



Department for
Science, Innovation
& Technology

Rt Hon Michelle Donelan MP
Secretary of State for
Science, Innovation and
Technology
100 Parliament Street
London SW1A 2BQ

The Rt Hon. the Baroness Stowell of Beeston MBE
Chair of Lords Communications and Digital Committee
House of Lords
London
SW1A 0PW

17th April 2024

Dear Baroness Stowell,

Thank you for the Communication and Digital Committee's recent report on Large Language Models. I note the Committee's conclusions and recommendations and have set out the Government's thinking on these issues in the attached report response.

We agree with the Committee that Artificial Intelligence (AI) is introducing significant change to the world and that large language models will introduce transformation comparable to the invention of the internet. As the PM has said "there is nothing in our foreseeable future that will be more transformative for our economies, our societies, and our lives than the development of Artificial Intelligence technology." This technology has huge potential to benefit the UK economy and address societal challenges. This is why we have chosen to introduce a pro-innovation regulatory approach, which we set out in the AI regulation white paper and the AI regulation white paper consultation response. To enjoy these benefits, we recognise that we also need to address the risks that this technology poses. And for this reason we have also placed a strong focus on AI safety.

I am pleased that the Committee supports the Government's overall approach to exploiting the opportunities and mitigating the risks of AI. However, we note the Committee's recommendations that we balance our activities in order to encourage innovation and opportunities, while also assessing and managing risks. Ensuring we get this balance right is a key priority for this Government. It is not only reflected by our regulatory approach, which drives both safety and innovation and builds public trust, but is also demonstrated by our significant investments in the AI ecosystem and the AI Safety Institute. Over the last decade, we've invested over £3.5 billion, including £1.5 billion in compute infrastructure and £100 million in the AI Safety Institute. To ensure that we're well placed to be an AI superpower and a global hub of AI innovation, we've also invested £290 million since 2018 to build a sustainable pipeline of skilled workers. It is also reflected by our work, which I and the Secretary of State for DCMS and I now lead, to look closely at the issues relating to copyright and AI, where we want to develop an approach that allows the AI and creative sectors to grow together in partnership.

We are determined to harness the power of AI, which is why the Prime Minister's Special Adviser on Business and Investment, Lord Petiglas, and I have convened the AI Opportunity Forum. Bringing together leading businesses, we are working together to drive AI adoption across the UK in a bid to boost productivity amongst companies of all sizes and in all sectors, to unlock further economic growth. We have also announced a new £7.4 million AI Upskilling Fund Pilot and a SME Digital Adoption Taskforce, to tackle future skills challenges and boost SME productivity, and the UK economy.

Furthermore, our recent Spring Budget set out our commitment to transforming public services and supporting research centres to progress AI public sector adoption and innovation.

Our pro-innovation pro-safety approach has won praise from both industry leaders and safety experts. From Google DeepMind and OpenAI to Britain's blossoming AI start-up scene, key players endorse this strategy. This success is further evidenced by Microsoft's decision to establish a new AI hub in London, citing the UK's vast pool of talent as a key factor.

I am grateful for the Committee's acknowledgement of the progress we have made, both domestically and as an international leader in AI. Our white paper consultation response sets out how we are proactively engaging with the international landscape to ensure the appropriate degree of cooperation required for effective AI governance. The Government agrees that we should not delay domestic action while awaiting international consensus, which is why we have already empowered our existing expert regulators to be able to address AI risks now.

Finally, I wish to reiterate the Government's pledge made in the Science and Technology Framework - to make the UK a science and technology superpower by 2030.

Yours sincerely,

A handwritten signature in black ink, appearing to read 'Michelle Donelan', with a stylized flourish at the end.

Rt Hon Michelle Donelan MP

Secretary of State

Department for Science, Innovation and Technology

GOVERNMENT RESPONSE – APRIL 2024
HOUSE OF LORDS COMMUNICATIONS AND DIGITAL COMMITTEE
LARGE LANGUAGE MODELS

INTRODUCTION

The Government thanks the Committee for its report, “Large Language Models”, and notes the findings.

We agree with the Committee that Artificial Intelligence (AI) is introducing change to the world and that large language models (LLMs) will introduce transformation comparable to the invention of the internet. This technology has the potential to benefit the economy and meet societal challenges. This is why we chose a pro-innovation regulatory approach, which we set out in the AI regulation white paper and the AI regulation white paper consultation response. To enjoy these benefits, we recognise that we also need to address the risks that this technology poses. And for this reason, we have also placed a strong focus on AI safety.

We are glad that the Committee supports our overall approach to balancing the opportunities and risks of AI and are grateful for the Committee’s acknowledgement of the progress we have made, both domestically and as an international leader in AI.

