing levels of automation in RAS to balance risk against advantages from machine capability – mass, tempo, pattern recognition and precision – within changing contexts will be how mission command is applied in an AI age.

4.10. The concept of an optimal span of command is driven by human cognitive loading and how many active elements an individual can control, even where the interpersonal demands like leadership are absent. If human operators are task-saturated piloting basic unmanned systems or managing unanticipated behaviours in technologically complex, but uncooperative systems, they might not have the mental capacity required to undertake higher-level thinking. Human multitasking has its limits, and those limits are often reached quickly.

Mental capacity and spans of command

Human working memory capacity is seven plus or minus two items. For dynamic active memory, this drops to two or three items.[62] Army Doctrine Publication *Land Operations* states that 'a span of command should not exceed five subordinate … groupings'. In 2016, the Defence Science and Technology Laboratory and Qinetiq undertook collaborative autonomy trials, looking at the ability of operators to cope with active engagement with up to four or five systems simultaneously, provided no one system required concentrated attention. When mental loading from one task significantly increased, the attention other areas received diminished significantly. For example, watching four video feeds from unmanned aerial vehicles circling areas of interest is sustainable, but trying to fly an unmanned aircraft system through complex terrain while avoiding threats is an intense activity that totally occupies an individual; at which point secondary or tertiary systems will be ignored; whether the operator intends to do so or not. However, if a swarm appears as a single entity from an operator's perspective, trials have demonstrated that operators can effectively control swarms of 80 or more unmanned aerial vehicles.[63]

4.11. The limits of human mental capacity mean the ability to dynamically vary the level of active control that operators exercise over systems becomes a fundamental enabler to tempo and team effectiveness. An ability to rapidly increase the amount of automated functionality used in RAS then allows the team to park RAS on lower risk tasks well suited to machine execution. As a safeguard, there must be automated alerts and warnings in place to attract human attention in sufficient time for orientation, action and decision, if required. This frees up the humans to focus on tasks of importance or those poorly suited to execution by machines alone, in particular, ambiguous or contextually dependent tasks.

Force designing for adaptability

4.12. Dynamic reorganisation will remain vital to a military's ability to adapt to new missions and changing circumstances.[64] Cheap, bespoke to task RAS are

62. Boff, K., *et al. Handbook of Perception and Human Performance*, 1988.
63. Lewis, M., *et al. Scaling-up Human Control for Large UAV Teams*, 2004.
64. For more details on the importance of agility in reorganisation see Joint Concept Note 2/17, *Future of Command and Control*, Chapter 2.

likely to offer opportunities to generate mass. However, if bespoke systems can only be controlled by set operators through a non-transferable control link, the RAS will only offer the team additional tools when that operator is positioned to act on the target. Therefore, RAS in a human-machine team will be most effective as a flexible pool of assets that a wide variety of individual operators can call upon.[65] Open architectures will be required to enable the dynamic adoption and reorganisation of RAS without the need to re-engineer control systems or retrain personnel for each change. Control interfaces must also be intuitive and impose low cognitive loads.

4.13. The combat cloud must be able to provide decision support information to those best prepared to decide and act.[66] The team or individual that has the greatest situational awareness must be able to assume control of the RAS best suited to the task and at the same time release unneeded systems. This will optimise the force's adaptability. Simple controls and policies will enable this adaptability. For example, pre-set limits fixing how much individuals can control systems; in this way an operations room watchkeeper should not be able to push a button to try and fly a complex airframe, but, they could, with permission, briefly take control of its electro-optical camera and quickly aim it and orient the pilot to a target.

4.14. No universal set of design principles for RAS is likely to be found. Individual technological assessments of systems must be judged against intended function within an anticipated operating environment in the same way as manned ships, aircraft or armoured vehicles. However, to judge the value of large numbers of lower cost systems requires us to change the idea of qualitative superiority from an attribute of the platform to an attribute of the force. In doing this, our assessments must also include a determination of how effectively human cognitive and physical ability is applied within a force design, and this measure is likely to correlate strongly with the force's adaptability. If the team can act rapidly and efficiently and, most importantly, if they can adapt effectively to changing circumstances, then the structure, policies and technical systems in the force are well designed.

