

Written Evidence Submitted by Google (GAI0099)

We have long since progressed beyond an era when advances in AI research were confined to the lab. AI has now become a real-world application technology and part of the fabric of modern life. Harnessed appropriately, Google believes AI can deliver great benefits for economies and society, and support decision-making which is fairer, safer and more inclusive and informed. But such promise will not be realised without great care and effort, which includes consideration of how its development and usage should be governed, and what degree of legal and ethical oversight — by whom, and when — is needed. To date, self- and co-regulatory approaches informed by current laws and perspectives from companies, academia, and associated technical bodies have been largely successful at curbing inopportune AI use. We believe in the vast majority of instances such approaches will continue to suffice, within the constraints provided by existing governance mechanisms (for example sector-specific regulatory bodies). However, this does not mean that there is no need for action by government. To the contrary, we believe governments and civil society groups worldwide have an integral role in the AI governance discussion. We therefore welcome the opportunity to respond to the committee's call for evidence on this important issue.

1. How effective is current governance of AI in the UK? What are the current strengths and weaknesses of current arrangements, including for research?

At Google we support a proportionate, risk-based approach to AI regulation. It is far too important not to regulate and too important not to regulate well. Google has long recognized the fact that AI carries risks, alongside immense benefits - it's why we established our [AI Principles](#) in 2018 and created a central Responsible Innovation team, to operationalize implementation.

Google takes a risk-based approach to AI governance as there can be no sensible "one size fits all" approach to AI governance. AI is a multi-purpose technology, which takes many forms and fulfils many purposes, spanning a wide range of risk profiles. Our internal AI Principles ecosystem supports employees in incorporating responsible practices into their work. At the core of this internal ecosystem is a three-tiered governance structure. It starts with our product teams themselves, which include dedicated user experience, privacy, and trust and safety experts who provide deep functional expertise consistent with the AI Principles. The second tier is a set of dedicated review bodies and expert teams. Finally, the central Responsible Innovation team is available to support implementation across the company, and all Google employees are encouraged to engage with the AI Principles review process throughout the project development lifecycle. Some product areas have also set up review bodies to address specific audiences and needs, such as enterprise offerings in Google Cloud, hardware in Devices and Services, and medical expertise in Health.

In the UK, there are already many existing regulations that apply to AI in different sectors. Many of these sectoral regulations are broad enough to apply to AI throughout the lifecycle, and establish judicial processes for resolving disputes. For instance, AI applications relating to healthcare fall within the remit of medical and health regulators, and are bound by existing rules associated with medical devices, research ethics, and the like. When integrated into physical products or services, AI systems are covered by existing rules associated with product liability and negligence. This can be found in

financial services where, for instance, AI applied in the insurance industry is subject to GDPR regulations, discrimination rules, and model risk management requirements.

However, there is room for further clarity in the UK's current legal frameworks to address the application of regulatory bodies to AI and regulatory overlaps. These can be addressed by additional guidance, and we welcome the government's recent proposals in their policy paper to put the cross-sector principles on a non-statutory footing and align with the government's UK Digital Strategy - that 'traditional' regulation can be complemented by non-regulatory tools. A self-regulatory or co-regulatory set of governance norms that could be applied flexibly and adaptively would enable policy safeguards while preserving the space for continued beneficial innovation.

This does not necessarily mean replacing existing laws, or developing new laws that are not tech-neutral. These norms would serve as a much needed guide to create cohesion. Further guidelines on best practices, hypothetical use cases and establishment of baseline responsible practices and standards in different industries could potentially address the risk of regulatory overlaps and contradictions. We also welcome the coordination of the Digital Regulation Cooperation Forum and their workstream on algorithmic processing, which seeks to understand the impact of algorithms across industry and regulatory remits.

2. What measures could make the use of AI more transparent and explainable to the public?

There are straightforward and effective steps users of AI can introduce to improve transparency and explainability. We have [previously called for](#) five areas where the government can collaborate with wider society and AI practitioners to clarify expectations about AI's applications on a context-specific basis, including fairness appraisal, safety considerations, human-AI collaboration, liability frameworks and explainability standards.

Transparency in AI should be a common good. Despite its potential to transform so much of the way we work and live, machine learning models are often distributed without a clear understanding of how they function. That is why we pioneered and continue to refine [Model Cards](#), which are documentation and transparency artifacts, and may vary by use case and risk level. As organisations and regulators consider how to approach these artifacts, it is important to ensure they are dynamic and may be updated over time as the technology matures and data inputs are adjusted.