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

REPORT EXECUTIVE SUMMARY

1. Prepare quickly

The UK must prepare for a period of protracted international competition and technological turbulence as it seeks to take advantage of the opportunities provided by LLMs.

2. Guard against regulatory capture

There is a major race emerging between open and closed model developers. Each is seeking a beneficial regulatory framework. The Government must make market competition an explicit AI policy objective. It must also introduce enhanced governance and transparency measures in the Department for Science, Innovation and Technology (DSIT) and the AI Safety Institute to guard against regulatory capture.

3. Treat open and closed arguments with care

Open models offer greater access and competition but raise concerns about the uncontrollable proliferation of dangerous capabilities. Closed models offer more control but also more risk of concentrated power. A nuanced approach is needed. The Government must review the security implications at pace while ensuring that any new rules support rather than stifle market competition.

4. Rebalance strategy towards opportunity

The Government’s focus has skewed too far towards a narrow view of AI safety. It must rebalance, or else it will fail to take advantage of the opportunities from LLMs, fall behind international competitors and become strategically dependent on overseas tech firms for a critical technology.

5. Boost opportunities

We call for a suite of measures to boost computing power and infrastructure, skills, and support for academic spinouts. The Government should also explore the options for and feasibility of developing a sovereign LLM capability, built to the highest security and ethical standards.

6. Support copyright

The Government should prioritise fairness and responsible innovation. It must resolve disputes definitively (including through updated legislation if needed); empower rightsholders to check if their data has been used without permission; and invest in large, high-quality training datasets to encourage tech firms to use licenced material.

7. Address immediate risks

The most immediate security risks from LLMs arise from making existing malicious activities easier and cheaper. These pose credible threats to public safety and financial security. Faster mitigations are needed in cyber security, counter terror, child sexual abuse material and disinformation. Better assessments and guardrails are needed to tackle societal harms around discrimination, bias and data protection too.

8. Review catastrophic risks

Catastrophic risks (above 1000 UK deaths and tens of billions in financial damages) are not likely within three years but cannot be ruled out, especially as next-generation capabilities come online. There are however no agreed warning indicators for catastrophic risk. There is no cause for panic, but this intelligence blind spot requires immediate attention. Mandatory safety tests for high-risk high-impact models are also needed: relying on voluntary commitments from a few firms would be naïve and leaves the Government unable to respond to the sudden emergence of dangerous capabilities. Wider concerns about existential risk (posing a global threat to human life) are exaggerated and must not distract policymakers from more immediate priorities.

9. Empower regulators

The Government is relying on sector regulators to deliver the white paper objectives but is being too slow to give them the tools. Speedier resourcing of Government-led central support teams is needed, alongside investigatory and sanctioning powers for some regulators, cross-sector guidelines, and a legal review of liability.

10. Regulate proportionately

The UK should forge its own path on AI regulation, learning from but not copying the US, EU and China. In doing so the UK can maintain strategic flexibility and set an example to the world—though it needs to get the groundwork in first. The immediate priority is to develop accredited standards and common auditing methods at pace to ensure responsible innovation, support business adoption, and enable meaningful regulatory oversight.

GOVERNMENT RESPONSE

1. FUTURE TRENDS

Government response to committee findings and recommendations 1 and 2

The Government agrees that AI, including LLMs, will have a transformative impact on our society and economy, both in the UK and across the world. LLMs have the potential to revolutionise the way we interact with information and technology and can be adapted to a wide range of downstream tasks. This includes the generation of content, translation, writing code and automating other tasks that previously required human effort. However, this technology can also present risks and uncertainty, and it is important that LLMs are developed and deployed responsibly, with due consideration of the wider impact they can have on society. That is why the Government is prioritising safety, for example in establishing the Central AI Risk Function (CAIRF) and setting up the AI Safety Institute (AIS), both housed in the Department for Science, Innovation and Technology (DSIT).

The opportunities for the UK are vast – with the UK AI market predicted to grow to over \$1 trillion by 2035¹. Use of AI will be transformational across the UK economy, from healthcare, to finance, to improving efficiency in the public sector. The UK is well placed to take advantage of the opportunities of LLMs, and to lead in global discussions around the mitigation of risks. We have a thriving AI sector, world-leading academic institutions, and well-established expert regulators. The Government is taking action to ensure safe and responsible innovation in this technology.

2. OPEN OR CLOSED MODELS

Government response to committee findings and recommendations 3 – 9

Innovation is at the heart of the Government's regulatory approach, which is enabled by a thriving and competitive market for AI. The Digital Markets, Competition and Consumers Bill will give the Competition and Markets Authority (CMA) the tools it needs to identify and address significant competition issues in AI and other digital markets. The CMA has published an initial review into the consumer and competition implications of AI Foundation Models, and Government will continue to work closely with the CMA in this area.

