LIBRARY BRIEFING

HOUSE OF LORDS

HL Bill 11 of 2023–24

Artificial Intelligence (Regulation) Bill [HL]

Author: James Tobin
Date published: 18 March 2024

The Artificial Intelligence (Regulation) Bill [HL] is a private member's bill proposed by Lord Holmes of Richmond (Conservative). It would establish a new body, the AI Authority, which would have various functions designed to help address artificial intelligence (AI) regulation in the UK. This would include a requirement for the AI Authority to ensure relevant existing regulators were taking account of AI; to ensure alignment in approach between these regulators; and to undertake a gap analysis of regulatory responsibilities with respect to AI. The AI authority would also have various other functions including monitoring economic risks arising from AI, conducting horizon-scanning of developing technologies, facilitating sandbox initiatives to allow the testing of new AI models, and accrediting AI auditors. In addition, the bill would introduce a set of regulatory principles governing the development and usage of AI.

The bill would represent a departure from the UK government's current approach to the regulation of AI. The government has said primary legislation will be necessary to regulate AI at some future point. However, it contends it is too soon in these technologies' evolution to legislate effectively and to do so now may be counterproductive. Ministers argue that existing sectoral regulators are best placed to regulate AI with support from a central function currently in development within the Department of Science, Innovation and Technology. In February 2024, the government announced a range of measures in support of its approach—building upon the white paper it issued on the regulation of AI in 2023—including the first iteration of guidance to regulators which includes voluntary regulatory principles. This approach has been welcomed by many including prominent technology companies such as Google and Microsoft. However, others such as the Ada Lovelace Institute have voiced concerns that relying on voluntary commitments from key AI developers rather than binding legal requirements, for example, is critically insufficient.

The UK's approach contrasts with that taken by the EU which is in the process of finalising wide-ranging legislation which will regulate the development and usage of AI across all member states. In the US, the regulation of AI is currently being examined at a federal and state level, with some states having already introduced legislation aimed at the regulation of AI particularly around privacy and accountability. The US has not legislated at a federal level. Instead, the White House issued an executive order in October 2023 setting out key principles and actions aimed at ensuring the safe development and usage of AI.

Table of Contents

1. Artificial intelligence: Rapid development

Artificial intelligence (AI) continues to develop rapidly with new tools being released on a regular basis capable of increasingly sophisticated tasks. For example, OpenAI's 'Sora' text-to-video model launched in February 2024, which can generate videos up to a minute long entirely from a text prompt.[1] However, the sophistication of these tools and their capabilities, not least the ability to generate highly realistic imagery and sophisticated text, has raised concerns over the impact on areas such as journalism and democracy.[2] Countries across the world are seeking to address how best to regulate the usage of these technologies, though key jurisdictions such as the UK, Europe and the US have taken different approaches, as examined in section 2 of this briefing.

AI is an umbrella term for several technologies, which can be broadly categorised as follows:

- **Narrow AI** is designed to perform a specific task (such as speech recognition), using information from specific datasets, and cannot adapt to perform another task. These are often tools that aim to assist, rather than replace, the work of humans.
- **Artificial general intelligence** (AGI, also referred to as 'strong' AI) is an AI system that can undertake any intellectual task/problem that a human can. AGI is a system that can reason, analyse and achieve a level of understanding that is on a par with humans; something that has yet to be achieved by AI. The US computer scientist Nils John Nilsson, for example, proposed that one way to test if a system had achieved AGI was if it could successfully learn the skills to perform the different jobs "ordinarily performed by humans", from "knowledge work" (such as a Library assistant) to "manual labour" (such as a roofer).
- **Machine learning** is a method that can be used to achieve narrow AI; it allows a system to learn and improve from examples, without all its instructions being explicitly programmed. It does this by finding patterns in large amounts of data, which it can then use to make predictions (for example what film or TV programme you might like to watch next on a streaming platform). The AI can then independently amend its algorithm based on the accuracy of its predictions.
- **Deep learning** is a type of machine learning whose design has been informed by the structure and function of the human brain and the way it transmits information. The application of deep learning can be seen in 'foundation models',

[1] Open AI, 'Sora', accessed 5 March 2024.
[2] Jeremy Hsu, 'Realism of OpenAI's Sora video generator raises security concerns', New Scientist, 17 February 2024.

> of which 'large language models' (LLMs) such as ChatGPT, are one example. The term refers to those models that are trained on very large, unlabelled datasets and which can be adapted to do a wide range of tasks, despite not having been trained explicitly to do those tasks. In other words, the model can take information it has learnt about in one situation and apply it to another, different situation. Sometimes LLMs are refined or 'fine-tuned' (trained using additional data) to achieve a specific goal. ChatGPT, for example, has been fine-tuned to allow users to ask it a question, or make a request, and for it to generate "human-like text" in response.[3]

The House of Lords Communications and Digital Committee said the world faces an "inflection point" on AI.[4] Speaking to the power of LLMs such as ChatGPT in particular, the committee said:

> [LLMs] will introduce epoch-defining changes comparable to the invention of the internet. A multi-billion pound race is underway to dominate this market. The victors will wield unprecedented power to shape commercial practices and access to information across the world.[5]

Box 1: Open vs closed AI models

A crucial part of the debate about AI is around the development of closed versus open models. The technology company Sharp has explained that in simple terms, **open AI** models are publicly accessible and can be modified by anyone. They are designed to be more flexible and adaptable, capable of learning and evolving over time. They are trained on publicly available data from across the internet, such as text articles, images, and videos. Open AI models have the source code openly shared so that people are encouraged to voluntarily improve its design and function. However, open AI models can pose a severe data security risk, and there are concerns that making source code available will allow bad actors to adapt and use them for nefarious purposes. In contrast, **closed AI** models are not publicly accessible although they may be trained on publicly available data. There is a high degree of risk control in comparison with open AI models, with closed models prioritising data security and confidentiality, maintaining strict control over internal data access and usage to safeguard sensitive information.