65. This will require an asset owner who can track these loaned assets across what may be a very agile command and control structure. This is especially likely to be the case in the land environment where multiple force elements will want access to a variety of unmanned aircraft system and unmanned ground system capabilities for limited windows of time at different places across the battlespace, often in rapid succession.
66. Deptula, D., Lieutenant General USAF (Retired). 'Evolving Technologies and Warfare in the 21st Century: Introducing the "Combat Cloud"'. Mitchell Institute Policy Papers, Volume 4, September 2016.

Experimentation, training, data generation and iterative improvement

4.15. To exploit developments in AI and robotics as they continue to emerge, we will need to adopt an aggressive strategy of iterative experimentation, prototyping, concept and technology development, and organisational refinement. High quality live and synthetic collective training and experimentation with AI systems will be essential to optimise our ability to create effective human-machine teams. Training and experimentation with real users will be vital for operators to understand the strengths, weaknesses and critical limitations of such AI systems while also providing vital data to improve AI responses, including about the human behaviours in the team. We must train and grow with our AI assistants such that the machine can tailor how it interfaces with us as individuals and with the wider team. Such collective training will need to be dynamic, varied, realistic, conducted against thinking opponents and act as surrogate warfare in which to experiment, develop and build collective trust and confidence. Such high-quality, human-machine team training will not just be required to train and develop the teams, but also to establish a better understanding of Defence's future requirements which are likely to change and evolve across all Defence's lines of development.[67]

High quality live and synthetic training with artificial intelligence will be essential to create effective human-machine teams

67. The Defence lines of development are: doctrine; information; equipment; personnel; infrastructure; logistics; training; organisation; and integration.

Section 4 – Trust, assurance and legislation

4.16. The increasing capabilities of robotic and AI systems will be limited not only by what can be done, but also by what actors trust their machines to do. There are multiple contributing elements to this assessment, such as individual operator confidence in a system, assurance regimes, and policies and legislation.

4.17. As RAS become more cost effective and develop a reliable track record so military forces will start to adopt them in significantly larger numbers. Trust and reliability are therefore key issues that drive the level of confidence, and hence the degree of automation we place in RAS. The fundamental factors affecting our trust in systems are listed below.

a. **Mechanical understanding.** The more we understand how a system works, the more comfortable we tend to be with it.

b. **Predictability.** If we can consistently anticipate how a system will behave, we will increase our confidence in using it and, in particular for AI, the systems tolerance to faults and erroneous data arising from real-world interactions.

c. **Familiarity.** Trust is emotional as well as intellectual; the more frequently we use or see a thing working effectively, the more likely we are to have confidence in it.

d. **Context.** Our trust in systems and their effectiveness is dependent on the context in which they are used. We normalise and adapt for this routinely; for example, driving more cautiously in icy road conditions.

4.18. Trust takes on added significance when seeking mass effect. As the ratio of RAS to operators increases, overall system trust declines rapidly if the reliability of automated functions or control systems decreases, and a lack of trust causes an increase in the cognitive workload for users.[68] As a consequence reliability and trust can become determining factors in effective RAS-to-human ratios for human-machine teams.

4.19. As AI becomes more capable we will have a greater spectrum of discretionary automation to balance against if we choose to use exclusively

68. Dixon, S.R., Wickens, C.D. (2003), 'Control of Multiple UAV's: A Workload Analysis'. Presented at the 12th International Symposium on Aviation Psychology.

human agency. Assuring non-deterministic systems designed to dynamically adapt and optimise decisions is inherently difficult. Achieving this requires understanding common AI errors, developing effective test strategies and managing AI adaptation. We must also be careful to avoid information being filtered by AI in such a way that only one rational decision is available to the operator, leading to the illusion of a human made decision. The development of appropriate standards and robust assurance and certification regimes will be critical, along with effective mechanisms to demonstrate meaningful human accountability.

