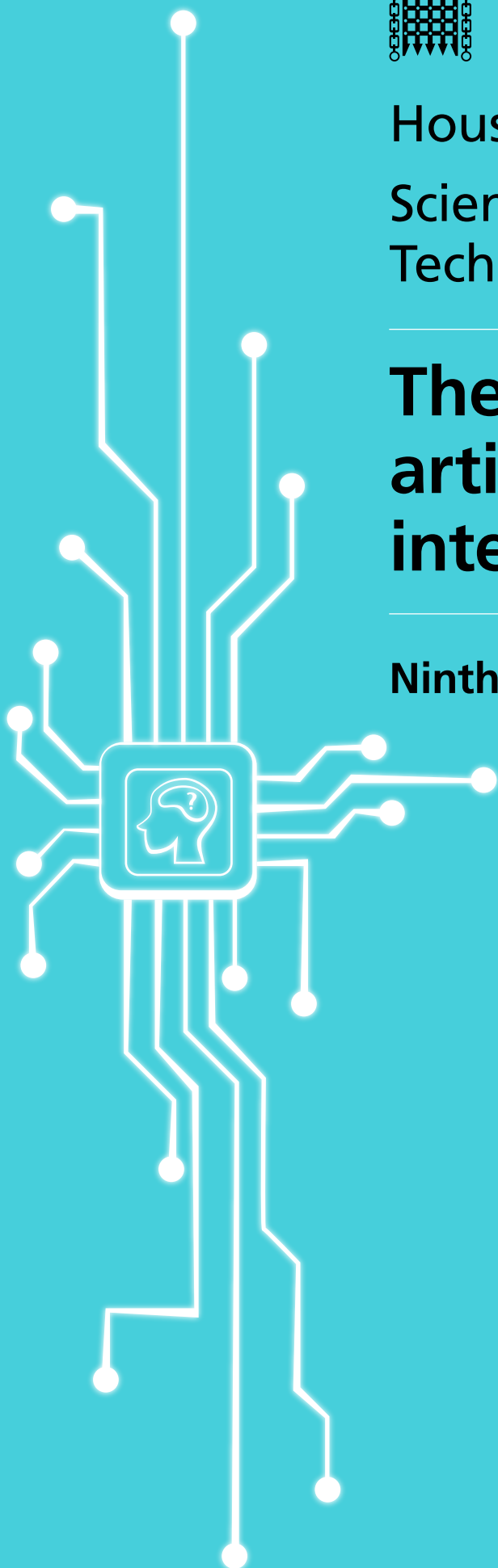




House of Commons
Science, Innovation and
Technology Committee

The governance of artificial intelligence: interim report

Ninth Report of Session 2022–23





House of Commons
Science, Innovation and
Technology Committee

The governance of artificial intelligence: interim report

Ninth Report of Session 2022–23

*Report, together with formal minutes relating
to the report*

*Ordered by the House of Commons
to be printed 19 July 2023*

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

[Greg Clark MP](#) (*Conservative, Tunbridge Wells*) (Chair)

[Aaron Bell MP](#) (*Conservative, Newcastle-under-Lyme*)

[Dawn Butler MP](#) (*Labour, Brent Central*)

[Chris Clarkson MP](#) (*Conservative, Heywood and Middleton*)

[Tracey Crouch MP](#) (*Conservative, Chatham and Aylesford*)

[Katherine Fletcher MP](#) (*Conservative, South Ribble*)

[Rebecca Long-Bailey MP](#) (*Labour, Salford and Eccles*)

[Stephen Metcalfe MP](#) (*Conservative, South Basildon and East Thurrock*)

[Carol Monaghan MP](#) (*Scottish National Party, Glasgow North West*)

[Graham Stringer MP](#) (*Labour, Blackley and Broughton*)

[Christian Wakeford MP](#) (*Labour, Bury South*)

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

© Parliamentary Copyright House of Commons 2023. This publication may be reproduced under the terms of the Open Parliament Licence, which is published at www.parliament.uk/site-information/copyright-parliament.

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

Committee staff

The current staff of the Committee are: Jessica Bridges-Palmer (Senior Select Committee Media Officer), Ian Cruse (Second Clerk), Arvind Gunnoo (Committee Operations Officer), Dr Claire Housley (Committee Specialist), Dr Claire Kanja (Committee Specialist), Dr Faten Hussein (Committee Team Leader (Clerk)), Dr Joshua Pike (Committee Specialist), Hafsa Saeed (Committee Operations Manager), and Ben Shave (Committee Specialist)

The following staff also worked for the Committee during this inquiry:
Gina Degtyareva (Former Senior Select Committee Media Officer)

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 inquiries is: 020 7219 2793; the Committee's e-mail address is: commonsitc@parliament.uk.

You can follow the Committee on Twitter using [@CommonsSITC](https://twitter.com/CommonsSITC).

Contents

Summary	3
1 Introduction	5
Our inquiry	5
Aims of this interim Report	5
2 A general-purpose technology	6
Foundation models and generative AI	6
A new technology?	6
Capability and the rate of development	7
3 Benefits	9
Everyday applications	9
Medicine and healthcare	9
Healthcare provision	9
Medical research	11
Education	11
Delivering future benefits	13
4 Twelve Challenges of AI Governance	14
1: The Bias challenge	14
2: The Privacy challenge	15
3: The Misrepresentation challenge	17
4: The Access to Data challenge	18
5: The Access to Compute challenge	18
6: The Black Box challenge	19
7: The Open-Source challenge	20
8: The Intellectual Property and Copyright Challenge	20
9: The Liability challenge	22
10: The Employment challenge	22
11: The International Coordination challenge	23
12: The Existential challenge	24
5 The UK Government approach to AI	26
The AI white paper	26
A UK-specific approach	27
The Digital Regulation Cooperation Forum	28
Foundation Model Taskforce	29

Chapter 6: The international dimension	31
The European Union	31
The United States	32
International coordination	33
6 Chapter 7: Conclusion and next steps	34
Conclusions and recommendations	35
Formal minutes	39
Witnesses	40
Published written evidence	42
List of Reports from the Committee during the current Parliament	46

Summary

Artificial intelligence (AI) has been the subject of public, private and research sector interest since the 1950's. However, since the emergence of so-called 'large language models' such as ChatGPT in particular, it has become a general-purpose, ubiquitous technology—albeit not one that should be viewed as capable of supplanting humans in all areas of society and the economy.

AI models and tools are capable of processing increasing amounts of data, and this is already delivering significant benefits in areas such as medicine, healthcare, and education. They can find patterns where humans might not, improve productivity through the automation of routine processes, and power new, innovative consumer products. However, they can also be manipulated, provide false information, and do not always perform as one might expect in messy, complex environments—such as the world we live in.

The recent rate of development has made debates regarding the governance and regulation of AI less theoretical, more significant, and more complex. It has also generated intense interest in how public policy can and should respond to ensure that the beneficial consequences of AI can be reaped whilst also safeguarding the public interest and preventing known potential harms, both societal and individual. There is a growing imperative to ensure governance and regulatory frameworks are not left irretrievably behind by the pace of technological innovation. Policymakers must take measures to safely harness the benefits of the technology and encourage future innovations, whilst providing credible protection against harm.

Our inquiry so far has led us to identify twelve challenges of AI governance, that policymakers and the frameworks they design must meet.

- 1) **The Bias challenge.** AI can introduce or perpetuate biases that society finds unacceptable.
- 2) **The Privacy challenge.** AI can allow individuals to be identified and personal information about them to be used in ways beyond what the public wants.
- 3) **The Misrepresentation challenge.** AI can allow the generation of material that deliberately misrepresents someone's behaviour, opinions or character.
- 4) **The Access to Data challenge.** The most powerful AI needs very large datasets, which are held by few organisations.
- 5) **The Access to Compute challenge.** The development of powerful AI requires significant compute power, access to which is limited to a few organisations.
- 6) **The Black Box challenge.** Some AI models and tools cannot explain why they produce a particular result, which is a challenge to transparency requirements.
- 7) **The Open-Source challenge.** Requiring code to be openly available may promote transparency and innovation; allowing it to be proprietary may concentrate market power but allow more dependable regulation of harms.

- 8) **The Intellectual Property and Copyright Challenge.** Some AI models and tools make use of other people’s content: policy must establish the rights of the originators of this content, and these rights must be enforced.
- 9) **The Liability challenge.** If AI models and tools are used by third parties to do harm, policy must establish whether developers or providers of the technology bear any liability for harms done.
- 10) **The Employment challenge.** AI will disrupt the jobs that people do and that are available to be done. Policy makers must anticipate and manage the disruption.
- 11) **The International Coordination challenge.** AI is a global technology, and the development of governance frameworks to regulate its uses must be an international undertaking.
- 12) **The Existential challenge.** Some people think that AI is a major threat to human life: if that is a possibility, governance needs to provide protections for national security.

In March 2023, the UK Government set out its proposed “pro-innovation approach to AI regulation” in the form of a white paper. It set out five principles to frame regulatory activity, guide future development of AI models and tools, and their use. These principles would not initially be put on a statutory footing but interpreted and translated into action by individual sectoral regulators, with assistance from central support functions.

The UK has a long history of technological innovation and regulatory expertise, which can help it forge a distinctive regulatory path on AI. The AI white paper should be welcomed as an initial effort to engage with this complex task, but its proposed approach is already risking falling behind the pace of development of AI. This threat is made more acute by the efforts of other jurisdictions, principally the European Union and United States, to set international standards.

Our view is that a tightly-focussed AI Bill in the next King’s Speech would help, not hinder, the Prime Minister’s ambition to position the UK as an AI governance leader. Without a serious, rapid and effective effort to establish the right governance frameworks—and to ensure a leading role in international initiatives—other jurisdictions will steal a march and the frameworks that they lay down may become the default even if they are less effective than what the UK can offer.

A summit on AI safety, expected in November or December, will also be key to delivering the Prime Minister’s ambition. The challenges highlighted in our interim Report should form the basis for discussion, with a view to advancing a shared international understanding of the challenges of AI—as well as its opportunities. Invitations to the summit should therefore be extended to as wide a range of countries as possible. A forum should also be established for like-minded countries who share liberal, democratic values, to ensure mutual protection against those actors—state and otherwise—who are enemies of these values and would use AI to achieve their ends.

1 Introduction

1. The rapid advancement of artificial intelligence (AI) has ushered in a new era of transformative technologies with far-reaching implications for society. As AI permeates various aspects of our lives, concerns regarding its governance and ethical considerations have become increasingly pertinent. This select committee report delves into the multifaceted landscape of AI governance, aiming to provide a comprehensive analysis of the existing frameworks, regulations, and ethical guidelines governing this powerful technology. By examining the benefits, risks, and potential social impacts, this report seeks to inform policymakers, stakeholders, and the public about the urgent need for a robust and transparent AI governance framework that upholds human rights, accountability, and societal well-being.

2. The above paragraph was authored not by the Science, Innovation and Technology Committee but by ChatGPT, using a simple prompt: write a 100-word introduction to a select committee report examining the governance of artificial intelligence. It captures some key themes that have emerged during our inquiry to date, and illustrates the fact that AI is now a general-purpose, ubiquitous technology—but, the above paragraph shows, not yet a perfect substitute for the way things are done now.

3. The recent rate of development has made debates regarding the governance and regulation of AI less theoretical, more significant, and more complex. We have therefore decided to publish an interim Report to outline our initial findings. Our inquiry continues and a further Report will be published in due course.

Our inquiry

4. We launched our inquiry on 20 October 2022, to examine: the impact of AI on different areas of society and the economy; whether and how AI and its different uses should be regulated; and the UK Government's AI governance proposals. We have received and published over 100 written submissions and taken oral evidence from 24 individuals, including AI researchers, businesses, civil society representatives, and individuals affected by this technology. We are grateful to everyone who has contributed to our inquiry so far.

Aims of this interim Report

5. This interim Report examines the factors behind recent AI developments, highlights the benefits offered by the technology, and identifies a series of challenges for policymakers. We examine how the UK Government has responded, and how this compares to other countries and jurisdictions.

- In Chapter 2, we consider the general-purpose nature of AI.
- In Chapter 3, we outline the benefits and risks of AI for two areas of society and the economy: medicine and healthcare, and education.
- In Chapter 4, we suggest challenges for policymakers that AI has created.
- In Chapter 5, we examine the UK Government's approach to AI.
- In Chapter 6, we consider the international dimension of AI governance.
- Finally, in Chapter 7, we outline the next steps for our inquiry.

