DRONE WARS

Human Security and Emerging Military Technology

None too clever?

Military applications of artificial intelligence

Note: The term 'drone' is used interchangeably with 'Unmanned Aerial Vehicle (UAV)' in this report.

Drone Wars UK is a small British NGO established in 2010 to undertake research and advocacy around the use of armed drones. We believe that the growing use of remotely-controlled, armed unmanned systems is encouraging and enabling a lowering of the threshold for the use of lethal force as well as eroding well established human rights norms. While some argue that the technology itself is neutral, we believe that drones are a danger to global peace and security. We have seen over the past decade that once these systems are in the armoury, the temptation to use them becomes great, even beyond the constraints of international law. As more countries develop or acquire this technology, the danger to global peace and security grows.

Published by Drone Wars UK

Written by Peter Burt
December 2021

Design by Chris Woodward
www.chriswoodwarddesign.co.uk

Drone Wars UK
Peace House, 19 Paradise Street
Oxford, OX1 1LD

www.dronewars.net
info@dronewars.net

Contents

Introduction

Artificial Intelligence (AI), automated decision making, and autonomous technologies have already become common in everyday life and offer immense opportunities to dramatically improve society. Smartphones, internet search engines, AI personal assistants, and self-driving cars are among the many products and services that rely on AI to function. However, like all technologies, AI also poses risk if it is poorly understood, unregulated, or used in inappropriate or dangerous ways. As well as transforming homes and businesses, AI is seen by the world's military powers as a way to revolutionise warfare and gain an advantage over enemies. Military applications of AI such as the DART logistics planning tool (see below) have entered everyday use over the past couple of decades and new systems with worrying characteristics are rapidly being rolled out.

This briefing, one of a series published by Drone Wars UK as part of our 'Future Wars' project, examines military applications of AI and describes how the UK's military is beginning to adopt AI technologies before going on to outline the various risks associated with them.

US Army illustration of ground troops on patrol with drones and autonomous systems. Credit: US Army

What is AI?

Academic studies generally acknowledge that there is no universally agreed definition of AI because of the diversity of approaches to research in this field. The UK government has characterised AI as "Technologies with the ability to perform tasks that would otherwise require human intelligence, such as visual perception, speech recognition, and language translation".[1] Discussion of the various technical and computer science approaches captured by this definition is beyond the scope of this briefing, although it is important to point out that AI systems usually have the capacity to learn or adapt to new information or stimuli.[2]

AI relies on the processing of large amounts of data to identify statistical patterns. Despite decades of research, until recently neither the computer capacity nor the extensive datasets needed to allow AI systems to function have been available. Exponential increases in computer processing speed and storage capacity, internet speed and 'cloud computing', together with mass digitisation and the collection of vast amounts of data, are now allowing AI to come of age.

AI is often categorised as being either 'narrow' or 'general'. Current applications are all examples of narrow AI, where machines perform a specific task for a specific purpose. The umbrella term 'computational methods' is perhaps a better way of describing such systems, which fall far short of human intelligence but have more generally applicable problem-solving capabilities than conventional software. As AI systems work by recognising statistical relationships within data sets, they are commonly employed for the following functions:

- Automating tasks.
- Processing complex or large datasets.
- Predicting behaviour.
- Flagging anomalies or events of interest.
- Data tagging and error correction.[3]

General AI is the hypothetical ability of a computer system to perform a range of cognitive functions and respond to a wide variety of input data and understand and solve any problem that a human brain can. Although this is a goal of some AI research programmes, it remains a distant prospect.[4]

AI does not operate in isolation, but functions as a 'backbone' in a broader system to help the system achieve its purpose. Users do not 'buy' the AI itself; they buy products and services that use AI or upgrade a legacy system with new AI technology. Autonomous systems, which are machines able to execute

1 'Industrial Strategy: Building a Britain fit for the future. Department for Business, Energy and Industrial Strategy, November 2017. P37. **https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf**

2 Michael Copeland: 'What's the Difference Between Artificial Intelligence, Machine Learning and Deep Learning?' Nvidia blog, 29 July 2016. **https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/**

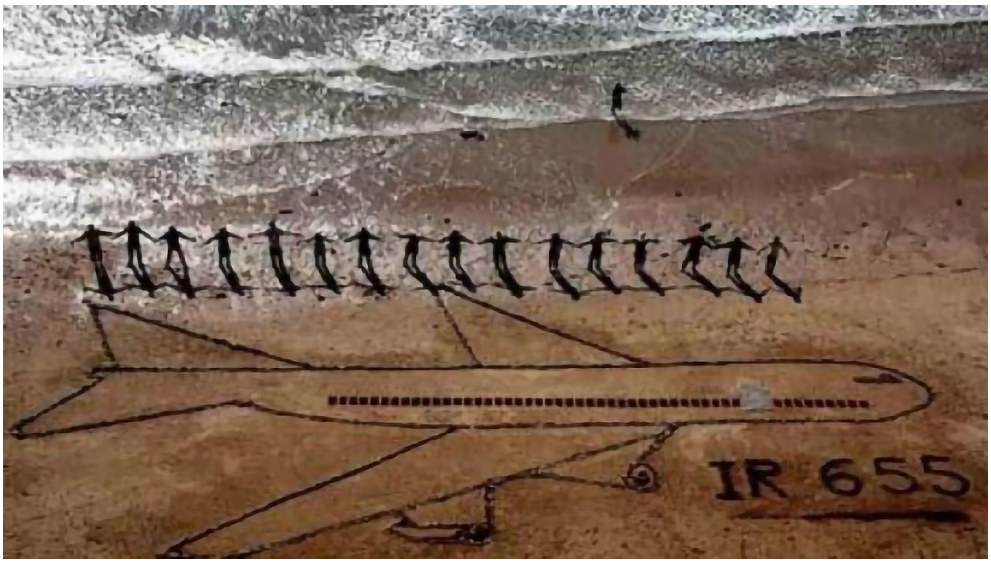
3 Lawrence Lewis: 'AI and Autonomy in War: Understanding and Mitigating Risks'. CNA Analysis & Solutions, August 2018. P6. **https://www.cna.org/CNA_files/PDF/DOP-2018-U-018296-Final.pdf**

4 Lawrence Lewis: 'AI and Autonomy in War: Understanding and Mitigating Risks', op cit. P8.

a task without human input, rely on artificial intelligence computing systems to interpret information from sensors and then signal actuators, such as motors, pumps, or weapons, to operate a mechanism which causes an impact on the environment around the machine.

Autonomous systems can work through what is sometimes described as *autonomy-at-rest* and *autonomy-in-motion*. Autonomy-at-rest describes systems that operate in software, or in the virtual world, whereas autonomy-in-motion describes systems that interact with the physical world. In a military context, both types of system can give cause for concern. Autonomy-at-rest systems could make critical decisions on the use of force which may have major consequences, while autonomy-in-motion systems could include lethal autonomous weapon systems (LAWS) – 'killer robots' able to make their own decisions on targeting and killing without any human input.

The use of automation and computing in military systems is not new. Automated weapons systems have been in existence since World War II, with advances in computer technology gradually contributing to their sophistication. In the 1940s some aircraft and air defence radars were fitted with transponders by which radar operators and radar systems themselves could determine whether the aircraft they were tracking were friendly or hostile. As technology advanced, air defence systems became able to identify aircraft and missiles by comparing their speed, radar profile, and head signature with a database, and were able to automatically correct course and home in on their targets using radar or heat-seeking sensors. Modern air defence systems such as the Patriot or Aegis missile systems, designed to operate against multiple high-speed incoming threats, are able to make targeting decisions for human approval or even engage targets without human supervision. This has proved to be deeply problematic: in such systems the human operator is reduced to playing a minimal but impossibly complex role, and automated air defence systems have been involved in a number of failures that have brought down civilian and military aircraft in friendly fire incidents.[5]



Vigil commemorating the 290 civilian victims of Iran Airlines flight 655, shot down by USS Vincennes in 1988. The ship's AEGIS air defence system gave operators little time to interpret the data it provided. Credit: Raheleh Zomorodinia

5 Ingvild Bode and Tom Watts: 'Meaning-Less Human Control: Lessons from Air Defence Systems on Meaningful Human Control for the Debate on AWS'. Drone Wars UK And University of Southern Denmark, February 2021. **https://dronewars.net/wp-content/uploads/2021/02/DW-Control-WEB.pdf**

The military and AI

Business and academia have led, and continue to lead, the development of AI since they are better placed to invest capital and access resources needed for research than the military and public sectors. As a result, it is likely that future military applications of AI will be adaptations of technologies developed in the commercial sector. Government research organisations such as the Defense Advanced Research Projects Agency (DARPA) and the Intelligence Advanced Research Projects Agency (IARPA) in the US, and the Defence Science and Technology Laboratory (DSTL) in the UK have a variety of AI projects which are intended to encourage collaboration with the commercial and academic sectors to adapt and employ AI and autonomous technologies for military purposes. Although this briefing draws principally on examples from the US and UK, China, Russia, and other military powers also have active programmes for developing military AI.

