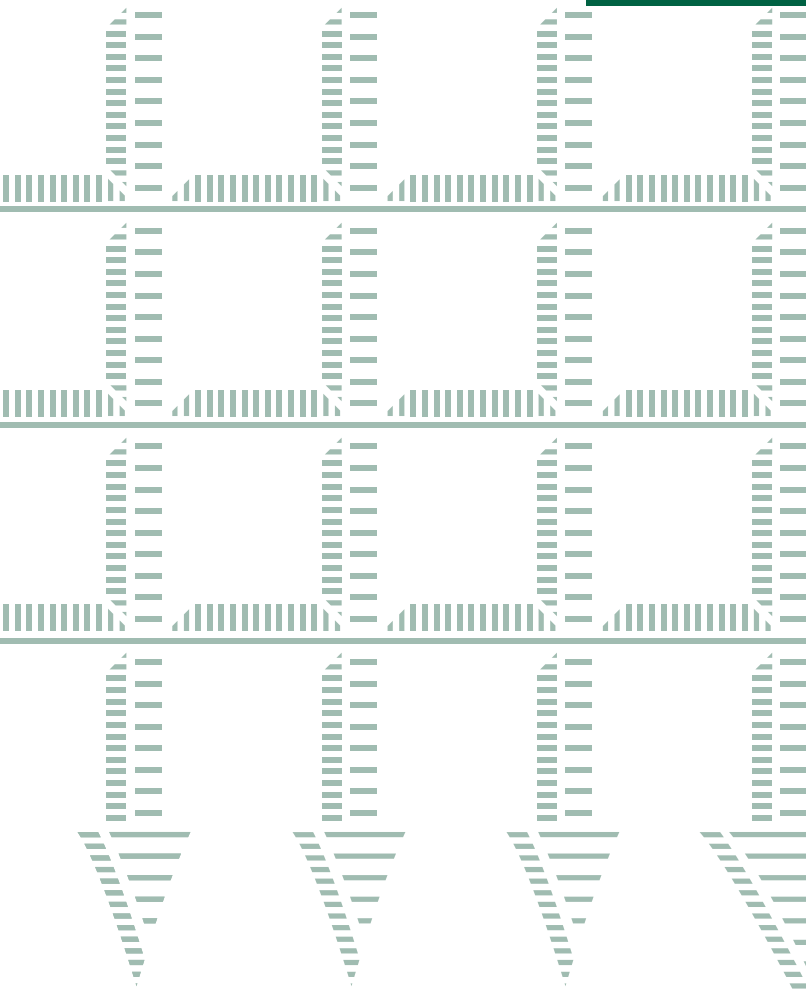


Science, Innovation and Technology Committee

Governance of artificial intelligence (AI): Government Response

First Special Report of Session 2024–25

HC 591



Science, Innovation and Technology Committee

The Science, Innovation and Technology Select Committee is appointed by the House of Commons to examine the expenditure, administration and policy of the Department for Science, Innovation and Technology, and associated public bodies. It also exists to ensure that Government policies and decision-making across departments are based on solid scientific evidence and advice.

Current membership

[Chi Onwurah](#) (Labour; Newcastle upon Tyne Central and West) (Chair)

[Emily Darlington](#) (Labour; Milton Keynes Central)

[George Freeman](#) (Conservative; Mid Norfolk)

[Dr Allison Gardner](#) (Labour; Stoke-on-Trent South)

[Tom Gordon](#) (Liberal Democrat; Harrogate and Knaresborough)

[Kit Malthouse](#) (Conservative; North West Hampshire)

[Jon Pearce](#) (Labour; High Peak)

[Steve Race](#) (Labour; Exeter)

[Dr Lauren Sullivan](#) (Labour; Gravesham)

[Adam Thompson](#) (Labour; Erewash)

[Martin Wrigley](#) (Liberal Democrat; Newton Abbot)

Powers

The Committee is one of the departmental select committees, the powers of which are set out in House of Commons Standing Orders, principally in SO No. 152. These are available on the internet via www.parliament.uk.

Publication

This Report, together with formal minutes relating to the report, was Ordered by the House of Commons, on 3 December 2024, to be printed. It was published on 10 January 2025 by authority of the House of Commons. © Parliamentary Copyright House of Commons 2025.

This publication may be reproduced under the terms of the Open Parliament Licence, which is published at www.parliament.uk/copyright.

Committee reports are published on the Committee's website at www.parliament.uk/CommonsSITC and in print by Order of the House.

Contacts

All correspondence should be addressed to the Clerk of the Science, Innovation and Technology Committee, House of Commons, London SW1A 0AA. The telephone number for general enquiries is 020 7219 5023; the Committee's email address is commonssitc@parliament.uk. You can follow the Committee on X (formerly Twitter) using CommonsSITC.

First Special Report

On 28 May 2024 the Science, Innovation and Technology Committee published its Third Report of Session 2023–24, [Governance of artificial intelligence \(AI\)](#) (HC 38). The Government Response was received on 25 November and is appended to this Report.

Introduction

The Government thanks the Committee for its report, “[Governance of AI](#)”, and notes the conclusions and recommendations.

We agree with the Committee that AI-specific legislation is required. We will shortly publish a consultation setting out our legislative proposals to establish binding regulations on the companies developing the most powerful AI models.

The Government’s response to each of the Committee’s recommendations is set out below.

Appendix: Government Response

Introduction

The case for AI

Government response to committee findings and recommendations
2 - 4

2. If governed appropriately, we believe that AI can deliver on its significant promise, to complement and augment human activity. The Government has articulated the case for AI: better public services, high quality jobs and a new era of economic growth driven by advances in AI capabilities. (Paragraph 22)

3. The Government is right to emphasise the potential societal and economic benefits to be won from the strategic deployment of AI. However, as our interim Report highlighted, the challenges are as clear as the potential benefits, and these benefits cannot be realised without public trust in the technology. (Paragraph 23)

4. The Government should certainly make the case for AI but should equally ensure that its regulatory framework addresses the Twelve Challenges of AI Governance that we have identified in our interim Report; and offer potential solutions to in this Report. (Paragraph 24)

Artificial Intelligence (AI) is at the heart of the UK Government's plan to kickstart an era of economic growth, transform how we deliver public services, and boost living standards for working people across the country. The Secretary of State for the Department of Science, Innovation, and Technology (DSIT) asked Matt Clifford, a leading tech entrepreneur and industry expert, to develop a plan to grow the UK's domestic AI sector and drive adoption of AI across the economy to boost growth and improve products and services. The AI Opportunities Action Plan sets out how we will achieve these goals, particularly through securing the necessary infrastructure, talent, and data access, as well as setting out the steps we will take to support AI adoption across the economy.

However, the full economic potential of AI can only be realised if businesses and consumers trust it. To support this, our intention is for legislation to establish binding requirements on the handful of companies developing the most powerful AI systems. This highly targeted proposed legislation would ensure the UK is prepared for this fast-moving technology. It would build on the voluntary commitments already secured at the Seoul and Bletchley AI Safety Summits and would strengthen the AI Safety Institute. The proposed legislation would support growth and innovation by reducing current regulatory uncertainty for AI developers, strengthening public trust and boosting business confidence.