2 A general-purpose technology

6. Artificial intelligence (AI), a broad term with no universally agreed definition,¹ has been discussed and debated since at least 1950, when Alan Turing posed a now-famous question: can machines think?² In this Chapter, we will highlight notable recent breakthroughs, consider some of the forces that have propelled this rapid rate of development, and the implications.

Foundation models and generative AI

7. AI development has focused increasingly on the ‘training’ and deployment of “... large, costly, wide-capability foundation models or general purpose AI systems (such as OpenAI’s GPT-3 or Google’s PaLM) which are then tailored to (or ‘finetuned’ for) particular tasks and application areas”.³ Terms such as foundation models, generative AI and large language models are often used interchangeably to refer to these models and tools, but they can broadly be defined as “... AI that is able to use tech in the place of a human being”.⁴

8. There is a growing number of these models and tools, but it is ChatGPT, launched in November 2022, that has sparked a global conversation. Hugh Milward, General Manager, Corporate, External and Legal Affairs at Microsoft UK, told us, shortly after Microsoft announced a “multibillion dollar” investment in ChatGPT developer OpenAI,⁵ that recent events had opened the door to “... a new industrial revolution”.⁶ The same analogy was also used by former UK Government Chief Scientific Adviser, Sir Patrick Vallance.⁷

A new technology?

9. Hugh Milward was among the contributors to our inquiry who pointed out that AI is not a new technology.⁸ In 1956, researchers convened at Dartmouth College, New Hampshire, to examine “... how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves”.⁹ Over the following decades, AI’s potential has been a focus for researchers, technology firms, and investors, as Jen Gennai, Director (Responsible Innovation) at Google, confirmed to us:

... it has had its ups and downs... it got unpopular for a while because people did not see where AI could help to solve some of the real problems and opportunities for economic development, commercial development or otherwise. We are now seeing more of that potential.¹⁰

1 [Qq. 2–3](#)

2 Alan Turing, “[Computing machinery and intelligence](#)”, *Mind*, vol. 59 (1950), pp 433–460

3 Mr Haydn Belfield, Dr Seán Ó hÉigeartaigh, Dr Shahar Avin, Giulio Corsi (University of Cambridge) and Prof José Hernández-Orallo (Universitat Politècnica de València) ([GAI0094](#))

4 [Q5](#)

5 [Microsoft and OpenAI extend partnership](#), Microsoft, 23 January 2023

6 [Q97](#)

7 [Q47](#)

8 [Q97](#)

9 [A proposal for the Dartmouth summer research project on artificial intelligence](#), J. McCarthy, M.L. Minsky, N. Rochester and C. Shannon, 31 August 1955

10 [Q98](#)

Capability and the rate of development

10. Professor Michael Osborne, Professor of Machine Learning at the University of Oxford, said that the capabilities of AI models and tools remained limited, and described them as “... very far from human-level intelligence, and even ChatGPT and similar large language models... still have really significant gaps in their understanding of the complexities of the real world”.¹¹

11. Professor Mihaela van der Schaar, Professor of Machine Learning, Artificial Intelligence and Medicine at the University of Cambridge, also suggested that the real world held challenges for models and tools capable of solving complex problems in more rigid environments such as board games: “... if we take something like the NHS, it is complex, messy data... the environment is changing according to rules we cannot really predict”.¹²

12. The ability to alter the performance of tools such as ChatGPT was also highlighted to us. Professor Sir Nigel Shadbolt, Professorial Research Fellow in Computer Science at the University of Oxford, described how “... ChatGPT is the thing you get in front of you, but there are lots of ways of getting into and behind those models and changing their behaviour, for sure”.¹³ AI models and tools can also ‘hallucinate’, offering incorrect answers,¹⁴ disseminating misinformation¹⁵ or revealing private data.¹⁶

13. The rate of development has nevertheless been notable and, as Sir Patrick Vallance told us, surprising:

I think everyone has been surprised by how much the large generative models have done things that people did not expect them to do. That is what is intriguing about it—very large datasets, very high compute power, and those models are turning out things that even people very close to the field thought, “Actually, I wasn’t sure it was going to do that”.¹⁷

14. Professor Osborne cautioned that “predicting the future is a mug’s game”, but also said that there had been “... repeated cases of the technology vastly exceeding what we had reasonably expected to be done” and developments that “... massively improved on what people thought would be possible in short order”.¹⁸

15. While AI is not a new technology, the rapidly acquired ubiquity of tools such as ChatGPT and the rate of development has come as a surprise to even well-informed observers. We are all now interacting with AI models and tools daily, and we are increasingly aware of these interactions.

11 [Qq. 9, 11](#)

12 [Q227](#)

13 [Q168](#)

14 [Disinformation Researchers Raise Alarms About A.I. Chatbots](#), New York Times, 8 February 2023

15 [ChatGPT is making up fake Guardian articles. Here’s how we’re responding](#), The Guardian, 6 April 2023

16 [Nvidia’s AI software tricked into leaking data](#), Financial Times, 9 June 2023

17 Oral evidence taken on 3 May 2023, HC (2022–23) 1324, [Q46](#)

18 [Qq. 4–5](#)

16. Nevertheless, the technology should not be viewed as a form of magic or as something that creates sentient machines capable of self-improvement and independent decisions. It is akin to other technologies: humans instruct a model or tool and use the outputs to inform, assist or augment a range of activities.

3 Benefits

17. The emergence of AI as a general-purpose, ubiquitous technology has affected many areas of society and the economy. In this Chapter we will consider the integration of AI models and tools into everyday devices, and the benefits it offers in two vital policy areas: medicine and healthcare, and education.

Everyday applications

18. AI models and tools are already widely used in consumer products such as smartphones, satnavs, and streaming service recommendations.¹⁹ Recently companies such as Google and Microsoft have announced a series of integrations into new and existing products, such as Bard,²⁰ Google Search,²¹ Bing,²² and Microsoft 365 Copilot—all with a view to increasing productivity.²³

19. Our inquiry has coincided with a period of intense competition to deliver AI-centred announcements across different sectors. Adrian Joseph, Chief Data and AI Officer at BT Group told us that “... we have been in that race for a very long time. This is not new. The big tech companies have been acquiring start-ups and investing in their own expertise for 10, if not 20, years”.²⁴ Below, we will consider two areas of society and the economy that have already benefited from the technology.

Medicine and healthcare

20. Medicine and healthcare is often said to be particularly well-placed to benefit from the use of AI models and tools—in the 2023 AI white paper, improvements in NHS medical care are listed among the key societal benefits.²⁵ We have heard that AI is already delivering benefits and has further potential in two areas: healthcare provision and medical research.

Healthcare provision

Diagnostics

21. AI can be used in healthcare as a diagnostic tool, capable of processing data and predicting patient risks.²⁶ Dr Manish Patel, CEO of Jiva.ai, described how his firm developed an algorithm to recognise potential prostate cancer tissue from MRI scans.²⁷ The Department of Health and Social Care has also invested in projects focused on

19 [A pro-innovation approach to AI regulation](#), GOV.UK, 29 March 2023

20 [What’s ahead for Bard: More global, more visual, more integrated](#), Google, 10 May 2023

21 [An important next step on our AI journey](#), Google, 6 February 2023

22 [Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web](#), Microsoft, 7 February 2023

23 [Introducing Microsoft 365 Copilot – your copilot for work](#), Microsoft, 16 March 2023

24 [Q108](#)

25 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 1

26 [Q251](#)

27 UK BioIndustry Association ([GAI0026](#))

using AI to detect other forms of cancer,²⁸ and established an AI Diagnostic Fund “... to accelerate the deployment of the most promising AI imaging and decision support tools” across NHS Trusts.²⁹

22. Dr Patel told us that a key advantage was the speed at which these tools could help medical professionals reach a diagnosis, avoiding longer waits and the associated emotional and financial costs.³⁰ Professor Delmiro Fernandez-Reyes, Professor of Biomedical Computing at University College London, said that it could also help relieve pressure on medical personnel, by augmenting their work, speeding up referrals and preventing diseases from worsening.³¹

23. Dr Patel noted that there is “... a very high barrier to entry” for companies offering AI diagnostic tools to healthcare providers, owing to the need for sufficiently representative training datasets, and a robust regulatory framework intended to ensure such tools are deployed safely.³² Professor Mihaela van der Schaar of the University of Cambridge pointed to a longstanding trend of bias in medical data: “... at times the data we collect—not only at one point, but over time as interventions are made—is biased”.³³ Dr Patel also said that the inherent bias within medical formulas such as the Body Mass Index³⁴ could not be automatically out-programmed by such tools.³⁵ Given this, he told us that the technology should be viewed as a way to augment rather than replace human expertise: “I don’t see that changing in the next 10 years. I think the medical community has to be confident that this technology works for them, and that takes time and it takes evidence”.³⁶

Improving existing processes

24. AI models and tools can also deliver benefits via the automation of existing processes—“... doing the dirty work” of improving logistics, as Professor Mihaela van der Schaar of The University of Cambridge phrased it.³⁷ She described how during the covid-19 pandemic tools were developed “... to predict how many beds and ventilators we would need and who would need them” and argued that pursuing similar efficiencies should be the primary use case for AI in medicine and healthcare.³⁸

25. Professor Michael Osborne of the University of Oxford described AI models and tools as “... a way to automate away much of the tedious admin work that plagues frontline workers in the NHS today, particularly in primary healthcare”, and said that the technology could help medical professionals process letters and manage data.³⁹

28 [£36 million boost for AI technologies to revolutionise NHS care](#), GOV.UK, 16 June 2021

29 [£21 million to roll out artificial intelligence across the NHS](#), GOV.UK, 23 June 2023

30 [Q65](#)

31 [Q231](#)

32 [Qq. 57, 63](#)

33 [Q259](#)

34 [An article in the British Journal of General Practice, published in 2010](#), described how the inventor of the BMI, Adolphe Quetelet, created the formula: “Adolphe Quetelet’s interest in the emerging discipline of statistics in the mid 1830s saw him collect data on men’s heights and weights at various ages. From this study, which he hoped would allow him to determine the ‘average’ man, he formulated what became known as the Quetelet formula, but which is now known as the BMI”.

35 [Q80](#)

36 [Q64](#)

37 [Q238](#)

38 [Qq. 238, 258](#)

39 [Q22](#)

26. AI models and tools can transform healthcare provision, by assisting with diagnostics and, perhaps more significantly, freeing up time for the judgement of medical professionals by automating routine processes.

Medical research

27. Our inquiry has also heard how AI models and tools can help deliver breakthroughs in medical research, such as drug discovery. Dr Andrew Hopkins, Chief Executive of Exscientia, a ‘pharmatech’ company,⁴⁰ described to us how it used the technology to “... design the right drug and select the right patient for that drug”, and how this allowed for a complexity of analysis beyond the cognitive and computational capabilities of human researchers.⁴¹ He said that this analysis could be applied to new drugs and existing drugs that had previously failed to pass clinical trials on efficacy grounds, with a view to repurposing them.⁴²

28. The global pharmaceutical company GSK described the impact of AI models and tools on medical research in similarly positive terms, and said that “... ultimately, AI will provide greater probability that the discovery and development of new medicines will be successful”.⁴³

29. The ability of AI models and tools to process substantial volumes of data, and rapidly identify patterns where human researchers might take months or be unable to, makes it a potentially transformational technology for medical research. Either through the development of new drugs, or the repurposing of existing ones, the technology could reduce the investment required to bring a drug to market; and bring personalised medicine closer to becoming a reality.

