Ministry of Defence – Written Evidence (AIW0035)

Preface

- Artificial Intelligence (AI) is a general-purpose enabling technology that has the potential to transform every aspect of Defence, from back-office corporate services to the delivery of military effect. As highlighted in the Integrated Review Refresh (March 2023), at a time when the geo-strategic landscape is increasingly complex and challenging, AI technologies are maturing rapidly and are becoming ever more widely available to allies and potential adversaries. We cannot shy away from the reality that future conflicts – in both the physical and virtual realms – will be AI-enabled. We have a duty to make the best use of these technologies to secure our national security and that of our allies. At the same time, we have a duty to use these technologies safely and responsibly, in line with our legal commitments and the values and standards of the society that we serve.

- The Defence AI Strategy (June 2022) set out how we will balance these duties through our ambition to become *"the world's most effective, efficient, trusted and influential Defence organisation for our size"*. The Strategy highlights the opportunities presented by AI – such as enabling quicker and better decision-making or automating 'dull, dirty and dangerous' tasks – while clearly articulating our approach to related risks and challenges. It sets out a framework to address key enablers (talent, digital and data infrastructures, policy), accelerate Research and Development to drive AI adoption at pace and scale, and to engage with partners across the AI ecosystem and internationally to promote values-based norms and standards for military AI.

- Alongside the Strategy, the MOD published its 'Ambitious, Safe, Responsible' policy, setting out a principles-based framework for the development and use of AI-enabled systems and capabilities. This approach – including the MOD's AI Ethics Principles – was developed in partnership with the UK's Centre for Data Ethics and Innovation through extensive consultation with a wide range of external experts, including ethicists, technologists, leading industry specialists and civil society groups. Many of these experts also sit on the MOD's AI Ethics Advisory Panel to provide ongoing independent advice on these issues. Although they form a single package, have equal status, are cross-referenced and are designed to be read in conjunction with one another, we published the Strategy and the 'Ambitious, Safe and Responsible' policy paper as separate documents to provide the agility to review our policy and

controls framework in response to any significant developments in this fast-moving technology field, and to promote engagement with our position on safe and responsible use (aspects of which might have been overlooked if integrated within the much more substantial Strategy document).

- Any use of AI to enhance Defence processes, systems or military capabilities is governed by our AI Ethics Principles. This is critical to retain the confidence of our people, our partners and our wider stakeholders (including Parliament and the general public) that Defence equipment is safe and reliable and would only be used responsibly in pursuit of legitimate military objectives. However, it is important to understand that AI is not a capability in and of itself but rather a component of a wider system or capability. As such, any use of AI would also be governed by our existing, robust framework of legal, safety and regulatory compliance regimes. We are embedding our AI Ethics Principles as an intrinsic part of these longstanding frameworks.

- For example, the MOD adheres to a strict legal review process for all new weapons systems as they are acquired or developed, in accordance with Article 36 of Additional Protocol I to the Geneva Conventions. The review process requires extensive evidence of performance against the capability's use case. It includes an assessment of whether the capability is inherently indiscriminate, capable of intentionally causing unnecessary suffering, or expected to cause excessive environmental damage.

- The UK is also a leading voice in international dialogues around the use of AI and autonomy in the Defence domain, including through the Group of Government Experts (GGE) on Lethal Autonomous Weapons Systems (LAWS) to the UN Convention for Certain Conventional Weapons (CCW). We are clear that International Humanitarian Law (IHL) provides a robust, technology-agnostic and effects-based legal framework for the regulation of armed conflict. It provides a comprehensive basis for the regulation of future technological development, in all its permutations, within the context of armed conflict. The UK has consistently advocated for the characterisation of autonomous systems in accordance with their ability to comply with rigorous principles of IHL.

- Defence is ambitious in our pursuit of strategic and operational advantage through AI. We are committed to the safe and responsible development and use of AI-enabled systems and capabilities and are actively promoting values-based norms and standards for military AI globally. We also recognise that the capabilities offered by these technologies are continuing to evolve and mature, and that we must

continue to engage proactively with the widest community of external experts to test thinking and refine our approach. We welcome the work of the House of Lords Select Committee on AI in Weapon Systems as a positive step in continuing this essential dialogue.

Definitions

1. How does the MoD define autonomous weapons systems (AWS); Why does the MoD not have an operational definition of AWS; Whether the MoD still stands by the definition it gave at the UN that fully autonomous systems would be those "capable of understanding, interpreting and applying higher level intent and direction"?

