AI ETHICS ADVISORY PANEL

Friday, 15ᵗʰ September 2023, 10:30 - 12:30 hrs (GMT)

MINUTE

Attendees:

- **Paul Lincoln**, MOD's 2ⁿᵈ Permanent Secretary (2PUS)

- **Professor Nick Colosimo,** Global Engineering Fellow & Technologist, BAE systems and Visiting Professor, Cranfield University (Centre for Autonomous & Cyber-Physical Systems)

- **Dr Darrell Jaya-Ratnam,** Managing Director, DIEM Analytics.

- **Richard Moyes,** Managing Director and co-founder, Article 36.

- **Professor Mariarosaria Taddeo,** Professor of Digital Ethics and Defence Technologies, Oxford Internet Institute, University of Oxford; Dstl Ethics Fellow, Alan Turing Institute

- **Professor David Whetham,** Professor of Ethics and the Military Profession, Kings College London

- **Professor Gopal Ramchurn,** Director, UKRI Trustworthy Autonomous Systems Hub and the University of Southampton

- **Professor Peter Lee,** Professor of Applied Ethics, University of Portsmouth.

- **Gareth Martin,** SPO National Security Director (attending in place of Paul Wyatt)

- **Cdre Rachel Singleton**, Head of Defence AI Centre (DAIC)

- **Felicity Burch,** Director for Centre for Data Ethics and Innovation (CDEI)

- **Dr Chris Moore-Bick**, Head of DST Policy and the Defence AI and Autonomy Unit (DAU)

- MOD Legal Adviser

- DAU Policy Advisers

- CDEI Policy Advisers

- DAIC representatives

- EC2 Spearhead representatives

1	**Welcome and Introduction from the Chair**
	MOD's 2nd Permanent Secretary (2PUS) welcomed members of the panel, making the following points:
	- Recognising that both the MOD and some panel members gave evidence to the House of Lords Special Inquiry Committee on AI in Weapon Systems, it was timely to hold this sixth meeting as interest in the panel's work is increasing. - The MOD wants to increase transparency about the work of the panel by proactively publishing minutes of its meetings, alongside considering an annual transparency report. - The Refreshed Defence Command Paper (DCPR) published in July this year makes a prominent mention of the Defence AI Strategy and confirms the importance we attach to the Strategy and the accompanying Ambitious, Safe and Responsible (ASR) policy.
2	**General Updates**
	Dr Chris Moore-Bick presented an update on the latest developments:
	As the department continues to implement the Defence AI Strategy, the DAU has:
	- Set up an AI Delivery Group, chaired by 2PUS, to drive AI Strategy implementation. - Iterated items brought to the panel during the previous meeting based on the panel's feedback. - Contributed to the NATO Data and AI Review Board (DARB) which was set up earlier this year as a key part of operationalising NATO's Principles of Responsible Use (of AI) and its AI and Data Strategy. - Engaged with other governmental departments to support the objectives of the AI Safety Summit. - Engaged with the House of Lords Special Inquiry Committee into AI in Weapon Systems, including through the evidence session with the Minister for Defence Procurement, 2PUS and the Deputy Commander Strategic Command.
3	**Presentation and Discussion on AI Ethics Implementation Use Cases**
	2PUS highlighted ongoing work to implement MOD's ASR policy. The DAU and CDEI briefed Members on workshops with three AI project teams to understand how the ASR principles were being operationalised and to identify good practice. CDEI presented their findings, highlighting the following points:
	- Case studies covered areas on Intelligence, Surveillance and Reconnaissance, the application of AI, ML and automation technologies for improving Command-and-Control and an exploratory project on predicting future redeployment of service personnel. All teams were clear that their use of AI was to enable operators to make better decisions rather than replace human decision-making. - All three teams took AI ethical risks very seriously. Some AI ethics principles were already well covered due to extant MOD structures and processes, whilst others needed specific attention. Teams felt the workshops were a valuable prompt to generate discussion about the principles. - CDEI's main recommendations were that Defence should develop further guidance on when and how to consider ethical issues in the development lifecycle. This might include

common tools, such as an AI Assurance Framework, Model/Dataset cards, and guidance on AI risk management and building the ASR principles into supplier relationships.

Panel members were broadly supportive of MOD's approach and felt that real use case studies provided a helpful means to understand and advise on the department's implementation of the AI Ethics principles. The following areas were discussed:

- Workshops were a helpful tool to promote engagement with the ASR principles, guard against 'ethics washing' and embed the right culture, something which prescriptive checklists on their own might not do.
- Risk management continuity must be maintained throughout the system lifecycle and whenever a project changes ownership (e.g. the handoff from development to operational use). AI Ethics Assurance must extend into industry.
- Human centric risks (e.g. abuse, misuse and disuse of AI-enabled systems) must be considered, along with wider risks such as data poisoning. Recurring themes must be tracked to understand systemic challenges.
- Teams may need support to understand potential trade-offs between AI ethics principles and how the principles will need to be balanced with military necessity.

The 2nd Permanent Secretary, CDEI and DAU thanked the panel for their feedback and noted that the department will focus on embedding ASR principles throughout the AI lifecycle, considering both specific risks and the 'system of systems' implications for human-machine teams.

4	**Any Other Business**
	International: Members noted that a number of nations and stakeholders were asking the UN Secretary General to seek views from states on the potential regulation of Lethal Autonomous Weapons Systems (LAWS). The MOD highlighted the UK's active participation in the UN Group of Government Experts on LAWS under the Convention for Certain Conventional Weapons.
	Transparency: It was agreed that the MOD will publish meetings of forthcoming EAP meetings and work towards a transparency report to capture the breadth of panel discussions.
	LLM Hackathon: Panel members were invited to take part in an upcoming LLM (Large Language Models) Hackathon, organised by the DAIC, that would encourage teams to build prototypes that address key Defence challenges.
	AI Practitioner's Handbook Website: A new website was in development that MOD hoped could be used as a one-stop-shop for publishing MOD guidance and policy on AI Ethics and wider training on AI for internal (and, where appropriate, external) audiences. This would help to improve transparency and understanding about the MOD's approach to safe and responsible AI. Panel members were invited to review the website as it develops to provide insight and challenge.