Education

30. Following the launch of ChatGPT in November 2022, its implications for education have been widely discussed, particularly its use by students. Joel Kenyon, a science teacher at a secondary school in London, told us that whilst he could not quantify the extent to which ChatGPT and other similar tools were being used by pupils, “... they are using it. There is no two ways about it”.⁴⁴ Daisy Christodoulou, a former English teacher and current Director of Education at No More Marking, an education software provider, said that there was “... an uneven distribution. There are some teachers and some students who have started using it a lot, and there are some who have not heard of it. It is spreading”.⁴⁵

31. An AI-accelerated shift towards personalised learning was also highlighted to us as a tangible benefit. Professor Rose Luckin, Professor of Learner Centred Design at University College London, said that there was evidence to show that “... students who might have been falling through the net can be helped to be brought back into the pack” with the help

40 [Exscientia: our mission](#), accessed 27 June 2023

41 [Qq. 174, 177](#)

42 [Qq. 180–183](#)

43 [GSK \(GAI0067\)](#)

44 [Q309](#)

45 [Q272](#)

of personalised AI tutoring tools.⁴⁶ The Prime Minister, the Rt. Hon. Rishi Sunak MP, told the Liaison Committee that “tutoring in the physical sense is hard to scale, but the technology allows us to provide that, and I think that would be transformational”.⁴⁷

32. We also heard from Mr Kenyon about how AI tools were useful time-savers in everyday tasks undertaken by teachers.⁴⁸ Ms Christodoulou said tools such as ChatGPT were particularly suited to certain tasks, such as text summarising: “If you take a complexish text where you would vouch for the accuracy—Wikipedia is a good example—pop it into ChatGPT and say, “Can you rewrite this so that it is appropriate for a 10- year-old?”, it does quite a good job”.⁴⁹

33. We heard different perspectives on the potential longer-term implications. Daisy Christodoulou recommended “... a good, hard look at how we assess. I think that ChatGPT has huge implications for continuous assessment course work. It is very hard to see how that continues”. She also highlighted ChatGPT’s ability to generate “... university-level essays. The point is that even if it can produce them only at the 50 to 60 percentile, by definition that will be good enough for 50% to 60% of students”.⁵⁰

34. Professor Luckin, however, was more positive. She argued that the emergence of AI, and its ability to compile information more efficiently than humans, had created an opportunity to move away from an information-based curriculum:

“... what are the real characteristics of human intelligence that we want our populations to have? Surely, they are not the ones that we can easily automate; surely, they are the ones that we cannot easily automate”.⁵¹

35. Dr Matthew Glanville, Head of Assessment Principles and Practice at the International Baccalaureate, said that the education qualification provider required its students to cite AI-generated content “... as they would reference any other material they have taken from a different source”.⁵² He said that AI would become “... part of our everyday lives... We really need to make sure that we support our students in understanding what ethical and useful approaches are”.⁵³

36. AI tools are already useful time-savers for education professionals, and whilst reliable data is hard to come by, it seems highly likely that the technology is this generation of students’ calculator or smartphone.

37. The benefits for time-pressed teachers using AI models and tools to help prepare lesson plans are clear, and increased availability of personalised learning and tutoring tools could benefit many pupils. However, widespread use of AI raises questions about the nature of assessment, particularly in subjects that rely heavily on coursework.

46 [Qq. 278–280](#)

47 [Oral evidence taken before the Liaison Committee on 4 July 2023, HC \(2022–23\) 1602, Q16 \(the Prime Minister\)](#)

48 [Qq. 296, 298](#)

49 [Q273](#)

50 [Q281](#)

51 [Q283](#)

52 [Q300](#)

53 [Q311](#)

38. *Education policy must prioritise equipping children with the skills to succeed in a world where AI is ubiquitous: digital literacy and an ability to engage critically with the information provided by AI models and tools.*

Delivering future benefits

39. Our inquiry has highlighted potential future benefits of AI across different areas of society and economic activity. Climate change,⁵⁴ antibiotic resistance,⁵⁵ the transition to driverless vehicles,⁵⁶ and the development of fossil fuel alternatives⁵⁷ are some of the global challenges that we have heard could be addressed using AI models and tools.

40. **The wide range of potential applications, and associated benefits, reflects the general-purpose nature of AI. As with previous technological innovations, the challenge for policymakers is translating this potential into reality, in a safe and sustainable way.**

54 [Q97](#)

55 [Using AI, scientists find a drug that could combat drug-resistant infections](#), MIT News, 25 May 2023

56 [Wayve \(GAI0061\)](#)

57 [UK BioIndustry Association \(GAI0026\)](#)

4 Twelve Challenges of AI Governance

41. The rapid development and deployment of AI models and tools has led to intense interest in how public policy can and should respond to ensure that the beneficial consequences of AI can be reaped at the same time as the public interest is safeguarded and, specifically, potential harms to individuals and society are prevented. In the UK, the Government published in March 2023 a white paper outlining its “pro-innovation approach to AI regulation”.⁵⁸ In the European Union a new legislative instrument, the Artificial Intelligence Act, has reached the final stage of development, negotiations between the Council of Ministers and the European Parliament.⁵⁹ Proposals for legislative action have also been put forward by United States Senate Majority Leader, Chuck Schumer.⁶⁰

42. In many jurisdictions—including our own—there is a sense that the pace of development of AI requires an urgent response from policymakers if the public interest is not to be outstripped by the pace of deployment. This is reinforced by the perception that the explosion of social media over the last 20 years took place before serious and coherent steps were taken to counteract harms—resulting in, for example, the Online Safety Bill still making its way through Parliament in 2023.⁶¹ By contrast, a more successful experience can be seen in the governance of other new fields of technology, such as the regulation of human fertilisation embryology in the UK following the Warnock Report in 1984.⁶²

43. Holding back a coherent policy response to AI is the reality that the optimal responses to all of the challenges AI gives rise to are not always—at this stage—obvious. So there is a growing imperative to accelerate the development of public policy thinking on AI so that it is not left irretrievably behind by the pace of technological innovation.

44. In this Chapter, based on evidence that we have taken before our inquiry, we set out twelve challenges that the governance of AI must meet. We say that the UK Government in responding to this Report and to its own white paper must set out how it will address each of these challenges.

1: The Bias challenge

AI can introduce or perpetuate biases that society finds unacceptable

45. Researchers and developers are reliant on data to test, train, operate and refine AI models and tools.⁶³ Professor Michael Osborne, a Professor of Machine Learning at the University of Oxford, told us that as datasets are compiled by humans, they contain

58 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023

59 [Artificial intelligence: in Europe, innovation and safety go hand in hand | Statement by Commissioner Thierry Breton](#), European Commission, 18 June 2023

60 [Sen. Chuck Schumer launches SAFE innovation in the AI age at CSIS](#), Center for Strategic and International Studies, accessed 27 June 2023

61 [A guide to the Online Safety Bill](#), GOV.UK, 16 December 2022

62 [Report of the Committee of Inquiry into Human Fertilisation and Embryology](#), Wellcome Collection, accessed 17 July 2023

63 Creative Commons ([GAI0015](#)), Institution of Engineering and Technology ([GAI0021](#))

inherent bias. He said that it was “... an illusion to think that data is neutral and objective”.⁶⁴ Adrian Joseph, Chief Data and AI Officer at BT Group, said there was “... bias in the data, in the algorithms and in the individuals that are creating some of the algorithms”.⁶⁵

46. The risks of encoding bias into AI models and tools are clear. Creative Commons, a non-profit organisation, said that if left unchecked it could “... replicate biases in society against minority and underrepresented communities, and lead to discrimination in critical areas affecting people’s lives...”.⁶⁶ If the bias in datasets used to train AI models and tools is not accounted for and addressed, they “... will faithfully reproduce that bias”, Professor Osborne told us.⁶⁷

47. Examples of where bias could have a particularly negative effect include recognition disparities between ethnic backgrounds by facial recognition tools used by law enforcement,⁶⁸ employment tools that associate women’s names with traditionally female roles,⁶⁹ the spread of politically-motivated disinformation and perpetuation of biased worldviews,⁷⁰ and racial bias in insurance pricing.⁷¹

2. The Privacy challenge

AI can allow individuals to be identified and personal information about them to be used in ways beyond what the public wants.

48. We have already outlined the reliance of AI models and tools on data, and some of the challenges this presents. A related privacy challenge also applies. Michael Birtwistle, Associate Director for AI and Data Law and Policy at the Ada Lovelace Institute, and a former UK Government official, told us that regardless of the sector, privacy should be “... an integral part of the balance of interests that you consider when you are deploying artificial intelligence”.⁷²

49. A balance between the protection of privacy and the potential benefit of deploying AI models and tools is particularly important in certain use cases, such as law enforcement. Lindsey Chiswick, Director of Intelligence at the Metropolitan Police, told us that it currently used two facial recognition technology techniques using NEC software:

... the AI element of that is where the algorithm is doing that biometric matching. Essentially, it is looking at the watchlist—looking at each image and taking a set of measurements of the face; that is the biometric bit—and comparing it to the measurements of the face of the image that the cameras then pick up.⁷³

64 [Q13](#)

65 [Q120](#)

66 Creative Commons ([GAI0015](#))

67 [Q10](#)

68 Liberty ([GAI0081](#))

69 Sage Group ([GAI0108](#))

70 [Q103](#)

71 Connected by Data ([GAI0052](#))

72 [Q538](#)

73 [Qq. 412, 422](#)

50. Ms Chiswick said that facial recognition technology was deployed on the basis of specific intelligence and that images were compared against either a bespoke watchlist unique to a particular deployment (for live facial recognition, or LFR) or images on the Police National Database (in the case of retrospective facial recognition).⁷⁴ She said that in 2023, the technology had been used six times, including during the coronation of King Charles III:

There were four true alerts. There were zero false alerts throughout those six deployments, and there have been two arrests made. The others were correct identifications but it was decided that arrest was not a necessary action in the circumstances.⁷⁵

51. Dr Tony Mansfield, Principal Research Scientist at the National Physical Laboratory (NPL) described to us how an evaluation of the accuracy of the facial recognition software used by the Metropolitan Police, undertaken by the NPL,⁷⁶ found that it could vary, depending on the ‘face-match threshold setting’ used:

We find that, if the system is run at low and easy thresholds, the system starts showing a bias against black males and females combined. There is some evidence for that—if the system is operated at certain thresholds, which I believe are outside those that the Met police has been deploying.⁷⁷

Big Brother Watch, a campaigning organisation, has argued that the NPL study provided evidence of “serious demographic accuracy bias”, in the form of:

... a statistically significant difference between the false positive rate of black and non-black subjects... specifically when the confidence threshold to generate a match was set below 0.6... Documents seen by Big Brother Watch show that the Met Police has frequently operated LFR below a 0.6 confidence threshold and set the threshold as low as 0.55, in 2017 and 2018, while its LFR policy suggests that the threshold is variable.⁷⁸

52. Whilst Ms Chiswick outlined the operational benefits, concerns were raised by other contributors to our inquiry. Big Brother Watch told us that whilst individual rights were protected by the Human Rights Act 1998, Data Protection Act 2018 and the Equality Act 2010, “... we have often found systems... which do not adequately respect the rights of individuals as set out by these pieces of legislation—for example, police forces’ use of live facial recognition”.⁷⁹ Michael Birtwistle of the Ada Lovelace Institute told us that in the absence of a comprehensive regulatory framework “... there is not a guarantee that facial recognition systems deployed by police will meet reasonable standards of accuracy or that their use will remain proportionate to the risks presented by them”.⁸⁰

74 [Qq. 412–414](#)

75 [Q452](#)

76 [Facial recognition technology in law enforcement equitability study: final report](#), National Physical Laboratory, 5 April 2023

77 [Qq. 453–454](#)

78 [Biometric Britain: the expansion of facial recognition software](#), Big Brother Watch, 23 May 2023, p. 67

79 [Big Brother Watch \(GAI0088\)](#)

80 [Q507](#)

53. Ms Chiswick emphasised that the decision to deploy this technology was not taken with “... a fishing expedition” in mind, and that the necessity and proportionality of each deployment was given careful consideration.⁸¹ She acknowledged public concerns around its use but confirmed that the Met was exploring other uses for AI to assist its operations.⁸²

3: The Misrepresentation challenge

AI can allow the generation of material that deliberately misrepresents someone’s behaviour, opinions or character

54. During recent years, controversies around ‘fake news’ have become increasingly frequent. The combination of data availability and new AI models and tools massively expands the opportunities for malign actors to ‘pass off’ content as being associated with particular individuals or organisations when it is in fact confected. Paul W. Fleming, General Secretary of the Equity trade union, described this process as:

... the creation of something completely new that never happened. It may be a series of static images that are then brought together into a video—an image of a politician slowly moved into saying something that they do not particularly want to say, or perhaps something that they do want to say.⁸³

55. The use of image and voice recordings of individuals can lead to highly plausible material being generated which can purport to show an individual saying things that have no basis in fact.⁸⁴ This material can be used to damage people’s reputations, and—in election campaigns—poses a significant threat to the conduct of democratic contests. Dr Steve Rolf, a researcher at the University of Sussex Business School, highlighted the potential for such material to impact “... democratic processes—for example, algorithmic recommendations on social media platforms that discourage wavering voters from turning out, thus tipping the balance in an election”.⁸⁵

56. Other uses of faked content can lead to fraud. For example, voice recognition technology is used extensively to verify people’s identity by many financial services providers in telephone transactions.⁸⁶ AI already makes it possible to reproduce a person’s voice patterns to speak words that they have never said, risking access through the voice recognition security barrier that has been imposed.⁸⁷