---

[3] House of Lords Library, 'Artificial intelligence: Development, risks and regulation', 18 July 2023. For further explanation of terms used in this area, see: Parliamentary Office of Science and Technology 'Artificial intelligence (AI) glossary', 23 January 2024.

[4] House of Lords Communications and Digital Committee, 'Large language models and generative AI', 2 February 2024, HL paper 54 of session 2023–24, p 3.

[5] As above.

Like many other commentators, the Communications and Digital Committee has said that it was optimistic about this new technology, saying it had the potential to deliver significant economic benefits and ground-breaking scientific advances. According to management consultancy PwC, the UK's GDP will be 10.3% higher in 2030 because of AI.[6] This is the equivalent of an additional £232bn. Recent analysis published by Microsoft also noted the potential benefits of the usage of AI in several areas to improve health outcomes, help conservation efforts, to drive smarter manufacturing and deliver educational benefits.[7]

However, the development of this technology also presents significant risks. Concerns have already been expressed over the use of copyrighted material to train existing and developing AI models.[8] There have been examples of fake AI-generated text and audio material involving high profile politicians such as the mayor of London, Sadiq Khan.[9] With the developing AI ability to generate realistic videos, it seems highly likely that fake videos will also soon emerge. There are also AI applications which can generate fake nude or sexually explicit imagery without the consent of those involved.[10] In addition, LLMs continue to experience an issue with hallucinations whereby they can generate false or misleading information in response to user prompts. There are concerns that as AI models develop in sophistication they will pose "credible threats" to public safety, societal values, open market competition and UK economic competitiveness.[11]

Noting these risks, in July 2023 the House of Commons Science, Innovation and Technology Committee identified 12 challenges of AI governance which it argued that policymakers and the regulatory frameworks they designed 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.

---

[6] PwC, 'The economic impact of artificial intelligence on the UK economy', June 2017.

[7] Microsoft, 'Governing AI: A blueprint for the UK', 15 February 2024.

[8] Creators Rights Alliance, 'Artificial intelligence and creative work', accessed 5 March 2024.

[9] BBC News, 'Sadiq Khan says fake AI audio of him nearly led to serious disorder', 13 February 2024.

[10] HL Hansard, 13 February 2024, cols 132–5.

[11] House of Lords Communications and Digital Committee, 'Large language models and generative AI', 2 February 2024, HL paper 54 of session 2023–24, p 3.

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.[12]

Similar sentiments were echoed by the House of Lords Communications and Digital Committee in January 2024 which said there was a "short window" for the UK to adapt to harness the benefits and address the challenges of AI. To that end, the committee published several recommendations on the regulation of AI, including the need to guard against 'regulatory capture' and to review both short-term and catastrophic risks:

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

[12] House of Commons Science, Innovation and Technology Committee, 'The governance of artificial intelligence: Interim report', 31 August 2023, HC 1769 of session 2022–23, pp 3–4.

- **Guard against regulatory capture**: There is a major race emerging between open and closed model developers. Each is seeking a beneficial regulatory framework. The government must make market competition an explicit AI policy objective. It must also introduce enhanced governance and transparency measures in the Department for Science, Innovation and Technology (DSIT) and the AI Safety Institute to guard against regulatory capture.
- **Treat open and closed arguments with care**: Open models offer greater access and competition but raise concerns about the uncontrollable proliferation of dangerous capabilities. Closed models offer more control but also more risk of concentrated power. A nuanced approach is needed. The government must review the security implications at pace while ensuring that any new rules support rather than stifle market competition.
- **Rebalance strategy towards opportunity**: The government's focus has skewed too far towards a narrow view of AI safety. It must rebalance, or else it will fail to take advantage of the opportunities from LLMs, fall behind international competitors and become strategically dependent on overseas tech firms for a critical technology.
- **Boost opportunities**: We call for a suite of measures to boost computing power and infrastructure, skills, and support for academic spinouts. The government should also explore the options for and feasibility of developing a sovereign LLM capability, built to the highest security and ethical standards.
- **Support copyright**: The government should prioritise fairness and responsible innovation. It must resolve disputes definitively (including through updated legislation if needed); empower rightsholders to check if their data has been used without permission; and invest in large, high-quality training datasets to encourage tech firms to use licenced material.
- **Address immediate risks**: The most immediate security risks from LLMs arise from making existing malicious activities easier and cheaper. These pose credible threats to public safety and financial security. Faster mitigations are needed in cyber security, counter terror, child sexual abuse material and disinformation. Better assessments and guardrails are needed to tackle societal harms around discrimination, bias and data protection too.
- **Review catastrophic risks**: Catastrophic risks (above 1,000 UK deaths and tens of billions in financial damages) are not likely within three years but cannot be ruled out, especially as next-generation capabilities come online. There are however no agreed warning indicators for catastrophic risk. There is no cause for panic, but this intelligence blind spot requires immediate attention. Mandatory safety tests for high-risk high-impact models are also needed: relying on voluntary commitments from a few firms would be naïve and leaves the government unable to respond to the sudden emergence of dangerous capabilities. Wider concerns

about existential risk (posing a global threat to human life) are exaggerated and must not distract policymakers from more immediate priorities.