AI specific legislation

Government response to Committee conclusions and recommendations 5 – 7

5. The next Government should stand ready to introduce new AI-specific legislation, should an approach based on regulatory activity, existing legislation and voluntary commitments by leading developers prove insufficient to address current and potential future harms associated with the technology. (Paragraph 33)

6. The Government should in its response to this Report provide further consideration of the criteria on which a decision to legislate will be triggered, including which model performance indicators, training requirements such as compute power or other factors will be considered. (Paragraph 34)

7. The next Government should commit to laying before Parliament quarterly reviews of the efficacy of its current approach to AI regulation, including a summary of technological developments related to its stated criteria for triggering a decision to legislate, and an assessment whether these criteria have been met. (Paragraph 35)

The Government welcomes the findings of the committee in relation to the need to legislate for the safety of AI. As announced at the King's Speech on 17 July 2024, the Government is proposing to establish appropriate legislation to place requirements on those working to develop the most powerful artificial intelligence models.

The Government intends to consult on these proposals before bringing forward legislation, which will include how the most powerful AI models will be captured. We look forward to feedback from the committee on these proposals and on our wider priorities for AI.

The role of regulators

Government response to Committee conclusions and recommendations 8 – 15

8. *We welcome confirmation that the Government will undertake a regulatory gap analysis to determine whether regulators require new powers to respond properly to the growing use of AI, as recommended in our interim Report. However, as the end of this Parliament approaches, there is no longer time to bring forward any updates to current regulatory remits and powers, should they be discovered to be necessary. This could constrain the ability of regulators to properly implement the Government's AI principles and undermine the UK's overall approach. (Paragraph 40)*

9. *The next Government should conduct and publish the results its regulatory gap analysis as soon as is practicable. If the analysis identifies any legislation required to close regulatory gaps, this should be brought forward in time for it to be enacted as soon as possible after the General Election. (Paragraph 41)*

10. *The general-purpose nature of AI will, in some instances, lead to regulatory overlap, and a potential blurring of responsibilities. This could create confusion on the part of consumers, developers and deployers of the technology, as well as regulators themselves. (Paragraph 45)*

11. *The steering committee that the Government has said it will establish should be empowered to provide guidance and, where necessary, direction to help regulators navigate any overlapping remits, whilst respecting the independence of the UK's regulators. (Paragraph 45)*

12. *The regulatory gap analysis being undertaken by the Government should identify, in consultation with the relevant regulators and coordinating entities such as the Digital Regulation Cooperation Forum and the AI and Digital Regulations Service, areas where new AI models and tools will necessitate closer regulatory co-operation, given the extent to which some uses for AI, and some of the challenges these can present—such as accelerating existing biases—are covered by more than one regulator. The gap analysis should also put forward suggestions for delivering this co-ordination, including joint investigations, a streamlined process for regulatory referrals, and enhanced levels of information sharing. (Paragraph 46)*

13. *The increasing prevalence and general-purpose nature of AI will create challenges for the UK's sectoral regulators, however expert they may be. The AI challenge can be summed up in a single word: capacity. Ofcom, for example, is combining implementation of a broad new suite of powers conferred on it by the Online Safety Act 2023, with formulating a comprehensive response to AI's deployment across its wider remit. Others will*

be required to undertake resource-intensive investigations and it is vital that they are able, and empowered, to do so. All will be required to pay greater attention to the outputs of AI tools in their sectors, whilst paying due regard to existing innovation and growth-related objectives. (Paragraph 55)

14. The announced £10 million to support regulators in responding to the growing prevalence of AI is clearly insufficient to meet the challenge, particularly when compared to the UK revenues of leading AI developers. (Paragraph 56)

15. The next Government must announce further financial support, agreed in consultation with regulators, that is commensurate to the scale of the task. It should also consider the benefits of a one-off or recurring industry levy, that would allow regulators to supplement or replace support from the Exchequer for their AI-related activities. (Paragraph 57)

The Government recognises that, beyond placing requirements on the development of the most powerful artificial intelligence models, there are a broad range of issues associated with AI development and deployment which require regulatory oversight.

In most cases, we believe that our existing expert regulators are best placed to apply rules to the use of AI in the contexts they know better than anyone else. The Government remains committed to a pro-innovation approach, with existing expert regulators addressing AI risks in their sectors, understanding where and how the product or service may be used. We are committed to ensuring that regulators have the right expertise and resources to make proportionate and informed regulatory decisions about AI in their sectors.

DSIT remains committed to providing £10m of funding to boost regulators' AI capabilities. This funding is one part of a broader programme of work to support regulators to adapt to the age of AI. The Government supports coordination, collaboration and knowledge exchange between regulators. For example, Government has provided £2m to support the DRCF's AI and Digital Hub, a cross-regulatory advice service which enables innovators to get joined-up advice on regulatory compliance.

The Government will continue to work with regulators to implement pro-innovation regulatory initiatives, including through our newly established Regulatory Innovation Office (RIO). The Regulatory Innovation Office was formally established on 8th October 2024. It will be the Government's primary lever for achieving its transformative ambitions in regulatory innovation.

The RIO will support regulators to update regulation, speeding up approvals, and ensuring different regulatory bodies work together smoothly. It will work to continuously inform the Government of regulatory barriers to innovation, set priorities for regulators which align with the Government's broader ambitions and support regulators to develop the capability they need to meet them and grow the economy.

The RIO will initially support the growth of four fast-growing areas of technology making a difference to people's lives before backing further technologies and sectors as the Office evolves. One of the four initial technologies that the RIO will focus on is Artificial Intelligence and Digital in healthcare, which is set to revolutionise healthcare delivery. RIO will support the healthcare sector to deploy AI innovations safely, improving NHS efficiency and patients' health outcomes.

AI in the public and private sectors

Government response to Committee conclusions and recommendations 16 – 20

16. AI can be used to increase productivity and augment the contributions of human workers in both the public and private sectors. We welcome the establishment of i.AI and the focus on AI deployment set out in the public sector productivity programme; as well as initiatives to increase business adoption such as the AI and Digital Hub. (Paragraph 73)

17. The next Government should drive safe adoption of AI in the public sector via i.AI, the National Science and Technology Council and designated lead departmental Ministers for AI. (Paragraph 74)

18. In its response to this Report, the Government should confirm the full list of public sector pilots currently being led or supported by i.AI, the criteria that determined i.AI pilot project selections, how it intends to evaluate their success and decide whether to roll them out more widely, and what other pilots are planned for the remainder of 2024. (Paragraph 75)

19. i.AI should undertake an assessment of the existing civil service workforce's AI capability, identify areas of the public sector that would benefit the most from the use of AI and where value for money can be delivered, set out how potential risks associated with its use should be mitigated, and publish a detailed AI public sector action plan. Progress against these should be reported to Parliament on an annual basis and through regular written or oral statements by Ministers. (Paragraph 76)

20. The requirement for Government departments to use the Algorithmic Transparency Recording Standard should be extended to all public bodies sponsored by Government departments, from 1 January 2025. (Paragraph 77)

Public sector adoption is a key part of the AI Opportunities Action Plan. The Plan will detail how we can reimagine our public services by ensuring the public sector takes advantage of the best emerging use-cases and tools. Further updates on this will be shared.

