07490 206 393
hello@kairoi.uk
www.kairoi.uk

Kairoi Ltd
71-75 Shelton Street
Covent Garden
London
WC2H 9JQ

Office for Artificial Intelligence
Department for Science, Innovation and Technology
100 Parliament Street
London
SW1A 2BQ

21st June 2023

RE: Consultation on policy proposals for UK's pro-innovation approach to AI Regulation

Kairoi Ltd ("we" hereafter) warmly welcomes the government's continued efforts to engage with the public, and the invitation to provide feedback on the *AI regulation white paper*[1] ("the white paper" hereafter). At Kairoi, we consult on AI ethics and research governance, and engaging with public policy on technological matters constitutes one of our four pillars, alongside communicating accurately about artificial intelligence (AI), identifying relevant technical standards, and involving the public.

The feedback focuses on three areas:

1. Ensuring certainty for UK innovators
2. Increasing cooperation with the broader AI ecosystem
3. Expanding the remit of education in AI

Short responses to the consultation questions 1-18 are provided in *§Appendix*.

Mitigating uncertainty

The principles introduced in the white paper reflect those of the OECD,[2] a welcome touchstone. The provision of definitions and rationales for each of the five principles is also welcome. The principles are proposed by the white paper to underpin a framework to be implemented by existing regulators.

The overarching path chosen – to delegate AI regulatory powers to existing bodies that best understand their own remits – is not *prima facie* problematic. Notwithstanding, **there**

---

[1] Department for Science, Innovation and Technology and Office for Artificial Intelligence (2023) AI regulation: a pro-innovation approach – policy proposals, *Open consultation*, https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals [accessed 21/06/2023]

[2] OECD (2019) Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO), *OECD Digital Economy Papers*, 291, OECD Publishing, Paris, DOI: 10.1787/d62f618a-en

AI Ethics & Research
Governance Consultancy

Kairoi Ltd, a Company Limited by Shares, Registered in England, No. 14317729
Registered Office: 71-75 Shelton Street, Covent Garden, London, WC2H 9JQ

appears to be a purposeful delay in the regulation of AI through this route. On the one hand, there are numerous stages between the principles and practical measures enacted by organisations. The OECD principles inform the white paper, which feeds into an iteratively developed framework ("the framework" hereafter), which regulators will integrate into future regulations, which will eventually inform industry-specific practices. Whether the principles are held cross-sectorially or not, the development of basic guidelines and, perhaps, laws that affect how AI is developed across sectors should not be delayed much further.

On the other hand, **parting from principles that are intended to be flexible undermines the white paper's interest in ensuring consistency across sectors**. Any one of the principles can be used to make this point, but *fairness* is the most obvious. *Fairness* can mean many things, and may be confused with concepts such as *equity*, *justice* or *equality*. The choice to delegate the application of "fairness" to sector-specific legislation opens the possibility of very different conceptions of the principle appearing in different laws. With this, we suggest developing more detailed definitions of the principles in the proposed framework. These definitions can then set a steady foundation on which legislators can build.

Whilst the white paper seeks to promote innovation across sectors in the UK, we should not shy away from centralising the establishment of a clear baseline for that innovation to be conducted responsibly. Furthermore, the flexible integration of principles into future legislation only increases uncertainty for innovators, for whom it is unclear what the principles mean and how they affect them. Innovation may be stunted without further clarification.

The uncertainty innovators face with regard to AI regulations in the UK also relate with the white paper's suggested definition of AI, which involves "adaptivity" and "autonomy." It is unclear whether this definition is sufficiently broad to capture even extant AI systems. For example, tools using "symbolic AI" do not clearly fall within the description of "adaptivity."

The case studies presented in *§3.1.* help illustrate what "adaptivity" and "autonomy" mean. However, "autonomy" seems too broadly applicable, with any tool that demonstrates some process between a human's input and its output seemingly falling into scope. Rather than "adaptivity" and "autonomy," what seems to be at the heart of the proposed definition is the lack of "explainability" often attributed to AI systems, as well as the possibility of such systems conducting tasks with little human oversight – which is a well-established aspect of AI.[3] To this effect, clearer rationale behind the definition ought to be provided.