- **Empower regulators**: The government is relying on sector regulators to deliver the white paper objectives [see section 2.1.1 of this briefing] but is being too slow to give them the tools. Speedier resourcing of government-led central support teams is needed, alongside investigatory and sanctioning powers for some regulators, cross-sector guidelines, and a legal review of liability.
- **Regulate proportionately**: The UK should forge its own path on AI regulation, learning from but not copying the US, EU and China. In doing so the UK can maintain strategic flexibility and set an example to the world—though it needs to get the groundwork in first. The immediate priority is to develop accredited standards and common auditing methods at pace to ensure responsible innovation, support business adoption, and enable meaningful regulatory oversight.[13]

2. Regulatory approaches to AI

So far, whilst there have been attempts to coordinate international action on AI, notably by the G7 and at the AI safety summit hosted by the UK in November 2023,[14] many nations and jurisdictions are taking different approaches to regulation. The UK's approach is examined below, as are the EU and USA's approaches as a means for comparison.

2.1 UK approach to AI regulation

2.1.1 Government white paper: March 2023

The government set out its approach to the regulation of AI in its March 2023 white paper 'A pro-innovation approach to AI regulation'.[15] Noting that across the world countries and regions were drafting the rules for AI, the white paper said that the UK "needed to act quickly to continue to lead the international conversation on AI governance and demonstrate

[13] House of Lords Communications and Digital Committee, 'Large language models and generative AI', 2 February 2024, HL paper 54 of session 2023–24, pp 4–5.

[14] European Commission, 'G7 leaders' statement on the Hiroshima AI process', 30 October 2023; and Department for Innovation, Science and Technology et al, 'The Bletchley declaration by countries attending the AI safety summit, 1–2 November 2023', 1 November 2023.

[15] Department for Science, Innovation and Technology and Office for Artificial Intelligence, 'A pro-innovation approach to AI regulation', March 2023.

the value of our pragmatic, proportionate regulatory approach".

To that end, the white paper did not propose the creation of a new AI regulator nor the introduction of primary legislation which would seek to introduce AI regulatory principles or structures. Instead, the government said it would put in place a new framework to bring "clarity and coherence" to the AI regulatory landscape underpinned by five principles to "guide and inform the responsible development and use of AI in all sectors of the economy". These principles were:

- safety, security and robustness
- appropriate transparency and explainability
- fairness
- accountability and governance
- contestability and redress

The white paper proposed that this approach would rely upon the UK's existing sectoral regulators supplemented by the following government-led "central functions" to provide support, coordination and coherence. Key areas for those central functions included: [16]

- monitoring, assessment and feedback
- supporting coherent implementation of the principles
- cross-sector risk assessment
- horizon scanning
- supporting innovators (including testbeds and sandboxes)
- education and awareness
- international interoperability

The white paper also did not rule out future legislation. It said that the government would not put these principles on a statutory footing initially, fearing that new "rigid and onerous" legislative requirements on businesses "could hold back AI innovation and reduce our ability

[16] House of Lords Communications and Digital Committee, 'Large language models and generative AI', 2 February 2024, HL paper 54 of session 2023–24, p 55.

to respond quickly and in a proportionate way to future technological advances".[17] Instead, the principles would be issued on a non-statutory basis and implemented by existing regulators. The white paper said this approach made use of regulators' domain-specific expertise to tailor the implementation of the principles to the specific context in which AI is used. However, the paper also added that "following this initial period of implementation", the government anticipated introducing a statutory duty on regulators requiring them to have due regard to the principles.

2.1.2 Government consultation response and 'phase 1' guidance to regulators: February 2024

The government consulted on the white paper's proposals, publishing its response on 6 February 2024.[18] The government said it had received "strong support" for its context-based framework relying on existing regulators and that several regulators were already acting in line with that proposed approach. It cited the examples of the Competition and Markets Authority's (CMA) review of foundation models and the updated guidance on data protection and AI by the Information Commissioner's Office (ICO). The government said it was asking a number of regulators to publish an update outlining their strategic approach to AI by 30 April 2024.[19]

The government also said it had started developing the central function to support effective risk monitoring, regulator coordination, and knowledge exchange. In initial guidance to regulators, published at the same time as the consultation response (and which the government refers to as 'phase 1' guidance), it said:

> To ensure a coherent and streamlined AI regulatory landscape, DSIT has started establishing a central function. The central function supports UK regulators' understanding of the AI risk landscape and will support them to conduct risk assessments by providing expert risk analysis, which is already underway within DSIT. This allows us to monitor risks holistically and identify any potential gaps in our approach that leave risk not adequately mitigated.

---

[17] Department for Science, Innovation and Technology and Office for Artificial Intelligence, 'A pro-innovation approach to AI regulation', March 2023.

[18] Department for Science, Innovation and Technology, 'A pro-innovation approach to AI regulation: Government response to consultation', 6 February 2024, CP 1019.

[19] The government wrote to a set of regulators to this effect on 15 February 2024: Department for Science, Innovation and Technology et al, 'Request for regulators to publish an update on their strategic approach to AI: Secretary of state letters', 15 February 2024.