In regard to ensuring the public sector benefits from the use of AI, the technology has significant potential to improve public services. AI will continue to form part of the Government's approach to digital service delivery, including through the new Digital Centre of Government housed within the DSIT. Government is shaping the new 'digital centre', which will expand the work of the Incubator for AI (i.AI) to harness the power of AI for the public good.

The projects being led or supported by i.AI, and currently being piloted by users, are listed as follows, with further details available on the i.AI website at ai.gov.uk:

- Caddy – a customer service copilot, currently being piloted in Citizens Advice Stockport, Oldham, Rochdale & Trafford (CASORT).
- Redbox – a secure Government chatbot, being trialled by over 600 users in the Cabinet Office
- Consult – A tool to draw out insights from large consultations
- Lex – A tool for navigating legislation data, being piloted by Government lawyers and legislation teams.

i.AI is also incubating other projects which will be announced in due course as they progress to pilot. As an incubator, i.AI does not routinely publish details of projects in their very early stages of scoping and development. By design i.AI incubates a large number of projects (around 50 have been incubated so far), assesses and evaluates them, and then down selects to deliver the programs with the strongest delivery and impact potential.

The team sources ideas for product development from ministerial priorities, departmental suggestions, external engagement, and internal ideation. Before proceeding, these ideas are prioritised by assessing and balancing deliverability, value for money and potential impact. i.AI has recruited a dedicated Impact and Evaluation team to embed evaluation into product development.

We will ensure there's correct governance in place to drive progress on building scientific and technological capabilities, including through ensuring ministerial buy-in across Whitehall through fora such as the Science and Technology Cabinet Committee.

The Government is committed to driving safe adoption of AI in the public sector. The Generative AI Framework was published in January 2024 to provide practical considerations for anyone planning or developing a generative AI solution. The Government is working on an update to support decision makers in Government and the wider public sector with relevant independent guidance. The updated AI Framework provides decision makers in Government and the wider public sector with the latest guidance on using AI safely and securely.

In regard to the Civil Service workforce AI capabilities, in August 2024, we published nine courses to Civil Service Learning covering a range of AI topics from introductory to advanced levels supplied by industry leaders, which will support all civil servants to learn about this important topic. This is an ongoing work stream. The Government will continue to assess the civil service workforce AI capabilities to build upon this.

In February 2024 we announced that the Algorithmic Transparency Recording Standard (ATRS) would now be made 'a requirement for all Government departments, with an intent to extend to the broader public sector over time'. The rollout of this mandatory requirement is well underway. We have initially prioritised central Government departments. This has then been followed by a priority group of around 85 arm's-length bodies (ALBs) which deliver public or frontline services, or which interact directly with the general public. We expect to publish the first collection of records drafted under this requirement on GOV.UK imminently.

We will also shortly publish on GOV.UK a scope and exemptions policy explicitly setting out the organisations for which use of the ATRS is currently mandatory. As a Data Standards Authority endorsed product, the ATRS remains recommended for use in the broader public sector, and we will continue to explore options for further embedding and enforcing its use.

The AI Safety Institute

Government response to Committee conclusions and recommendations 21 – 26

21. It is a credit to the commitment of those involved that the AI Safety Institute has been swiftly established, with an impressive and growing team of researchers and technical experts recruited from leading developers and academic institutions. (Paragraph 80)

22. *The next Government should continue to empower the Institute to recruit the talent it needs. (Paragraph 80)*

23. *Although the Institute is not a regulator, it has undeniably played a decisive role in shaping the UK's regulatory approach to AI. We commend the work of the Institute and its researchers in facilitating and informing the ongoing international conversation about AI governance. (Paragraph 89)*

24. *However, we are concerned by suggestions that the Institute has been unable to access as-yet unreleased AI models to perform the pre-deployment safety testing it was set up to undertake. If true, this would undermine the delivery of the Institute's mission and its ability to increase public trust in the technology. (Paragraph 90)*

25. *In its response to this Report, the Government should confirm which models the AI Safety Institute has undertaken pre-deployment safety testing on, the nature of the Governance of artificial intelligence (AI) testing, a summary of the findings, whether any changes were made by the model's developers as a result, and whether any developers were asked to make changes but declined to do so. (Paragraph 91)*

26. *The Government should also confirm which models the Institute has been unable to secure access to, and the reason for this. If any developers have refused access—which would represent a contravention of the reported agreement at the November 2023 Summit at Bletchley Park—the Government should name them and detail their justification for doing so. (Paragraph 92)*

Our intention is for legislation to put the AI Safety Institute (AISI) on a statutory footing and there will be further detail about this in the coming months. We are proud of the work of the Institute and what it has achieved in the last year. Putting the Institute on a statutory footing would strengthen its role leading voluntary collaboration with AI developers and leading international coordination of AI safety. This work is vital as the technology continues to develop at pace.

Over the last year, AISI has tested models from major labs both before and after deployment. It is actively engaged in safety testing of frontier AI models, working collaboratively with developers who have committed to responsible AI development.

Competition is fierce in the frontier AI market, and the nature and timing of releases are highly commercially sensitive. Because of this, and to maintain confidence in our collaborative relationships, it is often not appropriate or possible to provide commentary on individual models under evaluation. This approach helps maintain trust with industry partners and ensures we can fulfil our safety testing mandate. We will continue to share

appropriate information about our testing activities through official public announcements, particularly when developers consent to disclosure or when information is already in the public domain.

Earlier this year AISI agreed to work in partnership with the US AISI and build towards interoperability between the two Institutes. Building on this, several labs have provided access to their models pre-deployment. The Government can confirm that AISI's research encompasses critical areas including potential societal harms, misuse risks and autonomy risks. When safety concerns are identified through our testing processes, we work constructively with developers to address these issues. AISI maintains detailed records of all testing outcomes and subsequent modifications, though these details remain confidential to protect commercial interests and our working relationships with AI companies.

It is important to remember that it is the AI developers' responsibility to ensure their models are safe. Our focus remains on ensuring rigorous safety evaluations while respecting commercial confidentiality.

AISI continues to make progress in securing access agreements with frontier AI developers for both pre- and post- deployment testing. We have established productive working relationships with multiple developers since last year, and we are expanding these arrangements to enable joint testing capabilities with international partners, particularly through our collaboration with the US AI Safety Institute as outlined in the UK-US Memorandum of Understanding.