Expanding cooperation

It is worth noting that international efforts to define *artificial intelligence* are ongoing, with at least the EU on its part[4] and an EU-US collaboration[5] working on AI-related definitions. With this, **we recommend working more closely with regulators from across the globe to ensure not only interoperable definitions and rules, but to also facilitate international**

---

[3] McCarthy et al. (1955) A proposal for the Dartmouth summer research project on artificial intelligence, http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html [accessed 21/06/2023]
[4] Bertuzzi (2023) EU lawmakers set to settle on OECD definition for Artificial Intelligence, *Euractiv*, https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/ [accessed 21/06/2023]
[5] (2023) EU-U.S. Terminology and Taxonomy for Artificial Intelligence, https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence [accessed 21/06/2023]

expansion of UK-based innovators. Whilst the white paper has a section on global interoberability and international engagement (*§6.1.*), these components do not reflect in the white paper itself. The white paper must be more concrete with regards to developing interoperable AI regulation..

There has also been a recent shift from collaborating with organisations like those outlined in *§6* (e.g.: UNESCO, G7, Council of Europe Committee on AI…) to discussing AI with tech CEOs such as Sam Altman (OpenAI) and Demis Hassabis (Google DeepMind).[6] **It is imperative that the government not fall for the trap that regulating cutting-edge technologies – or designing "guard rails" – necessitates oversight from the organisations being regulated, which is cause for suspicion.**[7]

The recent U-turn from Sunak – from the white paper's light-touch approach to AI regulation to a more robust and collaborative approach[8] – is welcome, but the appointment of a tech entrepreneur who speaks of "God-like AI"[9] to lead the national "AI Foundation Model Taskforce"[10] is a worrying sign.

It is worth celebrating the excellent work already conducted within the UK on best practices in AI and the alignment with the white paper's context-based approach. Consider the Ada Lovelace Institute's focussed studies on recommendation systems in public services media[11] and algorithmic impact assessments in healthcare,[12] to mention just two. The Ada Lovelace Institute is a prime example of a hub of expertise in the space of responsible AI. Involving organisations such as theirs in the development of "guard rails" will enhance the credibility of regulatory efforts.

A further caution relates to the focus of technical standards, which are but one component of the AI assurance ecosystem. The white paper makes mention of the AI Standards Hub, which Kairoi advised on in April this year. However, this is the only area discussed of what the white paper calls "AI assurance techniques." As listed earlier on, at Kairoi, we also value accurate communications, robust governance and public engagement as key to ensuring best practices in AI. If the framework for future AI regulation in the UK views only technical standards as providing AI assurance, there is a risk of limiting its scope to only a few actors in the AI ecosystem.

---

[6] Webber (2023) UK to host major AI summit of 'like-minded' countries, *Politico*, https://www.politico.eu/article/u-k-to-host-major-ai-summit-of-like-minded-countries/ [accessed 21/06/2023]
[7] Perrigo (2023) Exclusive: OpenAI lobbied the EU to water down AI regulation, *TIME*, https://time.com/6288245/openai-eu-lobbying-ai-act/ [accessed 21/06/2023]
[8] Rajvanshi (2023) Rishi Sunak wants the UK to be a key player in global AI regulation, https://time.com/6287253/uk-rishi-sunak-ai-regulation/ [accessed 21/06/2023]
[9] Hogarth (2023) We must slow down the race to God-like AI, *Financial Times*, https://www.ft.com/content/03895dc4-a3b7-481e-95cc-336a524f2ac2 [accessed 21/06/2023]
[10] DSIT (2023) Tech entrepreneur Ian Hogarth to lead UK's AI Foundation Model Taskforce, https://www.gov.uk/government/news/tech-entrepreneur-ian-hogarth-to-lead-uks-ai-foundation-model-taskforce [accessed 21/06/2023]
[11] Jones et al. (2023) Inform, educate, entertain… and recommend? *Ada Lovelace Institute*, https://www.adalovelaceinstitute.org/report/inform-educate-entertain-recommend/ [accessed 21/06/2023]
[12] Groves (2023) Algorithmic impact assessment in healthcare, *Ada Lovelace Institute*, https://www.adalovelaceinstitute.org/project/algorithmic-impact-assessment-healthcare/ [accessed 21/06/2023]