1.1 Dialogues about the use of AI and autonomy in Defence systems (including weapons) are constantly evolving as our understanding of opportunities and risks posed by these technologies matures. This includes, in particular, international discussions about Lethal Autonomous Weapons Systems (LAWS) convened under the UN UN Convention for Certain Conventional Weapons (CCW). The MOD is an active and authoritative voice in these discussions, seeking to achieve a consensus that reflects the primacy of International Humanitarian Law (IHL) in governing the use of any weapons system while reflecting the complex and nuanced ways in which AI technologies may be used.

1.2 The MOD is, however, cautious about adopting a definition for 'Autonomous Weapon Systems' (AWS) because such terms have acquired a meaning beyond their literal interpretation. Autonomy is a behaviour and AI is a functionality that can enable that behaviour through technologies such as machine learning. Autonomy as a function can exist in a system or just one part of a system – it is not a weapon or capability in itself. There are different levels of autonomy and there may be autonomous functionality within part of a weapon system that in no way enables the weapon to operate without the absolute positive control of a human operator. An overly narrow definition could become quickly outdated in such a complex and fast-moving area and could inadvertently hinder progress in international discussions.

1.3 The MOD's approach toward the use of AI in weapons systems is set out in the 'Ambitious, Safe and Responsible' policy (ASR, June 2022). This explains that while we *"do not rule out incorporating AI within weapon systems… we are very clear that there must be context-appropriate human involvement in weapons which identify, select and attack targets".* ASR goes

on to explain: "*AI can enable systems – including weapons – to exhibit some measure of autonomy: deciding and acting to accomplish desired goals, within defined parameters, based on acquired knowledge and an evolving situational awareness. This potentially could lead to weapons that identify, select and attack targets without context-appropriate human involvement. That is not acceptable – the United Kingdom does not possess fully autonomous weapon systems and has no intention of developing them.*" This approach extends and supersedes the definition used in the Joint Concept Note on Human-Machine Teaming (published in May 2018).

1.4 The UK's statement as to what is meant by a fully autonomous lethal weapon system as reflected in the question above, was made in the informal sessions of the LAWS GGE in 2016. Though the statement itself is not incorrect, the development of technology and understanding of its applications in the military domain has progressed significantly in the intervening period. The MOD would, therefore, highlight the Joint Position submitted to the formal session of the LAWS GGE by Australia, Canada, Japan, the Republic of Korea, the United Kingdom, and the United States on 13 March 2023 which recognised that the research and development of new technologies in the field of artificial intelligence is progressing at a rapid pace, potentially enabling "novel and more sophisticated weapons with autonomous functions, including those weapon systems that, once activated, can identify, select, and engage targets with lethal force without further intervention by an operator ("autonomous weapon systems" for the purposes of the Joint Proposal, without prejudice to any other understandings of this or similar terms for other purposes)". This definition doesn't apply hard parameters to the scope of an autonomous system, and it does not allow any systems to escape definition by virtue of technological functioning not in current use. Nevertheless, it provides a clear exposition of the sort of systems which inform the debate within the LAWS GGE and provides a reference point for the examination of how such systems comply with the law.

Current Capability

2. What the current state of AWS development is, including the technical limitations of the AI underpinning AWS and who is developing AWS; Whether the MoD still stands by its commitment not to develop 'fully autonomous weapons systems'

2.1 The MOD has clearly set out (including through the ASR policy) that UK Defence does not possess and does not intend to develop fully autonomous weapon systems. We oppose the creation and use of weapons

systems that would operate without meaningful and context-appropriate human involvement. Human responsibility and accountability for decisions on the use of weapons systems cannot be transferred to machines.

2.2 AI Technologies are maturing at pace and this functionality has the potential to be incorporated into a wide range of systems, including weapons systems (both existing and emerging), to enable various degrees of autonomous or semi-autonomous behaviours. However, it is important to be clear that any any weapon system intended to be fielded by the UK military would be governed by the MOD's robust framework of legal, safety and regulatory compliance regimes, irrespective of the technology involved.

