Tomorrow's Affairs
Analysis of today
Assessment of tomorrow

By: Stephen Harwood

Proceed with Caution

'Proceed with Caution' is a warning. It signals danger ahead. It suggests that we slow down and take stock of where we are. However, do we heed this warning or continue without care?

'Proceed with Caution' is the title of an UK House of Lord's Report about the use of Artificial Intelligence (AI) in Defence, published 1st December 2023. It is the outcome of an inquiry by the UK House of Lords 'Artificial Intelligence in Weapons Systems Select Committee' which issued a call for evidence in March 2023. A core notion is the association of AI with Autonomous Weapon Systems (AWS).

The report's 98 pages span a range of topics which include defining what an AWS is, the benefits and risks associated with using AWS and whether use of AWS can be compliant with International Humanitarian Law (IHL). It concludes by considering the implications of the development of AWS upon UK Government Policy.

The report recognises the UK Parliament's central decision-making role in the development and use of AWS. This includes ensuring that there is public confidence in its development and also that any policy is ethically grounded.

Highlighted is the requirement that "the Government should ensure human control at all stages of an AWS's lifecycle". This draws upon the Ministry of Defence's principles for the use of AI in defence which includes responsibility and accountability as well as proactive mitigation of risk of bias and harm.

The report, as suggested by its title, supported other documents such as the policy paper on AI use in Defence, "Ambitious, Safe, Responsible", potentially offers comfort about the UK's use of AI in weaponry, and by implication the development and use of AWS.

Who is accountable if something goes 'wrong'?

There is reason to find comfort in AWS. The benefits of using AWS include fast, more efficient and better focused engagement, the reduced need for human combatants who potentially have the capability of inappropriate behaviour such as unwarranted destruction, revenge, dehumanisation, torture and murder.

However, there are risks. These include a lack of understanding of how the AI aspects of an AWS functions, and its unpredictability due to AI's capability for self-learning, the danger of malfunction and the potential for a cybersecurity breach.

We need to add the question of who is accountable if something goes "wrong", especially if there is violation of IHL. IHL comprises five core principles that must be complied with.

'Humanity' and 'Necessity' requires that any action is that which is sufficient to achieve its legitimate intent, but at the same time with minimal harm, loss of life and destruction.

This implies the "distinction" between legitimate military targets and illegitimate non-military targets (e.g. schools, dwellings, places of worship, museums, markets and hospitals).

> All feasible "Precautions"must be taken to protect civilians and civilian objects. If there are violations to IHL, then there is the question of how these violations are handled, which includes who is held accountable

This anticipates that there may be accompanying "collateral" loss of life, harm and destruction, but this must be "proportionate",' in other words, cannot be excessive relative to the clear military advantage to be expected.

Irrespective, all feasible "Precautions"must be taken to protect civilians and civilian objects. However, if there are violations to IHL, then there is the question of how these violations

are handled, which includes who is held accountable.

The assumption is that it is acceptable to use AI and hence AWS in conflicts. Indeed, the UK's apparent commitment to the ethical grounding of any policy relating to this could be viewed as supporting this.

However, there are two other views about AWS and its development and use. First is the argument, as expressed by the Stop Killer Robots coalition, that it is unethical, if not dehumanising, for a machine to decide who or who not to kill and as such should be regulated if not banned.

The counter-argument is that there is a need to pursue these developments as there will always be rogue nation-states and other actors who will ignore any agreements about how AI and AWS ought to be developed and used. This necessitates a superior capability and the requirement of AWS.

There is no better testbed for AWS than in a war

However, the UK is not alone in the development and use of AWS. The US is developing its own AI/autonomous systems (e.g. Project Maven).

However, a recent joint report by Human Rights Watch and the International Human Rights Clinic (IHRC) reveals that whilst the US Department of Defence updates Directives about its development and deployment of AWS, the US government does not have a government-wide policy that regulates how AWS might be used in situations that are not that of armed conflict, such as law enforcement and border control. It provides the example of the Central Intelligence Agency (CIA) and its use of armed drones.

This distinction between who within the US might use an AWS, triggers calls for a new international treaty to prohibit and regulate AWS and ensure that there is meaningful human control and hence clarification of who is accountable.

