Dr David Anderson, Intelligent Systems Research Group Head, University of Glasgow – University of Glasgow (AIW0041)

Introduction

Following a visit by members of the House of Lords select committee into AI in weapons systems to the University of Glasgow on 12th September 2023, the Chair on the day has asked those present to prepare a written submission. Following consideration of the six questions posed by the committee and considering the expertise of the UofG personnel who met with the committee, there now follows responses to those questions in which we feel confident to provide expert testimony.

Considered Responses

Question 1: What do you understand by the term autonomous weapons system (AWS)? Should the UK adopt an operative definition of AWS?

Response:

As there is the potential for confusion, ideally an unambiguous definition of Autonomous Weapons Systems (AWS) is desirable, one that ensures the distinction between AWS and Weapons Systems Employing Autonomy (WSEA). This definition should also make it crystal clear to all stakeholders, technical and policy, that autonomy is not binary – autonomy lies on a spectrum which goes from human operation (minimal autonomy) through automatic systems (partial autonomy) on to fully-autonomous systems. Only this final category of full autonomy contains the ethical and legal implications under consideration by the committee.

Many research & development teams throughout the UK are working on new defence system technologies that will incorporate Autonomy, Artificial Intelligence and Machine Learning algorithms. Inclusion of these algorithms will significantly enhance WSEA system performance, improve data handling, reduce battlefield uncertainty etc., to name only a few potential benefits. Such endeavours must not be jeopardised over imprecise discussions, specifically at policy level. To do so would put the UK military at a strategic disadvantage.

Question 2: What are the possible challenges, risks, benefits and ethical concerns of AWS? How would AWS change the makeup of defence forces and the nature of combat?

Response:

Fully autonomous weapons systems do fundamentally pose a potential ethical issue within the context of what is considered appropriate operational use. Giving an AI carte-blanche authority to decide where and when to execute lethal force is troubling and certainly not adhering to the MoD policy outlined in the recent "AI in defence" document. However, there may be some circumstances where lethal engagement is warranted, policing of a no-fly zone for example. Indeed, using autonomous aircraft for sorties of this kind has several operational advantages – no risk to UK pilots, extended endurance etc. Fundamentally a no-fly zone is the designation of an area where it is understood that lethal force has been authorised, it is well-regulated and there is minimal chance of innocent civilian presence. For AWS deployed on land, a similar exclusion zone concept where all non-UK or allied equipment is designated hostile would be much more difficult to implement in a manner guaranteed to produce no civilian casualties. We are then left with the question – is it ethical to deploy a fully autonomous AWS in this circumstance? Is there an acceptable level of risk to innocent or allied personnel we would be willing to accept to protect UK service personnel?

Taking a step back from fully autonomous AWS, there are significant benefits to be gained from integrating AI and ML technology into defence systems. For example, AI/ML are being used in sensor suite design and data integration with the objective of relaying much more accurate situational awareness to the warfighter/commander than ever before. Such systems will lead to decisions being taken that will reduce civilian casualties, blue-on-blue or blue-on-green engagements and enhance the effectiveness and survivability of our warfighters and assets in combat.

Question 3: What safeguards (technological, legal, procedural or otherwise) would be needed to ensure safe, reliable and accountable AWS?

Response:

Technological, legal and procedural safeguards must all be put in place prior to operational use of AWS. Considering the technological safeguards (the only one of the three I feel qualified to provide comment on) care needs to be taken when adopting AI/ML algorithms developed for other domains. For example, the backbone algorithm for most computing & data science AI applications is the deep neural network. DNNs are a class of universal function approximators that are adaptively trained to emulate a particular function via exposure to large amounts of training data. Therein lies the vulnerability of DNNs in the defence context – their ability to generalise is only as good as the training data used. If the training data

is not extensive, the DNN will respond unpredictably to operational edge-cases. There is also the danger of corruption in the training data, particularly if that data was obtained from open-source repositories. Fundamentally though, there is currently no agreed mathematical or practical methodology for guaranteeing deterministic behaviour from a data-driven AI, which is a major issue for safety-critical or offensive military systems.

Given the importance of deploying an AI that is trustworthy, safe and ethical, care should be taken in selecting which AI technologies are implemented on defence systems. One way of mitigating the impact of data dependency is to fuse the AI with existing physical models or constraints. Fields of research such as scientific machine learning (SML) or physics-informed neural networks (PINNs) provide a mechanism for constraining the AI to operate within the neighbourhood of conventional scientific processes, whether these be determined by partial differential equations or decision trees. Within the military context, SML has the potential to embed a failsafe into the AI, that ideally would ensure the AI is compliant with UK government policy and International Humanitarian Legislation.

Dr David Anderson
University of Glasgow
October 2023