



## AI ETHICS ADVISORY PANEL

Minute of the meeting held on Friday 26<sup>th</sup> March 2021

### Attendees

**Dominic Wilson**

MOD Director General for Security Policy (*Acting Chair*)

**Professor Nick Colosimo**

Executive Manager Future Capabilities, BAE systems

**Kata Escott**

MOD Director of Strategy

**Professor Peter Lee**

Professor of applied ethics, University of Portsmouth

**Richard Moyes**

Managing Director and co-founder, Article 36

**Professor Mariarosaria Taddeo**

Deputy Director of the Oxford Digital Ethics Lab

**Professor Dapo Akande**

Director of the Oxford Institute for Ethics, Law and Armed Conflict

**Dr Merel Ekelhof,**

Foreign Exchange Officer at the US Joint AI Center

**Dr Darrell Jaya-Ratnam**

Managing Director, DIEM Analytics

**Professor Dame Angela McLean**

MOD Chief Scientific Advisor (CSA)

**Professor Gopal Ramchurn**

Director at the Trusted Autonomous Systems Hub, University of Southampton

### Apologies

**Lt Gen Doug Chalmers**

Deputy Chief of Defence Staff for Military Strategy and Operations (DCDSMSO)

**Professor David Whetham**

Professor of Ethics and the Military Profession, Kings College London

**Tabitha Goldstaub**

Founder of CognitionX and chair of the AI Council

### Also attending

Hd DST Policy: Dr Chris Moore-Bick

CDEI: [REDACTED]

CDEI: Sam Cannicott

CIO-Data-COE-Analytics Asst Hd: [REDACTED]

DST-Policy-EmTechPol-AHd: [REDACTED]

DST-DAU-StratPol2: [REDACTED]

CDEI: [REDACTED]

CIO-AI-COE-Engage-Con1: [REDACTED]

DST-SDIP-3: [REDACTED]

### Introduction

1. The Chair welcomed members of the panel, introduced the participants, and provided an update on progress made on AI and wider Technology policy since the previous meeting. The Chair made the following points:
  - The UK Government recently published the Integrated Review of Security, Defence, Development and Foreign Policy. It was accompanied by a significant investment in

Defence, and in particular the promise of an AI Centre to accelerate adoption of these technologies.

- Among the Review's objectives are to sustain strategic advantage through Science and Technology, and to shape the open international order of the future; showing the importance of discussing AI Ethics at the panel, and ensuring our approach to AI reflects our values.

### **Policy Review: Updated Ethical Principles**

2. [REDACTED] from the DAU introduced an updated set of ethical principles based on the panel's previous feedback. In response, the panel discussed the following:

- Whether the principles document should contain a definition for AI, or whether this will be covered by the wider document within which the principles will be published.
- Whether references should be made in the principles to national laws as well as international laws of armed conflict.
- Whether the principles should rearticulate its references to Military Necessity, a concept key to IHL that could cause conflation between sets of principles.
- Whether references to Human Control in the second principle need further elaboration to set out the benefits of such control.
- Whether the specific focus on the unique respect given to humans within the Humanity principle should be lessened so that it can accommodate wider issues for humanity such as environmental damage, or non-kinetic battlefield effects.
- Whether types of AI might exist for which an ethical obligation to develop, if it can be demonstrated that ethical outcomes would result from its use.

[REDACTED]

### **Policy Review: Use Cases and Principles Implementation**

3. [REDACTED] and [REDACTED] of CDEI introduced a paper setting out the main ethical risks to be mitigated by the ethical principles, alongside how each principle could be applied to a set of use cases.

4. In response, the panel discussed:

- Whether the use case examples set out would be reasonable failure points, or whether they cover issues that would be managed by the existing set of regulations and rules governing Defence procurement and R&D.
- Whether the principles could inform a set of checklists that could be used to regulate procurement and industry collaboration with the MOD.
- Whether the risks are too focused on the operational side of capabilities, as opposed to the design or commissioning phase.
- Whether the human error element is well enough represented in the risks of AI systems, or whether risks of bad outcomes resulting from specific dangers of AI should be made more central to the paper.
- Whether the paper should elaborate the set of use cases to cover back office operations as well as operational tools.

**Open Discussion: Freedom of action and adopting the principles**

5. The Chair proposed a new agenda item based on earlier discussion around potential constraints for Defence introduced by adopting the principles. The panel discussed:
  - Whether the result of these principles is systems with scalable autonomy, which can act at different levels of autonomy relating to different contexts.  
[REDACTED]
  - Whether the MOD should consider the relationship between ethics and regulation, and the difference between “hard ethics” which underpins law, and “soft ethics” which underpins norms and decision making.

**Summary**

6. The Chair closed the meeting, thanking the participants and stated:
  - That the Secretariat will consider the panel’s comments and provide another updated draft of the principles with panellist’s feedback integrated.
  - That the MOD intends to publish its principles as part of the AI Strategy in the coming months, and may informally consult panel members ahead of this publication.
7. The next session of the Ethics Advisory Panel will be on the 5<sup>th</sup> July 2021.

**Action**

- MOD to consider new feedback on the Ethical Principles, and to further iterate a new set of principles for consideration by the panel.