81 [Q442](#)

82 [Qq. 441, 464](#)

83 [Q332](#)

84 [WARNING: Beware frightening new ‘deepfake’ Martin Lewis video scam promoting a fake ‘Elon Musk investment’ – it’s not real](#), Money Saving Expert, 7 July 2023

85 Dr Steve Rolf ([GAI0104](#))

86 [Hello, this is your bank speaking: HSBC unveils voice recognition](#), Financial Times, 19 February 2016

87 [You may have heard about AI defeating voice authentication. This research kinda proves it](#), The Register, 28 June 2023

4: The Access to Data challenge

The most powerful AI needs very large datasets, which are held by few organisations

57. Access to sufficient volumes of high-quality data is a priority for AI developers and researchers.⁸⁸ Dr Andrew Hopkins, Chief Executive of Exscientia, told us that 40% of the drug discovery firm’s employees were “... experimental biologists in the laboratory, generating data. You need to have high-quality data to generate your algorithms”.⁸⁹ The need for significant volumes of training data places developers with the most resources at an advantage. The Ada Lovelace Institute, a research institute, wrote that this explained the leading role of certain AI developers, thanks to their significant stores of data.⁹⁰

58. The data access challenge also raises competition concerns. The UK Competition and Markets Authority is currently reviewing “... the likely implications of the development of AI foundation models for competition and consumer protection”⁹¹ whilst Lina Khan, Chair of the United States Federal Trade Commission, has emphasised the importance of fair competition in the AI sector.⁹² The Ada Lovelace Institute proposed legislation to us that would mandate research access to Big Tech data stores, to encourage a more diverse AI development ecosystem.⁹³ Creative Commons also advocated for “... the creation of high quality, open data sets”.⁹⁴

5: The Access to Compute challenge

The development of powerful AI requires significant compute power, access to which is limited to a few organisations

59. AI developers and researchers require access to sufficient levels of compute⁹⁵ to power the development, training and refining of AI models and tools. Professor Sir Nigel Shadbolt, Professorial Research Fellow at the university of Oxford, said that the emergence of foundation models and generative AI could be partly attributed to an “... extraordinary increase” in compute availability.⁹⁶

60. However, vast amounts of compute power is costly and therefore disproportionately available to the largest players. Professor Shadbolt told us that university researchers risked being left behind private developers as compute requirements continued to grow.⁹⁷ The UK Government has announced plans to establish an Exascale supercomputer facility and an AI-dedicated compute resource to support research.⁹⁸ The Prime Minister, Rt.

88 ADS ([GAI0027](#)), Royal College of Radiologists ([GAI0087](#))

89 [Q198](#)

90 The Ada Lovelace Institute ([GAI0086](#))

91 [CMA launches initial review of artificial intelligence models](#), GOV.UK, 4 May 2023

92 [“We must regulate AI,” FTC Chair Khan says](#), Ars Technica, 3 May 2023

93 The Ada Lovelace Institute ([GAI0086](#))

94 Creative Commons ([GAI0015](#))

95 Defined in the independent [Future of Compute Review](#) commissioned by the UK Government and published in March 2023 as “... the systems assembled at scale to tackle computational tasks beyond the capabilities of everyday computers. This includes both physical supercomputers and the use of cloud provision to tackle high computational loads”.

96 [Q146](#)

97 [Q146](#)

98 [Plan to forge a better Britain through science and technology unveiled](#), GOV.UK, 6 March 2023

Hon. Rishi Sunak MP, has also confirmed that three AI labs—Google DeepMind, OpenAI and Anthropic—will “... give early or priority access to models for research and safety purposes”.⁹⁹

6: The Black Box challenge

Some AI models and tools cannot explain why they produce a particular result, which is a challenge to transparency requirements

61. As increased availability of data and compute have facilitated the development of new AI models and tools, these have increasingly become ‘black boxes’, that is, their decision-making processes are not explainable. The challenge, as put to us by Professor Michael Osborne of The University of Oxford, is “... to what degree an AI should be able to explain itself”.¹⁰⁰ The challenge is further complicated by the fact that the better an AI model or tool performs, the less explainable it is likely to be.¹⁰¹ Adrian Joseph of BT Group drew parallels with the human brain and described it as “the ultimate black box”.¹⁰²

62. When asked whether the AI models and tools used by Exscientia were inherently black boxes, Chief Executive, Dr Andrew Hopkins, said that they were not, and that knowing the “provenance” of outputs was valuable: “... connecting the dots between a prediction and the data that led to a prediction is vital for understanding in science, as much as it is for the general public”.¹⁰³

63. Professor Sir Nigel Shadbolt described the suggestion that future AI models and tools would inevitably be black boxes as “... a counsel of despair” and said that policymakers and wider society should “... demand that these systems begin to render some of the processes by which they came up with the output they do more transparent and more explicable”.¹⁰⁴

64. Greater explainability would also, according to the Public Law Project, help increase public trust in the deployment of AI models and tools in different sectors.¹⁰⁵ However, researchers and medical professionals at the Birmingham AI and Digital Healthcare Group, pointed out that other aspects of medicine were “... relatively black box” and that “... whilst greater explainability for all health interventions is desirable, it should not be an absolute requirement” if they are proven to be safe and effective.¹⁰⁶

99 [PM London Tech Week speech](#), GOV.UK, 12 June 2023

100 [Q40](#)

101 [Burgess Salmon LLP](#) (GAI0064)

102 [Q142](#)

103 [Q190](#)

104 [Q151](#)

105 [Public Law Project](#) (GAI0069)

106 [AI and Digital Healthcare Group, Centre for Regulatory Science and Innovation, Birmingham \(University Hospitals Birmingham NHS Foundation Trust/University of Birmingham\)](#) (GAI0055)

7: The Open-Source challenge

Requiring code to be openly available may promote transparency and innovation; allowing it to be proprietary may concentrate market power but allow more dependable regulation of harms

65. There are different views on whether the code used to run AI models and tools should be freely available—or open-source—for testing, scrutiny and improvement. Proponents of this approach have said that it can encourage innovation and prevent monopolisation by powerful, well-resourced players. The Tony Blair Institute for Global Change has pointed out that open-source innovation “underpins the entire Internet”.¹⁰⁷

66. Creative Commons told us an open-source approach can lower barriers to development, and that “... the capacity to develop and use AI... should be widely distributed, rather than concentrated among a narrow few”.¹⁰⁸ Professor Mihaela van der Schaar of the University of Cambridge said that, in settings where sensitive personal data are used, such as medicine and healthcare, the relevant authorities should encourage the use of open-source platforms that are open to inspection and robustly tested for safety.¹⁰⁹

67. However, others have highlighted the value of keeping code proprietary, and argued that doing so can also help protect against misuse, for example through the dissemination of misleading content.¹¹⁰ “The competitive landscape and safety implications” were cited by OpenAI (which, paradoxically, has not made its latest models and tools open-source) as reasons to limit public disclosure of information about “... the architecture (including model size), hardware, training compute, dataset construction, training method, or similar” of its most advanced model, GPT-4.¹¹¹ Chief Scientist Ilya Sutskever cited the potential for malign actors to cause harm using open-source models:

These models are very potent and they’re becoming more and more potent. At some point it will be quite easy, if one wanted, to cause a great deal of harm with those models. And as the capabilities get higher it makes sense that you don’t want to disclose them.¹¹²

8: The Intellectual Property and Copyright Challenge

Some AI models and tools make use of other people’s content: policy must establish the rights of the originators of this content, and these rights must be enforced

68. Whilst the use of AI models and tools have helped create revenue for the entertainment industry in areas such as video games and audience analytics,¹¹³ concerns have been raised

107 [A New National Purpose: AI Promises a World-Leading Future of Britain](#), The Tony Blair Institute for Global Change, 13 June 2023

108 Creative Commons ([GAI0015](#))

109 [Q258](#)

110 [Generative AI Systems Aren’t Just Open or Closed Source](#), Wired, 24 May 2023

111 [GPT-4 Technical Report](#), OpenAI, 14 March 2023

112 [OpenAI co-founder on company’s past approach to openly sharing research: ‘We were wrong’](#), The Verge, 15 March 2023

113 [Qq. 327, 329](#)

about the ‘scraping’ of copyrighted content from online sources without permission.¹¹⁴ Jamie Njoku-Goodwin, CEO of UK Music, told us that his industry operated on the “... basic principle that, if you are using someone else’s work, you need permission for that and must observe and respect the copyright”,¹¹⁵ but that new tools allowed this to be circumvented. He described the process as:

... people taking the work of creators, which has a copyright attached, feeding it into an AI, using that to generate so-called new works and then, potentially, being able to monetise off the back of that, without respecting or recognising the inputs.¹¹⁶

Ongoing legal cases are likely to set precedents in this area.¹¹⁷

69. Representatives of the creative industries told us that they hoped to reach a mutually beneficial solution with the AI sector, potentially in the form of a licensing framework for the use of copyrighted content to train models and tools.¹¹⁸ Jamie Njoku-Goodwin said that he and UK Music’s members would welcome greater engagement and that he had “... not had any evidence of companies coming to us to say, “We would like to seek a licence to use works to train AI”.”¹¹⁹ Dr Hayleigh Boshier, an intellectual property researcher at Brunel University, pointed out that AI and tech firms would also benefit from the enforcement of copyright and intellectual property rights.¹²⁰

70. The Intellectual Property Office, an executive agency of the UK Government, has begun to develop a voluntary code of practice on copyright and AI, in consultation with the technology, creative and research sectors.¹²¹ It has said that the guidance should “... support AI firms to access copyrighted work as an input to their models, whilst ensuring there are protections (e.g. labelling) on generated output to support right holders of copyrighted work”. A draft code is expected to be published by the end of July.¹²² The Government has said that if agreement is not reached or the code not adopted, it may legislate.¹²³

71. In February 2023, following criticism from the creative industries,¹²⁴ the Government withdrew a proposed text and database mining exception for any purpose, such as the development of AI models and tools. An exception for the purposes of non-commercial research has been in place since 2014.¹²⁵ These competing incentives define the intellectual property and copyright challenge. The Library and Archives Copyright Alliance told us that

114 [Streaming services urged to clamp down on AI-generated music](#), Financial Times, 12 April 2023

115 [Q329](#)

116 [Q329](#)

117 [Q379](#)

118 [Q350](#)

119 [Q348](#)

120 [Qq. 374, 390](#)

121 [The government’s code of practice on copyright and AI](#), GOV.UK, 29 June 2023

122 [HM Government response to Professor Dame Angela McLean’s Pro-Innovation Regulation of Technologies Review: Creative Industries](#), GOV.UK, 14 June 2023, p. 7

123 [The government’s code of practice on copyright and AI](#), GOV.UK, 29 June 2023

124 [Q341](#)

125 [Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation](#), GOV.UK, 28 June 2022

the decision to withdraw the proposed exception “... prevents the UK from capitalising on the diverse, agile and creative benefits that AI can bring to the UK’s economy, its society and its competitive research environment”.¹²⁶

9: The Liability challenge

If AI models and tools are used by third parties to do harm, policy must establish whether developers or providers of the technology bear any liability for harms done

72. A trend towards increasingly complex and international supply chains for AI models and tools, involving “... cloud-based services, servers, protocols, data centres, third-party data sources, and content delivery networks”,¹²⁷ has created a challenge over the determination of liability for unsafe or harmful uses of the technology, and compliance with governance requirements to mitigate risk.