For example, if a company were to develop a dog breed classifier, its model card might specify that the model is based on a convolutional neural network and that it outputs bounding boxes along with the breed's label. It could then build on this basic foundation with insight into the factors that help ensure optimal performance. What kind of photos tend to yield the most accurate results? Can it handle partially obscured dogs? What about dogs that are extremely close, extremely far away, or seen from unusual angles? Even simple guidelines like these can make a difference, helping leverage a model's capabilities and steer clear of its limitations.

We believe increased transparency for machine learning models can benefit everyone, which is why model cards are aimed at experts and non-experts alike. Developers can use them to design applications that emphasise a model's strengths while avoiding or informing end users of its weaknesses. For journalists and industry analysts, they might provide insights that make it easier to explain complex technology to a general audience. And they might even help advocacy groups better understand the impact of AI on their communities.

This is an effective practice that we expect will underpin future industry standards - used to promote necessary communication across stakeholders including users, developers, civil society stakeholders, and companies across the industry.

In addition to our Model Cards initiative, governments could help industry by providing examples and guidance for minimum standards of explainability across different AI application contexts. Standards and rules around the explainability of algorithms should differ significantly depending on context. Appropriate standards of explanation should not exceed what is reasonably necessary and technically feasible. As an analogy, society does not expect an airline to explain to passengers why a plane is taking a particular algorithmically determined flight path — a similarly pragmatic and context-specific approach should apply to explanations for AI. We recognise that it is not realistic for governments and civil society to provide guidelines in every instance. However, we believe governments, supported by other stakeholders, could assist by:

- Assembling a collection of best practice explanations along with commentary on their praiseworthy characteristics to provide practical inspiration.
- Providing guidelines for hypothetical use cases so industry can calibrate how to balance the benefits of using complex AI systems against the practical constraints that different standards of explainability impose.
- Describing minimum acceptable standards in different industry sectors and application contexts.

3. How should decisions involving AI be reviewed and scrutinised in both public and private sectors? Are current options for challenging the use of AI adequate and, if not, how can they be improved?

As highlighted, explainability and transparency are key to ensuring public scrutiny of decisions involving AI. In addition, in order for AI to be effectively reviewed and scrutinised, we must carefully consider how humans should be integrated into the AI process, and how AI systems are accountable to human beings. This goes beyond liability and establishes principles for how individuals and communities contribute to the responsible design, deployment, oversight and use of AI systems. We should also consider how performance of an AI-driven (or partially-driven) system compares to a human-driven system, rather than looking at AI errors in a vacuum.

Ultimately, AI systems and humans have different strengths and weaknesses. Selecting the most prudent combination comes down to a holistic assessment of how best to ensure that an acceptable decision is made, given the circumstances. However, making this determination is not straightforward. In some contexts, it is possible that a team of human and machine combined will perform better than either does alone. But in other situations it will be less clear-cut. Governments and regulators should work closely with industry and technical experts to understand how humans can provide safe and effective input, feedback, and control of AI systems.

People are central to an AI system's development and likely to remain so. From the beginning stages of problem and goal articulation, through to data collection and curation, and model and product design, people are the engine for the system's creation. Even with advanced AI systems able to design learning architectures or generate new ideas, the choice of which to pursue should still be overseen by human collaborators, not least to ensure choices fall within an organisation's legal and financial constraints. Similarly, people play a vital role in the upfront verification and monitoring of a system,

such as choosing which tests to run, reviewing results, and deciding if the model satisfies the performance criteria so as to enter (or remain) in real-world use. And of course, human users provide essential feedback to improve AI systems over time.

Governments may wish to identify red-line areas where human involvement is deemed imperative - especially for decision making in sensitive domains. For instance, for ethical reasons we would suggest that people should always be meaningfully involved in making legal judgments of criminality, or in making certain life-altering decisions about medical treatment. The industry as a whole would benefit from broad guidance as to what human involvement should look like — for example, an evaluation of common approaches to enabling human input and control, with commentary on which are acceptable or optimal, supplemented by hypothetical examples from different contexts.

4. How should the use of AI be regulated, and which body of bodies should provide regulatory oversight?

From our experience, we believe AI regulation should be principles-based and implemented on a non-statutory basis. Additionally, governments should take a multi-sectoral approach, given the application of many existing rules and regulations to AI. In cases where existing rules do not exist, or where there are gaps in governance structures, sectoral experts, academia, practitioners are well placed to identify emerging risks and assist in explaining different contexts. This allows the government to take steps to mitigate, in consultation with civil society and government.

We recommend clearly delineating responsibilities that align horizontally across regulators and detail what is in scope and determining a common lexicon, including definitions. A definition of AI could align with the OECD's definition: 'a machine-based system that can, for a given set of human defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments.' Existing legislation should be looked at first to identify gaps as well as rules that can be harnessed to effectively regulate AI. We believe the best approach is to build a resilient framework that can adapt to emerging technologies and their application. The framework should recognize the different roles that organisations play across the value chain: developers, data providers, integrators, deployers, end users, and many others. We are aware that ISO SC42 has several standards underway that seek to bring clarity across industry to the role of different actors in the AI life cycle.