4.20. Legal obligations and policies are unlikely to cede an opponent's military advantage in the near term. However, as future technologies emerge, particularly for systems supporting targeting and fires, we must consider the ethical and legal implications. Armed remote and automated systems must not only be trusted and safe, but also perform in such a way that they are seen to be safe and reliable by users and observers. Those developing such systems must ensure they are able to comply with international humanitarian law. Equally, legislative moves to encourage technology adoption within society must be scrutinised to ensure that in an ever more connected world, lines of accountability and responsibility are retained.

A Reaper MQ-9 remotely piloted air system prepares for take-off

4.21. Remote and automated systems are not single entities and AI encompasses an array of what are component-level technologies; furthermore we must remember that in evolution there is no single end point, there are trajectories and branches in multiple directions.[69] Moves to create legal obligations in advance of capabilities becoming technically possible, or even understandable, must be carefully and actively examined to ensure they are not unworkable or that they open legal avenues for others to misinterpret and misuse. To illustrate the difficulties in trying to define autonomy for regulatory purposes it is worth considering the problems faced by legislators in Nevada as they made laws to permit driverless cars to be used on public roads.[70] Initially they defined autonomous vehicles as those which substituted AI for human decision-making. Once the law was passed, it unintentionally placed heavy restrictions on commercial vehicle sales, due to the frequency with which modern cars functionally make substitutions for direct human control, such as crash avoidance systems and anti-lock brakes. The law was swiftly repealed.

4.22. In considering the future we must also remember that automation will increase across society, and where new technology is sufficiently safe and reliable, norms of trust and public appetites can be expected to follow. It may also turn out that in the future some highly automated weapons could actually be more able to comply with the Law of Armed Conflict principles of proportionality and distinction, rather than less able. If that does become the case, it may become difficult for a state to justify not using them.

Where automation increases across society, norms of trust can be expected to follow

69. More information is available at http://hplusmagazine.com/2011/01/19/what-technology-wants-what-kevin-kelly-says-interview-kevin-kelly/
70. Carlo, R. (2014), 'The Case for a Federal Robotics Commission', Centre for Technology at Brookings, page 6.

Key deductions and insights

- Human mental capacity is increasingly unable to cope with the data deluge involved in conflict, while computer algorithms are challenged by uncertainty and ambiguity in data and decision-making. Optimising human and machine capabilities in teams that maximise strengths and mitigate weaknesses is essential.

- Risk is assessed within context, and will remain a human responsibility. Mission command will change in an artificial intelligence (AI) age and will demand variable autonomy in remote and automated systems.

- High-quality live and synthetic collective training and above all experimentation with AI systems will be essential for us to learn how to optimise our ability to create effective human-machine teams.

- The increasing array of capabilities of robotic and AI systems will be limited by not only what can be done, but also by what actors trust their machines to do.

- The Ministry of Defence must continue to be proactive in considering legal, ethical and public concerns surrounding the use of robotics and AI.

Notes:

Deductions and insights

1. The following deductions and insights are those judged most critical to guide strategy, policy and force development for Defence and front line commands. They offer guidance on factors that will determine advantage in an era of robotics and artificial intelligence (AI) during conflict.

2. **The potential of artificial intelligence and protecting access.** The capability growth of remote and automated systems (RAS) is likely to be exponential rather than linear. While development may appear low in earlier years, huge advantage will be available to those able to exploit these foundational developments in later years. Gaining access to cutting-edge AI, by fair means or foul, offers the opportunity to achieve windows of technological advantage for states, companies and even individual actors. Defending such civil, commercial and military AI assets may become an issue of national security.

3. **Robotic and artificial intelligence systems are likely to revolutionise the battlespace.** AI-enabled tactical learning, combined with better detection, recognition and precision will increase lethality. It will offer opportunities to better exploit information to improve understanding, decision-making and tempo and will enable reduced headquarters size and more agile command and control. The larger volume of real-time data that is generated will be impossible to process without automated support. Deploying systems first enables an actor to establish a network and place sensors without interference or observation. AI will engage in high-speed battles of pattern detection and deception which will occur faster than human-operated defences alone.