73. The Trades Union Congress suggested to us that “... different actors influencing the technology at development, procurement and application” stages made identifying responsibility for discriminatory uses of AI models and tools challenging.¹²⁸ At deployment stage, the Ada Lovelace Institute said that obligations should rest on developers, providers (including intermediaries) and end users.¹²⁹

74. Developers also expressed support for distributed liability. DeepMind, since merged with the Google Research Brain Unit to form Google DeepMind,¹³⁰ said that responsibility should be spread across supply chains “... based on which aspect is most likely to lead to harm and an actor’s practical ability to comply given the structure of the market and the nature of their contribution”.¹³¹

10: The Employment challenge

AI will disrupt the jobs that people do and that are available to be done. Policy makers must anticipate and manage the disruption

75. Sir Patrick Vallance, former Chief Scientific Adviser to the UK Government, said to us that the increasing ubiquity of AI models and tools will have “... a big impact on jobs, and that impact could be as big as the industrial revolution was”.¹³² Professor Michael Osborne of the University of Oxford said that he expected “... tasks that involve routine, repetitive labour and revolve on low-level decision-making to be automated very quickly”.¹³³

76. Our inquiry heard different perspectives on automation. Paul W. Fleming, General Secretary of the Equity trade union, said that whilst some opportunities to earn were being reduced and jobs were being displaced, the technology’s rise offered: “... a whole

126 The Library and Archives Copyright Alliance ([GAI0120](#))

127 Dr Jennifer Cobbe and Dr Jatinder Singh, Compliant & Accountable Systems research group, Department of Computer Science & Technology, University of Cambridge ([GAI0106](#))

128 TUC ([GAI0060](#))

129 The Ada Lovelace Institute ([GAI0086](#))

130 [Announcing Google DeepMind](#), Google DeepMind, 20 April 2023

131 DeepMind ([GAI0100](#))

132 Oral evidence taken on 3 May 2023, HC (2022–23) 1324, [Q47](#)

133 [Q48](#)

new frontier of income potential and new work for our members. This is not something to be frightened or worried about”.¹³⁴ Hugh Milward of Microsoft UK was also not worried, and viewed AI as a “... co-pilot, not an autopilot. Its job is to augment the things that human beings are doing, rather than to replace the things that human beings are doing, and to really allow humans to be more human, in some respects”.¹³⁵

77. When assessing the potential impact of automation across the economy and society, Sir Patrick told us that it would be important to plan ahead, and ask:

... Which are the jobs and sectors that will be most affected, and what are the plans to retrain, or give people their time back to do a job differently? There will be jobs that can be done by AI, which can either mean that lots of people do not have a job or that, actually, lots of people have a job that only humans can do.¹³⁶

The Prime Minister, Rt. Hon. Rishi Sunak MP, has highlighted the socioeconomic risks created by “large-scale societal shifts” associated with the development of AI:

That does not mean that you should stand in the way of it, but it just means we should make sure that we are cognisant of it and provide people with the skills they need to flourish in a world that is being changed by technology.¹³⁷

11: The International Coordination challenge

AI is a global technology, and the development of governance frameworks to regulate its uses must be an international undertaking

78. Different jurisdictions have proposed different approaches to AI governance. The UK Government has expressed a preference for a “pro-innovation approach” in the AI white paper, published in March 2023.¹³⁸ The draft Artificial Intelligence Act currently being negotiated between European Union Member States and the European Parliament would implement a risk-based approach, with AI models and tools grouped into risk categories and some, such as biometric surveillance, emotion recognition and predictive policing, banned altogether.¹³⁹ In the United States, the White House has called on AI developers “... to take action to ensure responsible innovation and appropriate safeguards, and protect people’s rights and safety”,¹⁴⁰ whilst Senate Majority Leader Chuck Schumer has said that a proposed legislative approach will be outlined in the autumn of 2023.¹⁴¹

134 [Q327](#)

135 [Q134](#)

136 Oral evidence taken on 3 May 2023, HC (2022–23) 1324, [Q51](#)

137 [Oral evidence taken before the Liaison Committee on 4 July 2023, HC \(2022–23\) 1602, Q17](#) (the Prime Minister)

138 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023

139 [MEPs ready to negotiate first-ever rules for safe and transparent AI](#), European Parliament, 14 June 2023

140 [Readout of White House meeting with CEOs on advancing responsible artificial intelligence innovation](#), the White House, 4 May 2023

141 [Sen. Chuck Schumer launches SAFE innovation in the AI age at CSIS](#), Center for Strategic and International Studies, accessed 27 June 2023

79. Whilst there are divergent approaches currently being pursued, our inquiry has heard that the global implications of AI's emergence as a ubiquitous, general-purpose technology demand a coordinated response. As Professor Michael Osborne of the University of Oxford told us, internationally agreed and harmonised principles to inform AI governance were desirable "... because the problems that we face are very similar".¹⁴²

80. There are various international fora where such discussions could take place, and the development of a coordinated response will be shaped by geopolitical considerations. The Prime Minister, Rt. Hon. Rishi Sunak MP, has taken steps to position the UK as a leading actor in these discussions by announcing that it would host a global summit on AI safety later in 2023.¹⁴³ We nevertheless heard that the European Union and United States are likely to become the de facto global standard-setters in AI governance.¹⁴⁴

81. We will further explore the UK's regulatory approach to AI, international comparators, and initiatives to establish international coordination, in Chapters 5 and 6 of this interim Report.

12. The Existential challenge

Some people think that AI is a major threat to human life: if that is a possibility, governance needs to provide protections for national security

82. A related but separate challenge concerns the international security implications of AI's increasing prevalence, and debates over existential risks. Ian Hogarth, an investor and Chair of the UK Government's AI Foundation Model Taskforce,¹⁴⁵ has said that whilst it is difficult to predict when it will emerge, so-called "... God-like AI could be a force beyond our control or understanding, and one that could usher in the obsolescence or destruction of the human race".¹⁴⁶ Matt Clifford, an adviser to the UK Government on AI, has said that such a prospect would soon be realistic: "you can have really very dangerous threats to humans that could kill many humans, not all humans, simply from where we'd expect models to be in two years' time".¹⁴⁷

83. In a joint submission, Dr Jess Whittlestone of the Centre for Long-Term Resilience think-tank and Richard Moulange, a researcher at The University of Cambridge, suggested scenarios where the use of AI models and tools could threaten or undermine national and/or international security. These included the development of novel biological or chemical weapons, the risk of unintended escalation and the undermining of nuclear deterrence.¹⁴⁸

84. Appearing at the House of Lords Artificial Intelligence in Weapons Systems Committee in June 2023, former National Security Adviser, Lord Sedwill, said that whilst he did not expect AI to have the same impact on military doctrine as the development

142 [Q6](#)

143 [UK to host first global summit on artificial intelligence](#), GOV.UK, 7 June 2023

144 [British and Irish Law Education Technology Association \(GAI0082\)](#)

145 [Tech entrepreneur Ian Hogarth to lead UK's AI Foundation Model Taskforce](#), GOV.UK, 18 June 2023

146 [We must slow down the race to God-like AI](#), Financial Times, 13 April 2023

147 [AI systems 'could kill many humans' within two years](#), The Times, 5 June 2023

148 [Dr Jess Whittlestone and Richard Moulange \(GAI0071\)](#)

of nuclear weapons and subsequent development of the mutually assured destruction consensus,¹⁴⁹ it nevertheless represented “... the future of defence capability and the UK needs to be at the forefront of that”.¹⁵⁰

85. Whilst mutually assured destruction is widely accepted as valid military doctrine, there is disagreement as to whether the existential risks of AI highlighted by researchers and developers in May 2023 are realistic.¹⁵¹ Professor Michael Osborne said that he believed that such predictions could come to pass but suggested that the development of an international security framework to govern the development and use of nuclear weapons offered a template for mitigating the collective risks posed by AI.¹⁵²

86. Professor Osborne’s co-researcher Michael Cohen said that if a shared understanding of the risks could be reached through international fora, then:

“... the game theory isn’t that complicated. Imagine that there was a button on Mars labelled ‘geopolitical dominance’, but actually, if you pressed it, it killed everyone. If everyone understands that, there is no space race for it”.¹⁵³

It has been observed that reaching such a shared understanding and inspection framework comparable to those which govern the use of biological, chemical and nuclear weapons would present a significant diplomatic and technical challenge.¹⁵⁴

87. The 2023 AI white paper, discussed in more detail in Chapter 5, described the “existential risks” posed by AGI as “high impact but low probability”.¹⁵⁵ Joelle Pineau, vice-president of AI research at Meta, has warned against a focus on AGI, as it reduces the opportunity for “... rational discussions about any other outcomes. And that takes the oxygen out of the room for any other discussion, which I think is too bad”.¹⁵⁶ MIT Technology Review editor Will Douglas Heaven has suggested that “if something sounds like bad science fiction, maybe it is”.¹⁵⁷

88. ***The Government’s approach to AI governance and regulation should address each of the twelve challenges we have outlined, both through domestic policy and international engagement.***

149 [Oral evidence taken before the House of Lords Artificial Intelligence in Weapons Systems Committee on 8 June 2023, Q104 \(Lord Sedwill\)](#)

150 [Oral evidence taken before the House of Lords Artificial Intelligence in Weapons Systems Committee on 8 June 2023, Q107 \(Lord Sedwill\)](#)

151 [Statement on AI Risk](#), Center for AI Safety, 30 May 2023

152 [Q45](#)

153 [Q44](#)

154 [AI is supposedly the new nuclear weapons — but how similar are they, really?](#), Vox, 29 June 2023

155 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 50

156 [The Algorithm: how existential risk became the biggest meme in AI](#), MIT Technology Review, 19 June 2023

157 [The Algorithm: seven things to pay attention to when talking about AI](#), MIT Technology Review, 29 May 2023

5 The UK Government approach to AI

89. In March 2023, the UK Government set out its proposed “pro-innovation approach to AI regulation” in the form of a white paper,¹⁵⁸ and the Department for Science, Innovation and Technology (DSIT) is currently evaluating responses to an accompanying consultation.¹⁵⁹ In this Chapter we will examine the UK Government’s proposed approach to AI governance and regulation.

The AI white paper

90. The white paper aimed to ensure “... that regulatory measures are proportionate to context and outcomes, by focusing on the use of AI rather than the technology itself”.¹⁶⁰ It set out five principles to frame regulatory activity, guide future development of AI models and tools, and their use:

- Safety, security and robustness;
- Appropriate transparency and explainability;
- Fairness;
- Accountability and governance; and
- Contestability and redress.¹⁶¹

91. The white paper said that these principles would not initially be put on a statutory footing but interpreted and translated into action by individual sectoral regulators, with assistance from central support functions, initially delivered from within Government. Six proposed functions would cover:

- Monitoring and evaluation of the overall regulatory framework’s effectiveness and the implementation of the principles;
- Assessment and monitoring of risks across the economy arising from AI;
- Horizon scanning and gap analysis, including by convening industry, to inform a coherent response to emerging AI technology trends;
- Supporting testbeds and sandbox initiatives to help bring new technologies to market;
- Providing education and awareness to give clarity to businesses and ensure citizen participation in iteration of the framework; and
- Promoting interoperability with international regulatory frameworks.¹⁶²

158 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023

159 [AI regulation: a pro-innovation approach—policy proposals](#), GOV.UK, accessed 27 June 2023

160 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 5

161 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 6

162 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, pp. 6–7

92. The white paper outlined a series of deliverables on six-, twelve-, and twelve-month-plus timetables. Key deliverables included:

- Issue the cross-sectoral principles to regulators, together with initial guidance covering implementation (within six months);
- Assess the ability of key regulators to implement the principles, and how Government can best support them (within six months);
- Publish an AI Regulation Roadmap with plans for establishing the central functions (within six months);
- Encourage key regulators to publish guidance on how the cross-sectoral principles apply within their remit (within six to twelve months);
- Deliver a first iteration of the central support functions (twelve months or more);
- Publish a draft central, cross-economy AI risk register (twelve months or more); and
- Publish an updated AI Regulation Roadmap which will confirm plans for the future delivery of the central functions, including whether these will be overseen by Government or an independent body.¹⁶³

93. The white paper confirmed that whilst the Government did not intend to introduce AI-specific legislation immediately: “... when parliamentary time allows, we anticipate introducing a statutory duty on regulators requiring them to have due regard to the principles”.¹⁶⁴

A UK-specific approach

94. Contributors to our inquiry told us that the proposals in the white paper constituted a distinct, UK-specific approach to AI governance. Coran Darling, an Associate at law firm DLA Piper, described it as “a non-linear approach” that would provide “... the power and flexibility to regulators to distinguish their best practices and approach it on that basis”.¹⁶⁵ Michael Birtwistle, Associate Director of the Ada Lovelace Institute and a former UK Government official, said that there were “... a lot of advantages to the context-specific approach”, but that implementation of the principles would be the key challenge.¹⁶⁶

95. We also heard supportive sentiments from industry. Jen Gennai, Director (Responsible Innovation) at Google, told us that she favoured “... a principles-based approach that allows for support of innovation while ensuring guardrails”,¹⁶⁷ whilst Hugh Milward, General Manager of Corporate, External and Legal Affairs at Microsoft UK, said that the development of AI governance principles that could be applicable irrespective of how the technology evolved would put the UK “... in a really good place to get that generation of

163 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, pp. 72–73

164 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 6

165 [Q411](#)

166 [Q528](#)

167 [Q135](#)

innovation at the same time as guiding society”.¹⁶⁸ Google’s submission to our inquiry emphasised that AI “... is far too important not to regulate and too important not to regulate well”.¹⁶⁹

96. Asked whether the development of a UK-specific approach to AI governance would require primary legislation, the Prime Minister, Rt. Hon. Rishi Sunak MP, told the Liaison Committee:

I think what we need to do—and I think we can probably do lots of this without legislation—is sit down and figure out what safety features and guard rails we would like to put in place... it is too early to pre-empt what all that might look like, but you can imagine a world where at least the initial stages of that don’t require legislation, necessarily, but just require us to get in there and do safety evaluation on the models and have access to them.¹⁷⁰

We will consider the European Union and United States’ respective approaches to AI governance in Chapter 6 of this interim Report.