AI's characteristics make it attractive to the military. Features such as its rapid speed, ability to handle large and complex data sets to find patterns and undertake repetitive tasks accurately have shaped the roles to which it has been put by armed forces. Table 1 shows some of AI's attributes which are of particular interest to the military.

Table 1: Functions which support military uses of AI[6]

Rapid speed of analysis and action.
Performing simple automated tasks at scale.
Controlling robotic and autonomous systems.
Recognising patterns to predict future trends or detect anomalies.
Classifying and recognising objects and signals.
Optimising systems to achieve a goal.
Improving the quality of decision-making.

The next sections look at some of the many areas where AI is emerging as a military technology. In some of these applications AI is allowing, or may in future allow, the development of new products or systems. In other cases AI can be used to improve existing systems relatively cheaply in order to enable them to operate more efficiently, extend their lives, or give them new or improved roles. The US Air Force's U-2 spy plane, which first flew in 1955, has recently been flown by an AI co-pilot that is able to learn and adapt, unlike a conventional autopilot that merely obeys programmed routines to follow a planned route. The AI co-pilot, branded 'Artuμ' (Artoo) is an algorithm-based entity developed by the Air Forces' U-2 Federal Laboratory which, during a training flight, was able to control the aircraft's sensors and pass information about the location of missile launcher sites back to the human pilot. AI co-pilots are expected to regularly fly with human US Air Force pilots in the near future, and Artuμ was reported to have a good chance of being used in operations by the summer of 2021.[7]

6 Based on Forrest E. Morgan, Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, Derek Grossman: 'Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World'. RAND corporation, 2020. **https://www.rand.org/pubs/research_reports/RR3139-1.html**

7 Valerie Insinna: 'The tiny tech lab that put AI on a spyplane has another secret project'. Defense News, 11 February 2021. **https://www.defensenews.com/air/2021/02/11/the-tiny-tech-lab-that-put-ai-on-a-spyplane-has-another-secret-project/**

Military applications of AI

Remote working

Robotic and automated systems have traditionally been used to carry out 'dull, dirty, or dangerous' military activities to replace human operators. Autonomous technology can prevent the need for humans to enter risky situations in dangerous environments. Examples include explosives detection and disposal, mine sweeping operations at sea or on land, or reconnaissance in hostile environments.

Intelligence, Surveillance, and Reconnaissance

Intelligence gathering draws together large sets of data, including text documents, video and still imagery, electronic intercept information, and open source information from the internet. AI is beginning to play a role in processing and analysing this data.

An example of the use of AI in intelligence analysis is Project Maven – formally known as the Algorithmic Warfare Cross-Functional Team (AWCFT) – set up by the Pentagon in April 2017 to develop the use of AI for data analysis. By the end of that year, AI algorithms developed through Project Maven were in use in the field to assist with the analysis of drone video footage in the Middle East. The algorithms were 'trained' using thousands of hours of archived video footage, with objects labelled in advance by humans and covering different operating conditions, including different altitudes, object density, image resolution, and view angles. The project caused controversy for a number of reasons, not least because one of the contractors involved, Google, faced a backlash from employees concerned about the company's involvement in military work.[8]

Although Project Maven, as far as is known, focused on the use of AI for object identification, AI systems could be used to support a range of human activities at different stages in an intelligence analysis process. At an early stage in the process AI could be used to filter and triage material gathered in bulk, for example by using speech or signal recognition algorithms to 'clean up' data from noisy environments. At the next stage AI systems could assist in the analysis of raw data, for example through machine translation and summarisation of text, object identification from imagery, geo-locating images onto maps, or by fusing two-dimensional images to create three-dimensional models. Finally, AI could be used to analyse the behaviour of subjects of interest to derive insights about their activities, such as identifying a building's function based on a pattern-of-life analysis, and possibly also to predict future events and activities.[9]

8 Marcus Weisgerber: 'The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists'. Defense One, 21 December 2017. http://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/

Kate Conger and Dell Cameron: 'Google Is Helping the Pentagon Build AI for Drones'. Gizmodo, 6 March 2018. https://gizmodo.com/google-is-helping-the-pentagon-build-ai-for-drones-1823464533?rev=1520349331358

9 Alexander Babuta, Marion Oswald and Ardi Janjeva: 'Artificial Intelligence and UK National Security. Policy Considerations'. Royal United Services Institute, 27 April 2020. P11. https://rusi.org/sites/default/files/ai_national_security_final_web_version.pdf

Addressing some of these tasks, technology company Montvieux Ltd has developed a 'predictive cognitive control system' for the Ministry of Defence (MoD). The system is reportedly able to analyse complex data using deep learning neural networks to make confidence-based predictions of future events and outcomes which will be of direct operational relevance to the armed forces.[10]

Automation of organisational processes.

Armed forces, like all large organisations, rely on a large number of organisational, administrative, and data management processes to fulfil their aims. These are routine and often repetitive activities which may amount to a significant workload. The use of AI to automate these tasks has the potential to offer efficiency savings and improvements, freeing up staff time to deal with more complex matters. Such activities include elements of personnel management, logistics, and financial management and accounting.

Computer vision, natural language processing, and chatbots could assist in dealing with administrative tasks such as handling routine forms and enquiries, cross-referencing data (for example, in conducting security checks) and diary management. AI could also play a role in the automation of compliance and oversight processes, such as conducting audits of data management and detecting inappropriate or unlawful computer network use.[11]

AI could also assist with military logistics tasks. For example, the US Air Force uses AI to undertake predictive maintenance on a tailored as-need basis for each of its individual F-35 aircraft, rather than making repairs according to a maintenance schedule or when the aircraft has a fault. Real-time data from sensors embedded in aircraft engines and onboard systems provides data to a predictive algorithm which alerts technicians when an inspection or component replacement is needed.[12] At the beginning of 2021 the UK MoD announced a competition to seek ideas for technologies to enable the potential automation of military logistics chains across land, sea, and air.[13]

Cyber operations

Cyber security is an area where AI systems are being actively utilised. Cyber security threats such as malware attacks evolve rapidly and require a speed of response far greater than human decision-making allows. AI systems can proactively identify suspicious activity and respond to cyber-attacks in real time. AI cyber defence systems are able to identify elements of software that may be malicious without needing to rely on a database of known threats that have already been discovered. By scanning for suspicious patterns of behaviour as well as for potentially malicious codes, AI systems can spot signals indicating a novel cyber threat. They can also protect a network by identifying abnormal network traffic through making comparisons with system log data for normal traffic. Similar techniques can be used to detect insider threats and authenticate

10 'Intelligence technology to keep Joint Force Command one step ahead of adversaries'. Defence and Security Accelerator and Ministry of Defence, 17 July 2018. **https://www.gov.uk/government/news/intelligence-technology-to-keep-joint-force-command-one-step-ahead-of-adversaries**

11 Alexander Babuta, Marion Oswald and Ardi Janjeva: 'Artificial Intelligence and UK National Security. Policy Considerations'. Op cit. P9.