At this stage, we believe it would be counterproductive to identify specific developers with whom we have not yet secured testing access. We are actively engaged in complex negotiations regarding access to commercially sensitive systems and capabilities. These discussions require careful consideration of intellectual property, technical requirements, and security protocols. While some negotiations take longer than others, we are seeing positive movement toward expanded cooperation with frontier AI labs.

The international dimension

Government response to Committee conclusions and recommendations 27 – 31

27. In our interim Report we highlighted moves by both the United States and European Union to develop their own approaches to AI governance. The subsequent White House Executive Order and the EU AI Act are clear attempts to secure competitive regulatory advantage. (Paragraph 129)

28. It is true that the size of both the United States and European Union markets may mean that ‘the Washington effect’ and ‘Brussels effect’—referring to the de facto standardising of global regulatory approaches, potentially to the detriment of the UK’s distinct approach—will apply to AI governance. Nevertheless, the distinctiveness of the UK’s approach and the success of the AI Safety Summit have underlined the significance of its current and future role. (Paragraph 130)

29. Both the US and EU approaches to AI governance have their downsides. The scope of the former only imposes a requirement on Federal bodies and relies on voluntary commitments from leading developers. The latter has been criticised for its topdown, prescriptive approach and the potential for uneven implementation across different member states. (Paragraph 131)

30. The UK is entitled to pursue an approach that considers developments in other jurisdictions but does not unthinkingly replicate them. However, where there are lessons to be learned from other jurisdictions, the next Government should be willing to apply them. (Paragraph 132)

31. The UK has a long history of encouraging technological innovation by offering a stable, expert regulatory environment coupled with clear industry standards. The current Government is therefore right to have encouraged the growth of a strong AI sector in the UK, engaged with leading developers through the AI Safety Institute and future Summits, and participated in international standards fora. This international agenda should be continued by the next Government, and coupled with the swift establishment of a domestic framework that sufficiently addresses the Twelve Challenges of AI Governance highlighted in our interim Report. (Paragraph 133)

The Government is committed to making the UK a world leader in AI, to drive economic renewal, boost living standards, and deliver growth for people across the country.

We believe that coordinating with international partners is vital to effectively tackle cross-border challenges that AI poses. We will continue to engage closely with our international partners, including the US and EU, as we further develop our approach to AI governance including our legislative proposals.

As the Committee has noted, through the AI Safety Summit and AI Seoul Summit the UK has demonstrated international leadership on frontier AI safety, bringing together international partners to build consensus on the safe development and deployment of AI. We have also built world-leading state capacity in AI safety through our AI Safety Institute, which is furthering the science of AI safety through the Network of AISIs. We will continue to build on this work with our partners, both at the upcoming AI Action Summit in France, and in a range of multilateral fora.

We champion the multi-stakeholder, industry-led standards development process, where Government is one of the stakeholders in that ecosystem. We want to promote a robust and diverse digital standards ecosystem, strengthening and building international partners to foster collaboration and promote integrity in standards development.

In regards to standards specifically government recognises that AI standards, including those used in assurance and certification schemes, can help organisations put our proposed regulatory principles into practice, innovate responsibly, and build public confidence. Standards can also complement sector-specific approaches to AI regulation by providing common benchmarks and practical guidance to organisations.

The UK's AI Standards Hub is a partnership between the Alan Turing Institute, British Standards Institution and National Physical Laboratory, supported by DSIT. The UK's AI Standards Hub is built upon 4 pillars, Tracking and sharing information on AI standards; Convening, connecting, and community building; Education, training, and professional development; Thought leadership & international engagement

Twelve challenges of AI governance revisited

Government response to Committee recommendations 32 – 33

32. AI can entrench and accelerate existing biases. The current Government, future administrations and sectoral regulators should require deployers of AI models and tools to submit them to robust, independent testing and performance analysis prior to deployment. (Paragraph 140) 56 Governance of artificial intelligence (AI)

33. Model developers and deployers should be required to summarise what steps they have taken to account for bias in datasets used to train models, and to statistically report on the levels of bias present in outputs produced using AI tools. This data should be routinely disclosed in a similar way to company pay gap reporting. (Paragraph 141)

The Algorithmic Transparency Recording Standard (ATRS) is now mandatory for all Government departments, with an intent to extend to the broader public sector over time. It includes fields on risks and appropriate mitigations taken to ensure that tools perform appropriately and fairly, including impact assessments. The ATRS also includes fields requiring algorithmic tool deployers to summarise the data used to train the models underpinning their tools, including steps taken to ensure clean and complete data.

DSIT is working with industry and academia to develop robust tests for fairness and bias in AI systems. The Responsible Technology Adoption Unit, alongside the Information Commissioner's Office and the Equality and Human Rights Commission, is running the [Fairness Innovation Challenge](#). This is a grant challenge that has given over £465,000 of Government funding to support the development of socio-technical solutions to address bias and discrimination in AI systems. We are thrilled to have funded four projects to develop sociotechnical solutions to improve fairness in AI systems across four different sectors. Grant funding has been awarded to: The Open University (Higher Education); The Alan Turing Institute (Finance); King's College London (Healthcare); and Coefficient Systems Ltd (Recruitment).

DSIT is also developing the AI Management Essentials tool, a self-assessment tool for SMEs and startups to achieve baseline "responsible AI" practices in their organisation. The tool distils key principles from existing AI-related frameworks and standards including ISO/IEC 42001, the NIST Risk Management framework, and the EU AI Act. The tool will provide an open access, simplified baseline of requirements for responsible and trustworthy AI development and deployment, including requirements related to bias testing and reporting.

In addition, the Public Sector Equality Duty in the Equality Act 2010 requires public authorities, and those carrying out public functions, to have due regard to the need to eliminate discrimination, advance equality of opportunity, and foster good relations between different people. This includes when using AI to exercise public functions. This Government is committed to protecting and upholding the Public Sector Equality Duty, including in relation to AI.

Government response to Committee recommendations 34 – 35

34. Regulators and deployers should ensure that the right balance is maintained between the protection of privacy and pursuing the potential benefits of AI. Determining this balance will depend on the context in which the technology is being deployed, with reference to the relevant laws and regulations. (Paragraph 145)

35. Sectoral regulators should publish detailed guidance to help deployers of AI strike the balance between the protection of privacy and securing the technology's intended benefits. In instances where regulators determine that this balance has not been met, or where the relevant laws or regulatory requirements have not been met, it should impose sanctions or prohibit the use of AI models or tools. (Paragraph 146)

Data protection law requires organisations to apply data protection principles at each stage of the AI lifecycle when using personal data. These principles include, but are not limited to, accuracy, transparency, purpose limitation, data minimisation, confidentiality, and accountability. This means organisations must evaluate the risks to individuals, inform them of their rights arising from specific contexts, and put relevant safeguards in place. The principles-based and context-specific nature of data protection law allows for striking the right balance between protecting privacy and benefiting from what new technologies have to offer.