4. **Creating mass effect.** Novel combinations of human-machine teaming will present opportunities to augment manned platforms and create massed effect. Networked mass – large numbers of interconnected sensors and soldiers, vehicles, ships and aircraft – will contribute to resilient intelligence, surveillance and reconnaissance networks, understanding and enabling manoeuvre. Cheap, smart systems will provide resilience by absorbing casualties on a scale that will not be viable, or desirable, using a solely manned force and will also be used to overwhelm an opponent's defences. Such systems are likely to offer opportunities in mass which

are unaffordable with platforms of extremely high quality and cost. However, to realise advantage from cheap bespoke-to-task mass, demands shorter equipment life cycles, and far more agility in our procurement and logistic systems.

5. Optimising human-machine teaming. Optimising human-machine teaming requires an understanding of what humans are best at and what machines are best at in the era of narrow AI. The last roles likely to be automated will be where personnel conduct activities that demand contextual assessment and agile versatility in complex, cluttered and congested operating areas. Optimising how we use human mental and physical capacity within such a force will become a key factor in out-manoeuvring and out-thinking opponents. High quality live and synthetic collective training and experimentation will be vital for humans to understand the strengths, weaknesses and critical limitations of such AI systems while also providing vital data to improve AI responses, including about human behaviours in the team.

6. Trust and assurance for artificial intelligence. The increasing array of capabilities of robotic and AI systems will be limited by not only what can be done, but also by what actors trust their machines to do. The more capable our AI systems are, the greater their ability to conduct local processing and respond to more abstract, higher level commands. The more we trust the AI, the lower the level of digital connectivity we will demand to maintain system control. Developing appropriate standards and robust assurance and certification regimes will be critical, along with effective mechanisms to demonstrate meaningful human accountability. Although legal obligations and policies are unlikely to cede an opponent military advantage in the near term, as future technologies emerge, particularly for systems supporting targeting and fires, we must consider the ethical and legal implications.

7. Accessing skills and the race for technological advantage. The major strategic issue for all actors – nations or technology giants – is a chronic skills shortage. There is a significant shortage of skilled graduates, software engineers and computer technology staff with the necessary skills to develop the full breadth of possible AI-enabled technologies. Early investment in education to generate subject matter expertise may represent the critical long term source of economic and military advantage for a nation. For some technologies, such as lethal effects or stealth, only the military will lead primary investment and must continue to do so for disruptive advantage. However, investment and development in the commercial sector will exceed Government research investment for other applications. The ability to exploit commercial technology developments in Defence-industrial partnerships faster

than potential adversaries will be increasingly important to achieving technological superiority.

8. The new economics of warfare. Technical capabilities like precision, automated navigation, remote operation and image recognition will become cheap through exploiting commercially available systems. The cost of what were previously considered expensive precision warfare capabilities will fall and become more widely attainable, giving minor actors the ability to punch above their weight. Employing massed cheap systems will not be optimal in all cases; we will need to fight with the few and capable and the cheap and many in the right mix. To judge the value of large numbers of lower cost systems requires us to change the idea of qualitative superiority from an attribute of the platform to an attribute of the force. Approaches to human-machine teaming that adopt the automate what you can, leave the humans to fill in the remainder view are likely to build systems that are cheap, but neither resilient nor effective.

Notes:

Annex A

Understanding assessments of autonomy

A.1. Although Defence's endorsed definitions can be found in the Lexicon, the intended meaning of the terms autonomous system or autonomy in contemporary discussions about robotic and automated systems varies widely.[71] There is no clear, definable and universally agreed boundary between what constitutes automation and what is autonomous because the assessment of autonomy and the term's use is subjective and contextual. Because it is subjective and contextual it is frequently used to mean different things by different commentators.