12 Kelley M. Sayler: 'Artificial Intelligence and National Security'. Congressional Research Service, 26 August 2020. P10. **https://crsreports.congress.gov/product/pdf/R/R45178/9**

13 'Right on Time: £800k for military logistics innovation'. Defence and Security Accelerator and Defence Science and Technology Laboratory. 2 February 2021. **https://www.gov.uk/government/news/right-on-time-800k-for-military-logistics-innovation**

computer network users based on their 'behavioural biometrics', such as how they use a mouse or keyboard or write in a document.

However, AI can also be used to engage in offensive cyber warfare. AI can be used to identify weak spots in network defences, and also to design novel malware. During a 'Cyber Grand Challenge' organised by DARPA in 2016 contestants developed AI algorithms able to autonomously identify and patch vulnerabilities in their own software while simultaneously attacking other teams' weaknesses. The algorithms were able to find and fix the security vulnerabilities in seconds, compared to months using conventional cybersecurity approaches. Software developed as part of the challenge was capable of simultaneously undertaking both offensive and defensive roles, providing new capabilities to the user in real-life cyber warfare.[14]

Electronic warfare

Military operations and virtually all weapon systems depend on the electromagnetic spectrum for a wide variety of functions. These include the use of radio frequencies for communication, the use of microwaves for data transmission, radar and satellite communications, and the use of infra-red light for intelligence gathering and targeting.

During warfare, combatants aim to secure unimpeded access to the electromagnetic spectrum for friendly forces and deny enemies access to the spectrum. Electronic warfare is action taken using electromagnetic energy to control the electromagnetic spectrum, attack an enemy, and block enemy attacks. It has a number of elements:

- Electronic intelligence – gathering details of signals of interest and identifying their electronic signatures.
- Electronic support – sensing of the electromagnetic environment and developing situational awareness of potentially hostile emitters, for example that may be able to track a target or launch a missile.
- Electronic protection – preventing a receiver or sensor from being jammed or deceived, or using electromagnetic radiation to detect and defend against an incoming missile.
- Electronic attack – using electromagnetic energy to disrupt, deny, degrade, destroy, or deceive the enemy's signals.[15]

Advocates argue that AI has roles to play in all elements of electronic warfare, with its ability to deal with new and unexpected threats and the rapid speed at which it functions providing an advantage over conventional techniques. It could be used to detect weak but potentially important radio frequency signals from electromagnetic noise, or used to rapidly identify and characterise new unknown signals and automatically generate countermeasures 'on the fly'. AI methods could also be used to fuse data on electronic threats from a variety of sources to give a single overall picture of the situation in a visible format understandable to humans.

Within the next two years the US Air Force is aiming to apply cognitive AI and machine learning algorithms electronic warfare systems on board the F-15 combat aircraft. The Air Force wants to enable the aircraft's systems to respond to emerging threats quickly and provide rapid reprogramming and learning capabilities for the systems.[16]

14 Kelley M. Sayler: 'Artificial Intelligence and National Security'. Congressional Research Service. Op cit. P10.

15 John R. Hoehn: 'Defense Primer: Electronic Warfare'. Congressional Research Service, 29 October 2020. https://fas.org/sgp/crs/natsec/IF11118.pdf

16 John Keller: 'Air Force asks industry for artificial intelligence (AI) cognitive electronic warfare (EW) for F-15 jets'. Military and Aerospace Electronics, 15 March 2021. https://www.militaryaerospace.com/computers/article/14199230/electronic-warfare-ew-cognitive-artificial-intelligence-ai

Artists concept of the deployment of a swarm of Gremlin drones. Credit: DARPA

DARPA is also working on programmes to develop autonomous electronic warfare technologies. DARPA's Adaptive Radar Countermeasures (ARC) and Behavioural Learning for Adaptive Electronic Warfare (BLADE) programmes both aim to apply AI to characterise an electronic warfare threat 'on the fly' and devise a countermeasure to it. ARC software is a radar-jamming programme intended to isolate unknown radar signals in the presence of other signals, deduce the threat posed by these signals, and then synthesize and transmit countermeasure signals, all in real time. BLADE works in a similar way but targets wireless communications rather than radar and is intended to stop the flow of information by radio. Both systems are intended to be available by the late 2020s, and could be retrofitted to existing electronic warfare systems as well as incorporated into new systems. A third DARPA programme, the Spectrum Collaboration Challenge, aims to create a new kind of autonomous radio system which can collaborate with other radio systems to improve spectrum management. AI software automatically determines which radio frequencies are underused and allocates signals to them, rather than adding to bandwidth demands on already congested frequencies.[17]

Command and control and decision support

Command and control systems support operational commanders in directing tasks and monitoring the forces assigned to a mission, and help present information to a commander in an easily understood format to assist decision making. AI systems have long been used for such purposes by the US military. In the 1991 Gulf War the United States Transportation Command used the Dynamic Analysis and Replanning Tool (DART), developed by BBN Systems and Technologies and the ISX Corporation under a DARPA programme, to plan and solve problems during the transport of military supplies from bases in Europe to the Middle East.[18]

17 George I. Seffers: 'Smarter AI for Electronic Warfare'. Signal Magazine, 1 November 2017. https://www.afcea.org/content/smarter-ai-electronic-warfare

18 Sara Reese Hedberg: 'DART: Revolutionizing Logistics Planning'. IEEE Intelligent Systems, May / June 2002. https://www.gwern.net/docs/ai/2002-hedberg.pdf

In the near future it may be possible to use AI to pull together data from a range of sensors into a common operating picture for commanders. The common operating picture is a map which shows objects of interest on a battlefield, such as the positions of friendly and enemy forces, important infrastructure, and other relevant operational information. AI systems could allow information to be added to the map automatically in real time and possibly help resolve variances or gaps in input data which might give contradictory indications.

The military AI sector is also working to develop systems able to provide decision support to commanders. AI-based decision support systems are already used in applications such as medical diagnosis to analyse information and propose potential courses of action to human operators. With a military focus, such systems would be intended to help commanders respond to unfolding events on the battlefield, based on real time analysis of data and a knowledge of enemy combat doctrine. The system might present a menu of possible actions to the commander, with an indication and explanation of the likely consequences of each action. Although human judgement will be essential in command and control decision-making for some time yet, the speed and capacity of AI-based tools can help in eliminating peripheral tasks, allowing the commander to concentrate fully on tasks where humans outperform machines.[19]

The UK's armed forces are keen to explore the use of such systems to support their operations. Cervus Defence and Security has developed a data capture system for MoD which is able to analyse training and performance data such as shot accuracy and lethality to give a detailed overview of individual and collective performance. The system was trialled with the Parachute Regiment as part of a training exercise in Kenya in 2020.[20] During the summer of 2021 the British Army claimed to have used AI for the first time operationally during NATO's Exercise Spring Storm in Estonia. Soldiers taking part in the exercise used an AI engine which analysed data to provide information on the surrounding environment and terrain, assisting in planning military activities and managing command and control processes.[21]

Autonomous vehicles and military robotics

Remotely controlled aircraft and vehicles have been used effectively in combat over the last two decades, and systems such as the Lockheed Martin Desert Hawk drone and the Foster-Wheeler Talon tracked military robot were deployed with British ground troops in Afghanistan in 2010.[22] AI is now being increasingly used to allow such systems to operate autonomously. AI applications in this field are similar to those being developed for civil sector driverless vehicles, which use sensors and AI software to perceive the environment, recognize obstacles, fuse sensor data, navigate, and communicate with other vehicles. The technology is now mature enough that military autonomous vehicles are on the point of deployment by major military powers. MoD has funded Horiba Mira to develop military autonomous vehicles and navigation systems to deliver supplies

19 Raúl Valencia: 'Artificial Intelligence in Command and Control Systems'. GMV Blog, 16 June 2020. https://www.gmv.com/blog_gmv/language/en/artificial-intelligence-in-command-and-control-systems/

20 Defence and Security Accelerator (DASA) Annual review 2019/2020. https://spark.adobe.com/page/4Il1sNOtG21SH/

21 'Artificial Intelligence used on Army operation for the first time.' Ministry of Defence, 5 July 2021. https://www.gov.uk/government/news/artificial-intelligence-used-on-army-operation-for-the-first-time