The Digital Regulation Cooperation Forum

97. The AI white paper highlighted the UK’s “... high-quality regulators and our strong approach to the rule of law, supported by our technology-neutral legislation and regulations” as a key factor in the development of its AI sector.¹⁷¹ It said that whilst some regulators were sufficiently resourced to respond to the development of AI models and tools and that some mechanisms for regulatory cooperation are already in place, others had “... limited capacity and access to AI expertise. This creates the risk of inconsistent enforcement across regulators”.¹⁷²

98. The importance of regulatory capacity and coordination were highlighted by contributors to our inquiry, and the work of the Digital Regulation Cooperation Forum (DRCF), which brings together the Competition and Markets Authority, Information Commissioner’s Office, Ofcom and Financial Conduct Authority,¹⁷³ was highlighted as an example of best practice.¹⁷⁴ Katherine Holden of techUK, a trade body, suggested the establishment of “... formalised structures to co-ordinate approaches between regulators... an expanded version of the Digital Regulation Cooperation Forum [DRCF]”.¹⁷⁵ RELX, an information and analytics company, also said that “regulators should be expected to engage with one another via the DCRF, and this should apply to all regulators that are likely to be involved in issuing guidance or rules on AI”.¹⁷⁶

168 [Q135](#)

169 [Google \(GAI0099\)](#)

170 [Oral evidence taken before the Liaison Committee on 4 July 2023, HC \(2022–23\) 1602](#), Qq. 19–20 (the Prime Minister)

171 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 14

172 [A pro-innovation approach to AI regulation](#), CP 815, Department for Science, Innovation and Technology, 29 March 2023, p. 15

173 [Digital Regulation Cooperation Forum](#), GOV.UK, accessed 19 June 2023

174 [ADS \(GAI0027\)](#), [RELX \(GAI0033\)](#), [NCC Group \(GAI0040\)](#)

175 [Q82](#)

176 [RELX \(GAI0033\)](#)

99. We heard that the rate of development in the field has created challenges for even the best-equipped regulators. Professor Sir Nigel Shadbolt, Professorial Research Fellow in Computer Science at the University of Oxford, said regulators were “... trying to gear up for this world of AI... trying to understand what technical skills they need and what they would need to put into the process to enable their job to be easier”.¹⁷⁷ Michael Birtwistle of the Ada Lovelace Institute told us that ensuring properly-equipped regulators to the UK’s AI eventual governance framework would have a positive impact on public trust in AI models and tools: “... the importance and benefits of these technologies could be huge... we need public trust; we need those technologies to be trustworthy, and that is worth investing regulatory capability in”.¹⁷⁸

Foundation Model Taskforce

100. In addition to the AI white paper, the UK Government has announced the formation of an AI Foundation Model Taskforce, chaired by investor Ian Hogarth.¹⁷⁹ Bringing together experts from Government, industry and academia in a similar way to the successful covid-19 Vaccines Taskforce, it will have a mandate to “... carry out research on AI safety and inform broader work on the development of international guardrails, such as shared safety and security standards and infrastructure”.¹⁸⁰

101. The Taskforce will invest an initial £100 million “... in foundation model infrastructure and public service procurement, to create opportunities for domestic innovation”. Pilots involving the use of AI models and tools in public services are expected to launch later this year.¹⁸¹

102. The UK has a long history of technological innovation and regulatory expertise, which can help it forge a distinctive regulatory path on AI. The AI white paper should be welcomed as an initial effort to engage with a complex task. However, the approach outlined is already risking falling behind the pace of development of AI.

103. The UK Government’s proposed approach to AI governance relies heavily on our existing regulatory system, and the promised central support functions. The time required to establish new regulatory bodies means that adopting a sectoral approach, at least initially, is a sensible starting point. We have heard that many regulators are already actively engaged with the implications of AI for their respective remits, both individually and through initiatives such as the Digital Regulation Cooperation Forum. However, it is already clear that the resolution of all of the Challenges set out in this report may require a more well-developed central coordinating function.

104. The AI white paper is right to highlight the importance of regulatory capacity to the successful implementation of its principles. *The Government should, as part of its implementation of its proposals, undertake a gap analysis of the UK’s regulators, which considers not only resourcing and capacity, but whether any regulators require new powers to implement and enforce the principles outlined in the AI white paper.*

177 [Q154](#)

178 [Q523](#)

179 [Tech entrepreneur Ian Hogarth to lead UK’s AI Foundation Model Taskforce](#), GOV.UK, 18 June 2023

180 [Tech entrepreneur Ian Hogarth to lead UK’s AI Foundation Model Taskforce](#), GOV.UK, 18 June 2023

181 [Initial £100 million for expert taskforce to help UK build and adopt next generation of safe AI](#), GOV.UK, 24 April 2023

105. The Government is yet to confirm whether AI-specific legislation will be included in the upcoming King's Speech in November. This new session of Parliament will be the last opportunity before the General Election for the UK to legislate on the governance of AI. Following the Election it is unlikely that new legislation could be enacted until late 2025—more than two years from now and nearly three years from the publication of the white paper.

106. The Government has said in the AI white paper that it may legislate, at a minimum, to establish 'due regard' duties for existing regulators. That commitment alone—in addition to any further requirements that may emerge—suggests that there should be a tightly-focussed AI Bill in the new session of Parliament. Our view is that this would help, not hinder, the Prime Minister's ambition to position the UK as an AI governance leader. We see a danger that if the UK does not bring in any new statutory regulation for three years it risks the Government's good intentions being left behind by other legislation—like the EU AI Act—that could become the de facto standard and be hard to displace.

107. *In its reply to this interim Report, and its response to the AI white paper consultation, the Government should confirm whether AI-specific legislation, such as the introduction of a requirement for regulators to pay due regard to the AI white paper principles, will be introduced in the next Parliament. It should also confirm what work has been undertaken across Government to explore the possible contents of such a Bill.*

108. We welcome the establishment of a Foundation Model Taskforce, the appointment of Ian Hogarth as its chair, and the Government's stated intention for it to take a similar approach to the Vaccines Taskforce. This agile approach is necessary and proportionate to the importance of the issue. *The Government should confirm the Task Force's full membership, terms of reference, and the first tranche of public sector pilot projects, in its reply to this interim Report.*

Chapter 6: The international dimension

109. In Chapter 4 of this interim Report, we identified twelve challenges that the governance of AI must address. Many of these challenges are international in nature, and the importance of international coordination of AI governance and regulation were highlighted by contributors to our inquiry. In this Chapter, we will consider the governance approaches in the European Union (EU) and United States, and the role of the UK in international coordination initiatives.

The European Union

110. The European Union (EU) is one of two jurisdictions suggested to us as likely to play a leading role in shaping international AI governance and regulation, through an eventual EU AI Act proposed by the European Commission in 2021.¹⁸² Negotiations between the Commission, European Parliament and Member States over the final text of the Act are ongoing.¹⁸³

111. In contrast to the UK's proposed context-specific, principles-based approach, the draft EU AI Act takes a risk-based approach, with "... obligations for providers and those deploying AI systems depending on the level of risk the AI can generate".¹⁸⁴ It aims to set out a "technology-neutral, uniform definition of AI" and proposes grouping models and tools into risk categories, including unacceptable risk (which would be banned), high risk, and limited risk, with transparency obligations placed on generative AI tools, such as ChatGPT.¹⁸⁵ It has been reported that the final Act is unlikely to be agreed until later in 2023, and may not be fully in force until 2026.¹⁸⁶

112. Hugh Milward of Microsoft UK told us that the EU's proposed approach to AI governance provided "... a model of how not to do it",¹⁸⁷ whilst Katherine Holden of techUK, a trade body, also expressed reservations, and described it as "... a very centralised approach... it does not allow the opportunity for much flexibility. It is not particularly future-proofed, because they have a static list of high-risk applications".¹⁸⁸ In an open letter published in June 2023, executives from 150 businesses said that the EU's proposed approach would:

... jeopardise Europe's competitiveness and technological sovereignty without effectively tackling the challenges we are and will be facing. This is especially true regarding generative AI. Under the version recently adopted by the European Parliament, foundation models, regardless of their use cases, would be heavily regulated, and companies developing and implementing such systems would face disproportionate compliance costs and disproportionate liability risks.¹⁸⁹

182 [Proposal for a Regulation laying down harmonised rules on artificial intelligence](#), European Commission, 21 April 2021

183 [MEPs ready to negotiate first-ever rules for safe and transparent AI](#), European Parliament, 14 June 2023

184 [MEPs ready to negotiate first-ever rules for safe and transparent AI](#), European Parliament, 14 June 2023

185 [EU AI Act: first regulation on artificial intelligence](#), European Parliament, accessed 19 June 2023

186 [Europe takes another big step toward agreeing an AI rulebook](#), TechCrunch, 14 June 2023

187 [Q143](#)

188 [Q89](#)

189 [European VCs and tech firms sign open letter warning against over-regulation of AI in draft EU laws](#), TechCrunch, 30 June 2023

The United States

113. In October 2022, the White House Office of Science and Technology Policy published a non-binding ‘Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People’.¹⁹⁰ It offered “... a set of five principles and associated practices to help guide the design, use and deployment of automated systems to protect the rights of the American public”.¹⁹¹ The five principles were:

- **Safe and effective systems:** citizens should be protected from unsafe or ineffective systems;
- **Algorithmic discrimination protections:** citizens should not face discrimination by algorithms and systems should be used and designed in an equitable way;
- **Data privacy:** citizens should be protected from abusive data practices via built-in data protections and should have agency over how data about them is used;
- **Notice and explanation:** citizens should know that automated systems are being used, and understand how and why it contributes to outcomes that impact them; and
- **Human alternatives, consideration and fallback:** citizens should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems they encounter.¹⁹²

114. In May 2023, the White House called on AI developers “... to take action to ensure responsible innovation and appropriate safeguards, and protect people’s rights and safety”, following a meeting with leaders from four developers—Anthropic, Alphabet, Microsoft and OpenAI.¹⁹³ Senate Majority Leader Chuck Schumer has since outlined a ‘SAFE Innovation Framework for AI Policy’ and said that “... a new legislative approach for translating this framework into legislative action” will begin in the autumn of this year.¹⁹⁴

115. In an appearance at a subcommittee of the Senate Committee on the Judiciary examining privacy, technology and the law, OpenAI chief executive Sam Altman said that it was “... essential to develop regulations that incentivize AI safety while ensuring that people are able to access the technology’s many benefits”.¹⁹⁵

190 [Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People](#), White House Office of Science and Technology Policy, 20 October 2022

191 [Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People](#), White House Office of Science and Technology Policy, 20 October 2022, p. 4

192 [Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People](#), White House Office of Science and Technology Policy, 20 October 2022, pp. 5–7

193 [Readout of White House meeting with CEOs on advancing responsible artificial intelligence innovation](#), the White House, 4 May 2023

194 [Sen. Chuck Schumer launches SAFE innovation in the AI age at CSIS](#), Center for Strategic and International Studies, accessed 27 June 2023

195 [Written Testimony of Sam Altman, Chief Executive Officer OpenAI Before the U.S. Senate Committee on the Judiciary Subcommittee on Privacy, Technology, & the Law](#), 16 May 2023

International coordination

116. Contributors to our inquiry highlighted the geopolitical context of AI development. Adrian Joseph, Chief Data and AI Officer at BT Group, said that “... there is a risk that we in the UK lose out to the large tech companies, and possibly China, and are left behind”.¹⁹⁶ Microsoft’s Hugh Milward cited AI development as “... one example of an area where we have to stay ahead in ‘the West’ with values that are effectively pro-democracy”, when compared with states such as China.¹⁹⁷

117. Following a visit to the United States in June, Prime Minister Rishi Sunak MP announced that the UK would host a global summit on AI safety: “[it] will bring together key countries, leading tech companies and researchers to agree safety measures to evaluate and monitor the most significant risks from AI”.¹⁹⁸ Asked whether he intended to use the summit as a platform to establish an organisation comparable to the International Atomic Energy Agency, which includes a broad range of countries, including China and Russia,¹⁹⁹ he said:

... AI does not respect national borders, and I think we will all benefit from hearing and talking to each other in a conversation with the businesses themselves. That is really what this is about. We are a long way from anyone establishing an IAEA equivalent for AI. Those things are long into the distance, but in the first instance, just talking this through with like-minded countries seems a sensible step.²⁰⁰

118. Google told us that it supported Government plans “... to actively take a role in shaping global norms”.²⁰¹ We also heard that the UK could play a convening role and “... influence the global discussion on AI regulation in a way which is more compatible with the UK’s approach, delivers high standards and protections, and allows the UK to act as a bridge between different systems”.²⁰²

119. The Prime Minister was right to say that AI does not respect national borders, and we welcome the announcement of a global summit on AI safety in London. The challenges highlighted in our interim Report should form the basis for these important international discussions.

