November 2022

House of Commons Science and Technology Select Committee inquiry into AI governance

This submission is made on behalf of the UK Campaign to Stop Killer Robots (UK CSKR). We are a network of UK-based NGOs, tech experts and academics concerned with the risks associated with growing autonomy in weapons systems, including AI-enabled systems. The UK Steering Committee includes Amnesty International UK, Article 36, Drone Wars UK, United Nations Association UK and the Women's International League for Peace and Freedom UK. Each member of the UK CSKR may not necessarily endorse or take a position on all points made in this submission.

Our coalition is submitting evidence to this inquiry in the hope that our recommendations could help ensure efficient and ethically compliant AI governance in the UK, which we believe to be essential to ensure responsible and safe use of AI, especially in weapons systems that use AI. Our submission therefore relates mainly to military uses of AI and autonomous weapon systems (AWS) with AI components.

Are current options for challenging the use of AI adequate and, if not, how can they be improved? How effective is current governance of AI in the UK?

AI and its myriad applications has the potential to deliver significant positive societal outcomes. It also raises complex ethical dilemmas, particularly relating to AI's potential for becoming a decisive factor in warfare. This is new territory for the UK. Consistent with the UK's status as a democratic nation which emphasises its commitment to fostering an "open society,"[1] the UK should collaborate with the public to help resolve sensitive questions relating to military uses of AI.

It is therefore concerning to learn that the UK's Defence AI Strategy was formulated without public consultation or any attempt by ministers to lead a national conversation around this issue.[2] By excluding the possibility of public engagement and scrutiny of a strategy which guides AI use in one of the most controversial, high-stakes sectors, the UK has failed to live up to its own mantra of inclusivity and consultation, weakening the legitimacy of its AI governance frameworks.

The UK CSKR recommends that the Government;

- conducts a full public consultation on the use of AI in weapons systems;
- establishes mechanisms for the public, civil society and other stakeholders to be included in ongoing scrutiny of the implementation of the Defence AI Strategy and other relevant doctrines;

[1] The phrase "open societies" features 25 times throughout the 114-page Integrated Review. See UNA-UK and Rethinking Security's joint report for more info, December 2021, https://rethinkingsecurity.org.uk/wp-content/uploads/2021/12/briefing-external-consultation-and-the-integrated-review-09dec2021.pdf

[2] Freedom of Information Request submitted by UNA-UK, September 2021. Available at: https://una.org.uk/sites/default/files/2022-01/20210910_foi_-_ai_ethics_foi202108931_1.pdf

- develop an ethical code of conduct for emerging technology and weapons systems.

What are the current strengths and weaknesses of current arrangements, including for research?

At present, governance arrangements for AI are entirely voluntary and largely based on vaguely worded, unenforceable 'ethical codes'. Despite the immense impact that AI may have on future society, discussion on AI outside among the general public, politicians, and the media is, outside technical circles, largely superficial and is based almost entirely on perceived benefits that AI may bring. This ignores and obfuscates the potential harms.

Referring to the implementation of the government's Defence AI Strategy, as fulfilling its commitment that AI be used "safely, lawfully and ethically in line with the standards, values and norms of the society we serve," the government has highlighted "scrutiny and challenge" by its "Independent" Ethical Advisory Panel as a key measure.[3] This panel has no democratic mandate, no transparency and, alarmingly, contains some members whose departments are written into the AI Strategy as close government partners to "meeting our operational needs, and reflecting deeply-held values". This hardly inspires confidence regarding independent research, scrutiny and ethical safeguarding.

Additionally, despite increasing cross-party parliamentary interest in the issue of AI-enabled weapons systems, there has been little opportunity for debate or scrutiny of government policy, including the Defence AI strategy released in June 2022.

If the UK ambition to be a leader in ethical AI, as stated in the Integrated Review of Security, Defence, Development and Foreign Policy, is to be realised, and the UK position as a global tech superpower is not to be stymied, parliament and the Government needs to show leadership on this issue. In particular, they must support robust national and international legal regulation in this area. Rather, the nation is sleep-walking into a future where AI will be ubiquitous without regulation, preparation, or even an understanding of the technology.