22 Airforce Technology: 'Desert Hawk III Minature Unmanned Aerial Vehicle (MUAV)'. https://www.airforce-technology.com/projects/deserthawkuav/

Army Technology: 'TALON Tracked Military Robot'. https://www.army-technology.com/projects/talon-tracked-military-robot/

US Air Force dog Hammer encounters a Ghost Robotics Vision 60 semi-autonomous legged robot at Scott Air Force Base, December 2020. Credit: DoD

to front-line forces during combat. The company was awarded a contract to supply three of its Viking autonomous unmanned ground vehicles for trials with the British Army commencing in 2020.[23]

Drones able to fly autonomously are also under development. The US Air Force Research Laboratory has developed and flight tested the Skyborg Autonomy Core System, a software system able to demonstrate basic aviation capabilities and respond to navigational commands, which is intended to be capable of flying a loyal wingman drone. The US Air Force reportedly plans to have Skyborg powered aircraft available for operations in 2023.[24]

In the marine environment the Royal Navy has conducted advanced trials with an autonomous minesweeping craft and is also trialling Madfox (Maritime Demonstrator For Operational eXperimentation), an uncrewed surface vessel able to undertake surveillance and force protection work in a similar manner to the loyal wingman concept.[25] The Navy is also experimenting with an autonomous submarine, the Extra Large Uncrewed Underwater Vehicle (XLUUV) developed by MSubs Ltd, which will be available to trial on-board systems, sensors and payloads to help develop understanding of operating uncrewed underwater vehicle systems.[26]

23 Defence and Security Accelerator (DASA) Annual review 2019/2020, op cit.

24 Kyle Mizokami: 'The Air Force's AI Brain Just Flew for the First Time'. Popular Mechanics, 13 May 2021. https://www.popularmechanics.com/military/aviation/a36412460/air-force-ai-brain-first-flight-skyborg-details/

25 'Royal Navy gets first unmanned minesweeping system'. Ministry of Defence, Defence Equipment and Support, 5 May 2018. https://www.gov.uk/government/news/royal-navy-gets-first-unmanned-minesweeping-system

'New autonomous vessel delivered to Royal Navy'. Royal Navy, 26 March 2021. https://www.royalnavy.mod.uk/news-and-latest-activity/news/2021/march/26/210326-madfox-vessel

26 'Innovations to be tested on pioneering autonomous submarine'. Defence and Security Accelerator and Defence Science and Technology Laboratory, 16 February 2021. https://www.gov.uk/government/news/innovations-to-be-tested-on-pioneering-autonomous-submarine

Swarms

AI software is also under development to enable autonomous systems to act as an interconnected intelligent swarm. Inspired by swarms of insects, swarming machines are able to work in co-operation to overwhelm adversaries. Swarms operate autonomously, without central control, and individual component units are able to sense their local environment and other members of the swarm and co-operate with other members to perform a task. MoD has funded a consortium led by Blue Bear Systems Research, including Plextek DTS, IQHQ, Airbus and Durham University, to develop swarm technology. Blue Bear has developed a command and control system capable of managing a swarm and simultaneously handling different tasks. In trials the system was able to control a swarm of 20 drones, consisting of 5 different types and sizes of fixed wing drones with different operational capabilities.[27]

Lethal Autonomous Weapon Systems (LAWS)

LAWS are AI-powered weapon systems capable of identifying, engaging, and destroying a target with no human interaction. They are based on a combination of a sensor system which monitors the surrounding environment, an AI system which can identify an object as a potential target and decide on whether to engage it, and a weapon which can destroy the target. The underpinning technologies which, when combined, act as the 'building blocks' for a lethal autonomous weapon system are developing at a rapid pace. Crude systems may already have been used in combat, and advanced military nations will soon be able to mass produce them and deploy them on the battlefield.[28] In the view of advocates, autonomous weapon systems would have vastly improved reaction times when compared with conventional weapon systems, giving them decisive advantages in combat.

Weapon systems already exist in which an algorithm makes the decision to shoot, rather than a human. Those systems that can acquire and engage targets autonomously are mostly defensive systems, such as air defence systems or South Korea's SGR-A1 sentry robot.[29] These are intended to be operated under human supervision and to fire autonomously in situations where the time of engagement is deemed too short for humans to be able to respond. However, offensive weapons which are capable of acquiring and engaging targets autonomously are also emerging. Fire-and-forget munitions, such as the MBDA Brimstone missile used by the RAF and the IAI Harop loitering munition are able to select their own targets and show many of the features of a lethal autonomous weapon.

27 'Swarming drones concept flies closer to reality'. Defence Science and Technology Laboratory and Defence and Security Accelerator. 28 January 2021. **https://www.gov.uk/government/news/swarming-drones-concept-flies-closer-to-reality**

28 Zachary Kallenborn: 'Was a flying killer robot used in Libya? Quite possibly'. Bulletin of the Atomic Scientists, 20 May 2021. **https://thebulletin.org/2021/05/was-a-flying-killer-robot-used-in-libya-quite-possibly/**

'Autonomous weapons are already here. How do we control how they are used?'. World Economic Forum, 14 November 2017. **https://www.weforum.org/agenda/2017/11/autonomous-weapons-are-already-here-but-humans-are-still-needed-on-the-battlefield/**

'Off The Leash: The Development of Autonomous Military Drones in the UK'. Drone Wars UK, 10 November 2018. **https://dronewars.net/wp-content/uploads/2018/11/dw-leash-web.pdf**

29 Ingvild Bode and Tom Watts: 'Meaning-Less Human Control: Lessons from Air Defence Systems on Meaningful Human Control for the Debate on AWS', op cit.

Alexander Velez-Green: 'The Foreign Policy Essay: The South Korean Sentry – A "Killer Robot" to Prevent War'. Lawfare, 1 March 2015. **https://www.lawfareblog.com/foreign-policy-essay-south-korean-sentry%E2%80%94-killer-robot-prevent-war**

Information Operations

Social media platforms are increasing being used as a source of news and information, creating opportunities for those with malicious intent to spread false and misleading information with the aim of generating divisions and conflict, manipulating democratic processes, and targeting individuals to radicalise them or encourage them to disobey instructions. These operations can be conducted in tandem with 'on the ground' military operations to help achieve military objectives.

AI-enabled internet bots – software applications which run simple and repetitive automated tasks – can facilitate such campaigns by creating fake online identities and spreading information much faster than a person could. AI apps could also be used for mining data from social media to create a digital 'pattern of life' for individuals – including government officials, politicians, and members of the armed forces – for coercive purposes. AI generated imagery is able to create synthetic media – increasingly realistic photo and video footage, or 'deep fakes' – that could eventually be used to support online offensives intended to mislead and manipulate.

However, AI techniques might also be used to counter such operations. The US Air Force and Special Operations Command are working to develop a software tool able to automatically identify and assess suspected disinformation by assessing the credibility of an information source and comparing its content with other, credible sources.[30] A team from the University of Washington has created a system that can distinguish faked satellite images on the basis of their colour, edge clarity, and texture characteristics,[31] and DARPA has launched a Media Forensics project which seeks to "automatically detect manipulations, provide detailed information about how these manipulations were performed, and reason about the overall integrity of visual media."[32]

In the UK the MoD's Defence and Security Accelerator has commissioned RAND Europe to develop a method for detecting the malign use of information online as part of efforts to help MoD develop its behavioural analytics capability. The research team developed and applied a machine learning model to a known Russian 'troll' database to identify differences between authentic political supporters and trolls involved in online debates relating to the 2016 US presidential election. The model was 87 per cent effective at distinguishing trolls from authentic supporters.[33]

30 Patrick Tucker: 'Can AI Detect Disinformation? A New Special Operations Program May Find Out'. Defense One, 2 October 2020. https://www.defenseone.com/technology/2020/10/can-ai-detect-disinformation-new-special-operations-program-may-find-out/168972/

31 Patrick Tucker: 'A Better Way to Spot Deep-Faked Satellite Images'. Defense One, 23 April 2021. https://www.defenseone.com/technology/2021/04/better-way-spot-deep-faked-satellite-images/173586/

32 Matt Turek: 'Media Forensics (MediFor)'. DARPA. https://www.darpa.mil/program/media-forensics

33 'Using machine learning to detect malign information efforts online'. Rand Corporation, 23 June 2020. https://www.rand.org/randeurope/research/projects/using-machine-learning-to-detect-malign-information-efforts.html

Development of AI in the UK military

The UK government clearly attaches great importance to the development of Artificial Intelligence.