The Government agrees with the committee that open-source AI will have an important role in the UK economy in promoting competition and innovation. In the AI Regulation white paper consultation response, the Government noted that "openness in AI is, and will continue to be, critical to scientific progress, and we recognise that openness is core to our society and culture." However, there is a balance to strike as we seek to mitigate potential risks associated with open releases of the most powerful AI models with potentially dangerous capabilities. We are working closely with the open-source community to understand their needs and ensure that any future policy interventions seeking to mitigate risk from highly capable models are nuanced, targeted and designed to avoid or minimise negative impacts on valuable open-source activity.

We agree with the Committee that it is important to protect the integrity of our work and guard against any conflicts of interest. In line with DSIT's conflicts of interest policy, AISI requires all individuals joining the AI Safety Institute's Research Unit to declare any conflicts of interest. These conflicts are mitigated in line with the conflicts process agreed by the DSIT Permanent Secretary.

3. A PRO-INNOVATION STRATEGY

Government response to committee findings and recommendations 10 – 22

Overview

In the Government's Science and Technology Framework, we committed to making the UK an AI superpower, so that businesses and the whole of society can benefit from the

opportunities it presents. To do this, we have developed a widely praised pro-innovation regulatory approach and have invested heavily in the AI ecosystem. We are dedicated to building UK AI capabilities and ensuring that UK is the best place in the world to be an AI company or a business looking to harness the power of AI. We are also investing in the UK's science base and AI programmes, to ensure that researchers, start-ups and businesses across the UK can innovate to develop, design, and deploy AI. In total, the Government has invested over £3.5 billion in UK AI capability since 2014, which includes a £1.5 billion investment in compute, £290 million in AI skills and talent initiatives, £100m for the Bridge AI programme to support AI adoption, £370 million of investments into UK AI companies through the British Business Bank, £100 for the Alan Turing Institute, and £100 million for the AI Safety Institute.

Furthermore, the recent Spring Budget set out our commitment to transforming public services and supporting research centres to progress AI public sector adoption and innovation. The Chancellor announced an £800 million reform package that will free-up time for staff at the frontline of public services, including cutting result waiting times in the NHS and slashing admin tasks for the police.

The pace of this change will also bring challenges, and we share the Committee's view that it is critical for the UK to be prepared for this change. The UK hosted the AI Safety Summit in November 2023, in recognition that it was critical to secure international agreement on the opportunities and risks at the frontier of AI. However, our overall approach to AI regulation is one which is designed to be proportionate and pro-innovation.

Regulatory Approach

The Government is committed to delivering a regulatory approach that drives both safety and innovation. We recognise that to reap the benefits of AI, we must manage the risks, and create public trust in AI's safety and reliability, to encourage AI adoption, and therefore innovation and growth. To build that public trust, we need to have the right rules in place. However, it is also important that rules are designed in a way that is not overly burdensome, and is forward looking, adaptable and resilient. That is why the Government has adopted a principles-based regulatory framework, which builds on the strengths of our existing regulators and ensures that regulatory approaches are tailored to the risks and opportunities posed by AI in different sections of our economy.

In the AI regulation white paper consultation response, published in February 2024, the Government set out the case for some highly targeted technology-specific regulation. That highly targeted binding requirements on the developers of the most powerful AI systems, including some LLMs, may be required to effectively mitigate risks. However, LLMs have huge potential to drive growth and innovation, and the Government will not risk stifling this by rushing to legislate before we have a full understanding of the evidence on risks and potential mitigations. In the AI regulation white paper consultation response paper, we therefore set out the key questions regarding the regulation of the most powerful models, which we are exploring with the support of external experts.

AI Safety Institute

The Government agrees with the Committee's recommendation that a diverse range of skills and people is required to strike the right balance on AI. As part of AISI's recruitment strategy, we will continue to assess our research needs, and present capabilities, to understand what expertise we require to deliver our objectives and achieve our mission. AISI is dedicated to building new infrastructure to conduct necessary testing and evaluations of advanced AI. It has the largest technical team of its kind in the world, with 168 cumulative years of frontier AI experience. To date, we have onboarded 30 technical experts with a diverse set of skills and

expertise, including leading figures from industry and academia. AISI is committed to ensuring diversity and inclusion in all our work with this commitment reflected in our hiring practices to help attract a diverse workforce. We will continue to scale AISI's frontier AI capabilities at pace, with a view to build public trust, by appointing a diverse range of sociotechnical experts.

