



HOUSE OF COMMONS

# Science, Innovation and Technology Committee

## Oral evidence: Governance of artificial intelligence (AI), HC 38

Wednesday 8 November 2023

Ordered by the House of Commons to be published on 8 November 2023.

[Watch the meeting](#)

Members present: Greg Clark (Chair); Aaron Bell; Dawn Butler; Chris Clarkson; Tracey Crouch; Katherine Fletcher; Rebecca Long Bailey; Stephen Metcalfe.

Questions 653 - 755

### Witnesses

**I:** Matt Clifford, Prime Minister's representative, AI Safety Summit; and Emran Mian, Director General, Digital Technologies and Telecoms, Department for Science, Innovation and Technology.



## Examination of witnesses

Witnesses: Matt Clifford and Emran Mian.

Q653 **Chair:** The Science, Innovation and Technology Committee is very pleased to be looking at the AI safety summit that took place at Bletchley Park last week.

I am very pleased to welcome, to help us with that, Matt Clifford and Emran Mian. Matt Clifford was one of two representatives appointed by the Prime Minister for the safety summit. He is the chief executive of Entrepreneur First, a talent investor that connects entrepreneurs with investors and supports them in building technology start-ups. He is also chair of ARIA, the Advanced Research and Invention Agency.

Emran Mian is an official at the Department for Science, Innovation and Technology. He is director general for digital technologies and telecoms. He was previously director general for regeneration at the Department for Levelling Up, Housing and Communities, in which capacity, I should disclose, he worked with me when I was the Secretary of State there a while ago. He is a former director general at the Department for Education.

Welcome both. Thank you very much indeed for coming. Perhaps I can start with Mr Clifford. Your appointment for the purpose of the summit was a limited tour of duty, and I think you are coming to the end of it. Is that right?

**Matt Clifford:** That is right. Friday is my last day.

Q654 **Chair:** Well, we are very grateful to have you today. Was it always envisaged that you would finish with the summit?

**Matt Clifford:** That is right. I was just the Prime Minister's representative for the summit—then back to private sector life.

Q655 **Chair:** I am grateful to you for coming in at short notice, after a busy session last week. It was an AI safety summit. Do you have a definition of safety that governs the creation of the summit?

**Matt Clifford:** I think the way to think about this is to zoom out, for a second, and think about the context of where the technology is, why safety is important, and what it means. As you will all be aware, the last 18 months have seen an extraordinary level of investment in artificial intelligence and, as a result, interest in it from both the public and policymakers.

The way to think about the landscape in AI at the point when the Prime Minister decided to hold the summit, in the summer, is that there were a number of very powerful systems which were already accessible to consumers and businesses. Many of you will have used things like ChatGPT. The reason for holding the summit when we did, and for the focus on safety, is that we are standing right now at an inflection point



## HOUSE OF COMMONS

for the history and future of the technology. If you use the paid version of ChatGPT, the model that underlies it is called GPT-4. That was finished in the summer of 2022 by the company OpenAI. It is rumoured that it cost them about \$60 million to train that model. In 2024 they, and others, will release to the world new models that we expect will have had between 10 and 100 times that investment. I cannot think, after 12 years investing in emerging tech, of any other moment like this, in a comparably important technology.

I think that the Prime Minister's observation in the summer was that we already had powerful systems whose capabilities surprised their creators; that we are in the calm before the storm, before new systems are released in 2024 that could be many times more powerful; that we know that there is a positive relationship between investment in the systems and their capabilities, but neither we nor their creators can say in advance of their being made what they will be capable of—particularly in specific domains; and that, by default, the systems will be accessible and available to a very broad range of people.

The safety question comes out of saying that obviously, in general, we want to encourage innovation in technology, and there are clear opportunities in AI—this is not a doom and gloom story; but, equally, placing very powerful capabilities in the hands of arbitrary actors, some of whom may wish us harm, creates a huge public safety and, potentially, national security question.

That is before you even get into the, I suppose, more catastrophic risks of what happens when we put more and more of our societal and economic processes in the control of artificial intelligence systems.

What we mean by safety is really two things that directly address those two cases. The first is the question of what the capabilities of powerful AI systems are, who has access to them, and how we limit that access to the appropriate extent.