In September 2021 the UK government published its National AI Strategy, which sets out a ten year plan "to make Britain a global AI Superpower". The strategy presents proposals for investment in the AI sector, placing AI at the mainstream of the UK's economy and introducing it across all economic sectors and regions of the UK, and for governing the use of AI. With an emphasis on supporting the AI industry, the strategy leaves open questions as to how the government intends to manage the risks and harms presented by AI and how AI will be regulated. The UK's approach to AI regulation will be set out in a white paper to be published early in 2022.[34]

The 2021 Integrated Review sets out the government's aspiration to "be recognised as a Science and Tech Superpower, remaining at least third in the world in relevant performance measures for scientific research and innovation, and having established a leading edge in critical areas such as artificial intelligence".[35] The Defence Command Paper which accompanied the Integrated Review states that the MOD intends to invest £6.6billion over the next four years in defence research and development, focusing on emerging technologies in artificial intelligence, AI-enabled autonomous systems, cyber, space and directed energy systems.[36] A number of teams are taking forward the development of military AI systems within the MoD, including units in each of the armed forces, the Defence Science and Technology Laboratory (DSTL), and Defence Equipment and Support (DES).

Defence and security is also a core research area for the Alan Turing Institute, the United Kingdom's national institute for data science and artificial intelligence.[37] Strategic Command (responsible for defence intelligence and computing systems), DSTL, and GCHQ are listed as partners for the Institute's defence and security programme, which is funded by MoD and GCHQ. In June 2019 the Institute hosted a conference entitled 'The future of test & evaluation in defence: AI and autonomy', funded by the Ministry of Defence, and in October of the same year funded a workshop on 'Responsible Human-Machine Teaming' in

34 'National AI Strategy'. Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport, and Department for Business, Energy & Industrial Strategy. 22 September 2021. https://www.gov.uk/government/publications/national-ai-strategy

35 'Global Britain in a Competitive Age: the Integrated Review of Security, Defence, Development and Foreign Policy'. Cabinet Office, 16 March 2021. P7. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975077/Global_Britain_in_a_Competitive_Age-_the_Integrated_Review_of_Security__Defence__Development_and_Foreign_Policy.pdf

36 Ministry of Defence: 'Defence in a Competitive Age'. Command Paper CP 411, MArch 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974661/CP411_-Defence_Command_Plan.pdf

37 'Defence and security'. Alan Turing Institute. https://www.turing.ac.uk/research/research-programmes/defence-and-security

collaboration with DSTL.[38] Given that the government's planned new Advanced Research and Invention Agency (ARIA) is to be modelled on the US's Defence Advanced Projects Research Agency (DARPA), it can be anticipated that the new agency will also have a strong presence in military-oriented technology.[39]

The Ministry of Defence has established a Defence Artificial Intelligence and Autonomy Unit to address policy issues, and planned to publish an Artificial Intelligence and Autonomy Strategy in summer 2021 to enable rapid adoption of the technologies.[40] The direction of the strategy was outlined by General Sir Patrick Sanders, Commander of Strategic Command, in a speech at the Royal United Services Institute in May 2021:

> "It will begin by integrating existing digital technologies now – for example using machine learning and automation to support Intelligence analysis. It will be enabled by improving our digital infrastructure – the digital backbone – with a data strategy that enables data curation, data sharing and data exploitation, cloud services at Secret and Above Secret, and a common network architecture. It will lead to investing in more S&T in partnership with DSTL and to experimentation to ensure responsible development of AI enabled and autonomous systems."[41]

The MoD's equipment procurement organisation, Defence Equipment and Support (DES), has set up a special unit to develop autonomous technology. The Expeditionary Robotics Centre of Expertise has been launched by DES in collaboration with the British Army to bring together robotics and autonomous systems experts from across government, academia and industry to assess "unexplored, high-risk but rapidly maturing technologies".[42]

Each of the armed forces have their own experimental units for the development of new military technologies, such as AI. The Royal Navy has set up its own software development house, NELSON, and has established NavyX, an 'Autonomy and Lethality Accelerator' which has worked on the development of autonomous surface craft and drones.[43] The Navy is trialling AI software systems to help defend warships against missile attacks[44], and aspires to a fully uncrewed major warship within the next decade.[45]

The Army is introducing a Warfighting Experimentation Force, based on an infantry battalion with elements drawn from across the Army, which will trial

---

38 'The future of test & evaluation in defence: AI and autonomy'. Alan Turing Institute. **https://www.turing.ac.uk/events/future-test-evaluation-defence-ai-and-autonomy**

'Responsible Human-Machine Teaming'. Alan Turing Institute. **https://www.turing.ac.uk/events/responsible-human-machine-teaming**

39 'Advanced Research and Invention Agency (ARIA): policy statement'. Department for Business, Energy, and Industrial Strategy, 19 March 2021. **https://www.gov.uk/government/publications/advanced-research-and-invention-agency-aria-statement-of-policy-intent/advanced-research-and-invention-agency-aria-policy-statement**

40 'Science and Technology Strategy 2020'. Ministry of Defence. 19 October 2020. P15. **https://www.gov.uk/government/publications/mod-science-and-technology-strategy-2020**

41 'Commander of Strategic Command RUSI conference speech'. Ministry of Defence, Strategic Command, and General Sir Patrick Sanders KCB CBE DSO ADC Gen, 26 May 2021. **https://www.gov.uk/government/speeches/commander-of-strategic-command-rusi-conference-speech**

42 'DE&S-led expertise set to revolutionise development of UK military robotics'. Ministry of Defence and Defence Equipment and Support'21 May 2021. **https://www.gov.uk/government/news/des-led-expertise-set-to-revolutionise-development-of-uk-military-robotics**

43 Richard Brantingham: 'The NELSON Standards – A design system for the Royal Navy'. Ministry of Defence, 18 December 2019. **https://defencedigital.blog.gov.uk/2019/12/18/nelson-standards-creating-royal-navy-apps-with-a-consistent-look-and-feel/**

'NavyX'. Royal Navy. **https://www.royalnavy.mod.uk/navyx**

44 'Navy tests artificial intelligence against supersonic missiles'. Royal Navy, 29 May 2021. **https://www.royalnavy.mod.uk/news-and-latest-activity/news/2021/may/29/20210529-artificial-intelligence**

45 @NavyLookout Tweet, 19 May 2021. **https://twitter.com/NavyLookout/status/1395001077490241543**

new technology and its integration.[46] For the past ten years the Army has conducted regular Army Warfighting Experiment events on Salisbury Plain to test new technologies developed by DSTL and industry. These include "semi-autonomous uncrewed systems" and systems to reduce cognitive load for system operators "whilst not being fully autonomous".[47] The Royal Air Force has a Rapid Capabilities Office, which since 2017 has been exploring options for a Lightweight Affordable Novel Combat Aircraft (LANCA), a low-cost uncrewed combat aircraft designed to fly at high-speed alongside crewed aircraft as a 'loyal wingman' to provide support and protection and perform independent missions such as reconnaissance, electronic warfare, or combat missions. Spirit Aerosystems has been awarded a £30 million contract to lead 'Project Mosquito', intended to deliver a flight test programme by the end of 2023 in the hope that the drone will be in service by the end of the decade.[48] Strategic Command is also creating an artificial intelligence centre of excellence, as part of a wider 'digital foundry' initiative that extends across the Ministry of Defence.