As others do,[13] we have identified numerous actors from across the AI ecosystem who play a role throughout the AI lifecycle, from funders, universities and publishers,[14] to administrators and end-users.[15] Given the tremendous task at hand – the legislation of a technology that almost everybody uses – we must not burden technologists with all the responsibility of developing AI systems safely. Rather, the framework for future regulation must identify the roles of diverse actors to ensure that UK innovation in AI is responsible and truly cutting-edge.

Improving education

The white paper suggests the need for "providing education and awareness to give clarity to businesses and empower citizens to make their voices heard as part of the ongoing iteration of the framework." This is welcomed by Kairoi and acknowledges the broader AI ecosystem mentioned above. AI literacy is a naturally emergent field, and providing businesses and the general public with the tools to meaningfully engage with the framework's future iterations will be key to its success. But limiting the scope of "education and awareness" to this regulatory process will miss the mark.

Advancements in AI have wide-spread societal consequences, from perpetuating their developers' assumptions to generating and spreading misinformation. **As well as informing the general public to best engage with the development of the framework, education must provide the skills to critically evaluate novel technologies.** This is a step change from the vision for "Maths to 18"[16] insofar that we value critical thinking skills to be as important as technical know-how for innovation.

Critical thinking skills developed through sociology, music, philosophy, literature and other "non-STEM" fields of inquiry can support the current and future workforce to ascertain the usefulness of AI tools. Through a reflexive population, innovations that succeed in the UK will stand to public scrutiny when exporting abroad.

As explained earlier, there is a broad range of actors in the AI ecosystem, as well as a wide range of activities for businesses to consider when developing AI systems responsibly. The AI assurance market – beyond establishing technical standards – can be informed by governance, risk and compliance consultancies that inform diverse parties on best practices throughout the AI lifecycle.[17] At Kairoi, for example, we have openly developed a series of resources for organisations to use generative AI tools safely, and to recruit technologists and other staff who are aware of the potential of advancements in AI.[18]

---

[13] National Institute of Standards and Technology (2023) AI Risk Management Framework, *Trustworthy & Responsible AI Resource Center*, https://airc.nist.gov/AI_RMF_Knowledge_Base/AI_RMF [accessed 21/06/2023]

[14] Petermann et al. (2022) Looking before we leap, *Ada Lovelace Institute*, https://www.adalovelaceinstitute.org/report/looking-before-we-leap/ [accessed 21/06/2023]

[15] Kherroubi Garcia (2023) Another piece of the AI ethics puzzle, *Kairoi*, https://kairoi.uk/blog/another-piece-of-the-ai-ethics-puzzle/ [accessed 21/06/2023]

[16] Prime Minister's Office (2023) Prime Minister outlines his vision for Maths to 18, https://www.gov.uk/government/news/prime-minister-outlines-his-vision-for-maths-to-18 [accessed 21/06/2023]

[17] Raghunathan (2023) The Ethical AI Startup Ecosystem 05: Governance, Risk, and Compliance (GRC), *Montréal AI Ethics Institute*, https://montrealethics.ai/the-ethical-ai-startup-ecosystem-05-governance-risk-and-compliance-grc/ [accessed 21/06/2023]

[18] Kairoi Ltd (2023) Resources, *GitHub repository*, https://github.com/KairoiAI/Resources [accessed 21/06/2023]

Concluding

There are still many questions with regard to how the white paper will inform a framework that will support legislators that will inform practice. There are also questions about the impact of the white paper on the international standing of UK innovations, with other regions' legislative efforts not getting very much attention.