2.3 All military systems must comply with our legal obligations, be safe to operate and operate safely within the accepted risk tolerance for the specific use case. We recognise that the nature of AI technologies can present fundamentally different testing and assurance challenges when compared to traditional physical and software capabilities, not least as it can be technically challenging to explain the basis for some of a system's decisions. We are actively examining our processes and compliance regimes to ensure that we can meet these challenges, including by ensuring that key safety standards are defined, achieved and maintained via MOD Regulators while providing agile alternative risk-based approaches to support operational requirements where required and appropriate.

2.3 Finally, it should be noted that all guided weapons and their associated targeting systems in, or about to enter, service are enabled to some degree by software algorithms that respond to particular inputs and produce consistent outputs. Such algorithms could be used in functions that assist target identification, sensor management, weapon guidance, route planning, weapon arming and fusing. The MOD does not regard these as 'fully autonomous weapons systems' because human control and decision-making is ensured at various stages. For example, depending on the weapon system, a human operator may decide when to turn the system on or off and when to fire it, as well as programmming the threat data, approving the route the weapon takes, setting the limits on the flight path etc.

International Humanitarian Law (IHL) and Future Operating Environment

3. What is the MoD's position on the sufficiency of existing International Humanitarian Law (IHL) in governing the use of AWS

and whether other international agreements or accords might be needed and how likely these are to come to fruition?

3.1 International Humanitarian Law (IHL) is principle-based and regulates how a weapon is used and its effect. It imposes positive obligations that take account of the core principles of IHL when engaging in military activity – necessity, distinction, humanity, proportionality – as well as an obligation to apply feasible precautions. IHL is technology agnostic which means it applies to all weapons systems, autonomous or not. This means that IHL is not just sufficient but is in fact the most appropriate means of regulating means and methods of warfare.

3.2 There is no other legal framework we could create now which is likely to obtain the same degree of respect and compliance as IHL. Rather, the UK considers that developing international consensus on interpretation of the law along with a body of best practice for the development and use of these technologies is a necessary step towards driving behavior in this space. This involves focusing on characteristics of systems and methods of deployment. In our view, this is the most effective way to ensure systems with autonomy are developed and used responsibly and lawfully.

4. What assessment has the MoD made of the impact AWS may have on warfare including issues such as: the speed of escalation, the number and nature of casualties, and their use by non-state actors?

4.1 AI has potential to be used to enhance a very wide range of defence systems and military capabilities, though these technologies are immature in many areas and we (and our allies) continue to undertake R&D and studies to understand the most effective and impactful applications. It is not possible to model the net effects that autonomous weapons systems may have on warfare in any meaningful way given the broad range of possibilities and future scenarios.

4.2 The Defence AI Strategy recognises, in principle, benefits, risks and implications arising from ubiquitous use of AI in military systems. Beneficial effects could include improved misssion planning, enhanced logistics management, better informed and quicker decision-making, more precise delivery of military effects (and hence reduced collateral damage), and protecting our people from harm by automating 'dirty and dangerous' tasks. Potential risks from widespread use of AI-enabled capabilities on the battlefield might conceivably include misunderstandings at times of heightened tension owing to unexpected AI behaviour, including malfunctions or as a result of cyber-attacks. Similarly, proliferation of

advanced AI solutions has potential to increase the threats from non-state actors either through direct technology transfers from hostile states or through repurposing commercial technologies.

4.3 The MOD maintains several expert functions to horizon scan, monitor technology trends and ensure that the UK Armed Forces are prepared to adapt to future threats and policy challenges. This includes 'futures' assessment functions in MOD Head Office, the Development, Concepts and Doctrine Centre (DCDC), the Defence Science and Technology Laboratory (Dstl), Defence Intelligence and the Front-Line Commands. We also work closely with partners across government, with NATO and with our allies to drive a common perspective on these issues and explore options to mitigate key risks.

Shaping Global AI Developments

5. Providing context on the MOD's contribution to international forums.

5.1 The MOD works with a wide range of international stakeholders to understand the opportunities and risks posed by AI-enabled military systems and capabilities. This includes bilateral partnerships (such as the AI Cooperation Statement signed with the US Department of Defence in May 2021); defence engagement with allies through NATO, 5-Eyes forums and the 16 nation AI Partnership for Defence; and active support to multilateral discussions in forums such as the Organisation for Economic Co-operation and Development (OECD), Council of Europe, UNESCO and the Future Tech Forum. Most recently, we supported the first intergovernmental summit on 'Responsible Use of AI in the Military Domain' (REAIM 23), hosted by the Dutch and South Korean governments in the Hague in February 2023.