The second is to what extent the actions that these systems take in response to user input align with the intentions of the user. These are political, economic and societal questions. They are also technological questions.

Going into the summit we took a quite broad view of how we might address those questions, and what questions were raised.

On the summit agenda there were roundtables on misuse—looking specifically at how these powerful systems might aid bad actors who wish to do us harm—but also sessions on what we call risks from the integration into society. Some people call those short-term risks—I think that 2024 is not far away and we should think of all these things as short term—but they include how bias and discrimination in the systems might affect the way society works.



Q656 **Chair:** They are dimensions of safety, in your view.

**Matt Clifford:** We very much see them as dimensions of safety, along with things like electoral interference and misinformation. To be clear, a lot of these are not new risks. They are risks exacerbated, rather than created, by AI, but, again, because of the unknown increase in capability that we expect over the next 12 months it feels very important to look at existing risks that might be amplified into a difference of kind as well as magnitude.

Q657 **Chair:** To be clear, the risk of unacceptable bias being practised by an AI is a safety risk, in your view.

**Matt Clifford:** I think it is a safety risk. Emran is probably a better person to speak to the broader policy agenda, but with what we looked at in the summit we were really saying that we expect these systems to be broadly integrated into society—into all our social and economic processes—over some period of time. Therefore, anything that either puts the public at risk or causes a risk of social harms: the Bletchley declaration that the country has signed up to talks about catastrophic risks such as whether the technology can help a terrorist to build a bioweapon in ways that were previously inaccessible, but also about societal harms, including ways in which existing social processes might be damaged, or existing social harms might be exacerbated, by widespread adoption of these systems.

Q658 **Chair:** Thank you. My colleagues will come on to some of that and we may have some questions on policy, for Mr Mian, as the permanent official, but going back to the summit, and reflecting on last week, what do you see as its significant accomplishments?

**Matt Clifford:** If you go back to where we were in the summer, the first thing that I would say is that the biggest significance of the summit was putting AI safety on the international agenda. It was the first international meeting of its kind. Obviously, AI has featured on the agendas of many fora—maybe most prominently the G7, but there is a UN process going on and there are a number of things in the OECD, including GPAI, a sponsored AI-focused organisation. There is a line in the G20 communiqué.

It is not as if AI has been absent from international debate, but putting AI safety at the heart of that agenda is itself an accomplishment. Candidly, if you had asked us in the summer whether we would have been happy with that as the main outcome, we probably would have. Because of the timing issue that I mentioned it was important to do it quickly; but I would point to four more substantive outcomes that are probably at the upper end of what we hoped for in the summer.

The first is the Bletchley declaration itself. That is the first shared statement on what frontier AI is—which was a contentious issue before we started the process—what its risks are, and how we might work together to mitigate them. That was signed by all 28 countries that



## HOUSE OF COMMONS

attended, including, for the first time, the US and China, so a majority of the world's population and GDP was covered by the countries that signed it. Obviously, there are well documented differences of view between us and China, and the US and China, on a broad range of issues; but it felt important to be able to find some common ground. We obviously had to keep the declaration relatively high-level, but bringing China to the table to talk about technology, when it is on most measures the second-biggest power, is a huge step forward for international governance of AI.

The second thing I would point to is the "State of the Science" report that the Ministers of all 28 countries agreed to produce, in advance of the next summit. We are slightly careful about the analogy, because there are obviously ways in which it does not fit, but it is inspired by the IPCC and the idea that, for us to make progress collectively and internationally, we need to move towards something of a consensus on where the technology is and where the science is. This is the first time that there will have been a track 2 process across countries on this.

I think it was an achievement for the UK to persuade Yoshua Bengio, who is arguably the godfather of AI and is a winner of the Turing prize, which is the equivalent of the Nobel prize in computer science, to chair that work. The report will be an input to the next summit. There will be a small secretariat, I think, in DSIT, that will support that work. This, I think, is another big achievement that will hopefully become an important benchmark or foundation for future discussions. That is the second thing I would point to.

Q659 **Chair:** Briefly, on that, the report would be for the next summit, in the Republic of Korea.