The Defence Science and Technology Laboratory (DSTL), the MoD's research arm, set up the AI Lab, intended to be a flagship for AI, machine learning and data science, in May 2018 at its Porton Down headquarters. The AI Lab aims to enhance and accelerate the application of AI-related technologies to defence and security challenges.[49] DSTL also supports the Defence and Security Accelerator (DASA), a funding stream for innovation in defence and security. Among the AI initiatives funded by DASA is TheiaView, a flexible object detection and classification toolset for tracking prioritised threats in real-time video which can detect objects and inform robots to take action based on machine learning.[50] Other recent AI projects funded through DASA include ARGA, a software tool which uses machine learning and natural language processing to support authors as they write text, and a machine learning tool to analyse the condition of military assets and equipment and predict potential component failure.[51]

What is missing in all this, so far, is a consideration of the ethical issues involved in the use of AI. The Integrated Review and other government statements leave no doubt that the government attaches immense importance to the military applications of AI and intends to race ahead with its development. However, although publications outlining doctrine on the use of automated systems have

46 'Warfighting Experimentation Force unveiled'. Ministry of Defence, 23 March 2021. https://www.gov.uk/government/news/warfighting-experimentation-force-unveiled

47 'Science enables soldiers and aircrew to partner with smart machines on Salisbury Plain'. Defence Science and Technology Laboratory and Defence and Security Accelerator, 23 September 2020. https://www.gov.uk/government/news/science-enables-soldiers-and-aircrew-to-partner-with-smart-machines-on-salisbury-plain

48 '£30-million injection for UK's first uncrewed fighter aircraft'. Ministry of Defence, 25 January 2021. https://www.gov.uk/government/news/30m-injection-for-uks-first-uncrewed-fighter-aircraft

49 'Flagship AI Lab announced as Defence Secretary hosts first meet between British and American defence innovators'. Ministry of Defence, 22 May 2018. https://www.gov.uk/government/news/flagship-ai-lab-announced-as-defence-secretary-hosts-first-meet-between-british-and-american-defence-innovators

50 'Innovation for a Safer Future. DASA Strategy 2021-2024'. Defence and Security Accelerator, 4 May 2021. P26. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/982893/DASA_-_Innovation_for_a_Safer_Future_Strategy_2021-2024_No_Annex_Hi-res.pdf

51 'Intelligent support to human authors for faster and better exploitation of unstructured text'. Defence and Security Accelerator, 17 January 2019. https://www.gov.uk/government/news/intelligent-support-to-human-authors-for-faster-and-better-exploitation-of-unstructured-text

'Pre-emptive strikes on military equipment faults'. Defence and Security Accelerator, 17 January 2019. https://www.gov.uk/government/news/pre-emptive-strikes-on-military-equipment-faults

been published,[52] to date the MoD has remained silent on the ethical framework governing the use of its AI and autonomous systems, despite already having taken significant decisions on the future use of military AI.

The MoD plans to publish its Defence AI Strategy towards the end of 2021. The strategy is expected to set out a set of high level ethical principles and has been prepared following discussion with selected experts from academia and industry, although no open consultation on ethical and other issues associated with military uses of AI has yet taken place. One of the principal purposes of the strategy will be to reassure industry and the public that MoD is a responsible partner for collaboration on AI projects.

Despite this, major questions remain unanswered. Under what circumstances will British armed forces employ AI technology? What degree of human control does the government think appropriate? How will risks be addressed? And how will the UK demonstrate to allies and adversaries that the UK intends to take a principled approach to the use of military AI technology?

52 'Joint Concept Note 2/17. Future of Command and Control'. Ministry of Defence, 8 September 2017. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/643245/concepts_uk_future_c2_jcn_2_17.pdf

'Joint Concept Note 1/18. Human-Machine Teaming'. Ministry of Defence, 18 May 2018. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/709359/20180517-concepts_uk_human_machine_teaming_jcn_1_18.pdf

Risks posed by military AI systems

The goal of using AI and autonomous systems for military purposes is controversial. Even military sources acknowledge that there are significant risks involved, and that it would be unwise to rush into the indiscriminate development and deployment of these capabilities.

Clearly there are different elements of risk associated with each of the different military applications of AI which are described in the previous section. An algorithm sorting through data as part of a back-office operation at the Ministry of Defence headquarters would raise a different level of issues and concerns and require a different level of scrutiny than an autonomous weapon system.

Looking forward over a foreseeable period of the next ten years or so, threats posed by the more extreme but remote prospect of the development of a super-intelligent AI system posing an existential risk to humanity can be discounted. Nevertheless, AI systems currently in development undoubtedly pose threats to lives, human rights and well-being.

A particular concern is that governments will rush into developing military uses for AI without paying sufficient attention to the seriousness of these risks. In competing to gain the military benefits of AI soonest, states might not put proper precautions in place, giving rise to a 'race to the bottom'. As a result, military applications of AI could reduce, rather than increase, security.

The risks posed by military AI systems can be grouped into three categories: ethical and legal, operational, and strategic.[53]

Ethical and legal risks

Compliance with the laws of war

Military AI and robotic systems must be capable of being used in compliance with international humanitarian law and international human rights law, which govern the conduct of war. On the practical level, it is far from clear that robotic systems, and in particular autonomous weapons, would be able to meet the standards set by the laws of war on making lethal decisions and protecting non-combatants.

During warfare, commanders who plan and decide upon an attack have legal obligations to distinguish between combatants and non-combatants; ensure that damage caused is proportional to the military objectives, and take precautions to protect civilians. Civilian loss of life through 'collateral damage' is only lawful if it is proportionate to the military objective, and commanders can be held to account for any violations.

53 Forrest E. Morgan, Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, Derek Grossman: 'Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World', op cit. P30.

Buildings bombed during attacks on Raqqa, Syria, in 2018. The use of AI and autonomous systems to conduct warfare could make it impossible to attribute responsibility for civilian casualties and war crimes. Credit: Amnesty International

However, the judgements needed to distinguish combatants from non-combatants, assess the proportionality of an attack, and take precautions, are largely qualitative decisions based on experience, common sense and an understanding of context that robotic systems do not have. The engineering and programming challenges of designing computer systems able to operate in compliance with the laws of war in complex, cluttered, and dynamic environments remain immense. For these reasons, it seems that human control over the selection and attack of targets will continue to be necessary to guarantee compliance with the laws of armed conflict.

Autonomous weapon systems might also fail the test of compliance with the "principles of humanity and the requirements of the public conscience" – the 'Martens Clause' included in the 1899 Hague Conventions. As machines are unable to show compassion or empathy, it has been argued that they would lack inherently human restraints on killing and be unable to rise to the challenge of considering the moral aspects of killing. Should machines be permitted to decide whether human beings live or die? Could the principles of humanity and the dictates of public conscience ever permit human life to become valued merely in terms of collateral damage in the calculations of a machine which is not accountable for its actions?

Accountability

Concerns have been raised that the use of AI systems may lead to an 'accountability gap' should violations of humanitarian law occur. It is nor clear who would be held responsible if things went wrong – yet it makes no sense to punish a computer if it operates unpredictably and as a result war crimes are committed.

If an autonomous weapons system is able to select and attack targets independently, it might be difficult to hold the individuals who deployed it or programmed it liable for breaching the laws of armed conflict should the weapon act unpredictably. The complexity of AI systems may make it hard to decide whether a crime was the responsibility of the operator, programmer, or manufacturer, or was just an unpredictable accident with no one ultimately at fault.

Human rights and privacy

AI systems also pose potential threats to human rights and individual privacy. Features of AI which enable data mining, persistent surveillance, analysis of physical traits such as face or gait recognition, and intrusion into online activity provide governments and big business with wide scope for monitoring and controlling populations and abusing human rights.

Information operations that spread false information and seek to manipulate public opinion or create division may also cause broader harms such as the downplaying of objective facts and information in decision-making and the undermining of trust that is essential for democracy to function. The development of AI technology for explicitly military purposes will exacerbate the risks of these undesirable impacts.

Inappropriate use

Introduction of a new and poorly understood system into combat may result in it being used in a way which results in poor decisions or inappropriate engagements, or even in it being misused with malicious intent. Technologies developed for one set of tasks may then be adopted in other fields for missions not envisaged by the designers or proponents, regardless of training, procedures and safeguards to ensure compliance with international humanitarian law and control the conditions under which they can be safely operated. Armed forces under pressure in battle environments may be tempted to modify technologies to overcome safety features and controls.