> The central function also catalyses the development of regulators' skills and expertise in AI. In its white paper consultation response, DSIT announced a £10mn package to boost regulators' AI capabilities. The central function will work closely with regulators in the coming months to identify the most promising opportunities to leverage this funding and will continue to support regulators to future-proof their AI capabilities, as AI technologies and the broader context in which they are used continue to change.
>
> Further key roles of the central function include supporting increased coherence across regulators, promoting information sharing and working with regulators to analyse and review potential gaps in existing regulatory powers and remits. However, we also note that some of these principles may not be relevant for specific regulators.[20]

The response said the role of the central coordination function would be set out further in 'phase 2' guidance which the government expected to publish by summer 2024. The government also said a third phase would follow, which will "involve collaborative working with regulators to identify areas for potential joint tools and guidance across regulatory remits".[21] It said that "through this process, we will aim to work with regulators to identify where additional information, resources are needed, and where appropriate, collaborate on joint solutions—for example, encouraging multi-regulator guidance".[22]

The initial guidance to regulators includes the five "pro-innovation" regulatory principles contained within the 2023 white paper (safety, security and robustness; appropriate transparency and explainability; fairness; accountability and governance; and contestability and redress) on a voluntary basis. The government has said that it does not want this approach to be prescriptive but rather up to individual regulators to decide how best to address the issue:

> [This guidance] is not intended to be a prescriptive guide on implementation as the principles are voluntary and how they are considered is ultimately at regulators' discretion. Elements of this guidance may not be applicable to regulators who adopt a 'technology agnostic' approach to regulation as long as these regulators are satisfied that their regulatory framework adequately covers issues relating to AI adoption.[23]

---

[20] Department for Science, Innovation and Technology, 'Implementing the UK's AI regulatory principles', February 2024, p 5.
[21] As above.
[22] As above.
[23] As above, p 4.

The government's consultation response also announced a new commitment by UK Research and Innovation (UKRI) that future investments in AI research will be leveraged to support regulator skills and expertise; and a £9mn partnership with the US on responsible AI as part of its 'International science partnerships fund'.[24]

In the consultation response, the government said that it remained convinced that its approach was the right one for the moment, and whilst it remained too early for legislation, mandatory measures would ultimately be required:

> Our overall approach—combining cross-sectoral principles and a context-specific framework, international leadership and collaboration, and voluntary measures on developers—is right today as it allows us to keep pace with rapid and uncertain advances in AI. However, the challenges posed by AI technologies will ultimately require legislative action in every country once understanding of risk has matured. In this document, we build on our pro-innovation framework and pro-safety actions by setting out our early thinking and the questions that we will need to consider for the next stage of our regulatory approach. Recognising there are no easy answers, we will work closely with civil society, industry, and international partners to examine these issues, and will be transparent in sharing early expert views on them.
>
> As AI systems advance in capability and societal impact, it is clear that some mandatory measures will ultimately be required across all jurisdictions to address potential AI-related harms, ensure public safety, and let us realise the transformative opportunities that the technology offers. However, acting before we properly understand the risks and appropriate mitigations would harm our ability to benefit from technological progress while leaving us unable to adapt quickly to emerging risks. We are going to take our time to get this right—we will legislate when we are confident that it is the right thing to do.[25]

Recognising the challenges posed by the rapid development of general-purpose AI, the government said it was setting out a "pro-innovation case for further targeted binding requirements on the small number of organisations developing highly capable general-purpose AI systems to ensure that they are accountable for making these technologies sufficiently safe".[26] Again, ministers asserted this could be done whilst allowing regulators to

---

[24] Department for Science, Innovation and Technology, 'A pro-innovation approach to AI regulation: Government response to consultation', 6 February 2024, CP 1019.

[25] As above, p 7.

[26] As above.

provide effective rules for the use of AI within their respective remits.

The paper also said that the government would continue its work "to address the key issues of today, from electoral interference to discrimination to intellectual property law, and the most pressing risks of tomorrow, such as biosecurity and AI alignment". This included activities to support regulator capabilities and coordination as highlighted above, such as a new steering committee with government and regulator representatives to support coordination across the AI governance landscape, and targeted consultations on the cross-economy AI risk register.

On the usage of regulatory sandboxes and testbeds—allowing new technologies to be tested under the supervision of/by regulators and which is a focus of the Artificial Intelligence (Regulation) Bill [HL]—the consultation document noted strong support for a regulatory sandbox for AI. As a result, the government said it would fund the pilot of a multi-regulator advisory service offered by the Digital Regulation Cooperation Forum (which was itself established to ensure greater cooperation on online regulatory matters), called the 'AI and Digital Hub'. This hub brings together four of what the government calls the "most critical regulators of AI and digital technologies", including the CMA, the ICO, Ofcom, and the Financial Conduct Authority. The government says that this hub, backed by £2mn in government funding, will provide tailored support to businesses so they can meet requirements across various sectors while "safely innovating".[27]

2.1.3 Other key UK developments

The consultation response published on 6 February 2024 also said that the government was providing "wider support for the AI ecosystem".[28] It noted its previous commitment to provide £1.5bn to build the next generation of supercomputers in the public sector and also announced a further £80mn boost in AI research through the launch of nine new research hubs across the UK to "propel transformative innovations" and "harness the power of AI in everything from mathematics to healthcare". The paper also noted the AI safety summit hosted by the government in November 2023 where leaders discussed and agreed actions to address emerging risks posed by the development and deployment of the most powerful AI systems. Leading AI developers set out the steps they are already taking to make models safe and committed to sharing the most powerful AI models with governments for testing so that

---

[27] Department for Science, Innovation and Technology, 'New advisory service to help businesses launch AI and digital innovations', 19 September 2023.

[28] Department for Science, Innovation and Technology, 'A pro-innovation approach to AI regulation: Government response to consultation', 6 February 2024, p 7, CP 1019.