**Matt Clifford:** I think that is the plan, yes. To be clear, the Governments will nominate experts, but we anticipate that they will be experts from outside government—largely academics. That is important because it means that all 28 countries can participate in an equal way.

The third thing, which the Prime Minister announced in his closing press conference, last Thursday, is what I think is a landmark agreement about the role of the public sector in evaluating and testing AI systems and AI models for the most extreme risks. I have already referenced ChatGPT. When OpenAI finished building GPT-4 last summer they then took six months to do testing, largely internally but with a few external contractors, basically to say, "What are the capabilities of this model, and is it dangerous?" I would say that that was the responsible thing to do.

I would, however, say that there are two extreme limitations in that approach to AI safety. One is that there are certain risks, most obviously those related to national security, that private sector companies cannot adequately test for. They do not have the knowledge of the most extreme cyber, bio and other threats to be able to do that testing.



The second is that the approach only worked because at the time OpenAI was the only significant player producing a chat interface on top of one of the large language models. I would argue that, today, when it completes the next version, the commercial pressure to release it early will be very large. Each month of testing is a month when it is not in the hands of consumers.

The agreement that the Prime Minister announced last Thursday is between nine of the most prominent creators of large language models, and the G7 countries, plus other like-minded partners—Singapore, Korea and Australia—and it basically says that the companies agree that before they release the next generation of their models, which we expect in 2024, they will work with Governments, including national security actors, to do pre-deployment testing of those models for the most extreme risks.

This is a huge step forward. It creates a common baseline, ultimately for the world but certainly, now, for the western countries, that asserts the role of the state, and hopefully over time starts to create some common standards about what it means to say that a model is safe before it is released.

The final thing, which may sound facetious but which I think is a big thing, is that we have two more summits. That is important because it means the issue stays on the international agenda, going forward. You can think of the technology as being on an exponential curve that shows no sign of flattening off, at least in terms of investment. So having six-month and 12-month checkpoints at which we can ask how we are living up to the aspirations of the Bletchley declaration, where the science is, and how successfully we have done the pre-deployment testing of AI systems, feels like an important step forward in its own right.

**Chair:** Thank you. That is a very helpful summary. It has provoked supplementaries, immediately, from my colleagues—starting with Aaron Bell, and then Stephen Metcalfe.

Q660 **Aaron Bell:** Thank you, Chair, and thank you, Matt, for your time. Personally, I think it was sensible to invite China, but there seems to be some—I do not know whether it is misreporting, or the way it was reported, about how they were, on the AI model testing point: initially they were not signing up to the AI model testing thing, and it was assumed that they were not at the meetings on Thursday. They then said that they were—that the Minister was—but they have not signed up to it. Are you able to shed any light on the disagreements that they had on that point, about why they were unable to sign up to that part of the communiqué?

**Matt Clifford:** Yes, actually there was no disagreement on that. We made a decision early on that day one would be broad and inclusive, and focused on the general approaches to identifying and mitigating risk. China was fully included in that. We then realised that there was so much



## HOUSE OF COMMONS

material on that that, on day two, we would have a ministerial track, chaired by the Secretary of State, that would continue the discussions at country-only level. China was involved in that, and participated fully. The core output of that day-two session among the Ministers was the “State of the Science” report. China signed up to that and to the communiqué.

It was always the intention that there would be a “like-minded-only” session on day two. That is less about a disagreement, and more a jurisdictional issue. The Chinese AI companies do not serve customers outside China, and none of the western AI companies serves customers inside China. This is really about countries coming together with aligned values and long-standing security co-operation to work with companies that operate in their jurisdictions to reach a common agreement on what is quite intrusive safety testing. We judged early on that it was never going to be appropriate to have that conversation with China, not least because the companies are in two separate universes. There was no disagreement there. They fully participated and signed up to the Bletchley declaration and the “State of the Science” report. Obviously, there is a lot of diplomatic work to be done in the run-up to the Korean summit, but we would expect them to participate in that as well.

Q661 **Aaron Bell:** So they did not participate in the public events on the Thursday, but they have said that their Minister was present.

**Matt Clifford:** That is right.

Q662 **Aaron Bell:** Was there a separate discussion about the AI model testing?

