AI ETHICS ADVISORY PANEL

Record of Decisions (RODs)
Meeting of Tuesday 5th July 2022

<u>Attendees</u>

Laurence Lee (*Chair*)
Second Permanent Secretary for Defence

Professor Nick Colosimo
Executive Manager Future Capabilities, BAe Systems

Dr Merel Ekelhof
Exchange Officer at the US Joint AI Center (acting in personal capacity)

Michael Gibson
Policy Deputy Head, Strategy & Policy, Defence Science & Technology (DST)

Dr Darrell Jaya-Ratnam
Managing Director, DIEM Analytics

Professor Nick Jennings
Chair of the AI Council. Vice-Chancellor and President of Loughborough University.

Professor Peter Lee
Professor of Applied Ethics, University of Portsmouth

Professor Dame Angela McLean
MOD Chief Scientific Advisor (CSA)

Dr Chris Moore-Bick
Head, DST Policy
MOD

Richard Moyes
Managing Director and co-founder, Article 36

Air Vice Marshal Al Smith
Assistant Chief of Defence Staff (ACDS), Military Strategy, Defence Engagement, MOD

Professor Mariarosaria Taddeo
Deputy Director of the Oxford Digital Ethics Laboratory

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

Paul Wyatt
Director General Security Policy, MOD

<u>Apologies</u>

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

Sam Cannicott,
Deputy Director, Centre for Digital Ethics and Innovation (CDEI)

Air Marshal Ian Gale,
Director General, Joint Force Development (JFD), MOD,
.

Tabitha Goldstaub
Founder of Cognition X and chair of the AI Council

Professor Gopal Ramchurn
Director at the Trusted Autonomous

Polly Scully
Strategy Director, MOD

Systems Hub, University of Southampton

Lt Gen Roly Walker
Deputy Chief of Defence Staff for Military
Strategy and Operations, MOD

<u>Also attending</u>

███████████████	███████████
Ethics Policy Lead, AI and Autonomy, MOD	Secondee, Centre for the Study of Existential Risk (CSER), University of Cambridge

<u>Reflections on the launch of the Defence AI Strategy and the 'Ambitious, Safe, Responsible' policy</u>

1. The launch was well received by the Defence and Technology sectors and signifies an important milestone on the journey, which will now focus on implementation.   Defence is keen to build on the momentum from launch and widen engagement with industrial and academic sectors.

2. The launch was well received among Defence suppliers but not widely known outside Defence circles. Members made comments about supporting investment, how the Ethics Principles will be translated into practice, interoperability and managing divergences with partners, and how Defence will lead on norm forming. They also discussed the subtle change of focus to 'context-appropriate human involvement', noting that 'bounded autonomy' would be a key mitigation to help manage concerns around adaptive systems

3. The issue of Lethal Autonomous Weapons (LAWS) was being reviewed by the UN Group of Government Experts on the Convention on Conventional Weapons (CCW). The UK needs to lead in shaping the global landscape on accepted norms, but one panel member felt that its position was cautious and thus not world leading. It was suggested that the UK needs to build confidence in suggesting solutions. Richard Moyes commented that norm setting would provide more certainty for industry and the UK could gain geopolitical advantage from leading on norms/standards. Saying but not doing makes us vulnerable to processes developed by other states.

4. AVM Al Smith said that Defence was conscious of the need to ensure interoperability, though subtle divergencies will exist to reflect national requirements within alliances and partnerships. Five Eyes and NATO have agreed ethics guidance, that sets baseline standards.

<u>Implementing AI Ethics Principles across Defence</u>

███

██

███

██
██████████████████████████████
██
██████████████████████████

6. ██ in the context of organisational cultures and operationalisation of the ethics principles. Assigning responsibilities throughout system lifecycles will be challenging, and the need for ethics training and education will be crucial to embed desired behaviours across Defence. Use cases would be subject to wider partner discussion, including developers, manufacturers and operators. Different projects and programmes would be reviewed at varied stages of the system life cycle.

7. The Armed Forces need to set the right policy, permissions and constraints frameworks that comply with our legal and ethical obligations but also do not impede our ability to fight effectively. Norms will evolve as result of action and reaction and this may pose some risks, with respect to escalation.

8. Pete Lee pointed to creating legal and ethical checklists for practitioners; noting the EPSRC work on responsible research and innovation offered tools that could be leveraged. Darrell Jaya-Ratnam observed that policy was playing catch up to practice.

9. It was necessary to understand what has been tried in the past (success or failure) and have a compendium of precedent. It must also reflect on what users consider most material, i.e., autonomy closest to effect has the greatest impact. Nick Colosimo suggested a matrix setting lifecycle against DLODs to provide a 'heatmap of compliance'. In response, Mariarosaria Taddeo thought it was important not to be too prescriptive, rather providing a methodology with guidance on trade-offs and options. David Whetham argued that principles couldn't be seen as obstacles

10. Members discussed lessons from Article 36 compliance processes, benefits of multi-disciplinary teams and working with projects from the conceptual stage. But they also noted that incremental adaptation may not work as there are key differences for learning systems. Whereas Article 36 is about the impact on the human body, AI reviews would be more around discrimination, algorithmic bias etc.

11. Each operator needs to know or understand to use a system ethically, balancing risk with operational necessities. Rules/guidance on red lines are needed, but also understanding what they can trade in/out to meet their requirement.

ACTION POINT1: DAU to provide EAP members for the next meeting, a paper as to its plans and progress as to ethics implementation.

---

12. Every operator prior to deployment needs to be aware of appropriate use. But it was important to beware of false comfort derived from adapting norms.

Future role of the Panel (15 minutes)

████The panel discussed the future role of the panel following the publication of Ambitious, Safe, Responsible and the AI Strategy and, whether the ██████████████████████
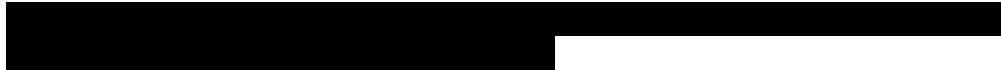
13. The consensus was broadly that it is more appropriate for the panel to continue as a community of expertise, as this safeguards the independence of the panel and its members and ensures, flexibility and open discussion. The existing model fits comfortably with evolving activity and the focus should remain on exploring questions to addressing immediate needs. Furthermore, if the panel were to adapt more formal policy making powers, then members would have to declare an interest in companies for which they were associated.

14. There was also discussion of the ethical connections with other technologies such as human argumentation technology, which may require similar structures of governance and accountability envisaged for AI.

ACTION POINT 2: DAU to provide for the next EAP meeting, possible options for its future role.

██████████████
Secretariat for the AI Ethics Advisory Panel
Defence Autonomy Unit