"we can ensure safety today and prepare for the risks of tomorrow".[29]

To support this aim, the government has created an AI Safety Institute to lead evaluations and safety research in the UK government.[30] Ministers have said that the AI Safety Institute will work in collaboration with partners across the world, including in the US, in developing a coherent and collaborative approach to international governance. The institute recently published its third progress report which included an update on recruitment.[31]

In July 2023, the Committee on Standards in Public Life wrote to UK regulators asking them for an update on how they are adapting to the challenges posed by AI.[32] The responses received from those regulators have now been published.[33] The committee also wrote to government departments and public bodies with similar questions, and their responses are also available.[34] On 6 March 2024, the committee published an update of progress made by regulators, government departments and public bodies against the recommendations the committee made in its 2020 report 'Artificial intelligence and public standards'.[35]

2.1.4 Reaction to the UK government's approach

The announcement of the government's consultation response was accompanied by several supportive comments, particularly from those in the technology industries. For example, Hugh Milward, vice-president, external affairs Microsoft UK, welcomed the government's response to the white paper consultation. He said that seizing "AI's potential to grow our economy, revolutionise public services and tackle major societal challenges" would require "responsible and flexible regulation that supports the UK's global leadership in the era of AI".[36]

---

[29] Department for Science, Innovation and Technology, 'A pro-innovation approach to AI regulation: Government response to consultation', 6 February 2024, p 7, CP 1019.

[30] Department for Science, Innovation and Technology and AI Safety Institute, 'AI Safety Institute: Overview', updated 17 January 2024.

[31] AI Safety Institute, 'AI Safety Institute: Third progress report', 5 February 2024.

[32] Committee on Standards in Public Life, 'Letter from Lord Evans to regulators asking them for an update on how they are adapting to the challenges of AI', 4 July 2023.

[33] Committee on Standards in Public Life, 'Responses from regulators on how they are adapting to the challenges posed by AI', updated 6 March 2024.

[34] Committee on Standards in Public Life, 'Responses from public bodies on how they are adapting their governance processes for AI', updated 6 March 2024.

[35] Committee on Standards in Public Life, 'AI and public standards: An update on progress made against our 2020 recommendations', 6 March 2024.

[36] Department for Science, Innovation and Technology and UK Research and Innovation, 'UK signals step change for regulators to strengthen AI leadership', 6 February 2024.

Similarly, Lila Ibrahim, chief operating officer at Google DeepMind, welcomed the balance she said the government had struck between "supporting innovation and ensuring AI is used safely and responsibly".[37] She said the hub and spoke model would help the UK "benefit from the domain expertise of regulators, as well as provide clarity to the AI ecosystem". She expressed support for the commitment to provide regulators with further resources.[38]

However, others have been more critical of the government's approach. For example, Michael Birtwistle, associate director (Law and policy) at the Ada Lovelace Institute, an independent research institute on AI and data, said that the government should be given credit for strengthening its initial approach but that much more needed to be done:

> The government should be given credit for evolving and strengthening its initially light-touch approach to AI regulation in response to the emergence of general-purpose AI systems. Ministers are right to acknowledge that AI is already causing harm in many everyday contexts and poses a broad range of risks to society. The government's work to build in-house expertise on AI through the establishment of the central AI risk function and the AI Safety Institute, as well as its development of standards on algorithmic transparency and AI management, are promising first steps. However, much more needs to be done to ensure that AI works in the best interests of the diverse publics who use these technologies.
>
> We are concerned that the government's approach to AI regulation is 'all eyes, no hands': it has equipped itself with significant horizon-scanning capabilities to anticipate and monitor AI risks, but it has not given itself the powers and resources to prevent those risks or even react to them effectively after the fact. While an uplift in regulatory funding is welcome, £10mn falls well short of the hundreds of millions pounds per annum that we allocate to safety in other critical industries.
>
> Unless binding legislation is brought forward, the government's approach to regulating AI will remain reliant on the goodwill of powerful AI companies like Microsoft, Google and Meta. Voluntary commitments to good practice are not enough: the evidence shows that only hard rules enshrined in law will incentivise developers and deployers of AI to comply and empower regulators to act.[39]

---

[37] Department for Science, Innovation and Technology and UK Research and Innovation, 'UK signals step change for regulators to strengthen AI leadership', 6 February 2024.

[38] As above.

[39] Ada Lovelace Institute, 'Ada Lovelace Institute statement on the UK's approach to AI regulation', 7 February 2024.

2.2 Comparator regulatory regimes in the EU and USA

2.2.1 EU AI Act

The European Union is legislating to introduce its first binding legal regulatory framework on AI. The European Commission has said the AI Act aims to provide AI developers and deployers with "clear requirements and obligations regarding specific uses of AI".[40] At the same time, the act seeks to reduce administrative and financial burdens for business, particularly for small and medium-sized enterprises (SMEs). The European Commission also notes that the AI Act is part of a wider package of policy measures to support the development of trustworthy AI, which also includes the AI innovation package and the 'Coordinated plan on AI'.[41]

The commission says that the proposed legislation will do the following:[42]

- address risks specifically created by AI applications
- prohibit AI practices that pose unacceptable risks
- determine a list of high-risk applications
- set clear requirements for AI systems for high-risk applications
- define specific obligations for deployers and providers of high-risk AI applications
- require a conformity assessment before a given AI system is put into service or placed on the market
- put enforcement in place after a given AI system is placed into the market
- establish a governance structure at European and national level

The act is founded upon a 'risk-based' approach to AI whereby such technologies will be categorised according to four levels of risk—unacceptable, high, limited and minimal—and subject to corresponding levels of restrictions.