We also note the new British initiative, known as the AI Standards Hub, that is being led by The Alan Turing Institute in partnership with the British Standards Institution and the National Physical Laboratory. We hope that their work to ensure that a broad range of stakeholders are contributing to and benefiting from standards will help to further advance the development and use of trusted AI.

In regulating AI, we can look to recent innovations that support the ethical development of transformative technologies. Regulatory sandboxes support an environment where industry and regulatory engagement can improve product credibility. Having regulation present as part of the design process will have a significant impact on the development of the AI regulatory framework as it will be more straightforward for stakeholders in this space to anticipate the policy and regulatory direction of travel. Our Privacy Sandbox commitments with the CMA to replace third party cookies on Chrome with Privacy Sandbox tools offer a good template of innovative procedural reforms.

Here too the coordination of the Digital Regulation Cooperation Forum will be important, to ensure that regulators work together so that the principles of AI regulation are applied consistently across different contexts.

5. To what extent is the legal framework for the use of AI, especially in making decisions, fit for purpose? Is more legislation or better guidance required?

As noted, there are already many laws and legal frameworks that are applicable to AI. However, we agree with the government's proposal to create guidance for regulators that would achieve cohesion and consistency across sectors. A self-regulatory or co-regulatory set of governance norms that could be applied flexibly and adaptively would enable policy safeguards while preserving the space for continued beneficial innovation. This does not necessarily mean replacing existing laws, or developing new laws that are not tech-neutral.

Specifically regarding legal frameworks for AI decision-making, while we strongly agree that organisations should be responsible for the decisions they make, attributing legal liability can be less clear-cut due to the complexity of AI. Frameworks should ensure they balance safety and innovation appropriately. We suggest that regulators consider who should be responsible for what in a complex AI ecosystem and evaluate potential weaknesses in existing liability rules based on real-world examples and explore complementary rules for specific high-risk applications; consider sector-specific safe harbour frameworks and liability caps in domains where there is a worry that liability laws may otherwise discourage societally beneficial innovation; and explore insurance alternatives for settings in which traditional liability rules are inadequate or unworkable.

6. What lessons, if any, can the UK learn from other countries on AI governance?

It is important that, as governments around the world develop regulations and norms for AI, they work closely with other governments and international institutions to ensure that their approaches are aligned. Otherwise, we are likely to end up with a global patchwork that would slow the pace of AI development and create significant barriers to trade and competition, while also risking a race to the bottom.

We support the actions of countries, such as the UK, that are taking a considered approach to AI governance with guidance for existing regulators. Coadec has undertaken research into the AI ecosystem, asking startups for their opinions of the UK AI regulation proposal, who thought this could provide "an opportunity for the UK to cement its status as one of the most attractive places for AI innovation". We believe this is especially true if the EU implements recently-introduced proposals to bring general purpose AI into the scope of the AI Act, which would undermine the AI Act's risk-based approach.

We recommend designing and implementing regulations based on internationally recognized standards and principles, which can serve as the basis for robust self- and co-regulatory regimes, as guideposts for regulators, and as regulatory standards themselves if incorporated by reference. Global regulatory buy-in can support the development and adoption of effective standards for AI.

Memoranda of Understanding or Free Trade Agreements can also help to promote collaboration between countries and ensure consistency between rules. This can already be seen, for example, in an MoU between Australia and Singapore, and FTAs, such as the Digital Economy Partnership Agreement between Singapore, Chile and New Zealand.

Countries that are vocal in international fora, such as Japan, contribute to ensuring that there is consistency in AI rules globally. The UK should ensure that it is also a worthwhile partner in relevant fora, such as the OECD and G7 (particularly under Japan's G7 presidency). Amplifying the UK's voice will be particularly important as each country develops its own approach to AI governance and there is the risk of very different standards and approaches being introduced across borders. We support the government's plans to actively take a role in shaping global norms, and a flexible approach will help in navigating cultural and legal differences. SMEs will be most impacted if this is not done right, so straightforward and proportionate regulation that gives latitude for startups and other British companies at the forefront of AI technology will help encourage innovation safely.

We believe it is crucial for policy stakeholders worldwide to engage in this important conversation. As AI technology evolves and our own experience with it grows, we expect that the global community as a whole will continue to learn and additional nuances will emerge, including a fuller understanding of the trade-offs and potential unintended consequences that difficult choices entail.

(November 2022)