It is worth noting that in contrast to the UK Government, the tech sector widely supports the creation of new legal frameworks to regulate autonomy in weapons systems, including to protect intellectual property developed for peaceful purposes from being misused through incorporation into AWS.

What measures could make the use of AI more transparent and explainable to the public?
How should decisions involving AI be reviewed and scrutinised in both public and private sectors?

Parliamentarians and the media urgently need to take a more critical view of AI, and seek information about the technology from a variety of sources, including critical voices. At present the narrative on AI is driven almost entirely by industries and researchers who advocate wholesale uptake of the technology. The government should drive a 'national conversation', which engages the general public on the subject of potentially disruptive new technologies such as AI, and invite a wide variety of voices to participate.

[3] Defence: Artificial intelligence - question for Ministry of Defence, July 2022. Available at: https://questions-statements.parliament.uk/written-questions/detail/2022-07-21/HL1998

In order for the use of AI to be more transparent and compliant with ethical principles, a credible and inclusive public consultation must be conducted. This would allow the public and parliament to examine the military, ethical, humanitarian and legal implications of advances in AI. Such consultation should:

- Interrogate the fundamental ethical question relating to our relationship with technology - how far should society be prepared to go with respect to outsourcing military operations to algorithms, sensors, AI and autonomous technologies?
- Assess the present state of technological developments, the prospects of deployment of AI-powered AWS, and the inherent risks represented by AWS.
- Explore the efficacy of existing international law for regulating use of AI-enabled autonomous weapons, assess the progress of international negotiations towards a new treaty to regulate AWS and the UK's role in them.
- Provide in-depth assessment on the adequacy of the MoD national strategy for the deployment of AI and accompanying policy statement .
- Assess the consequences of remote and AI-enabled autonomous weaponry in furthering power asymmetries in regional and global conflicts.
- Highlight various tech sector perspectives regarding the risks posed by dual-use AI technology, and their export
- Examine whether the anticipated effect of the UK working toward internationally agreed limits on autonomy in weapons systems stymie innovation, as the Government claims; or whether it would in fact protect research, development and industrial supply chain, (as the recent AI Foreign Affairs Committee Inquiry found)[4]
- Raise awareness of the moral and ethical issues relating to the uses of AI in military contexts and contribute to the emerging national conversation on this issue
- Formulate recommendations to the UK government aimed at ensuring an ethical and responsible use of AI and autonomous technology, particularly when applied to targeting and weapon systems in general.

To what extent is the legal framework for the use of AI, especially in making decisions, fit for purpose? Is more legislation or better guidance required?

The legal framework for the use of AI is extremely narrow, restricted and based largely on guidance. As such, it is not fit for the purpose of protecting the public from the potential harms posed by AI systems.

As part of the Defence AI Strategy, the Government unveiled a new policy on autonomous weapons. Published as an annex on "Lethal Autonomous Weapons Systems", to an accompanying policy paper, the Government recognises that systems which identify, select and attack targets without "context-appropriate human involvement" would be unacceptable.

This is significant, as it continues to recognise that there is a line that should not be crossed in terms of machines making decisions in combat. It is also positive that the new UK position recognises that ethical principles, and the law, are important for determining where this line

[4] Encoding values: Putting tech at the heart of UK foreign policy, July 2022, Available at: https://committees.parliament.uk/committee/78/foreign-affairs-committee/news/171953/encoding-values-putting-tech-at-the-heart-of-uk-foreign-policy/

should be drawn. It also rightly asserts that mechanisms of human authority, responsibility and accountability will always be present when force is used.