Operational risks

Operational risks are risks associated with the failure of AI in military applications in unintended or unanticipated ways. Some of these risks are associated with different forms of bias.[54]

Technical sources of bias

Autonomous systems must be 'trained' using a set of data against which it will compare data from the real world in which it operates. If the training data set is inappropriate – for example, is incomplete, of low-quality, or irrelevant or inaccurate – then biases will be introduced which will affect the operation of the system when it is deployed. AI systems are only as good as their training data and a small amount of corrupted training data can have large impacts on the performance of the system. Even under the best of conditions poorly trained AI systems may produce outputs that discriminate against particular groups.[55]

Another source of bias is known as processing bias, and occurs when an algorithm is itself biased in the way that it transforms data. This may occur, for example, when algorithms employ statistical models to estimate or predict outputs in situations where the model does not give an accurate representation of real-world data.

Biases can arise from inappropriate use or deployment of a system. Algorithmic systems are developed and trained for particular purposes in particular contexts, and inevitably have a narrow window of operation. Taking the system outside this

54 'Algorithmic Bias and the Weaponization of Increasingly Autonomous Technologies'. United Nations Institute for Disarmament Research, 22 August 2018. https://unidir.org/files/publications/pdfs/algorithmic-bias-and-the-weaponization-of-increasingly-autonomous-technologies-en-720.pdf

55 Arthur Holland Michel: 'Known Unknowns'. United Nations Institute for Disarmament Research, 17 May 2021. https://unidir.org/known-unknowns

context and beyond its tested capabilities may result in unexpected and possibly dangerous behaviour. In a combat environment, which is inevitably complex and rapidly-changing, autonomous systems will face new situations which hinder the collection of data and do not match the circumstances against which the system was developed and tested. As a result, they will be liable to failures which are impossible to anticipate.

Human sources of bias

Bias may result when humans misuse a system or misunderstand its output. These biases relate to the degree of trust the operator has in the reliability of the system. Bias can occur when the user fails to interpret the output of an AI system correctly, and thus acts erroneously. If the user does not understand exactly the meaning of the output, or is misdirected by unreliable features of the output, then mistakes will be made.

One such challenge is overconfidence in the output of an autonomous system. During the 2003 Iraq War highly automated Patriot air defence systems shot down a Royal Air Force Tornado aircraft, killing the crew, when the system incorrectly identified the aircraft as hostile and the human operator accepted the incorrect identification and authorised the engagement. This kind of 'automation bias' occurs if operators place too much trust in the capabilities of an automated system.

On the other side of the coin, accidents can also arise in the opposite situation, where operators under-trust a system and place insufficient reliance on an automated process. This is most likely to happen with a system which is known to generate 'false positive' errors. In such situations operators may ignore important information provided by the system or override its actions without justification. A tragic example of this is the shooting down of an Airbus A300 passenger aircraft, Iran Air Flight 655, and the loss of 290 lives on 3 July 1988 by an automated Aegis air defence system on the warship USS Vincennes. Contrary to information provided by the system, the ship's crew decided that radar signals were from an attacking combat aircraft and decided to shoot it down.[56]

Other elements of human operation can cause bias. Humans have a tendency to simplify situations they encounter, and are poor at accurately assessing risks, particularly in new situations. In some situations confirmation bias (a tendency to look for evidence to support a preference and discount contrary evidence) and projection bias (overestimating the degree to which other people think and behave like we do) can influence decision-making.

The risks of such accidents intensifies in systems which are so complex that their outputs are unexplainable, and in which operators cannot easily determine why their systems have made particular decisions or are behaving in particular ways. At the same time overconfidence – believing that we have a better understanding of how a system works than we actually do – can also lead to mistakes in understanding the reasons for an AI system's behaviour and in predicting how it may behave in future.

56 Vincent Boulanin and Maaike Verbruggen: 'Mapping the development of autonomy in weapon systems'. Stockholm International Peace Research Institute, November 2017. P40. **https://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_1.pdf**

Elisa B. Kania: 'The critical human element in the machine age of warfare'. Bulletin of the Atomic Scientists, 15 November 2017. **https://thebulletin.org/critical-human-element-machine-age-warfare11277#**

Manipulation with malicious intent

Military AI systems, like all networked systems, are vulnerable to attacks from malicious actors who may attempt to jam, hack, or spoof the system. Systems are not only vulnerable to hackers who may attempt to gain direct access to manipulate or take control of the system, but are also vulnerable to other types of attack. These include data-poisoning attacks, in which the training data is manipulated in order to upset the functioning of the system. It may also be possible to spoof a system by tricking it into making errors by presenting it with falsified data in the field. In some cases AI systems have been fooled by changing just one pixel in an image[57], creating vulnerabilities that could be exploited by an opponent using counter-AI capabilities. Autonomous systems can also be jammed – it is possible to prevent them from communicating with other parts of the network, causing operational functions and impeding their function.

Inherent system features

Machine learning systems may be 'black-box' systems, for which it is not possible for humans to observe or understand how decisions are made. It is not possible to reprogramme such a system to change its mode of operation. If a machine learning system is permitted to continue to learn in an unsupervised manner it can potentially evolve, meaning that the system in operation will not be identical to a system that was tested and may make different decisions.

Field data issues may also influence the operation of an AI system. Conflicting data may complicate a decision. Data may not conflict, but may appear to suggest what is actually not the case, and the system may fail to differ an anomaly from a threat.

Strategic risks

Thresholds

Although advocates of military applications of AI argue that they may lower the risk of harm to military personnel, this introduces a new risk that leaders will resort to using autonomous military systems in conflict rather than pursuing non-military options to resolve differences, lowering the threshold for the use of force. Although fewer soldiers may be in the battle zone, civilian resident populations will still be present and will continue to face the risks of warfare, transferring the costs of war from soldiers to civilians.

Escalation management

The use of non-human systems in warfare increases the potential for escalation through a number of mechanisms. Firstly, if war becomes more likely as a result of the use of AI and autonomous technology, the increased frequency of military action itself increases the risks of escalation. Secondly, if commanders believe that the use of autonomous technologies means that their own soldiers will face less harm in conflict, they may be tempted to take greater risks and act more aggressively, fuelling escalation. Finally, the speed at which military action can be executed – at machine speed, rather than within human time frames – creates the risk that an unintended 'flash war' may suddenly develop. The space for deliberation and negotiation will decrease, leading to the possibility of rapid accidental escalation, particularly in crisis situations where the risk of a misjudgement or false move might have severe consequences.

57 BBC News: 'AI image recognition fooled by a single pixel change'. 3 November 2017. http://www.bbc.co.uk/news/technology-41845878

Combinations of emerging military technologies could increase the risks of escalation even further. Hypersonic technology which allows missiles to travel at many times the speed of sound, when combined with AI decision-making technology, would be particularly destabilising and dangerous and might reduce reaction times so much as to prevent the possibility of any human intervention. Cyberwarfare could see similar developments, where AI software systems could not only provide defence against a cyber attack but also launch an automated retaliatory attack.[58]

Arms racing and proliferation

The pursuit of military AI with the aim of gaining an advantage over potential adversaries already appears to be causing arms racing, with major and regional powers racing to develop their capabilities in order to stay ahead of their rivals. This competition has no absolute end goal, merely the relative goal of staying ahead of the other competitors, but it increases the risks of misunderstanding and escalation.

Many AI applications have dual uses, meaning that civilian technologies can be applied to military uses, and may be readily available from commercial or open sources. Different technologies may be combined for malicious uses. As capabilities improve and costs drop AI and dual use technologies will become more readily available to minor powers, non-state actors, and criminal groups.

Strategic Stability

Should advanced AI systems develop to the point that they are able to predict enemy tactics or the deployment of forces, this could have highly destabilising consequences, particularly in the domain of nuclear deterrence. If AI can predict the location of an adversary's forces, an aggressor might be tempted to launch a pre-emptive attack with the intention of destroying them without fear of retaliation. Even a perception that forces might be vulnerable in this way might encourage a state to undertake a 'use them or lose them' first strike to avoid the possibility that they might be unable to use their forces later in a conflict.