We have partnered with societal researchers and ethicists where we draw on best practice from the wider UK ecosystem including academia and industry. AISI works with the Responsible Tech Adoption Unit (RTAU), who lead the Government's work to enable trustworthy innovation using data and AI. The RTAU brings together policymakers, industry, civil society, and academia – including a vast array of AI ethics experts from different backgrounds, both in the UK and internationally.

Compute

Across 2023, the Government committed £1.5 billion to develop cutting-edge compute infrastructure in the UK, which underpins the development and deployment of AI technologies. This includes investment in an Exascale supercomputer, which will be hosted at the Edinburgh Parallel Computing Centre (EPCC), and the delivery of the AI Research Resource (AIRR). We have invested £300 million in the AIRR, which will comprise of the UK's two fastest supercomputers; Isambard-AI in Bristol and Dawn in Cambridge. These facilities utilise some of the most energy efficient supercomputing technologies available today, such as direct liquid cooling. To ensure that our investments in compute continue to utilise green and renewable sources of energy, we have been actively engaging with the National Grid to ensure that grid connections are supporting the delivery of compute infrastructure.

In the Autumn Statement we confirmed our plans to further expand the AIRR by investing an additional £500 million, to ensure that the UK continues to provide cutting-edge compute to researchers and businesses. We are working with UK Research and Innovation (UKRI) to deliver the AIRR expansion, which will include a competitive procurement process.

Recognising that government needs to work with partners to ensure investments in compute can be sustained, we have been developing partnerships with industry and international counterparts. At the Global Investment Summit in November 2023, Microsoft committed £2.5 billion to build cutting-edge AI infrastructure in the UK. Google has also invested \$1 billion (£790m) into a 33-acre data centre in Hertfordshire. Earlier this year, the Government signed a Memorandum of Understanding with the Canadian Government focused on compute collaboration. We continue to work with industry and governments across the world to grow our partnerships in this area.

Skills and the Labour Market

AI has the potential to be a net creator of jobs and have a positive impact on economic growth. There are over 200,000 vacancies in the UK demanding AI skills and the jobs created by the AI economy are 7% more productive than the average UK job (CBI, 2023). AI has the potential to allow people more time to work on the most rewarding aspects of their roles by alleviating some of the administrative burdens like paperwork or inputting data.

The Department for Education (DfE) has published initial work on the impact of AI on UK jobs, sectors, qualifications, and training pathways. We can be confident that we will need new AI-related skills through national qualifications and training provision. The Government has invested £3.8 billion in higher and further education in this parliament to make the skills system employer-led and responsive to future needs. Along with DfE's Apprenticeships and Skills Bootcamps, the new Lifelong Learning Entitlement reforms, and Advanced British Standard,

will put academic and technical education in England on an equal footing and ensure our skills and education system is fit for the future.

The Government is committed to ensuring that the UK has AI skills at all levels. Since 2018, we have funded a £290 million package of AI skills and talent initiatives to make sure that AI education and awareness is accessible across the UK. This includes funding 24 AI Centres for Doctoral Training, which will train over 1,500 PhD students. UKRI have recently funded a new cohort of 12 UKRI AI Centres for Doctoral Training, ensuring that students will continue to be recruited through such centres annually for the next five years. Inflationary pressures, and an increase to student stipends, have meant that the total number of students which public funds will cover through this route will decrease. However, UKRI are working across all centres to increase private sector support and additional studentships, which we hope will maintain the total number of PhDs funded via this route.

We are also working with Innovate UK and the Alan Turing Institute to develop guidance that sets out the core AI skills people need, from 'AI citizens' to 'AI professionals'. We published draft guidance for public comment in November 2023 and we intend to publish a final version, and a full skills framework, in Spring 2024. Initiatives such as the flexible AI upskilling fund pilot, announced in the Spring Budget, are the latest in a line of measures to support workers to develop the skills they need to use and take advantage of AI. Our current investments are ensuring the UK has a strong pipeline of highly skilled individuals who will have the knowledge to apply AI in specific sectors and drive UK productivity. We will continue to work with the sector, to gain more understanding of how and where additional investments will be required to meet the needs of the UK. For example, through the recently announced joint DSIT-DfE review of AI skills needs, which will examine how to tool the UK's skills system to identify and fill skills gaps as AI technology develops. We have also raised support for university commercialisation to its highest level yet and want to support more UK universities to realise the potential of their research in the real world.

The Independent Review of University Spinout Companies was published in November 2023 and the Government accepted all the Review's recommendations. The Review set out best-practice policies on licensing and equity share, and we are asking universities to review their policies and report against these best-practices by the end of 2024. The Review also identified the need to support spin-out accelerator programmes. The Government is currently mapping the existing landscape of support for spinouts and deep-tech start-ups. Using outputs from the mapping exercise, we will develop proposals to fill any gaps in support and better coordinate existing public funding.