However, the position provides no indications of how "context appropriate human involvement" is to be assessed, applied, taught or understood. In this new formulation, "context appropriate human involvement" could mean almost anything. Such ambiguity, using unpredictable technology, is unacceptable. Given the lack of clarity or further detailed scrutiny (including by parliament) of the limits the UK would apply, this could be equivalent to the UK public being told by the military, "leave it to us" to determine what is appropriate. This is emblematic of our concerns that this policy position, and the wider Defence AI Strategy, was formulated without public consultation or any attempt by ministers to have a national conversation on this issue.

The government has stated that its priority with respect to AWS is "setting clear international norms for the safe and responsible development and use of AI, to ensure compliance with International Humanitarian Law through meaningful and context-appropriate levels of human control".[5] In our view, this would be best achieved through a new international legal instrument to prohibit and regulate autonomy in weapons systems.

What's missing?

In the international discussion on AI-enabled AWS there is widespread recognition that certain factors are necessary for meaningful human control, or sufficient predictability, over autonomous systems, including:

- human control over the duration and geographic scope of an autonomous system's operation – these are vital to making judgements about the possible effects of a system's use.
- human understanding of the target profiles a system uses – this is vital to understanding what will trigger the system to apply force, including the possibility of false positives (targeting things that are not the intended targets) – and this is also an area where AI and machine learning raise particular concerns.

Regrettably, these factors are not mentioned in the government's policy.

The International Committee of the Red Cross and civil society organisations have urged states to reject autonomous systems that would apply force directly to people. However, there is no indication in the UK policy position of whether the UK considers it acceptable to allow machines to identify and automatically fire on human beings. These questions require further attention and the formulation of detailed policy. In this regard, it is encouraging that according to a recent parliamentary answer the government considers that "systems that are designed to identify people as targets based on only biometrics, and perceived gender, race, and age are very unlikely to comply with the requirements of International Humanitarian Law."[6]

[5] Autonomous Weapons: Ethics - Question for Ministry of Defence, July 2022. Available at: https://questions-statements.parliament.uk/written-questions/detail/2022-07-21/HL2031

[6] Autonomous Weapons: Ethics - Question for Ministry of Defence, November 2022. Available at: https://questions-statements.parliament.uk/written-questions/detail/2022-11-07/HL3230

Recommendations

- The UK should join over 70 states in supporting the creation of a new legally binding treaty to regulate autonomy in weapons systems. This is supported by tech leaders, the UN Secretary General, the ICRC, and the global Stop Killer Robots campaign representing 180 organisations in 66 countries. This should include:
    - A general obligation to ensure meaningful human control over the use of force
    - Prohibitions on AWS that target people and autonomous weapon systems that cannot be used with meaningful human control
    - A commitment that all AWS that are not prohibited are subject to positive obligations to ensure meaningful human control.
- Until the time that this treaty has entered into force and ratified by the UK, the UK should pass national legislation to this effect. This could be a stand alone piece of legislation or within the context of an Algorithms Act which places limits on the types of functions that can be delegated to AI systems.
- The UK should significantly tighten its arms export controls to better regulate the export dual-use components which could be incorporated into AWS.

How should the use of AI be regulated, and which body or bodies should provide regulatory oversight?

As yet, there is no formal national regulatory body for the use of AI systems. A new regulator should be established to undertake this role. Given that the regulation of AI and computer-based systems is in its infancy, such a regulator should initially focus on understanding and developing its role, as well as building capacity to establish an effective regulatory framework. In due course, legislation should be enacted to formally implicate this framework, which should be binding on all researchers, developers and users of AI, as well as aim to protect the public and society from the harms of AI and ensure that AI systems remain secure. This will require the regulator to take a risk-based approach, focusing on AI technologies with the highest potential for harm. Given the power and size of the tech sector, the regulator should have wide-ranging powers and the ability to impose heavy sanctions, including the ability to impose unlimited fines and ban organisations and individuals from involvement in the sector. In order to maintain public trust in AI technology, the regulator should operate to high standards of openness and transparency.