120. The summit should aim to advance a shared international understanding of the challenges of AI—as well as its opportunities. Invitations to the summit should therefore be extended to as wide a range of countries as possible. Given the importance of AI to our national security there should also be a forum established for like-minded countries who share liberal, democratic values, to be able to develop an enhanced mutual protection against those actors—state and otherwise—who are enemies of these values.

196 [Q109](#)

197 [Q119](#)

198 [UK to host first global summit on artificial intelligence](#), GOV.UK, 7 June 2023

199 [List of member states](#), International Atomic Energy Agency, accessed 7 July 2023

200 [Oral evidence taken before the Liaison Committee on 4 July 2023, HC \(2022–23\) 1602](#), Q18 (the Prime Minister)

201 Google ([GAI0099](#))

202 Mr Haydn Belfield, Dr Seán Ó hÉigeartaigh, Dr Shahar Avin, Giulio Corsi (University of Cambridge) and Prof José Hernández-Orallo (Universitat Politècnica de València) ([GAI0094](#))

6 Chapter 7: Conclusion and next steps

121. There is as little consensus about how AI will evolve as there has been excitement and hyperbole following its rise to ubiquity. AI cannot be un-invented. It has and will continue to change the way we live our lives. Humans must take measures to safely harness the benefits of the technology and encourage future innovations, whilst providing credible protection against harm.

122. Some observers have called for the development of certain types of AI models and tools to be paused, allowing global regulatory and governance frameworks to catch up. We are unconvinced that such a pause is deliverable. When AI leaders say that new regulation is essential, their calls cannot responsibly be ignored—although it should also be remembered that is not unknown for those who have secured an advantageous position to seek to defend it against market insurgents through regulation.

123. *The twelve Challenges of AI Governance which we have set out must be addressed by policymakers in all jurisdictions. Different administrations may choose different ways to do this.*

124. We believe that the UK's depth of expertise in AI and the disciplines which contribute to it—the vibrant and competitive developer and content industry that the UK is home to; and the UK's longstanding reputation for developing trustworthy and innovative regulation—provides a major opportunity for the UK to be one of the go-to places in the world for the development and deployment of AI. But that opportunity is time-limited. Without a serious, rapid and effective effort to establish the right governance frameworks—and to ensure a leading role in international initiatives—other jurisdictions will steal a march and the frameworks that they lay down may become the default even if they are less effective than what the UK can offer. *We urge the Government to accelerate, not to pause, the establishment of a governance regime for AI, including whatever statutory measures as may be needed.*

Conclusions and recommendations

A general-purpose technology

1. While AI is not a new technology, the rapidly acquired ubiquity of tools such as ChatGPT and the rate of development has come as a surprise to even well-informed observers. We are all now interacting with AI models and tools daily, and we are increasingly aware of these interactions. (Paragraph 15)
2. Nevertheless, the technology should not be viewed as a form of magic or as something that creates sentient machines capable of self-improvement and independent decisions. It is akin to other technologies: humans instruct a model or tool and use the outputs to inform, assist or augment a range of activities. (Paragraph 16)

Benefits

3. AI models and tools can transform healthcare provision, by assisting with diagnostics and, perhaps more significantly, freeing up time for the judgement of medical professionals by automating routine processes. (Paragraph 26)
4. The ability of AI models and tools to process substantial volumes of data, and rapidly identify patterns where human researchers might take months or be unable to, makes it a potentially transformational technology for medical research. Either through the development of new drugs, or the repurposing of existing ones, the technology could reduce the investment required to bring a drug to market; and bring personalised medicine closer to becoming a reality. (Paragraph 29)
5. AI tools are already useful time-savers for education professionals, and whilst reliable data is hard to come by, it seems highly likely that the technology is this generation of students' calculator or smartphone. (Paragraph 36)
6. The benefits for time-pressed teachers using AI models and tools to help prepare lesson plans are clear, and increased availability of personalised learning and tutoring tools could benefit many pupils. However, widespread use of AI raises questions about the nature of assessment, particularly in subjects that rely heavily on coursework. (Paragraph 37)
7. *Education policy must prioritise equipping children with the skills to succeed in a world where AI is ubiquitous: digital literacy and an ability to engage critically with the information provided by AI models and tools.* (Paragraph 38)
8. The wide range of potential applications, and associated benefits, reflects the general-purpose nature of AI. As with previous technological innovations, the challenge for policymakers is translating this potential into reality, in a safe and sustainable way. (Paragraph 40)

Twelve Challenges of AI Governance

9. *The Government's approach to AI governance and regulation should address each of the twelve challenges we have outlined, both through domestic policy and international engagement.* (Paragraph 88)

The Government approach to AI

10. The UK has a long history of technological innovation and regulatory expertise, which can help it forge a distinctive regulatory path on AI. The AI white paper should be welcomed as an initial effort to engage with a complex task. However, the approach outlined is already risking falling behind the pace of development of AI. (Paragraph 102)
11. The UK Government's proposed approach to AI governance relies heavily on our existing regulatory system, and the promised central support functions. The time required to establish new regulatory bodies means that adopting a sectoral approach, at least initially, is a sensible starting point. We have heard that many regulators are already actively engaged with the implications of AI for their respective remits, both individually and through initiatives such as the Digital Regulation Cooperation Forum. However, it is already clear that the resolution of all of the Challenges set out in this report may require a more well-developed central coordinating function. (Paragraph 103)
12. The AI white paper is right to highlight the importance of regulatory capacity to the successful implementation of its principles. *The Government should, as part of its implementation of its proposals, undertake a gap analysis of the UK's regulators, which considers not only resourcing and capacity, but whether any regulators require new powers to implement and enforce the principles outlined in the AI white paper.* (Paragraph 104)
13. The Government is yet to confirm whether AI-specific legislation will be included in the upcoming King's Speech in November. This new session of Parliament will be the last opportunity before the General Election for the UK to legislate on the governance of AI. Following the Election it is unlikely that new legislation could be enacted until late 2025—more than two years from now and nearly three years from the publication of the white paper. (Paragraph 105)
14. The Government has said in the AI white paper that it may legislate, at a minimum, to establish 'due regard' duties for existing regulators. That commitment alone—in addition to any further requirements that may emerge—suggests that there should be a tightly-focussed AI Bill in the new session of Parliament. Our view is that this would help, not hinder, the Prime Minister's ambition to position the UK as an AI governance leader. We see a danger that if the UK does not bring in any new statutory regulation for three years it risks the Government's good intentions being left behind by other legislation—like the EU AI Act—that could become the de facto standard and be hard to displace. (Paragraph 106)

15. *In its reply to this interim Report, and its response to the AI white paper consultation, the Government should confirm whether AI-specific legislation, such as the introduction of a requirement for regulators to pay due regard to the AI white paper principles, will be introduced in the next Parliament. It should also confirm what work has been undertaken across Government to explore the possible contents of such a Bill.* (Paragraph 107)
16. We welcome the establishment of a Foundation Model Taskforce, the appointment of Ian Hogarth as its chair, and the Government's stated intention for it to take a similar approach to the Vaccines Taskforce. This agile approach is necessary and proportionate to the importance of the issue. *The Government should confirm the Task Force's full membership, terms of reference, and the first tranche of public sector pilot projects, in its reply to this interim Report.* (Paragraph 108)

The international dimension

17. The Prime Minister was right to say that AI does not respect national borders, and we welcome the announcement of a global summit on AI safety in London. *The challenges highlighted in our interim Report should form the basis for these important international discussions.* (Paragraph 119)
18. *The summit should aim to advance a shared international understanding of the challenges of AI—as well as its opportunities. Invitations to the summit should therefore be extended to as wide a range of countries as possible. Given the importance of AI to our national security there should also be a forum established for like-minded countries who share liberal, democratic values, to be able to develop an enhanced mutual protection against those actors—state and otherwise—who are enemies of these values.* (Paragraph 120)

Conclusion and next steps

19. There is as little consensus about how AI will evolve as there has been excitement and hyperbole following its rise to ubiquity. AI cannot be un-invented. It has and will continue to change the way we live our lives. Humans must take measures to safely harness the benefits of the technology and encourage future innovations, whilst providing credible protection against harm. (Paragraph 121)
20. Some observers have called for the development of certain types of AI models and tools to be paused, allowing global regulatory and governance frameworks to catch up. We are unconvinced that such a pause is deliverable. When AI leaders say that new regulation is essential, their calls cannot responsibly be ignored—although it should also be remembered that is not unknown for those who have secured an advantageous position to seek to defend it against market insurgents through regulation. (Paragraph 122)
21. *The twelve Challenges of AI Governance which we have set out must be addressed by policymakers in all jurisdictions. Different administrations may choose different ways to do this.* (Paragraph 123)

22. We believe that the UK's depth of expertise in AI and the disciplines which contribute to it—the vibrant and competitive developer and content industry that the UK is home to; and the UK's longstanding reputation for developing trustworthy and innovative regulation—provides a major opportunity for the UK to be one of the go-to places in the world for the development and deployment of AI. But that opportunity is time-limited. Without a serious, rapid and effective effort to establish the right governance frameworks—and to ensure a leading role in international initiatives—other jurisdictions will steal a march and the frameworks that they lay down may become the default even if they are less effective than what the UK can offer. *We urge the Government to accelerate, not to pause, the establishment of a governance regime for AI, including whatever statutory measures as may be needed.* (Paragraph 124)

Formal minutes

Wednesday 19 July 2023

Greg Clark, in the Chair

Dawn Butler

Chris Clarkson

Tracey Crouch

Katherine Fletcher

Rebecca Long-Bailey

Stephen Metcalfe

Graham Stringer

Draft Report (*Governance of artificial intelligence: interim Report*), proposed by the Chair, brought up and read.

Ordered, That the draft Report be read a second time, paragraph by paragraph.

Paragraphs 1 to 124 read and agreed to.

Summary agreed to.

Resolved, That the Report be the Ninth Report of the Committee to the House.

Ordered, That the Chair make the Report to the House.

Ordered, That embargoed copies of the Report be made available, in accordance with the provisions of Standing Order No. 134.

Adjournment

Adjourned till Wednesday 6 September 2023 at 9.20am.

Witnesses

The following witnesses gave evidence. Transcripts can be viewed on the [inquiry publications page](#) of the Committee's website.