5.2 The UK actively supports the NATO Emerging and Disruptive Technologies agenda, was instrumental in the development of the NATO AI Strategy (October 2021) and contributes significant legal and technical expertise to support ongoing activities to develop NATO AI standards and best practice toolkits. That strategy includes the NATO Principles of Responsible Use, which the UK considers to be complementary to our own AI ethics principles. In addition, the UK supports NATO's broader efforts to develop and field AI-enabled military capabilities, including through Science and Technology (S&T) and joint capability development programmes and by hosting, as of 30 March 2023, the first Regional Office of NATO's Defence Innovation Accelerator for the North Atlantic (DIANA).

5.3 The MOD also supports UK government engagement with related international initiatives. In October 2022, the UK joined 69 other states in signing a Joint Statement on LAWS at the UN General Assembly. The UK is also an active participant in UN-convened discussions on LAWS held under the auspices of the Convention on Certain Conventional Weapons (CCW). In addition to providing policy, technical and legal expertise to the discussions, the UK has contributed two national working papers (including proposing a clear framework with actionable guidance for policy, technical, and military stakeholders) and sponsored a Joint Proposal with the US, Australia, Canada, Japan and the Republic of Korea on Principles and Good Practices on Emerging Technologies in the Area of LAWS in March 2023.

Human-Centricity

6. What does the MoD define as context-appropriate human involvement? How does the MoD seek to maintain meaningful control when reacting at speed or when operators do not have sight of the data used by the system to make decisions?

6.1 As we have clearly set out in the Ambitious, Safe and Responsible (ASR) policy, we strongly believe that AI within defence systems – including in weapons – can and must be used lawfully, safely and ethically. In particular, there must be context-appropriate human involvement in weapons which identify, select and attack targets in order to satisfy fundamental principles of International Humanitarian Law as well as our own values and standards as expressed in the Defence AI Ethics Principles. However, it is important to understand that the nature and degree of human oversight or control over an AI-enabled system or capability will vary depending on a range of factors (e.g. context, function, level of autonomy) and may change over the course of a system's lifecycle. By way of example, the level and nature of human involvement appropriate for a target identification system operating as part of a human-machine team may be very different from that required for a defensive system that is designed to defend a maritime platform against hypersonic weapons.

6.2 For these reasons, we believe it is more helpful to consider autonomous functions and the nature of human involvement within and alongside those functions – a system of systems approach – rather than a weapon system in isolation. Human involvement is not limited to real-time supervision but could extend to setting a systems' operating parameters or other technical or procedural controls. Such controls will need to be established on a case-by-case basis, and any mitigations would be thoroughly assessed for legal compliance as part of standard Article 36

Weapons Reviews. In the case of a weapon system that involves highly automated decision-making processes where the decision to attack may be made or supported by autonomous or automated processes, there remains a need for a conscious accountable human actor to be satisfied that the relevant legal requirements of that decision are being met. The term 'context-appropriate human involvement' refers to the level of human involvement necessary to ensure that this fundamental requirement is being met in the particular context of deployment and having regard to the nature of the automated capability being deployed.

6.3 Good human machine interface (HMI) design and effective training will be essential, especially when users are required to step in and intervene at speed. The ASR principles on *Human-centricity* and *Understanding* underpin this, requiring decision-makers to develop appropriate trust in the system and sufficient system understanding to allow them to understand its likely behaviors in a given situation. Where speed is critical, prior bounding of behaviours might be the appropriate means to deliver both speed of response and control without adversely impacting operational performance (e.g. trading autonomy for more predictable automation behaviours to manage uncertainty).

7. What is the MoD's position on ensuring human-in-the-loop decision making and where accountability would lie for the actions of AWS?

7.1 Accountability always applies to a human and cannot be delegated to a machine. Defence already has significant expertise in understanding how accountability works on a systemic basis and how to assign accountability to the right level. We apply accountabilities through duty holding and through rules of engagement in order to apply military personnel and capabilities in a manner that achieves military strategic effect. New technological capabilities are adopted within that system of accountabilities.

7.2 We are clear that we do not necessarily need real-time supervision and the means for intervention (i.e. have a "human-in-the-loop") in every scenario, but that human responsibility for AI systems must be clearly established, ensuring accountability for their outcomes, with clearly defined means by which human control is exercised throughout the system lifecycles and irrespective of the use case. This has been set out in our ASR policy under the principle of *Responsibility*. It will be essential that personnel across the system lifecycle who are involved in the decisions to field and use AI enabled systems understand what they are accountable for and have the training needed to exercise that accountability.