Adoption

The Government is committed to accelerating AI adoption both in the private and public sector. There is considerable work underway to deliver on the UK Science and Technology Framework's 'innovative public sector' commitment, and to support widespread AI adoption in the private sector.

Through the new AI Opportunity Forum, the Government is working with pioneering AI companies to join forces with business leaders to drive wider adoption across the economy. We are also stepping up our plans to accelerate the rollout of AI across the public sector. The Department for Business and Trade (DBT) have been considering how Government can best support SMEs to adopt basic digital technologies which improve firm level productivity and businesses' ability to grow. DBT is setting up an industry-led taskforce to rapidly explore how best to support SMEs to adopt digital technology.

In February 2024, the Cabinet Office published a new framework, which will implement principles for government departments on the responsible use of generative AI. Written in

collaboration with industry, the framework also looks to upskill civil servants through free generative AI courses to ensure public servants have a robust set of skills when working with AI.

We are also taking steps to ensure AI is adopted across the public sector. The Central Digital and Data Office (CDDO) is focused on three core AI priorities which will be underpinned by specific commitments:

1. Developing shared AI infrastructure and solutions for public service challenges.
2. Creating guidance and frameworks for responsible and effective AI adoption in the public sector.
3. Building Public Sector AI skills and capability to deliver at scale.

The Deputy Prime Minister has also announced the creation of an 'Incubator for AI' team (i.AI) in the Cabinet Office, reporting directly to Downing Street. The team comprises of deep technical specialists and AI experts, and will work with CDDO, the Government Digital Service (GDS), DSIT, other government departments, and wider public sector organisations, to deliver meaningful improvements in the short term, while setting the public sector up for success in the longer term.

The current market for LLM tools is immature and is evolving very rapidly. As such, our current assessment is that it is not suitable, and indeed may not be feasible, to develop a sovereign LLM at this stage. The UK government is working to improve open access models for use in government and we will continue to monitor this nascent market and the need for a sovereign LLM, including the not insignificant cost implications of the infrastructure needed to sustain such technology.

4. RISK

Government response to committee findings and recommendations 23 – 36

The Government agrees with the Committee that risks may arise from both open and closed AI models. Mitigating the risks of AI, including those that may be catastrophic, is a responsibility the Government takes very seriously. Mitigating AI risks is a cross-Government effort, with risk owners distributed across many government departments.

Central Functions

In the AI regulation white paper, the Government proposed a set of central functions to bring coherence to the regulatory framework, to monitor and evaluate it, and spot regulator gaps. The purpose of the central function includes risk assessment and monitoring; regulator coordination; regulator capabilities and powers; monitoring and evaluation; support for innovators; and research, engagement and public trust.

DSIT's Central AI Risk Function (CAIRF) is responsible for identifying and assessing AI risks and coordinates the cross-government effort to mitigate those risks, working in partnership with Government departments, agencies, regulators and AI experts. CAIRF provides a single Government view of the risks posed by AI and is a driving force for mitigating those risks across Whitehall. It does this by identifying, assessing, and preparing for AI risks so that the Government can develop effective mitigations.

The team comprises a range of skilled experts from policy officials to lawyers, economists, data scientists, machine learning experts and engineers with experience in AI, risks, and regulation. This interdisciplinary composition ensures we can make the most progress on AI risk and advocate effectively across Government and in industry. Working together with AISI

and cross-government stakeholders, CAIRF looks holistically at the full range of relevant risk factors, including all types of AI systems, technological and societal vulnerabilities across the development and deployment lifecycle, and the impacts downstream.

A crucial part of CAIRF's work is developing and running the AI Risk Register. It is designed to be comprehensive, covering both catastrophic and immediate risks and across the breadth of risks associated with AI, such as example misuse and bias. Catastrophic risks are those that, if not mitigated, could lead to widespread and extreme impacts. CAIRF is working closely with the Cabinet Office to ensure the AI Risk Register is consistent with the National Security Risk Assessment

Engagement on Risk

The Government is carrying out a considerable amount of cross-government work on how to mitigate AI-related risk, which we are doing in partnership with industry. This is central to the CAIRF model. DSIT, in collaboration with National Cyber Security Centre (NCSC), is undertaking work on the cyber security of AI and will be launching a public consultation this spring on a draft voluntary Code of Practice and global standard.

The Government is working extensively to ensure the UK can rapidly respond to threats to our democratic processes, including from AI, through our Defending Democracy Taskforce. The Government regularly meets with social media companies to understand the policies and processes they have in place, and action they will be taking to protect the integrity of the next election. The Online Safety Act will soon make social media platforms legally responsible for removing illegal mis and disinformation and enforcing their terms of service. We are also engaging extensively with broadcasters and news publishers on the impact of AI on journalism, given its vital role in the provision of accurate information and trustworthy news.