Conclusion: Accidents will happen

Accidents are inevitable in complex systems, and this includes even basic automated systems. As systems become more complex and 'tightly coupled' with other systems, creating 'systems of systems', the risks of unintended accidents increase. For complex AI and autonomous systems operating at machine speed it becomes increasingly challenging for human operators to predict problems and monitor and supervise and monitor systems. Problems could be magnified if systems are fielded before being adequately tested, and by specific factors relating to the military context, such as the secrecy of technology and systems, or the use of swarms, where an error by one unit is shared with others. Human supervision does not necessarily guarantee that accidents involving automated systems will be avoided, and human input can actually add to problems if personnel are not properly trained or disciplined, if information provided by the system is too complex for an operator to interpret rapidly, or if operators are facing stressful situations.

58 Kalev Leetaru: 'Will Hypersonic Weapons Finally Push Us Towards AI-Powered Missile Defense?'. Forbes, 23 June 2019. https://www.forbes.com/sites/kalevleetaru/2019/06/23/will-hypersonic-weapons-finally-push-us-towards-ai-powered-missile-defense/#520c8ce04871

Mitigating the impacts of AI

Weapon systems that operate with varying degrees of autonomy in their critical functions already exist, and new advances are constantly emerging. Despite the military advantages that AI-enabled systems are supposed to bring, as we have seen, there are significant risks associated with their use.

A report by Harvard University's Ash Center for Democratic Governance and Innovation on the ethical use of artificial intelligence in government operations warns that machines, at least for now, should not be allowed to make key decisions in relation to individual humans[59]. The report recommends that "AI should only be used for analysis and process improvement, not decision support, and human oversight should remain prevalent". Artificial intelligence systems "should not be tasked with making critical government decisions about citizens" – which would clearly include decisions by the military over their life or death.

Various mitigation measures have been proposed to control the use of AI in high-risk applications. The European Union has published a draft regulation on AI to guarantee safety and human rights whilst allowing the uptake of AI technology.[60] The draft regulation follows a risk-based approach. Systems with an unacceptable risk, which are considered a clear threat to the safety, livelihoods and rights of people will be banned. This would include lethal autonomous weapon systems. High-risk AI systems, including all remote biometric identification systems, will be subject to strict obligations before they can be put on the market and would be recorded on a database maintained by the European Commission. Systems with limited risks will have specific transparency obligations to ensure that users understand that they are interacting with a machine.

The United Nations Secretary General has called for LAWS to be prohibited under international law,[61] and the International Committee for the Red Cross has recommended that states adopt new legally binding rules to ban LAWS and ensure that weapons remain under human control at all times[62]. Such rules should apply to all systems that apply force based on processing sensor inputs and should outlaw systems which target people and systems which cannot be meaningfully controlled by a human, including opaque technologies that are too complex to be understood. Remaining sensor-based systems should be subject

59 Modhana Ranindranath: 'Agencies Should Watch Out for Unethical AI, Report Warns'. Nextgov.com, 23 August 2017. http://www.nextgov.com/cio-briefing/2017/08/agencies-should-watch-out-unethical-ai-report-warns/140461

60 European Commission: 'Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence'. 21 April 2021. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

61 'Autonomous weapons that kill must be banned, insists UN chief'. United Nations, 25 March 2019. https://news.un.org/en/story/2019/03/1035381

62 'ICRC position on autonomous weapon systems'. International Committee of the Red Cross, 12 May 2021. https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems

to positive obligations, for example on the location and duration of use and their target specifications, to protect the existing laws of war from erosion.[63]

Public debate over policy on new military technologies, including AI and autonomous systems, is important, and can help develop a robust set of ethical principles to govern their use and development. The US Department of Defense has shown considerable interest in the development of autonomous weaponry but, following a study by the Defence Innovation Board, has published a set of ethical principles to determine how the US military will use AI[64]. MoD and the UK government could – and should – do likewise, approaching these issues with a sense of responsibility and urgency so that technological advances are not allowed to outpace ethical deliberation. A set of ethical principles for the use of AI by the UK military would address important issues such as:

- How decision making processes can be made transparent and easy to explain and understand.
- How the human-machine teaming approach advocated by the MoD will work and guarantee safeguards.
- Who should be held responsible for decisions made and actions undertaken by AI-enabled systems.
- Which high-consequence uses, for example nuclear command and control, and decisions on the use of lethal force, should not be entrusted to machine-based systems.

63 'Regulating Autonomy in Weapons Systems'. Article 36, October 2020. https://article36.org/wp-content/uploads/2020/10/Regulating-autonomy-leaflet.pdf

64 'DOD Adopts Ethical Principles for Artificial Intelligence'. US Department of Defense, 24 February 2020. https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/

'AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense'. Defense Innovation Board, 31 October 2019. https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF

Conclusions: under human control?

This briefing sets out the various military applications which have been envisioned for AI, and also highlights their potential for causing harm to individuals. As we have seen, proposals for mitigating the risks posed by military AI systems are based on the principle of ensuring that AI systems remain at all times under human supervision.

But will it even be possible to retain human control over advanced AI systems? In conversation together, Nobel Laureates Kazuo Ishiguro and Venki Ramakrishnan appeared to think not. Ishiguro warned that "Human beings are just so far behind that there is no way that you can keep a human in the loop. It will be, indeed, like having some kind of retired nightwatchman trying to supervise a stadium full of rioting football fans", while Ramakrishnan pointed out that "AI could come up with the next big idea, like democracy, or communism, or Nazism, or money, or the joint stock company. Once it understands how to manipulate human emotions, we've got much bigger things to be concerned about".[65]

The potential for harm from AI extends beyond individuals. The widespread use of AI has deep social implications and the potential to fundamentally reshape the world of the future. Writing in 'Wired' magazine, Jaron Lanier and Glen Weyl observe that ""AI" is best understood as a political and social ideology rather than as a basket of algorithms. The core of the ideology is that a suite of technologies, designed by a small technical elite, can and should become autonomous from and eventually replace, rather than complement, not just individual humans but much of humanity." They point out that this ideology "has strong resonances with other historical ideologies, such as technocracy and central-planning-based forms of socialism, which viewed as desirable or inevitable the replacement of most human judgement/agency with systems created by a small technical elite."[66] Viewed in this context, the development of military AI can only exacerbate the very real tensions between militarism and democracy.

This should raise serious questions about how and why AI is being used, especially in its military applications. The answers are not reassuring. There are many technological, legal, and ethical barriers to overcome before AI can be widely deployed but the world's major powers, in their rush to take the lead in military AI, are racing ahead to introduce such technology before addressing these concerns. Reading the rhetoric in the Integrated Review and from military commentators, it is hard to avoid the conclusion that much of the motivation for the UK's interest in military AI is a response to developments in the US, Russia, and China: 'If they do it, then we don't want to be left behind'. Despite this, the security challenges that the UK faces are very different to those faced by these other nations, raising the question what is the problem we are trying to solve,

65 'Kazuo Ishiguro and Venki Ramakrishnan: imagining a new humanity'. Financial Times, 26 March 2021. https://www.ft.com/content/eca7988d-2961-4b27-9368-ff58c966e969

66 Jaron Lanier and Glen Weyl: 'AI is an Ideology, Not a Technology'. Wired, 15 March 2020. https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/

The world's major military powers are racing to introduce AI technology before addressing the serious concerns raised. Credit: DoD

and is there a better way to solve it? Drone Wars UK believes that the use of advanced technology is not the only – or even the best – way of addressing security concerns, and that the UK should move towards alternative approaches for maintaining national security based on placing the protection and well-being of people at the heart of security policy.[67]

As yet, there is little public appreciation of the changes and risks that society is facing as a result of advances in AI and robotics. This briefing is, in its own small way, intended as part of the wake-up call. AI can and should be used to improve conditions in the workplace and services to the public, and not to increase the lethality of war-fighting.

67 For more information please see **https://rethinkingsecurity.org.uk/**

Shining a spotlight
on military drones