7.3 The Secretary of State's Policy Statement on Health, Safety & Environmental Protection states that all TLB Holders are Senior Duty Holders and responsible for managing Risk to Life Activities in their areas of responsibility. Additionally MOD policy in relation to health and safety reflects the general UK statutory duty held by employers and employees in respect to third parties who are affected by the performance of their duties, so far as is reasonable. This requirement is reflected in a developed policy and governance structure that provides a good framework for the inclusion of AI regulation. AI is a functionality which needs to be addressed and regulated on a functional basis. Given the different use cases of AI-enabled systems in different contexts there cannot be a single gateway, but governance structures need to be clear so that everyone in the organisation understands how to manage and understand the additional functionality AI introduces.

7.4. At an organisational level, all Defence Top Level Budgets and Executive Organisations are required to designate an accountable officer responsible for AI Ethics implementation. This individual will be responsible for ensuring that effective policies, processes and governance arrangements are in place to embed AI ethics in their organisation.

Key enablers - Skills, Training and Data

8. How is the MoD considering issues such as impact of AWS on the defence workforce, including:

8.1. The in-house skills needed to develop AWS:

8.1.1 The Defence AI Strategy explains how, over time, we will need to incubate and uplift skills across the whole of the Department and UK Armed Forces to drive the effective exploitation and use of AI technologies. The skills requirements raised by any potential future adoption of autonomous weapons systems is a subset of this wider drive.

8.1.2 Leaders must be equipped with a level of understanding that is sufficient to navigate the hype, seize opportunities and act as intelligent customers. There is a need to improve AI literacy across professional communities (particularly policy, legal and commercial staff) and for deep subject matter expertise in coding and engineering disciplines across the Whole Force. A well-trained user base will be required to ensure that new AI-enabled capabilities are deployed effectively and safely. We are clear that whilst developers of AI systems need deep specialist knowledge, we are not

expecting the whole force to need such a deep technical knowledge – but they do need to understand broadly how the AI systems they operate work, their limitations, and how they should responsibly and effectively use them.

8.1.3 We are currently developing a Defence AI Skills Framework which will identify key skills requirements across Defence. This will be overseen by our Head of AI Profession, who will sit within the Defence AI Centre (DAIC) and will also be responsible for developing our recruitment and retention offer; setting standards for delivery team skills; and creating AI career development and progression pathways, with options for skilled generalists as well as deep specialists skills within the AI Professional framework to develop in-house or procure, assure, and employ solutions developed in partnership.

8.2: The skills needed by operators who work alongside AWS:

8.2.1 Training personnel in the use of new Defence equipment and weapons systems is not a new challenge for the Department. Defence is bound by UK law and has robust compliance regimes. Defence activities also include those that are inherently dangerous and require additional risk management beyond that of our statutory obligations. Dependent on the specifics of the AI system under operation, core skills requirements may include understanding how to use the system within any bounds of operation (including risks or limitations), how to maintain and , to sustain it, any requirements for ongoing assurance and verification of outputs, and appropriate methods of disposal – i.e. a whole lifecycle approach.

8.2.2 AI-enabled systems, and their outputs, must be appropriately understood by relevant individuals to facilitate effective and ethical decision-making in Defence. As such, our personnel must have an appropriate, context-specific understanding of the AI-enabled systems they operate and work alongside (potentially including highly autonomous weapons capabilities in the future). This level of understanding will naturally differ depending on the knowledge required to act ethically in a given role and with a given system.

8.2.3 We will need to ensure that our people train with AI-enabled systems under realistic conditions that test nominal (i.e. the normal or expected) operating conditions and system behaviours, as well as in scenarios or environments that may challenge AI capabilities. This will be critical to ensure that system operators, planners and commanders develop an understanding of system performance in a range of conditions and situations and are able to calibrate their trust accordingly, knowing when and how to

place constraints on use or implement sensible precautions. This will need to occur regularly, especially after retraining or updating the AI Agents.