The Home Office is responsible for criminal and national security risks such as election security, child sexual abuse, fraud, terrorism, and weakened law enforcement, and the impact that widespread AI adoption has on the proliferation of these risks. The Home Office is working closely with industry, international and wider government stakeholders to develop mitigations, including for disinformation, cyber and novel weaponry risks that they support.

This includes working with industry to evaluate existing detection functions; develop deepfake watermarking tools; assess existing regulatory powers; and embed ethical responsibility through safety-by-design. The Home Office is also working with law enforcement and intelligence agencies to harness the great benefits of AI to prevent criminal misuse and aid the identification of illegal AI-generated material.

The Home Office is working in partnership with DSIT and industry to consider how the technology impacts crime and national security. This includes work to ensure law enforcement retain necessary lawful access to monitor and identify criminal misuse of AI and protect the British public from harm.

Data Protection

Data protection is enshrined in UK law, in the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018 (DPA2018) which implements the EU GDPR in the UK. This sets the statutory framework for the treatment of personal information in the UK. The UK GDPR stipulates in Article 28 what needs to be in a contract between a data controller and a data processor, or between the latter and a sub-contractor. These detailed contractual requirements are already therefore established in law.

The Department for Health and Social Care recognises that ensuring public confidence in new technologies and the use of their data is essential. The Department will work with NHS and public authorities such as the ICO, to ensure, through a proportionate and effective framework of regulation, licensing, contractual and statutory safeguards, that people's data is subject to the highest standards of data protection, and there is transparency as to how this is ensured by emerging technology providers.

AI Safety Institute

In advance of the AI Safety Summit, DSIT published a discussion paper on the capabilities and risks of Frontier AI². AISI is building an in-house capability to evaluate the safety of models. The first stage of this will focus on, as set out in the paper, the cross-cutting risks to

misuse (cyber security, biological, and chemical risks), loss of control of AI and societal harms from AI. This will be enhanced by the International AI Safety report which will build consensus around the latest research and opinions from the world's leading experts, led by Yoshua Bengio. This report will synthesise the best of existing research and insight on advanced AI capability and risks and identify areas of research priority. To ensure a holistic view of the wider risk landscape, the AI Safety Institute is continuing to work closely with the Central AI Risk Function and their established AI Risk Register.

5. INTERNATIONAL CONTEXT AND LESSONS

Government response to committee findings and recommendations 37 – 40

The Government agrees with the Committee on the importance of pursuing a bespoke regulatory approach, tailored to the unique context of the UK. Every country will have their own domestic approach to AI, and what works for the European Union (EU), or United States of America (US) and other countries won't necessarily work for the UK.

However, it is also important to recognise the inherently global nature of AI, meaning effective international coordination will be critical in some areas. The UK is an established global leader in AI and in developing our own approach to AI regulation, we will both influence and respond to international developments. Our white paper consultation response sets out how we will work closely with international partners – building on the work of existing multilateral and multistakeholder initiatives – to ensure the appropriate degree of cooperation required for effective AI governance. This includes forums such as the G7, G20, OECD, Global Partnership on AI (GPAI), United Nations, Council of Europe, and Standards Development Organisations (SDOs), among others. Furthermore, since the publication of the AI Regulation White Paper in March 2023, we have made great strides forward in building bilateral partnerships to further shared interests in AI – signing bilateral agreements with Japan, the United States, Singapore, the Republic of Korea and Canada.

There are a range of issues where international collaboration will be key to a successful and effective international governance regime. This includes crucial work to encourage the development and uptake of global digital technical standards, which the UK is promoting through our AI Standards Hub. It also includes international collaboration on the evaluation of LLMs, in which AISI has a leading international role. AISI is engaging with comparable global Institutes and governments, so that our efforts to ensure the safe and responsible development of frontier AI benefit from international expertise. The UK recently announced a partnership with the United States which will see us work together to develop tests for the most advanced AI models. AISI is also closely engaging with other international partners including the EU, Japan and Republic of Korea, as we develop a network of international

partners. We will continue to drive forward the international conversation- including through future AI Safety Summits - to promote safe, secure, and trustworthy AI, achieve appropriate levels of coherence with other regulatory regimes, and minimise potential barriers to trade.

The Government agrees that we should not delay domestic action while awaiting international consensus, which is why we have already empowered our existing expert regulators to be able to address AI risks now. Alongside our AI Regulation white paper consultation response, I sent letters to regulators that play a critical role in AI. I asked that they publish an update by 30 April 2024, outlining their strategic approach to AI and the steps they are taking in line with the expectations in the AI Regulation white paper. We also continue to act quickly to establish the central function in DSIT, to bring coherence to the regulatory landscape, address regulatory gaps, and ensure a comprehensive understanding of risk.