On those technologies posing an 'unacceptable risk', the European Commission stated that all "AI systems considered a clear threat to the safety, livelihoods and rights of people will be

[40] European Commission, 'AI Act', 6 March 2024.
[41] European Commission, 'Commission launches AI innovation package to support artificial intelligence startups and SMEs', 24 January 2024; and 'Coordinated plan on artificial intelligence', 6 March 2024.
[42] European Commission, 'AI Act', 6 March 2024.

banned".[43] It said this would include, for example, social scoring by governments and toys using voice assistance that encourages dangerous behaviour.

AI technologies designated as high-risk include those used in the following areas:

- critical infrastructures (for example transport), that could put the life and health of citizens at risk
- educational or vocational training, that may determine the access to education and professional course of someone's life (for example scoring of exams)
- safety components of products (for example AI application in robot-assisted surgery)
- employment, management of workers and access to self-employment (for example CV-sorting software for recruitment procedures)
- essential private and public services (for example credit scoring denying citizens opportunity to obtain a loan)
- law enforcement that may interfere with people's fundamental rights (for example evaluation of the reliability of evidence)
- migration, asylum and border control management (for example automated examination of visa applications)
- administration of justice and democratic processes (for example AI solutions to search for court rulings)[44]

The European Commission states that such high-risk AI systems will be subject to strict obligations before they can be put on the market, including:

- adequate risk assessment and mitigation systems
- high quality of the datasets feeding the system to minimise risks and discriminatory outcomes
- logging of activity to ensure traceability of results
- detailed documentation providing all information necessary on the system and its purpose for authorities to assess its compliance
- clear and adequate information to the deployer

[43] European Commission, 'AI Act', 6 March 2024.
[44] As above.

- appropriate human oversight measures to minimise risk
- high level of robustness, security and accuracy[45]

The commission adds that all remote biometric identification systems are considered high-risk and subject to strict requirements. The use of remote biometric identification in publicly accessible spaces for law enforcement purposes is, in principle, prohibited. The commission adds that narrow exceptions would be strictly defined and regulated by the act, such as when necessary to search for a missing child, to prevent a specific and imminent terrorist threat or to detect, locate, identify or prosecute a perpetrator or suspect of a serious criminal offence. Those usages are subject to authorisation by a judicial or other independent body and to appropriate limits in time, geographic reach and the databases searched.

Limited risk refers to the risks associated with "lack of transparency in AI usage". The commission says that the AI Act introduces specific transparency obligations to ensure that humans are informed when necessary, fostering trust. For instance, when using AI systems such as chatbots, humans should be made aware that they are interacting with a machine so they can take an informed decision to continue or step back. Providers will also have to ensure that AI-generated content is identifiable. Further, AI-generated text published with the purpose of informing the public on matters of public interest must be labelled as artificially generated. This also applies to audio and video content constituting deep fakes.

The AI Act allows the free use of minimal-risk AI. This includes applications such as AI-enabled video games or spam filters.

Regarding general-purpose AI/large language models, the AI Act introduces transparency obligations "to enable a better understanding of these models and additional risk management obligations for very capable and impactful models". These additional obligations include self-assessment and mitigation of systemic risks, reporting of serious incidents, conducting test and model evaluations, as well as cybersecurity requirements.

The EU also aims to make the act 'future-proof' by enabling ongoing quality and risk management so that the rules could adapt to technological change. The European AI Office, established in February 2024 within the European Commission, will oversee the AI Act's enforcement and implementation with EU member states.

---

[45] European Commission, 'AI Act', 6 March 2024.

In December 2023, the European Parliament and the Council of the European Union reached a political agreement on the AI Act. This was followed in February 2024 by EU deputy ambassadors unanimously agreeing the text following negotiations between representatives of the European Council, members of the European Parliament and European Commission officials.[46] The European Parliament approved the AI Act on 13 March 2024.[47] The legislation will enter into force 20 days after its publication in the official journal, and will be fully applicable two years later, with some exceptions: prohibitions will take effect after six months, the governance rules and the obligations for general-purpose AI models become applicable after 12 months and the rules for AI systems—embedded into regulated products—will apply after 36 months.[48]

To facilitate the transition to the new regulatory framework, the commission has also launched the 'AI pact'.[49] This is a voluntary initiative that seeks to support the future implementation and invites AI developers from Europe and beyond to comply with the key obligations of the AI Act ahead of them coming into force.

2.2.2 US approach to AI regulation

The United States is examining the regulation of AI at both a state and federal level.[50] Since 2019, 17 states have enacted 29 bills focused on regulating the design, development, and use of artificial intelligence.[51] These bills primarily address two regulatory concerns: data privacy and accountability. Legislatures in California, Colorado and Virginia have established regulatory and compliance frameworks for AI systems.

At a federal level, the White House issued an executive order on AI in October 2023.[52] That order directed new standards for AI safety and security, including the following actions:

- require that developers of the most powerful AI systems share their safety test results and other critical information with the US government

---

[46] Politico, 'EU countries strike deal on landmark AI rulebook', 2 February 2024.
[47] European Parliament, 'Artificial Intelligence Act: MEPs adopt landmark law', 13 March 2024.
[48] European Commission, 'AI Act'6 March 2024.
[49] European Commission, 'AI Pact', 6 March 2024.
[50] Thomson Reuters, 'Legalweek 2024: Current US AI regulation means adopting a strategic—and communicative—approach', 11 February 2024.
[51] Council of State Governments, 'Artificial intelligence in the states: Emerging legislation', 6 December 2023.
[52] The White House, 'Fact sheet: President Biden issues executive order on safe, secure, and trustworthy artificial intelligence', 30 October 2023.