8.3: The impact of AWS on the morale and self-worth of those working alongside AWS, and any support systems necessary:

8.3.1 The Defence AI Strategy rightly recognised that our people are our finest asset, that the real world impact of military action demands applied human judgement and accountability, and that we must support any staff that are affected by the adoption of AI. Moreover, The 'Ambitious, Safe and Responsible' (ASR) principle on *Human-centricity* clearly sets out that *"the impact of AI-enabled systems on humans must be assessed and considered, for a full range of effects both positive and negative across the entire system lifecycle"*. Whether they are MOD personnel, civilians, or targets of military action, humans interacting with or affected by AI-enabled systems for Defence must be treated with respect and it may be unethical to use certain systems where negative human impacts outweigh the benefits.

8.3.2 Conversely, there may be a strong ethical case for the development and use of an AI system where it would be demonstrably beneficial or result in a more ethical outcome. For example, we can reduce the need for humans to undertake repetitive and monotonous tasks – such as operators having to constantly monitor the output of CCTV camera sensors – in order to enhance outcomes while also enabling human operators to focus on more rewarding activities. As an example, the SAPIENT system (Sensing for Asset Protection with Integrated Electronic Networked Technology) is intended to make low-level decisions autonomously, such as deciding which direction to look or zoom in, in order to help a person to fulfil a higher-level objective.

8.3.3 The impact of AI technologies on our people is likely to vary significantly depending on the function and context and will need to be assessed on a case-by-case basis. When we introduce AI into a system or process, we will ensure that its operators are trained to use the AI and have sufficient understanding of its risks, benefits and limitations. Rather than feeling replaced by a machine, this upskilling process should increase the operator's confidence and make their job more rewarding.

8.4: The training needed for those soldiers who may be faced with enemy AWS:

8.4.1 It is standard practice for service personnel to be trained on context-specific threats as part of pre-deployment training before entering the operational theatre. This may include additional role-dependent training

needs specific to particular adversary capabilities. The MOD will keep these training requirements under review as the deployment of AI-enabled battlefield systems increases and our understanding of adversarial capabilities and military concepts and doctrines improves.

9. What challenges are associated with the underlying data sources including the sufficiency of training data and securely transferring and processing data?

9.1 AI systems are often critically dependent on access to (large scale) structured and labelled data. The Defence AI Strategy recognises that this is a particular issue for the MOD as data can often be stove-piped and badly curated, making it challenging, time consuming and cost intensive to access sufficient levels of machine-ready data to train AI models. Data ownership and the ability to share data can also present significant challenges; the MOD does not always own the data that it needs and there can sometimes be cultural, security and commercial challenges in sharing data more widely. For example, due to the classified nature of Defence's work it can sometimes be difficult to share relevant data sets with small-and-medium sized companies and innovators.

9.2 The MOD is working to tackle these challenges through implementation of the Defence Digital and Data Strategies, transforming into a data-driven organisation, breaking down silos and improving data sources. The Defence Data Framework has been developed to transform Defence's culture, behaviours and data capabilities. The Defence AI Centre is also examining options to make representative datasets available to our supplier base in key areas, and to use synthetic data in less data rich environments. We are also working closely with our allies and partners to maximise data-sharing opportunities.

Procurement and working with other Sectors

10. What level of engagement does the MoD have with the private sector who are developing AWS; How is the MoD engaging with organisations developing AI for civilian purposes which may have defence applications?

10.1 While the MOD does not rule out incorporating AI within weapons systems, we have been very clear that there must be context-appropriate human involvement in any systems that could identify, select or attack targets. Transparency, insofar as possible, will be key to working effectively

with the private and academic sectors: we are open about the range of areas where we see potential opportunities to use AI to enhance military capabilities, including a subset at the kinetic end of the spectrum. We are very clear with potential partners that we take our legal safety and ethical responsibilities very seriously and are challenging our academic and industry partners to assists us to ensure that our technical and policy approaches are effective and robust. We understand and accept that some partners will not want to work with us on certain military applications of AI; we hope that transparency and dialogue ensure that we can maintain trusting and constructive relationships on those areas where we *can* work together.

10.2. The Defence AI Strategy sets out how we will need to work with the widest range of partners across the UK AI ecosystem to realise our aim to and adopt and exploit AI at pace and scale. Most new and emerging technologies like AI are inherently dual use, and innovation is found across a broad ecosystem of university spinouts and fast moving small-to-medium scale tech enterprises. We accept that we must modernise and streamline our approaches to industry engagement to reach out to these communities and incentivise them to collaborate with us.