The Information Commissioner's Office (ICO), the UK's independent data protection regulator, plays an important role in the regulation of AI models, which are generally trained on high volumes of data – often including personal data. The ICO monitors the effects of AI on people and society using sources including its own casework, stakeholder engagement and wider intelligence gathering. The ICO has already published guidance and resources on AI and data protection law, including an AI toolkit to help organisations identify and mitigate risks, and recently conducted a consultation series on how aspects of data protection law should apply to the development and use of generative AI models.

As outlined by ICO guidance, the UK GDPR requires organisations to put in place appropriate technical and organisational measures to implement the data protection principles effectively. This is known as data protection by design and by default and is applicable to AI systems processing personal data. The data protection framework also allows the ICO to carry out consensual audits to assess whether controllers or processors are complying with good practice in the processing of personal data. Further, the ICO has a range of enforcement powers available against data protection breaches such as serving an assessment notice when carrying out investigations to help understand how organisations use and store data, through to enforcement notices or penalty notices.

Government response to Committee conclusions and recommendations 36 and 37

36. We welcome the Government amendment to the Criminal Justice Bill as a necessary step towards ensuring the UK's legal framework reflects the current state of technological development and protects citizens, primarily women and girls, from the consequences of AI-assisted misrepresentation, including deepfake pornography. (Paragraph 150)

37. Should the Bill's remaining stages fail to be completed prior to the dissolution of Parliament, the next Government must introduce similar provisions as soon as is practicable after the General Election. (Paragraph 150)

The Government refuses to tolerate violence against women and girls (VAWG) and will deliver the manifesto commitments to ban the creation of sexually explicit deepfakes and halve VAWG within a decade. DSIT is working with the Home Office and the Ministry of Justice to identify the most appropriate legislative vehicle to introduce this measure, which will ensure those who create these images without consent face the appropriate punishment.

Government response to Committee recommendation 38

38. The Government and regulatory authorities, informed by the work of the Defending Democracy Taskforce, should safeguard the integrity of the upcoming General Election campaign in its approach to the online platforms that host deepfake content which seeks to exert a malign influence on the democratic process. If these platforms are found to have been slow to remove such content, or to have facilitated its spread, regulators must take stringent enforcement action—including holding senior leadership personally liable and imposing financial sanctions. (Paragraph 154)

The Government has in place established systems and processes to protect the democratic integrity of the UK. Alongside this, the Defending Democracy Taskforce works to reduce the risk to the UK's democratic processes, institutions and society and ensure that these are secure and resilient to threats of foreign interference. The Taskforce brings together Ministers, operational agencies, and other partners to work on the full range of threats facing the UK's democracy and improve confidence in the integrity and security of our elections.

In 2023, the Defending Democracy Taskforce set up the Joint Election Security and Preparedness Unit (JESP) as a permanent function dedicated to protecting UK elections and referendums. JESP stood up an Election Cell for the 2024 UK General Election. The cell coordinated a wide range of teams across Government to respond to issues, including AI-generated mis- and disinformation, as they emerged. DSIT met regularly with the major social media platforms in the run up to and during the election period, to discuss what action they were taking to protect the integrity of the UK's democratic processes. This helped to ensure a broadly robust response from platforms to content that may undermine the election, including through media literacy initiatives and improvements in their Terms of Service.

The Online Safety Act has also introduced measures to combat these threats. Under the Act, platforms will have duties to implement systems and processes to mitigate illegal content, including illegal mis- and disinformation, and will be required to take steps to remove in-scope content if they become aware of it on their services. This includes content which constitutes Foreign Interference, which has been added as a

priority offence and captures a wide-range of state-sponsored mis- and disinformation and state-linked interference online aimed at the UK. AI-generated material, including deepfakes, is captured by the Act where it constitutes user generated content that is illegal or harmful to children. Services above a designated threshold will also need to remove this content where it is prohibited in their terms of service.

Government response to Committee recommendation 39

39. A cross-Government public awareness campaign should be launched to inform the public about the growing prevalence of AI-assisted misrepresentation, the potential consequences, what the Government is doing to address the Challenge, and what steps individuals can take to protect themselves online. (Paragraph 155)

The Government is working to make the internet safer. Doing this effectively requires a broad toolkit – using the Online Safety Act to ensure platforms take steps to limit harmful content online, but also ensuring both children and adults have the knowledge and skills to navigate the online world. This includes navigating the risks presented by emerging technology such as AI.

Media literacy can help tackle a wide variety of online safety issues for all internet users, including children. It includes understanding that online actions have offline consequences, being able to contribute to a respectful online environment, and being able to engage critically with online information. It is a key tool for building people’s resilience to misinformation, disinformation and AI-generated deepfakes.

The Online Safety Act updated Ofcom’s statutory duty to promote media literacy. Under the Act, Ofcom is required to bring about better understanding of ways in which members of the public can keep themselves and others safe, including by encouraging the development and use of technologies that help users protect themselves from harmful content, including AI-generated mis- and disinformation. Ofcom is also required to raise awareness of the nature and impact of mis- and disinformation, to ensure the public can ‘establish the reliability, accuracy and authenticity’ of content found on regulated services. These duties are already in force.

Between 2022 and 2024, DSIT has provided almost £3million in grant funding for a range of projects, including educational interventions designed to empower users with the skills and knowledge they need to make safe and informed choices online. In 2024, this covered almost £0.5 million in funding to scale up two media literacy programmes which, between them, provide media literacy training and support to teachers, children aged 11-16, other professionals working with families and parents/carers.

The Government has also established an independent Curriculum and Assessment Review that will consider the key digital skills needed for future life and the critical thinking skills needed to ensure children are resilient to misinformation and extremist content online.

Government response to Committee conclusions and recommendations 40 and 41

40. At the so-called ‘frontier’ of AI a small group of leading developers are responsible for and accruing significant benefits from the development of advanced models and tools—thanks in part to their ability to access the necessary training data. This potential dominance is arguably to the detriment of free and open competition. (Paragraph 159)

41. As the regulator responsible for promoting competitive markets and tackling anticompetitive behaviour, the CMA should identify abuses of market power and use its powers to stop them. This could take the form of levying fines or requiring the restructuring of proposed mergers. (Paragraph 160)

The Government is committed to ensuring that fair and open competition drives growth and innovation in the AI ecosystem. The Digital Markets, Competition and Consumers Act equips the Competition and Markets Authority (CMA) with new, faster and more effective tools to address significant competition issues in AI markets, as well as in broader digital markets where AI plays a role.

The regime allows the CMA to impose significant penalties on firms that fail to comply, ensuring robust accountability. Additionally, Strategic Market Status (SMS) firms will be subject to new merger reporting requirements, which will enable proactive oversight of market consolidation and maintain competitive dynamics.

The CMA, as an independent regulator, will determine how, when, and if to exercise these new powers. The Government is working closely with the CMA to ensure the digital markets measures commence in January 2025. The CMA is using its existing markets and mergers tools to analyse AI markets and assess competition impacts with this emerging sector.