Wednesday 25 January 2023

Professor Michael Osborne, Professor of Machine Learning and co-founder, University of Oxford and Mind Foundry; **Michael Cohen**, DPhil candidate in Engineering Science, University of Oxford [Q1–54](#)

Mrs Katherine Holden, Head of Data Analytics, AI and Digital Identity, techUK; **Dr Manish Patel**, CEO, Jiva.ai [Q55–96](#)

Wednesday 22 February 2023

Adrian Joseph, Chief Data and AI Officer, BT Group; **Jen Gennai**, Director, Responsible Innovation, Google; **Hugh Milward**, General Manager, Corporate, External and Legal Affairs, Microsoft UK [Q97–143](#)

Professor Dame Wendy Hall, Regius Professor of Computer Science, University of Southampton; **Professor Sir Nigel Shadbolt**, Professorial Research Fellow in Computer Science and Principal, Jesus College, University of Oxford [Q144–173](#)

Wednesday 08 March 2023

Professor Andrew Hopkins, Chief Executive, Exscientia [Q174–222](#)

Professor Delmiro Fernandez-Reyes, Professor of Biomedical Computing, University College London, Adjunct Professor of Paediatrics, University of Ibadjan; **Professor Mihaela van der Schaar**, John Humphrey Plummer Professor of Machine Learning, Artificial Intelligence and Medicine, and Director, Cambridge Centre for AI in Medicine, The University of Cambridge [Q223–262](#)

Wednesday 29 March 2023

Professor Rose Luckin, Professor of Learner Centred Design, University College London, Director, Educate; **Daisy Christodoulou**, Director of Education, No More Marking [Q263–294](#)

Dr Matthew Glanville, Head of Assessment Principles and Practice, The International Baccalaureate; **Joel Kenyon**, Science Teacher and Community Cohesion Lead, Dormers Wells High School, Southall, London [Q295–326](#)

Wednesday 10 May 2023

Jamie Njoku-Goodwin, CEO, UK Music; **Paul Fleming**, General Secretary, Equity [Q327–373](#)

Coran Darling, Associate, Intellectual Property and Technology, DLA Piper; **Dr Hayleigh Boshier**, Senior Lecturer in Intellectual Property Law, Brunel University [Q374–411](#)

Wednesday 24 May 2023

Lindsey Chiswick, Director of Intelligence, Metropolitan Police; **Dr Tony Mansfield**, Principal Research Scientist, National Physical Laboratory [Q412–506](#)

Michael Birtwistle, Associate Director, AI and data law & policy, Ada Lovelace Institute; **Dr Marion Oswald**, Senior Research Associate for Safe and Ethical AI and Associate Professor in Law, The Alan Turing Institute and Northumbria University

[Q507-538](#)

Published written evidence

The following written evidence was received and can be viewed on the [inquiry publications page](#) of the Committee's website.

GAI numbers are generated by the evidence processing system and so may not be complete.

- 1 ACT | The App Association ([GAI0018](#))
- 2 ADS ([GAI0027](#))
- 3 AI & Digital Healthcare Group, Centre for Regulatory Science and Innovation, Birmingham (University Hospitals Birmingham NHS Foundation Trust/University of Birmingham) ([GAI0055](#))
- 4 AI Centre ([GAI0037](#))
- 5 AI Governance Limited ([GAI0050](#))
- 6 Abrusci, Dr Elena (Lecturer, Brunel University London); and Scott, Dr Richard Mackenzie-Gray (Postdoctoral Fellow, University of Oxford) ([GAI0038](#))
- 7 Academy of Medical Sciences ([GAI0072](#))
- 8 Ada Lovelace Institute ([GAI0086](#))
- 9 Alfieri, Joseph ([GAI0062](#))
- 10 Alliance for Intellectual Property ([GAI0118](#))
- 11 Assuring Autonomy International Programme (AAIP), University of York; McDermid, Professor John; Calinescu, Professor Radu; MacIntosh, Dr Ana; Habli, Professor Ibrahim; and Hawkins, Dr Richard ([GAI0044](#))
- 12 BCS - Chartered Institute for Information Technology ([GAI0022](#))
- 13 BILETA ([GAI0082](#))
- 14 BT Group ([GAI0091](#))
- 15 Belfield, Mr Haydn (Academic Project Manager, University of Cambridge, Leverhulme Centre for the Future of Intelligence & Centre for the Study of Existential Risk); igeartaigh, Dr Seán Ó hÉ (Acting Director and Principal Researcher, University of Cambridge, Centre for the Study of Existential Risk & Leverhulme Centre for the Future of Intelligence); Avin, Dr Shahar (Senior Research Associate, University of Cambridge, Centre for the Study of Existential Risk); ndez-Orallo, Prof José Herná (Professor, Universitat Politècnica de València); and Corsi, Giulio (Research Associate, University of Cambridge, Leverhulme Centre for the Future of Intelligence)) ([GAI0094](#))
- 16 Big Brother Watch ([GAI0088](#))
- 17 British Standards Institution (BSI) ([GAI0028](#))
- 18 Burges Salmon LLP ([GAI0064](#))
- 19 CBI ([GAI0115](#))
- 20 CENTRIC ([GAI0043](#))
- 21 Carnegie UK ([GAI0041](#))
- 22 Center for AI and Digital Policy ([GAI0098](#))
- 23 Chiswick, Lindsey (Director of Intelligence, Metropolitan Police) ([GAI0121](#))
- 24 Clement-Jones, Lord (Digital Spokesperson for the Liberal Democrats, House of Lords); and Darling, Coran ([GAI0101](#))

- 25 Cohen, Michael (DPhil Candidate, University of Oxford); and Osborne, Professor Michael (Professor of Machine Learning, University of Oxford) ([GAI0046](#), [GAI0116](#))
- 26 Collins, Dr Philippa (Senior Lecturer in Law, University of Bristol); and Atkinson, Dr Joe (Lecturer in Law, University of Sheffield) ([GAI0074](#))
- 27 Committee on Standards in Public Life ([GAI0110](#))
- 28 Compliant & Accountable Systems Research Group, Department of Computer Science & Technology, University of Cambridge; and Compliant & Accountable Systems Research Group, Department of Computer Science & Technology, University of Cambridge ([GAI0106](#))
- 29 Connected by Data ([GAI0052](#))
- 30 Copyright Alliance ([GAI0097](#))
- 31 Creative Commons ([GAI0015](#))
- 32 Crockett, Professor of Computational Intelligence Keeley (Professor of Computational Intelligence, Manchester Metropolitan University) ([GAI0020](#))
- 33 DeepMind ([GAI0100](#))
- 34 Department for Digital, Culture, Media and Sport; and Department for Business, Energy and Industrial Strategy ([GAI0107](#))
- 35 Edwards, Professor Rosalind (Professor of Sociology , University of Southampton); Gillies, Professor Val (Professor of Social Policy and Criminology , University of Westminster); Gorin, Dr. Sarah (Assistant Professor , University of Warwick); and Ducasse, Dr. H el ene Vannier (Senior Research Fellow , University of Southampton) ([GAI0035](#))
- 36 Employers Lawyers Association ([GAI0031](#))
- 37 Equity ([GAI0065](#))
- 38 Fotheringham, Kit (Postgraduate Researcher, University of Bristol) ([GAI0042](#))
- 39 GSK ([GAI0067](#))
- 40 Google ([GAI0099](#))
- 41 Hopgood, Professor Adrian (Professor of Intelligent Systems, University of Portsmouth) ([GAI0030](#))
- 42 Imperial College London Artificial Intelligence Network ([GAI0014](#))
- 43 Information Commissioner’s Office (ICO) ([GAI0112](#))
- 44 Institute for the Future of Work ([GAI0063](#))
- 45 Institute of Physics and Engineering in Medicine (IPEM) ([GAI0051](#))
- 46 Leslie, Professor David (Director of Ethics and Responsible Innovation Research, The Alan Turing Institute; and Professor of Ethics, Technology and Society, Queen Mary University of London) ([GAI0113](#))
- 47 Liberty ([GAI0081](#))
- 48 Library and Archives Copyright Alliance ([GAI0120](#))
- 49 Loughborough University ([GAI0070](#))
- 50 Mason, Mr Shane (Freelance consultant) ([GAI0006](#))
- 51 Microsoft ([GAI0083](#))
- 52 Minderoo Centre for Technology and Democracy, University of Cambridge ([GAI0032](#))

- 53 NCC Group ([GAI0040](#))
- 54 NICE; HRA; MHRA; and CQC ([GAI0076](#))
- 55 National Physical Laboratory ([GAI0053](#))
- 56 Oswald, Dr Marion ([GAI0012](#))
- 57 Oxford Internet Institute ([GAI0058](#))
- 58 Oxford Internet Institute, University of Oxford; University of Exeter ([GAI0024](#))
- 59 Patelli, Dr Alina (Senior Lecturer in Computer Science, Aston University) ([GAI0095](#))
- 60 Protect Pure Maths ([GAI0117](#))
- 61 Public Law Project ([GAI0069](#))
- 62 Publishers Association ([GAI0102](#))
- 63 Pupils 2 Parliament ([GAI0096](#))
- 64 Queen Mary University London ([GAI0073](#))
- 65 RELX ([GAI0033](#))
- 66 Reed, Professor Chris (Professor of Electronic Commerce Law, Centre for Commercial Law Studies, Queen Mary University of London) ([GAI0059](#))
- 67 Richie, Dr Cristina (lecturer, TU Delft) ([GAI0001](#), [GAI0002](#))
- 68 Rolf, Dr Steve (Research Fellow, The Digital Futures at Work (Digit) Centre, University of Sussex Business School) ([GAI0104](#))
- 69 Rolls-Royce plc ([GAI0109](#))
- 70 Sage Group ([GAI0108](#))
- 71 Salesforce ([GAI0105](#))
- 72 Sanchez-Graells, Professor Albert (Professor of Economic Law, University of Bristol Law School) ([GAI0004](#))
- 73 School of Informatics, University of Edinburgh ([GAI0079](#))
- 74 Scott, Mr. Michael (Chair of Trustees, Home-Start Nottingham) ([GAI0005](#))
- 75 Sense about Science ([GAI0078](#))
- 76 TUC ([GAI0060](#))
- 77 Tang, Dr Guan H (Senior Lecturer, Centre for Commercial Law Studies, Queen Mary University of London) ([GAI0077](#))
- 78 TechWorks ([GAI0068](#))
- 79 Tessler, Leonardo (PhD in law Candidate, University of Montreal) ([GAI0092](#))
- 80 The Alliance for Intellectual Property ([GAI0103](#))
- 81 The Institution of Engineering and Technology (IET) ([GAI0021](#))
- 82 The LSE Law School, London School of Economics.; and The LSE Law School, London School of Economics. ([GAI0036](#))
- 83 The Nutrition Society ([GAI0007](#))
- 84 The Royal Academy of Engineering ([GAI0039](#))
- 85 The Royal College of Radiologists (RCR) ([GAI0087](#))
- 86 Thorney Isle Research ([GAI0016](#))

- 87 Tripathi, Mr Karan (Research Associate , University of Sheffield); and Tzanou , Dr Maria (Senior Lecturer in Law, University of Sheffield) ([GAI0047](#))
- 88 Trustpilot ([GAI0054](#))
- 89 Trustworthy Autonomous Systems Hub; The UKRI TAS Node in Governance & Regulation; and The UKRI TAS Node in Functionality ([GAI0084](#))
- 90 UK BioIndustry Association ([GAI0026](#))
- 91 UK Dementia Research Institute ([GAI0111](#))
- 92 UKRI ([GAI0114](#))
- 93 United Nations Association UK; Article 36; Women’s International League for Peace and Freedom UK; and Drone Wars UK ([GAI0090](#))
- 94 University of Glasgow ([GAI0057](#))
- 95 University of Sheffield ([GAI0017](#))
- 96 University of Surrey ([GAI0075](#))
- 97 Wayve ([GAI0061](#))
- 98 Which? ([GAI0049](#))
- 99 Whittlestone, Dr Jess (Head of AI Policy, Centre for Long-Term Resilience); and Moulange, Richard (PhD student , MRC Biostatistics Unit, University of Cambridge) ([GAI0071](#))
- 100 Wudel, Alexandra (Political Advisor, German Parliament); Gengler, Eva (PhD Student, FAU Nürnberg); and Center for Feminist Artificial Intelligence ([GAI0013](#))
- 101 Wysa Limited ([GAI0093](#))
- 102 medConfidential ([GAI0011](#))
- 103 techUK ([GAI0045](#))

List of Reports from the Committee during the current Parliament

All publications from the Committee are available on the [publications page](#) of the Committee's website.

Session 2022–23

Number	Title	Reference
1st	Pre-appointment hearing for the Executive Chair of Research England	HC 636
2nd	UK space strategy and UK satellite infrastructure	HC 100
3rd	My Science Inquiry	HC 618
4th	The role of Hydrogen in achieving Net Zero	HC 99
5th	Diversity and Inclusion in STEM	HC 95
6th	Reproducibility and Research Integrity	HC 101
7th	UK space strategy and UK satellite infrastructure: reviewing the licencing regime for launch	HC 1717
8th	Delivering nuclear power	HC 626

Session 2021–22

Number	Title	Reference
1st	Direct-to-consumer genomic testing	HC 94
2nd	Pre-appointment hearing for the Chair of UK Research and Innovation	HC 358
3rd	Coronavirus: lessons learned to date	HC 92

Session 2019–21

Number	Title	Reference
1st	The UK response to covid-19: use of scientific advice	HC 136
2nd	5G market diversification and wider lessons for critical and emerging technologies	HC 450
3rd	A new UK research funding agency	HC 778