10.3. A range of work is underway in this area. Defence's CommercialX function is streamlining our approach to contracting with digital suppliers and, through implementation of the Defence and Security Industrial Strategy (DSIS), we are identifying ways to break down structural and procedural obstacles that can inhibit new suppliers (particularly small-and-medium sized companies) from working with us. We are actively fostering a more dynamic and integrated relationship with the AI ecosystem by cultivating a Defence and National Security AI Network through the Defence AI Centre (DAIC). This involves working closely alongside international partners, wider government, industry, academia, and other partners to develop our approach, drawing from good practice, solutions, and frameworks developed elsewhere. As examples, the DAIC held an industry day in March 2023 with over 200 people and the AI Fest event in April 2023 included over 1000 delegates from across Industry, Academia, and Government.

11. How does the MoD manage the procurement of AWS given their rapid pace of development and the speed at which these systems become obsolete?

11.1 It is important to be clear that although the MOD does not rule out incorporating AI within weapons systems – subject to context appropriate human Involvement – we are not in the process of procuring "autonomous weapons systems" as implied.

11.2 Defence has a deservedly strong reputation for taking a robust approach to safety and legal and regulatory compliance; this extends to the acquisition of any AI-enabled technologies. As described in our 'Ambitious, Safe and Responsible' policy, we intend that our approach will enable – rather than constrain – the adoption and exploitation of AI-enabled solutions and capabilities across Defence. We will empower teams developing and delivering concepts, technologies, and solutions to explore ambitious but responsible ideas and use cases. We will provide them with clear frameworks to support the early identification and resolution of safety, legal and ethical risks; this will give them the confidence to explore the full potential of the technology while complying with policy and other essential requirements. We will encourage them to identify wider factors impeding their progress – such as policy or process – in the expectation that appropriate solutions will be identified and implemented rapidly.

11.2 As explained above, we are also examining our processes and compliance regimes to ensure that we can meet the challenges of accelerating technological change. We will ensure that key safety standards are defined, achieved and maintained for new core capabilities via MoD Regulators while providing agile alternative risk-based approaches to support operational requirements where required and appropriate.

Financial

12. What assessment has the MoD made of the financial impact of developing and deploying AWS, particularly the trade-off between current high development costs and future low deployment costs?

12.1. The MOD is currently not developing or deploying any autonomous weapons systems. As with any other Defence capability, the decision whether to pursue such systems as a solution to a particular military requirement would include business cases that consider options from a through life perspective (i.e. both acquisition and support) together with an assessment of effectiveness as set against alternative options. Early-stage experimentation and exploitation work will be pursued via S&T arrangements.

External Challenge

13. What is the extent of internal and external challenge to the MoD's work on AWS, particularly how they work with their in-house AI Ethics Advisory panel?

13.1 We understand that the use of AI in the military is controversial. Internal and external challenge are critical to ensure that we understand the range of inter-related issues and can provide assurance that any AI-enabled system or military capability is safe, trustworthy and ethical. This includes close working with a diverse range of experts (ethicists, technologists, academics, legal advisers, industry specialists, frontline military personnel etc), partnering with international allies and external organisations such as the Centre for Data Ethics and Innovation (CDEI), and engagement with a wide range of civil society stakeholders and interest groups. This engagement helps to ensure that our approach to the adoption of AI technologies is transparent, inclusive and founded on a robust evidence base.

13.2 The MOD AI Ethics Advisory Panel (EAP) brings together a group of experts from Defence, Academia, Industry and Civil Society to advise the Second Permanent Secretary on the safe and responsible use of AI in Defence. The current membership is listed in Annex B of the 'Ambitious, Safe and Responsible' policy. The EAP is intended to meet regularly and, while it does not have formal decision-making powers (as this preserves members' independence which allows them to voice their opinions freely), the EAP's advice has been critical to the development of Defence's AI Ethics Principles and will be no less central in ensuring effective implementation over the coming months and years. We continue to seek challenge from the panel to identify gaps or vulnerabilities within Defence where ethical standards might be particularly at risk.

13.3. Other examples of our ongoing engagement with external stakeholders and civil society groups include recent workshops with the International Committee for the Red Cross to test current policy positions, regular discussions in the context of the UN GGE on LAWS, and working with Team Defence and a group of industry stakeholders to understand how industry would design an AI Assurance framework for Defence.

Ministry of Defence
June 2023