Government response to Committee conclusions and recommendations 42 and 43

42. AI models and tools rely on access to high-quality input data. The phrase ‘garbage in, garbage out’ is not new, but it is particularly applicable to AI. (Paragraph 161)

43. The potential for human error and bias notwithstanding, deployers should not solely rely on outputs produced with AI tools to determine their decision-making, particularly in areas that could affect the rights and standing of the individuals or entities concerned, such as insurance decisions or recruitment. These algorithmic decisions should always be reviewed and verified by trained humans, and those affected should have the right to challenge these decisions—a process that should also be human-centred. (Paragraph 161)

AI-driven automated decision-making can play an important role in delivering effective services and improving productivity. The UK's data protection framework strikes a balance between enabling the best use of the technology whilst providing safeguards for individuals when they matter most.

Individuals have rights under the UK's data protection law, where they have been subject to automated decision-making without human involvement, including profiling. These rights apply to all forms of automated decision-making regardless of the technology being used. Where solely automated decisions are based on the use of personal data and have a legal or a similarly significant effect on individuals, data protection law requires organisations to offer a range of safeguards to individuals. These safeguards include informing them of the automated decision, enabling them to contest the decision, and the right to obtain human review.

In addition to these safeguards, data protection principles of accuracy, transparency, purpose limitation, data minimisation, confidentiality, and accountability apply to all processing of personal data, including where forms of automated decision-making is used.

Within the public sector specifically, the Algorithmic Transparency Recording Standard (ATRS) requires Government departments, and in time the broader public sector, to proactively publish information about how and why algorithmic tools are embedded in broader decision-making processes. This includes explanations of when and where human review takes place.

In the Generative AI framework for HMG, we recommend maintaining appropriate human involvement in automated processes, and that developers and users of AI tools ensure that there is a human-in-the-loop who can oversee outputs when generative AI is in use in situations with a high impact.

Government response to Committee recommendation 44

44. The Government and future administrations should support the emergence of more AI startups in the UK by ensuring they can access the high-quality datasets they need to innovate. This could involve facilitating access to anonymised public data from data.gov.uk, the NHS and BBC via a National Data Bank, subject to appropriate safeguards. (Paragraph 162)

DSIT will be establishing a National Data Library, aiming to transform the way the Government manages our national data assets and to unlock the full value of our public data assets. This will provide simple, ethical, and secure access to public data assets, giving researchers and businesses powerful insights that will drive growth and transform people's quality of life through better public services and cutting-edge innovation, including AI. To deliver this vision we will need to build trust with data owners, users, and the wider public. This will include ensuring that data – particularly sensitive, personal data – is held securely. Work is underway to design and implement the library, with decisions to be made in due course.

The AI Opportunities Action Plan, produced by Matt Clifford, will set out the foundations needed for growing the AI sector and boosting AI adoption across the economy, and data access is a key part of that. More detailed recommendations will be available following the publication of the Action Plan.

Government response to Committee conclusions and recommendations 45 and 46

45. We welcome the Government's moves to establish a dedicated AI Research Resource and a cluster of supercomputers but are concerned that it has yet to set out further details of how researchers and startups will be able to access the compute they need to maximise the potential benefits of AI across society and the economy. (Paragraph 165)

46. The Government, or its successor administration, should publish an action plan and proposed deliverables for both the AI Research Resource and its cluster of supercomputers, and further details of the terms under which researchers and innovative startups will be able to access them. It should also undertake a feasibility study into the establishment of a National Compute Cluster that could be made available to researchers and startups. (Paragraph 166)

The Government recognises the importance of compute in enabling the research and innovation that drives growth and increases opportunities for people across the UK. DSIT is working alongside UKRI to develop a long-term

plan for the UK's compute needs that enables us to meet our AI and R&D ambitions and this will be in place ahead of the multiyear spending review in the Spring.

DSIT will shortly announce details of how researchers will access the AIRR (AI Research Resource). It will have a bespoke access model that reflects the unique needs of AI research and the Government's desire for compute to support the Government's missions and priorities.

Government response to Committee recommendation 47

47. The Black Box Challenge is one of the most paradigm-shifting consequences of AI, as it upends our well-established reliance on explainability and understanding. Given the complexity of currently available and in all likelihood future models, the starting point should be an acknowledgement how little we can understand about how many AI models produce their outputs, an acceptance that new ways of thinking will be required, and a regulatory approach that accounts for the impossibility of total explainability. (Paragraph 170)

While providing transparency and explainability on a technical level may be challenging in the context of Black Box systems, the Algorithmic Transparency Recording Standard (ATRS) helps teams to provide process-level transparency around algorithmic tools they are developing or using. This includes information about the design, development and deployment practices behind an algorithmic tool, and the mechanisms used to demonstrate that the solution is responsible and trustworthy. The ATRS aims to make sure that information about algorithmic solutions used by Government and the public sector are clearly accessible to the public on GOV.UK.

Government response to Committee recommendations 48 and 49

48. The regulators charged with implementing the Government's high-level AI governance principles should, in their approach to these models and tools, prioritise testing and verifying their outputs, as well seeking to establish—whilst accepting the difficulty of doing so with absolute certainty—how they arrived at them. (Paragraph 171)

49. The open-source approach has underpinned many technological breakthroughs, including the Internet and AI. Whilst some providers of products and services, such as AI models and their applications, will want to keep elements of their offerings proprietary, a healthy AI marketplace should be sufficiently diverse to support both 'open' and 'closed' options. The volume of investment flowing into AI developers of all types of models, rather than one or the other, is evidence of this market diversity. (Paragraph 175)

AI openness has helped support innovation, transparency, and AI safety research. We are committed to defending the importance of openness and supporting the UK's open-source ecosystem while also taking steps to improve AI safety.

Government response to Committee recommendation 50

50. When procuring AI models for deployment in the public sector the Government and public bodies should utilise those best suited to the task. (Paragraph 176)

Procuring AI should follow the same competitive processes to select the best supplier that can deliver against the requirements and provide best value. DSIT and Crown Commercial Services will be examining how existing procurement frameworks can be updated to make them better at supporting departments to select the right AI solution for their needs. (Paragraph 176)

HMG has set up AI lots in its frameworks, administered by CCS, including a dynamic purchasing system (set up in 2020, running until 2026) to which additional suppliers and products can be added at any time, with minimal rework. CCS and DSIT will be reviewing the routes for buyers to access the rapidly evolving market for these products and services, and where possible simplifying the process for buyers, rationalising the number of discrete frameworks, to allow for improved analysis of the use of these mechanisms to access the market. They will also gain feedback from existing buyers to improve the experience.

Government response to Committee recommendation 51

51. The Government should in its response to this Report tell us how it will ensure law enforcement and regulators are adequately resourced to respond to the growing use of AI models and tools to generate and disseminate harmful and illegal content. (Paragraph 180)

An annual industry fee will fund Ofcom's costs of regulation under the Online Safety Act. This will provide Ofcom with adequate resources to effectively exercise their online safety functions, including responding to developing challenges such as the growing use of AI.