- develop standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy
- protect against the risks of using AI to engineer dangerous biological materials by developing strong new standards for biological synthesis screening
- protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content
- establish an advanced cybersecurity program to develop AI tools to find and fix vulnerabilities in critical software
- order the development of a national security memorandum that directs further actions on AI and security

The order also contained several measures on privacy, including:

- protect Americans' privacy by prioritising federal support for accelerating the development and use of privacy-preserving techniques
- strengthen privacy-preserving research and technologies
- evaluate how agencies collect and use commercially available information and strengthen privacy guidance for federal agencies to account for AI risks
- develop guidelines for federal agencies to evaluate the effectiveness of privacy-preserving techniques

The order also contained measures on advancing equity and protecting civil rights; the impact of AI on consumers, patients and students; supporting workers; promoting innovation and competition; ensuring responsible and effective government use of AI; and "advancing American leadership abroad". Under this latter heading, the order references the expansion of bilateral, multilateral, and multi-stakeholder engagement, and the acceleration, development and implementation of "vital AI standards with international partners and in standards organisations, ensuring that the technology is safe, secure, trustworthy, and interoperable".

The order set a deadline of 90 days on several key actions to address some of what the White House described as "AI's biggest threats to safety and security".[53] These included setting key disclosure requirements for developers of the most powerful systems; assessing

[53] The White House, 'Fact sheet: President Biden issues executive order on safe, secure, and trustworthy artificial intelligence', 30 October 2023.

AI's risks for critical infrastructure; and hindering foreign actors' efforts to develop AI for harmful purposes. The White House announced in January 2024 that all executive agencies had completed the 90-day actions tasked by the order.[54] Those actions included the agencies having done the following:

- Used Defense Production Act authorities to compel developers of the most powerful AI systems to report vital information, especially AI safety test results, to the Department of Commerce. These companies now must share this information on the most powerful AI systems, and they must likewise report large computing clusters able to train these systems.
- Proposed a draft rule that proposes to compel US cloud companies that provide computing power for foreign AI training to report that they are doing so.
- Completed risk assessments covering AI's use in every critical infrastructure sector.

In the US budget announced in March 2024, President Biden also announced significant new funding aimed at the development of responsible AI.[55] It included:

- over $3bn across agencies to responsibly develop, test, procure, and integrate transformative AI applications across the federal government
- $300mn in mandatory funding to increase agency funding for AI, both to address major risks and to advance its use for public good.
- $65mn investment with the Department of Commerce to support the AI Safety Institute's mandate to ensure AI models are safe, secure, and trustworthy, as well as other critical AI work.

3. Artificial Intelligence (Regulation) Bill [HL]: Clause by clause

The Artificial Intelligence (Regulation) Bill [HL] contains nine clauses and is intended to introduce a range of provisions to provide for the regulation of AI technologies in the UK.

---

[54] The White House, 'Fact sheet: Biden-Harris administration announces key AI actions following President Biden's landmark executive order', 29 January 2024.

[55] The White House, 'Fact sheet: The President's budget advances President Biden's unity agenda', 11 March 2024.

Clause 1 of the bill would provide the secretary of state with the power to create an AI Authority through delegated legislation. Such an AI Authority would have the following functions:

- ensure that relevant regulators take account of AI
- ensure alignment of approach across relevant regulators in respect of AI
- undertake a gap analysis of regulatory responsibilities in respect of AI
- coordinate a review of relevant legislation, including product safety, privacy and consumer protection, to assess its suitability to address the challenges and opportunities presented by AI
- monitor and evaluate the overall regulatory framework's effectiveness, including the extent to which they support innovation
- assess and monitor risks across the economy arising from AI
- conduct horizon-scanning to inform a coherent response to emerging AI technology trends, and support testbeds and sandbox initiatives to help AI innovators get new technologies to market
- accredit independent AI auditors
- provide education and awareness to give clarity to businesses and to empower individuals to express views as part of the iteration of the overall regulatory framework
- promote interoperability with international regulatory frameworks

Clause 1 would also allow the secretary of state to amend these functions by regulation. The secretary of state could also dissolve the AI Authority "following consultation with such persons as he or she considers appropriate". This measure, and other clauses in the bill which would provide the government with the explicit ability to amend significant parts of the legislation, was noted by the House of Lords Delegated Powers and Regulatory Reform Committee as detailed in section 4 of this briefing.

Clause 2 of the bill would set out a number of principles which the AI Authority must have regard to on the regulation of AI, in dealing with businesses seeking to develop and deploy AI products, and the output and design of AI systems themselves. Those principles, which mirror in part those contained within the government's white paper, are as follows:

(a) regulation of AI should deliver—

(i) safety, security and robustness;

(ii) appropriate transparency and explainability;

(iii) fairness;

(iv) accountability and governance;

(v) contestability and redress;

(b) any business which develops, deploys or uses AI should—

(i) be transparent about it;

(ii) test it thoroughly and transparently;

(iii) comply with applicable laws, including in relation to data protection, privacy and intellectual property;

(c) AI and its applications should:

(i) comply with equalities legislation;

(ii) be inclusive by design;

(iii) be designed so as neither to discriminate unlawfully among individuals nor, so far as reasonably practicable, to perpetuate unlawful discrimination arising in input data;

(iv) meet the needs of those from lower socio-economic groups, older people and disabled people;

(v) generate data that are findable, accessible, interoperable and reusable;

Clause 2 also includes a further principle (principle (d)), which states that a burden or restriction imposed on a person, or on the carrying on of an activity in respect of AI should be "proportionate to the benefits". As part of making such a determination, consideration would be given to the nature of the service or product being delivered; the nature of risk to consumers and others, and whether "the cost of implementation is proportionate to that level of risk"; and whether the burden or restriction "enhances UK international competitiveness".