Military uses of AI represent areas with the highest potential for harm, and must therefore be closely regulated. As we argue below, many AI applications also have military-civilian dual uses. The MoD and the armed forces should therefore come under the regulatory vires of the same national regulatory agency as the civilian AI sector. The MoD will argue that it should be allowed to establish its own internal regulatory arrangements for AI, but this would be insufficient. Current in-house regulation of high-hazard military activities (e.g. by the Defence Nuclear Safety Regulator) is untransparent and unaccountable, and lacks the resources and diversity of skills available to equivalent civilian regulatory agencies. In addition, the Secretary of State for Defence faces a conflict of interest between allowing ongoing operations and enforcing regulatory requirements. There should be no difficulty in ensuring that personnel employed by a national regulatory agency have adequate security clearances and follow appropriate procedures to ensure national security and defence interests are not compromised.

In the absence of appropriate regulation, the direct development of autonomous military technologies and military AI is being driven by arms companies. As these companies manufacture and sell highly harmful high-technology products, there is a need for the government to keep their activities under scrutiny and monitor and regulate their exports. Despite existing export controls on military and dual-use technology, there has been a concerningly consistent pattern over many years of the supply of British military equipment to abusive regimes. We consider that the system of export controls needs to be tightened in general, together with specific attention to prevent sophisticated technologies with potential military uses from falling into the wrong hands. We also consider that funding to arms companies for the development of new military equipment should be conditional on their products and activities remaining within strict ethical limits.

It should be noted that universities play an important role in the development of new AI technologies, increasingly in partnership with industry. It is important that they too are subject to regulation, and that there are controls on the use of intellectual property developed by universities, the nature of their research activities, funding relationships and the involvement of overseas partners and students in research programmes specifically relating to technologies with potential for harm. The government should encourage efforts to build an ethical research culture with complete transparency to students and faculty.

While military-specific AI technology may be developed by weapons companies, many vital technologies are increasingly 'dual use' – they have the potential to be applied for both beneficial or harmful purposes, which is particularly concerning. There should be a responsibility on researchers and developers of new technologies to identify both the beneficial and malicious uses of technology they develop and to introduce meaningful measures to mitigate against the risk of its malicious use. We see from the US Department of Defence's Project Maven[7] (a controversial AI programme facilitated by Google), and others, that there is great concern among engineers[8] about issues such as the lack of regulatory framework, transparency around final use or consistency with corporate ethical codes of conduct. Crucially, there is little assurance that technology or code produced within UK companies may be ring-fenced from military purposes. This could create an apprehensive workforce.[9].

The nature of multinational companies working on dual-use technologies in the global supply chain means that technology, as well as expertise, crosses national borders and can be sold by industry to any potentially hostile actor or for unforeseen use, with profit and harms dislocated.

The government should take action to ensure that UK-based companies are responsible for any liabilities arising from the use of technologies they have developed, both at home and overseas. This includes ensuring that they have adequate reserves and insurance to cover liabilities. Regulation to prevent and control harms arising from the development of emerging technologies should specifically cover harms overseas. Incentivising business to develop along ethical guidelines could create safer and more profitable outcomes.

What lessons, if any, can the UK learn from other countries on AI governance?

[7] Project Maven to Deploy Computer Algorithms to War Zone by Year's End. Available at:

[8] Google employees ask tech giant to pull out of Pentagon AI project.Available at: https://www.defense.gov/Explore/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/ https://globalnews.ca/news/4124514/google-project-maven-open-letter-pentagon/

[9] Workers in the AI sector are quitting over ethical concerns. Available at: https://tech.newstatesman.com/business/tech-workers-ai-sector

The European Union has published a draft regulation on AI to guarantee the safety and rights of people and businesses whilst allowing uptake of AI technology. The draft regulations follow 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. 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. Although the UK is no longer a member of the EU, it shares the same democratic and rights-based values as the EU and is affected by its regulations, the UK should thus adopt legislation providing at least the same levels of protection as the EU's regulations. All AI systems, including high-harm military systems, should be covered by the legislation. There should be no national security exemption to the legislation: although it would be necessary to adapt procedures and processes to maintain security, this can be done straightforwardly within the terms of regulatory legislation.