Ofcom has been provided with the necessary powers to respond to the use of AI where that use is in scope of the Online Safety Act. The Act regulates AI generated content where it constitutes illegal content or content which is harmful to children and is either user-generated content on an in-scope service or is discoverable within a click of search results. Ofcom will also

have robust enforcement powers to issue substantial fines, implement business disruption measures, or even initiate prosecutions against senior managers in exceptional circumstances.

The growing use of AI models and tools by criminals to produce illegal content (e.g. child sexual abuse material) or content that is used to facilitate crime (e.g. fraud) is a genuine concern. There are also risks that the prevalence of this offending gets worse and expands to other areas of the criminal justice system such as the creation of false alibis and evidence. The Government is investing in deepfake detection capability, looking closely at the applicability of the criminal law to AI-enabled offending and investing in policing skills and training. The Government is also looking at measures which might restrict the ability for AI tools to be used by criminals.

Government response to Committee recommendations 52 and 53

52. The growing volume of litigation relating to alleged use of works protected by copyright to train AI models and tools, and the value of high-quality data needed to train future models, has underlined the need for a sustainable framework that acknowledges the inevitable trade-offs and establishes clear, enforceable rules of the road. The status quo allows developers to potentially benefit from the unlimited, free use of copyrighted material, whilst negotiations are stalled. (Paragraph 185)

53. The current Government, or its successor administration, should ensure that discussions regarding the use of copyrighted works to train and run AI models are concluded and an implementable approach agreed. It seems inevitable that this will involve the agreement of a financial settlement for past infringements by AI developers, the negotiation of a licensing framework to govern future uses, and in all likelihood the establishment of a new authority to operationalise the agreement. If this cannot be achieved through a voluntary approach, it should be enforced by the Government, or its successor administration, in co-operation with its international partners. (Paragraph 186)

The Government recognises the Committee's call to implement a sustainable framework in this area. We believe both in human-centred creativity and the potential of AI to unlock new creative frontiers and agree that all will benefit from greater clarity over copyright. Finding a balance between these concerns is a complex matter.

The application of copyright law to AI is disputed in the UK and around the world. Addressing uncertainty about the copyright framework for AI in the UK is a priority for DSIT and DCMS Ministers. As the Committee notes, any

resolution to these issues may require Government intervention. That is why the Government intends to launch a consultation soon that will aim to promote continued growth in the UK AI sector and creative industries.

The Government remains committed to engaging closely with stakeholders on this issue, building on recent roundtables that were held between DSIT and DCMS Ministers with representatives from the creative industries and AI sector. As this is a global issue, international cooperation will be crucial, and the Government is committed to working closely with international partners to push progress in this area.

Government response to Committee recommendation 54

54. Nobody who uses AI to inflict harm should be exempted from the consequences, whether they are a developer, deployer, or intermediary. The next Government together with sectoral regulators publish guidance on where liability for harmful uses of AI falls under existing law. This should be a cross-Government undertaking. Sectoral regulators should ensure that guidance on liability for AI-related harms is made available to developers and deployers as and when it is required. Future administrations and regulators should also, where appropriate, establish liability via statute rather than simply relying on jurisprudence. (Paragraph 189)

Given the pace of AI technology's development, it's important to set clear expectations for the behaviour of frontier AI developers and ensure they are trusted by the public. Our proposed legislation would reduce regulatory uncertainty for AI developers, strengthen public trust and boost business confidence.

These proposed targeted requirements for the developers of the most powerful AI systems would complement the UK's existing regulatory framework, helping to build trust in AI and drive adoption across the country. Our intention is to make sure our statute book is fit for the age of AI and that accountability is assigned appropriately. We will also ensure that our existing expert regulators have the right expertise and resources to make proportionate and informed regulatory decisions about AI in their sectors.

Government response to Committee recommendation 55

55. AI is already changing the nature of work, and as the technology evolves this process is likely to accelerate, placing some jobs at risk. At the same time, there are productivity benefits to be won, provided people are equipped with the skills to fruitfully utilise AI. This is a process that should begin in the classroom, and through greater prioritisation of initiatives such as the Lifetime Skills Guarantee and digital Skills Bootcamps. (Paragraph 193)

The Government is committed to creating an agile and responsive skills system which delivers the skills needed to support a world-class workforce in STEM sectors and drive economic growth. The AI Action Plan will consider how we can strengthen our AI skills and talent base to unlock the potential of AI across the country.

DFE's Curriculum and Assessment Review will ensure every young person gets the opportunity to develop skills prized by employers, including digital. The Government has established an independent Curriculum and Assessment Review, covering ages 5 to 18, chaired by Professor Becky Francis CBE, an expert in education policy.

The review group will publish an interim report in the new year setting out their interim findings and confirming the key areas for further work. The final review with recommendations will be published in autumn 2025.

The Government will also bring forward a comprehensive strategy for post-16 education to break down barriers to opportunity, support the development of a skilled workforce, and drive economic growth through our industrial strategy.

Meeting the skills needs of the next decade is key to delivering the Government's regional and national missions. The Industrial Strategy was launched at the International Investment Summit on 14th October and highlighted 'Digital and Technologies' as a critical growth sector.

Given AI's potential to boost productivity and its impact on the labour market, it is essential our workforce acquires the necessary AI-related skills. Skills England will address these national and regional skills gaps by aligning training with employer needs, collaborating with regional entities, and ensuring skills are prioritised in policy decisions. We will build a highly skilled workforce, preparing both young people and adults for a technology-driven world.

The growing demand for digital literacy, data analysis, and AI skills is evident. However, creativity, critical thinking, and emotional intelligence remain crucial. Equipping people with adaptability and a commitment to lifelong learning will help them navigate technological changes throughout their careers.

We are currently engaging multiple stakeholders to ensure our education system meets future skills needs. Collaboration with Government, industries, employers, training providers, and the education sector is vital to adapt to and leverage AI's impact on the labour market.

Government response to Committee recommendation 56

56. The current Government, or its successor, should commission a review into the possible future skills and employment consequences of AI, along the lines of the 2017 Taylor Review of modern working practices which examined the landscape, suggested ideas for debate and has resulted in legislative change. It should also in its response to this Report tell us how it will ensure workers whose jobs are at risk of automation will be able to retrain and acquire the skills necessary to change careers. (Paragraph 194) Governance of artificial intelligence (AI).

The Unit for Future Skills (which has now moved into Skills England Analysis and Insight) previously released a report on the impact AI will have on workers in the UK ([GOV.UK Impact of AI on UK jobs and training](#)). The recently published Skills England report also set out the risks of AI ([Skills England: driving growth and widening opportunities](#)) particularly noting; “11% of tasks in the UK economy are exposed to existing generative AI and this figure could increase to 59% if companies integrate AI more deeply”.

Both national and local plans have AI at the forefront of their thinking. AI is an area of focus both in terms of utilising the benefits of AI while ensuring workers are sufficiently upskilled or retrained in light of automation.