The Government also agrees that it would be too early to bring forward new primary legislation aimed solely at LLMs. In our response to the AI white paper, we set out that we will not “rush to regulate” while we still do not fully understand the risks or the effectiveness of potential mitigations. Doing so would risk implementing the wrong measures that insufficiently balance addressing risks and supporting innovation. However, we anticipate that all jurisdictions will, in time, want to place targeted binding requirements on the design, development, and deployment of the most capable AI systems to ensure accountability is allocated across the value chain in a way that is effective to mitigate risks.

6. MAKING THE WHITE PAPER WORK

Government response to committee findings and recommendations 41 – 45

Overview

Our response to the AI regulation white paper consultation sets out our progress in establishing the central function in DSIT. Several aspects of the central function have already been established, including CAIRF, which will be consulting in the spring on its risk framework. We are continuing to work with regulators as we develop and refine the supporting roles that the function will play to support the excellent work that regulators are already doing to govern AI safely.

Working with Regulators

We recognise the urgency of establishing the central function, which is why we are building it iteratively to ensure it is robust and will work effectively. We prioritised the establishment of risk assessment and monitoring activities in the first instance and our next steps are to formally establish regulator coordination structures. It is also important to recognise that there has already been significant work to coordinate across the regulatory landscape in relation to AI, through initiatives such as the Information Commissioner’s Office’s working group on AI, the Digital Regulation Cooperation Forum’s joint programme of work, and the collaboration between key financial regulators.

Alongside the establishment of the wider central function, we committed to establishing a steering committee which will support and guide the activities of a formal regulator coordination structure within government. We will be providing an update on the implementation of the central function in the spring. Furthermore, we look forward to key regulators publishing their AI strategies by the end of April; this will drive transparency around the actions regulators are taking to address AI risks and opportunities within their remits and will inform our ongoing regulatory policymaking.

The Government agrees on the importance of ensuring that regulators have the right resources to deliver the regulatory framework. That is why in our AI regulation white paper consultation response, we announced £10 million of funding to prepare and upskill regulators to address the risks and harness the opportunities of this defining technology. The fund will support the work that many regulators are already doing to develop cutting-edge research and practical tools to monitor and address risks and opportunities in their sectors, from telecoms and healthcare to finance and education. However, we recognise that this funding will not be a 'fix all' solution. It supplements existing work such as underway by government including exploring how to give regulators more flexibility on pay and conditions to recruit specialist skills, as recommended by the Government Chief Scientific Advisor, Dame Angela McLean. We also continue to explore and develop additional options to upskill regulators on AI.

We also recognise the need to assess the existing powers and remits of the UK's regulators to ensure they are equipped to address AI risks and opportunities in their domains and implement the principles in a consistent and comprehensive way. In the consultation response, we have committed to work with government departments and regulators to analyse and review potential gaps in existing regulatory powers and remits.

We know that backing regulators with the rights skills and tools is key to ensuring they can implement our regulatory framework and respond to AI risks and opportunities in their remits. That is why, alongside the consultation response, we published initial guidance for regulators to support them to implement our AI regulatory principles in their remits. The guidance sets out practical considerations for regulators, including how they could apply their existing powers in the context of AI and the guidance emphasises the need for regulator collaboration to respond to cross-cutting AI use cases. As set out in the AI regulation white paper consultation response, we will collaborate with regulators to iterate and expand our initial cross-sectoral guidance on implementing the principles, with further updates planned by summer.

Government response to committee findings and recommendations 46 - 53

The Government agrees that while voluntary measures are important in helping us to make AI safer now, the challenges posed by AI will ultimately require legislative action in every jurisdiction once understanding of risk has matured. In light of this, the AI regulation white paper consultation response set out the potential case for future binding measures – such as transparency and accountability – carefully targeted to apply to the developers of highly capable general-purpose AI systems. To effectively tackle the risks that these models can contribute to, we recognise that some binding measures will ultimately be required in the future, and we are working with a wide range of stakeholders to develop our understanding of the risks and effectiveness of mitigations. While voluntary measures help us make AI safer now, the intense competition between companies to release ever-more-capable systems means we will need to remain highly vigilant to meaningful compliance, accountability and effective risk mitigation.

We have been working with regulators and a range of legal experts to explore ways that liability is currently allocated through the AI value chain. The government recognises the need for more work, but our early findings show that accountability is not allocated across the supply chain in a way that leads to effective risk mitigation. Businesses using AI may be 'on the hook' for risks that they can't fully understand or mitigate since those risks emerge further up the supply chain during the training and development of the models. We think that this could be impacting adoption across the economy as well as leaving risks unmitigated. Initial findings suggest that the scope of existing laws and regulatory powers means that it can be hard to hold AI developers to account for some of the risks their models can create or contribute to.