Notwithstanding, the world of AI is moving fast, and a great deal has changed since the white paper was published. This simply emphasises the need to collaborate closely with diverse stakeholders in future iterations of the framework. Once again, we appreciate the invitation to contribute through the open consultation. The rapid pace of activity surrounding AI also calls for expanding the understanding of the AI assurance market to include non-technical experts, and governance, risk and compliance consultancies. Such an expansion also ensures that regulation is not developed to meet the requirements of those being regulated, but to hold them accountable and to ensure that UK innovation is the best for the country and the world.

Sincerely,

Ismael Kherroubi Garcia, FRSA

Founder & CEO, Kairoi

Appendix

1. **Do you agree that requiring organisations to make it clear when they are using AI would improve transparency?**
   Yes

2. **Are there other measures we could require of organisations to improve AI transparency?**
   Yes, require justifying why less transparent technical solutions are pursued and, where advisable, integrate symbolic AI components to enable greater transparency.

3. **Do you agree that current routes to contest or get redress for AI-related harms are adequate?**
   It is unclear.

4. **How could current routes to contest or seek redress for AI-related harms be improved, if at all?**
   The ICO must continue [their efforts](#) and ensure alignment with the AI Standards Hub.

5. **Do you agree that, when implemented effectively, the revised cross-sectoral principles will cover the risks posed by AI technologies?**
   If implemented effectively by legislators through collaboration with experts, academia, industry and charities, yes.

6. **What, if anything, is missing from the revised principles?**
   Clarity.

7. **Do you agree that introducing a statutory duty on regulators to have due regard to the principles would clarify and strengthen regulators' mandates to implement our principles while retaining a flexible approach to implementation?**
   No, it may undermine the proposed definitions as regulators look to how other regulatory bodies interpreted the principles.

8. **Is there an alternative statutory intervention that would be more effective?**
   An independent body must be instituted to enable consistency and keep the framework up-to-date to account for the latest laws and practices.

9. **Do you agree that the functions outlined in Box 3.1 would benefit our AI regulation framework if delivered centrally?**
   Yes.

10. **What, if anything, is missing from the central functions?**
    Broader educational efforts.

11. **Do you know of any existing organisations who should deliver one or more of our proposed central functions?**
    It is unclear.

12. Are there additional activities that would help businesses confidently innovate and use AI technologies?

Co-development of internal policies.

1. If so, should these activities be delivered by government, regulators or a different organisation?

It should be a cooperative effort involving regulators, businesses and expert consultancies like Kairoi.

13. Are there additional activities that would help individuals and consumers confidently use AI technologies?

Broader educational efforts.

1. If so, should these activities be delivered by government, regulators or a different organisation?

These should be independently designed and delivered by charities like We and AI or the Royal Society of Arts, Manufacturing and Commerce.

14. How can we avoid overlapping, duplicative or contradictory guidance on AI issued by different regulators?

A central function must ensure this does not happen – the regulators must not work in silos.

15. Do you agree with our overall approach to monitoring and evaluation?

In principle, yes.

16. What is the best way to measure the impact of our framework?

Through a seamless portal for communicating between the diverse parties.

17. Do you agree that our approach strikes the right balance between supporting AI innovation; addressing known, prioritised risks; and future-proofing the AI regulation framework?

It is unclear.

18. How can we avoid overlapping, duplicative or contradictory guidance on AI issued by different regulators?

It is unclear.

19. As a regulator, what support would you need in order to apply the principles in a proportionate and pro-innovation way?

Technical expertise.

20. Do you agree that a pooled team of AI experts would be the most effective way to address capability gaps and help regulators apply the principles?

Yes, but the experts must be from diverse fields of AI, including non-STEM experts.

21. Which non-regulatory tools for trustworthy AI would most help organisations to embed the AI regulation principles into existing business processes?

Consultancy services and guidelines.

22. Do you have any other thoughts on our overall approach? Please include any missed opportunities, flaws, and gaps in our framework.

See above letter.