Clause 2 also contains the power for the secretary of state to amend the above principles by regulations.

Clause 3 would require the AI Authority to collaborate with relevant regulators to construct regulatory sandboxes for AI. Such sandboxes are defined in the bill as an

arrangement by one or more regulators which:

> (a) allows businesses to test innovative propositions in the market with real consumers;
>
> (b) is open to authorised firms, unauthorised firms that require authorisation and technology firms partnering with, or providing services to, UK firms doing regulated activities;
>
> (c) provides firms with support in identifying appropriate consumer protection safeguards;
>
> (d) requires tests to have a clear objective and to be conducted on a small scale;
>
> (e) requires firms which want to test products or services which are regulated activities to be authorised by or registered with the relevant regulator before starting the test.

Again, clause 3 would provide the secretary of state with the power to amend the definition of such regulatory sandboxes through regulations.

Clause 4 of the bill would require that any business which develops, deploys or uses AI must have a designated AI officer. The duties of such an individual would be to:

- ensure the safe, ethical, unbiased and non-discriminatory use of AI by the business
- ensure, so far as reasonably practicable, that data used by the business in any AI technology is unbiased

The bill would enable the government to create such a requirement through regulations following consultation with the AI Authority. The duties of such AI officers could also be amended through delegated legislation.

Clause 5 concerns transparency, intellectual property obligations, and labelling. It would require that the government—following consultation with the AI Authority—introduce regulations to provide that any person involved in training AI must supply to the AI Authority a record of all third-party data and intellectual property (IP) used in that training, and assure the AI Authority that they use all such data and IP by informed consent and that they have complied with all applicable IP and copyright obligations. Any person supplying a product or service involving AI would also have to give customers "clear and unambiguous

health warnings, labelling and opportunities to give or withhold informed consent in advance". Further, any business which develops, deploys or uses AI would have to allow independent third parties accredited by the AI Authority to audit its processes and systems.

The bill states that regulations under this section may provide for informed consent to be either express (opt-in) or implied (opt-out), and different provision could be made for different cases.

Clause 6 of the bill states that the AI Authority must implement a programme for "meaningful, long-term public engagement" about the opportunities and risks presented by AI. It would also place a requirement on the AI Authority to consult the general public and such persons as it considers appropriate as to the most effective frameworks for public engagement.

Clause 7 provides detail on the interpretation of the terminology used in the bill, including that the bill would apply to both narrow and deep learning/generative AI.

Clause 8 would provide that regulations made under the bill may create offences and require payment of fees, penalties, and fines. It also states that regulations made under the powers in clauses 1 or 2 would be subject to the affirmative procedure. Regulations made under clause 3 would be subject to the negative procedure.

Clause 9 includes general provisions, including that the bill would apply to all of the UK.

4. Delegated Powers and Regulatory Reform Committee scrutiny of the bill

The House of Lords Delegated Powers and Regulatory Reform Committee reported on the Artificial Intelligence (Regulation) Bill [HL] on 14 December 2023.[56] The committee noted that the proposed legislation is a framework bill which "leaves much of the detail to be included in regulations made by ministers". Specifically, the committee cited the following measures allowing the secretary of state to amend various provisions by means of delegated

[56] House of Lords Delegated Powers and Regulatory Reform Committee, 'Fourth report of session 2023–24', 14 December 2023, HL Paper 31 of session 2023–4, p 1.

legislation:

- Clause 1(1) and (2) requires ministers to set up the AI Authority with 11 specific functions. But clause 1(3) allows ministers to rewrite these functions in regulations, and to dissolve the AI Authority.
- Clause 2(1) sets out 14 specific regulatory principles to which the AI Authority must have regard. But clause 2(2) allows ministers to rewrite these regulatory principles.
- Clause 3 requires the AI Authority to collaborate with relevant regulators to construct "regulatory sandboxes" for AI. This term is given an elaborate definition in clause 3(2). But regulations under clause 3(3) can rewrite the definition.

The committee said that as a result it drew the attention of the House to "the fact that ministers are given wide powers to rewrite significant parts of the bill".[57]

[57] House of Lords Delegated Powers and Regulatory Reform Committee, 'Fourth report of session 2023–24', 14 December 2023, HL Paper 31 of session 2023–4, p 1.

About the Library

A full list of Lords Library briefings is available on the Library's website.

The Library publishes briefings for all major items of business debated in the House of Lords. The Library also publishes briefings on the House of Lords itself and other subjects that may be of interest to members.

Library briefings are produced for the benefit of Members of the House of Lords. They provide impartial, authoritative, politically balanced information in support of members' parliamentary duties. They are intended as a general briefing only and should not be relied on as a substitute for specific advice.

Every effort is made to ensure that the information contained in Lords Library briefings is correct at the time of publication. Readers should be aware however that briefings are not necessarily updated or otherwise amended to reflect subsequent changes.

Disclaimer

Authors are available to discuss the contents of the briefings with Members of the House of Lords and their staff but cannot advise members of the general public.

Any comments on Library briefings should be sent to the Head of Research Services, House of Lords Library, London SW1A 0PW or emailed to hlresearchservices@parliament.uk.