The Government is engaging legal experts to review liability for risks or harms from highly capable general-purpose models. The Government committed in the AI regulation white paper consultation response to continue developing our domestic policy position on AI regulation, including the questions on liability. The Government also committed to publishing an update on our work on new responsibilities for developers of highly capable general-purpose AI systems by the end of the year.

The Government agrees that technical standards, in particular global standards, are key to promoting best practice in technological deployment and development. They ensure products and services are secure, safe, and reliable while also promoting global interoperability. We recently set out our approach to digital standards, including AI standards, in a speech by Minister for Technology and the Digital Economy, Saqib Bhatti. We believe that the multi-stakeholder, industry-led model for standards development is the best way to realise robust standards that are fit for the future.

We are working to strengthen the global digital standards ecosystem and increase the UK's contribution and leadership. This includes initiatives to make standards development and adoption more accessible to stakeholders, including the small businesses who are so vital to our economy, through programmes like the AI Standards Hub. The AI Standards Hub was established by DSIT, Alan Turing Institute, British Standards Institution and National Physical Laboratory in 2022. It is a world first initiative that helps stakeholders navigate and participate in the development of global digital standards for AI.

7. COPYRIGHT

Government response to committee findings and recommendations 54 – 61

The UK has world-leading protections for copyright and intellectual property which are an important enabler for the success of our creative industries, and the Government is committed to ensuring the continuation of a robust copyright framework that rewards human creativity. The basic position under copyright law is that making copies of protected material will infringe copyright unless it is licensed, or an exception applies. However, this is a complex and challenging area, and the interpretation of copyright law and its application to AI models is disputed; both in the UK and internationally.

As set out in the Government's response to the AI Regulation White Paper consultation, the Intellectual Property Office (IPO) engaged stakeholders on copyright and AI as part of a working group aiming to agree a voluntary code. The IPO sought for rights holders and AI companies to come together and agree a way forward on the key issues. However, many of the issues relating to copyright and AI are challenging and the working group was not able to reach a consensus.

The Government is determined to enable our AI and creative industries sectors to continue to thrive together, with creators leveraging advances in AI to drive forward new products, experiences, growth and jobs. We are encouraged to see AI-driven tools being used to support creative processes, in areas such as video and film editing, music composition and generation, and virtual and augmented reality. We believe this demonstrates that it is possible for content owners and AI developers to create value for each other.

AI developers are clear that they must be able to access a wide range of high-quality data to train cutting-edge models in the UK. Rights holders have some concerns about copyright protections in the context of generative AI and have called for assurance that their ability to retain autonomy and control over their valuable work will be protected. The Government is

committed to supporting, not undermining human creativity, and developing an approach to AI and copyright that allows the AI and creative sectors to grow together in partnership.

The Committee's report noted uncertainty over the copyright legal framework and its application, which has in several cases led to litigation. We agree that copyright law should provide certainty and remain fit for the future. However, it would not be appropriate for the Government to comment on ongoing court cases. These cases are for the courts to decide on and must be allowed to conclude independently.

Examples of market-led approaches to data licensing are emerging across the AI and creative sectors, with deals being struck to enable access to high quality data on mutually agreed terms. Collective licensing is one approach that can provide at-scale access to data and can enable creators to be remunerated for their work. The Government is also working with creative and journalism sectors to better understand opportunities for greater access to expanded, high quality data sources at the scales needed for LLM training.

Transparency over the use of copyrighted material to train models and regarding attribution of outputs is important. The Government believes that greater transparency is needed from AI developers in relation to the data used to train LLMs so that rights holders can better understand whether content they produce is used as an input into AI models. We are separately engaging with stakeholders to understand broader perspectives in relation to transparency about the purposes of web crawlers. Transparency over training data has a particular significance in an IP context, but also in a broader safety context, for example, transparency about bias in training datasets. Greater transparency is also needed around the attribution of outputs. In its response to the AI regulation white paper consultation, the Government confirmed that it intends to progress work on transparency – working with both rights holders and AI companies to understand what is technically feasible and what is proportionate.

DSIT, DCMS and the IPO are working collaboratively to build on the progress made through the IPO's working group and voluntary code process. As set out in the DSIT AI white paper consultation response, the Government is working at both official and Ministerial level to further understand views across the creative, press and AI sectors. The Government will also work with international counterparts to understand their emerging approaches and identify opportunities for collaboration.

Footnotes

1. United Kingdom Artificial Intelligence Market, US International Trade Administration, 2023
2. Capabilities and Risks from Frontier AI, Department for Science, Innovation and Technology, 2023