> Indeed, there is possibly no better testbed for AWS than in a war

Whilst other countries are developing their own systems, such as China, Russia, India, Iran, South Korea, Australia and Turkey the critical issue is how well an AI regulated (i.e. [semi-]autonomous) weapon system performs in a war zone. Indeed, there is possibly no better testbed for AWS than in a war.

On the 24th February, 2022, Russia invaded Ukraine. This has provided a window into how newer forms of technology can be used in major conflicts, especially AI.

Indeed, Ukraine is reported to have used AI for different military related objectives, such as documenting war-crimes, cybersecurity defence and with the use of domestically produced drones to identify and attack targets.

Moreover, this has provided a live lab for foreign companies and Ukrainian start-ups to innovative and prove their technologies. Moreover, AI aids the production of misinformation.

"The Gospel"

However, another war materialised on the 7th October, 2023, which 115 days later was described by Ajith Sunghay, Head of the UN Human Rights Office, as "a massive human rights crisis and a humanitarian disaster" in an area that has been referred to as a "mass assassination factory".

This is the 'Operation Iron Swords' offensive against Gaza, a besieged strip of land which comprises 365km2 and home to a population of 2.3 million people and is predominantly urban.

Whatever the view about what has happened before and on that date, since that date, the

deaths, destruction, displacement, and deprivation of so many civilians, especially women and children in Gaza, invites many questions about the attacking force: the Israeli Defense Forces (IDF). This is action against a Hamas military wing that is estimated to comprise 30,000.

However, our interpretation of what is going on is blighted by a question of what information to believe, particularly given, in wartime, the classified nature of information and the persuasive role of misinformation, claims and counterclaim.

A report by the Israeli-Palestinian +972 Magazine and the Hebrew-language news site Local Call reveals that the IDF are using a system named 'Habsora' or, in English, "The Gospel".

The English name means a revelation from heaven, a truth to be implicitly received, or a preached doctrine of salvation, inviting questions about the invoked symbolism of the name.

"The Gospel" is an AI enabled system which is supported by an extensive database about potential targets. This includes data which allows the calculation of the number of civilians that are likely to be killed, which is 'knowingly approved'.

The report reveals that there are four categories of targets: tactical (military), underground (tunnels), 'power' (e.g. residential towers, with the aim to generate 'civil pressure' against Hamas) and homes (of 'operatives').

AI is expected to reduce collateral damage, but "Proceed with Caution" is the best strategy for any person or nation state

Any alleged connection with Hamas can be a reason for a strike, which, by being disproportionate and overwhelming, has the intention to scare civilians and build an anti-Hamas reaction.

Moreover, it is reported that IDF is committed to IHL, attacking only military targets. Whilst there are claims that warnings have been given to evacuate buildings, there are counterclaims of no warnings.

Further, many homes have been attacked that have no Hamas association. This is aside from humans being misinterpreted as targets resulting in civilian deaths.

Nevertheless, it is argued that Hamas operates from within the human shield of the civilian population, explaining the collateral damage. Concerned about the use of "The Gospel", the Association for Civil Rights in Israel (ACRI) has filed a Freedom of Information request regarding the manner of its use.

Where does this leave us?

IHL governs the use of AI and, by default, AWS. The UK recommendation is to 'Proceed with Caution' under an ethical banner. The US military has established Directives regulating its deployment, but these do not apply to other US government organisations. The Russia-Ukraine war reveals how an AI grounded war might unfold.

However, given that AI underpins how Israel is attacking Hamas, then there are serious concerns about the importance of 'meaningful human control' and what this entails.

AI is expected to reduce collateral damage, but despite its use in Gaza, Gaza has been described by the UN Secretary General António Guterres as "an entirely man-made disaster" for civilians who are "enduring horrifying levels of hunger and suffering" in a

"mass assassination factory".

Perhaps "Proceed with Caution" is the best strategy for any person or nation state.

Dr. Stephen Harwood has had many lives including that of geophysicist, writer, educator, change agent, entrepreneur, and director. He founded TechnoForeSight to promote and advise upon developments in the future of technology, work, and education. Dr. Stephen Harwood used to teach at the University of Edinburgh Business School, from where he was awarded his Doctorate. He also writes, with publications spanning a range of topics including technology futures, ERP implementation, maker spaces, and sustainability. His work is grounded in the theoretical domains of cybernetics and systems thinking