Advances in AI have the potential to increase productivity and create new high value jobs in the UK economy and the new Industrial Strategy “will aim to secure investment into crucial sectors of the economy to drive long-term sustainable, inclusive and secure growth” (SE report, September 2024). The Industrial Strategy will also include a skills plan for each sector, and Digital has been included as an Industrial Strategy sector. This skills plan will set out how the sector will adapt to AI and ensure a fully equipped and trained workforce. This is currently being drafted and is expected to be published in Spring 2025.

Local Skills Improvement Plans have also reviewed the impact of AI and are preparing for AI at the local level; “almost all LSIP areas covered AI and automation to some degree, recognising that it is an area they will continue to engage on with employers” (SE report, September 2024).

The Government will also bring forward a comprehensive strategy for post-16 education to break down barriers to opportunity, support the development of a skilled workforce, and drive economic growth through our industrial strategy.

Earlier this year [2024], DSIT commissioned Ipsos Mori to undertake an exploratory research project to assess how AI skills needs may change up to 2035, depending on how the technology develops. The outcomes of the report will help guide our future policymaking.

Government response to Committee recommendations 57 – 59

57. We welcome the organisation of the AI Safety Summit at Bletchley Park and commend the Government on bringing many key actors together. We look forward to subsequent Summits and hope that the consensus and momentum delivered at Bletchley Park can be maintained. (Paragraph 202)

58. However, looking beyond the AI safety discussion, we do not believe that harmonisation for harmonisation's sake should be the end goal of international AI governance discussions. A degree of distinction between different regulatory regimes is, in our view, inevitable. Such distinction may be motivated by geopolitics, but it may also simply be a case of healthy economic competition. (Paragraph 203)

59. Future AI Safety Summits must focus on the establishment of international dialogue mechanisms to address current, medium- and longer-term safety risks presented by the growing use of AI; and the sharing of best practice to ensure its potential benefits are realised in all jurisdictions. This should not set us on the road to a global AI governance regime—we are unconvinced that such a prospect is either realistic or desirable. (Paragraph 204)

To improve international co-ordination on mitigating the most severe AI risks, the UK has expanded efforts to contribute to the AI global governance landscape, co-ordinating an international response. These efforts have been in collaboration with global stakeholders, feeding into different channels and mechanisms, such as the Hiroshima AI Process, and the Council of Europe Convention on AI.

Within this wider international governance landscape, we are in agreement with the committee that the AI Summit series is now a key contributor to this objective. The Government agrees that a key aim moving forwards should be to sustain the consensus and momentum delivered at Bletchley Park, as we have been with the Seoul Summit and as we support the France AI Action Summit.

The inaugural Bletchley Park AI Safety Summit took place in November 2023 to explore and build consensus on international action which promotes safety at the frontier of AI. It was followed by the AI Seoul Summit in May 2024. The summit series has elevated the global conversation on AI safety. It helped further consensus on enabling factors to unlock AI opportunities through the safe development and deployment of the most advanced AI systems. This includes emerging global scientific agreement on the risks and capabilities of frontier AI, safety testing agreements for the third-party evaluation of frontier AI systems, and voluntary agreements from global frontier AI organisations to publish safety frameworks pursuant to

the Frontier AI Safety Commitments. We agree with the Committee that international fora should be used to address AI risks, including potential existential risks, and AI Summits are an appropriate place to do this.

We are also pleased to take forward action to assess key frontier AI risks outside of the Summit series, including through evaluations conducted by the UK AI Safety Institute. The international network of AI Safety Institutes was agreed at the Seoul Summit between key partners and looks to build global best practice around AI safety evaluations among countries with established or emerging AISI capabilities. The establishment of the Network demonstrates the ongoing effort to share information about frontier AI models, their limitations, capabilities and risks to promote the safe, secure and trustworthy development of AI internationally. The establishment of AI Safety Institutes presents an opportunity to understand a spectrum of frontier AI risks and share best practice internationally. This work, paired with our commitment to international collaboration and through the Summit series, will drive forward the international conversation on AI risks, including appropriate international preparedness in the case those critical risk thresholds are triggered.

Whilst the UK's priority remains to continue the momentum we have built around international co-operation on frontier AI safety risks, the Summit series has never been narrowly focused. For example, while technical agreements on particular frontier AI risks were established at Bletchley and Seoul, a wider range of risks, and opportunities were also discussed to reflect that frontier AI systems can contribute to a variety of issues, including those that can impact global inclusion, innovation and access, as well as supporting progress towards the Sustainable Development Goals.

We welcome the continued discussions which will take place at the France AI Action Summit in February 2025, as France has suggested a wide range of AI topics that represent the global impact of AI. This includes looking at how the world can harness the opportunities of AI in the public interest, e.g. strengthening access to data and compute.

Finally, in the context of an increasingly populated and complex international AI governance landscape (e.g., network of AISIs, UN bodies and initiatives, G7 Hiroshima Process, OECD-GPAI integrated partnership, Council of Europe Convention on AI), we believe, in agreement with the committee, that the AI Summits will continue to play a meaningful role in driving diverse international collaboration on AI safety among Governments, companies, academia and civil society to maximise the benefits of AI. We agree that harmonisation for harmonisation's sake is not the aim, but rather that we should seek coherent, coordinated approaches to ensure we are effectively working together to mitigate the most critical frontier AI safety risks.

Government response to Committee recommendation 60

60. The debate over the existential risk—or lack of it—posed by the increasing prevalence of AI has attracted significant attention. However, the Government’s initial assessment, that such existential risks are high impact but low probability, appears to be accurate. Nevertheless, given the potential consequences should risks highlighted by the AI Safety Institute and other researchers be realised, it is right for Governments to continue to engage with experts on the issue. (Paragraph 209)

DSIT has established the Central AI Risk Function (CAIRF), which sits within the AI Policy Directorate and works hand in hand with the AI Safety Institute. It is the central coordination function on AI for the Government and aims to reduce the likelihood and impact of AI-related risks. This includes existential risks.

As part of the Government’s AI risk assessment, management and mitigation efforts, CAIRF has engaged extensively with experts in a wide range of domain areas. We have established a global Network of Experts and will continue to draw on their expertise while developing policy on AI risk.

The upcoming legislative proposals on Frontier AI Safety will be an opportunity to further engage with experts on the issue.

Government response to Committee recommendations 61 – 62

61. When implementing the principles set out in the AI White Paper regulatory activity should be focused on here-and-now impacts. Assessing and responding to existential risk should primarily be the responsibility of the UK’s national security apparatus, supported by the AI Safety Institute. (Paragraph 210)

62. Should the acuteness of existential AI risk be judged to have increased, discussions regarding the implications and possible response should take place in international fora, such as AI Safety Summits. (Paragraph 211)

The AI White Paper was published during the last administration and acknowledged the importance to tackle risks with widespread impact. The Government is proposing to establish the appropriate legislation to place requirements on those working to develop the most powerful artificial intelligence models. As part of this, the Government will consult on how best to establish a regulatory regime that will address the most immediate risks.

The CAIRF works closely with the UK national security community while considering AI risks with national security implications.