a. Tactical and technical contexts. In this context autonomy is an emergent property, where inherent technical capabilities combine to supersede the demands of the task and environment beyond the point that continuous active human control is required. In colloquial military technology discussions, robotic systems are often described as autonomous where they have long endurance, react effectively to external stimuli and require little or no human oversight for the duration of their mission. It is important to understand that this use of the term is frequently independent of the complexity of the system's programming. Even very simple mechanical devices, such as landmines, can sometimes fulfil these criteria.

b. Ethical or legal contexts. In the context of legislative or ethical debates autonomy is often used by participants to describe elements with agency and independent decision-making powers. Something far beyond the ability of simple mechanical devices. Arguably this description is reserved for sentient entities. For the duration of the era of artificial narrow intelligence (the scope of this publication and the technologically foreseeable future) no machine possesses ethical or legal autonomy; all machines will function because of some human initiation to undertake a task. For clarity, the Ministry of Defence's

71. Joint Doctrine Publication (JDP) 0-01.1, *United Kingdom Supplement to* NATO*Term.*

position, reiterated in September 2017, is 'that we do not operate, and do not plan to develop, any lethal autonomous weapons systems (while accepting that systems such as Phalanx CIWS are very highly automated)'.

c. Relative autonomy descriptors. Where autonomy is often more useful as a concept is when considered as a relative capability to accomplish a task, rather than as a binary autonomous or automated judgement. Understanding that autonomy can be variable by altering limits on programming across different functions will be conceptually important in the future. In particular as we increasingly choose how much artificial intelligence operation or direct human control we wish to employ in balancing risk and efficiency across different tasks.

Lexicon

Part 1 – Acronyms and abbreviations

AI	artificial intelligence
API	application programming interface
APT	advanced persistent threat
C4I	command, control, communications, computers and intelligence
DEW	directed energy weapons
EMS	electromagnetic spectrum
GDP	gross domestic product
ISR	intelligence, surveillance and reconnaissance
JCN	joint concept note
JDN	joint doctrine note
JDP	joint doctrine publication
MOD	Ministry of Defence
NATO	North Atlantic Treaty Organization
OODA	observe, orient, decide, act
RAS	remote and automated system
UAS	unmanned aircraft system
UAV	unmanned aerial vehicle
UGS	unmanned ground system
US	United States of America

Part 2 – Terms and definitions

This section is used to list definitions and descriptions which may be helpful to the reader.

advanced persistent threat
A set of stealthy and continuous hacking processes often orchestrated by human targeting a specific entity. APT consists of three major components: advanced, persistent, and threat. Advanced signifies sophisticated techniques to exploit vulnerabilities in systems. Persistent indicates that an external command and control is continuously monitoring and extracting data from the target. Threat indicates the intent to attack as vulnerabilities are identified. (Musa, S., 'Advanced Persistent Threat', Academia.edu. 2009)

artificial intelligence
The performance by computer systems of tasks normally requiring human intelligence, such as translations between languages. (*Concise Oxford English Dictionary*, 12th Edition)

automated system
In the unmanned vehicle or platform context, an automated or automatic system is one that, in response to inputs from one or more sensors, is programmed to logically follow a predefined set of rules in order to provide an outcome. Knowing the set of rules under which it is operating means that its output is predictable. (JDP 0-01.1)

autonomous system
An autonomous system is capable of understanding higher-level intent and direction. From this understanding and its perception of its environment, such a system is able to take appropriate action to bring about a desired state. It is capable of deciding a course of action, from a number of alternatives, without depending on human oversight and control, although these may be present. Although the overall activity of an autonomous unmanned system will be predictable, individual actions may not be. (JDP 0-01.1)

command and control capability
A dynamic and adaptive socio-technical system configured to design and execute joint action. (JCN 2/17)

information advantage
The competitive advantage gained through the continuous, adaptive, decisive and resilient employment of information and information systems. (JCN 1/18)

single information environment
A logical construct whereby assured information can pass unhindered from point of origin to point of need. The single information environment will include a single intelligence environment. (Defence Information Strategy, 2017)

Designed by the Development, Concepts and Doctrine Centre

Published by the Ministry of Defence
This publication is also available at www.gov.uk/mod/dcdc