**Matt Clifford:** There were two tracks on Thursday—one chaired by the Prime Minister, with like-minded countries only, about security co-operation, and one chaired by the Secretary of State. Ministers from all 28 countries were invited to that. China did participate. That was about the “State of the Science” report. That was not a change of plan. There was always going to be a G7-plus-plus track and a broad track, on that.

Q663 **Aaron Bell:** Where does this leave us collectively, globally, if China is not signing up to AI model testing, even though there may be different jurisdictions, as you said, and at the moment Chinese technology is not being purchased by western companies or individuals? That could change. What has China actually said about AI model testing and its willingness to go down that route that the rest of the world signed up to?

**Matt Clifford:** I am not an expert on this, but I did a lot of work with the Chinese Government, and some track 2 work during the summer preparation. The legislation and regulation that is already in place in China is much more stringent than anything we have in the west. That is partly for the obvious political reasons that if you do not have freedom of speech you need to lock down the models a lot more effectively than you do in the west.

I would not want to speculate about what they will or will not do but, in terms of the misuse harms that we worry about a lot, which is really



## HOUSE OF COMMONS

where the security testing is focused, we want to know whether you can use the models to execute a large-scale cyber-attack, for example. Can you use them to facilitate end-to-end planning of a biological attack? These are open questions.

Our view is that today the risk, from today's models, is relatively low. Again, the question for those models in China will be dealt with by the Chinese Government, obviously. My non-expert view of where their legislation is today is that they have pretty stringent controls on their companies already. Yes, I would hope that through processes like the AI safety summit we can reach some common international standards, but, given the fact that we have non-aligned systems and do not have broad co-operation with them on security matters, it is unlikely that they would ever come under the same testing regime, although the standards of those regimes may align over time. That would be a good thing.

**Chair:** That is helpful. Thank you very much indeed. Stephen, you had a supplementary.

Q664 **Stephen Metcalfe:** Yes. A quick technical question: when you talk about models and the current iteration, which is a \$60 million investment, is that a fixed model that does not learn on the hoof, as it were?

Depending on the answer to that, when Governments are given access to pre-release versions of a model to test it in a way that commercial companies cannot—presumably because they have additional knowledge and information about potential for gaming the system—is there any danger that the system will learn from that testing?

**Matt Clifford:** A great question. It is a very good and complicated question. Right now, if you use ChatGPT, it does not learn between sessions from your questions or answers. It does what is called in-context learning. It takes the inputs and outputs from the session, and the reason it feels like a continuous conversation is that it is using them.

Today, there is not a risk of your teaching the thing how to make a bomb, through a ChatGPT session, and its taking that knowledge.

However, there is a mode of operations that applies to large language models, called fine tuning. You take a base model that has had a lot of money invested in it, and that has been trained on a very broad set of data. Then, you do the fine tuning, usually on a domain-specific set of data, to optimise for domain-specific tasks. Those could, for example, be in the biology domain. They could be in the cooking domain. It can be a very benign thing.

The companies have started to announce—in fact, OpenAI, the maker of ChatGPT, announced this on Monday—that they are making it easier for people to do fine tuning. Most fine tuning will be commercially valuable and completely benign. There is, of course, the risk of non-benign—indeed, malign—fine tuning.





## HOUSE OF COMMONS

I think that the broad principle on how you test for this is that Governments need to be given access to the model in whatever mode it will be made available to the public. In many cases, the companies would, I think, say, "That does not go far enough," but, at the minimum, if the model is going to be served through a chatbot, Governments should be able to test it through a chatbot. If the model is going to be open source, so that anyone can do anything with it, the testing really needs to happen on the full model. If the model can be fine-tuned, the testing needs to be done in that mode, as well.

There is a broad principle and there is agreement on that, but you are right that probably the hardest bit of the puzzle is that this is adversarial work. I do not mean that we are against the companies; I mean more that we have to anticipate what a bad actor might do adversarially.

Q665 **Stephen Metcalfe:** What you do not want to do is give the bad actor an advantage, by understanding the inputs that you put into that model.

**Matt Clifford:** No, and we would not do that. I cannot go into great detail on this, but even in the agreements that we have between what is becoming the AI Safety Institute in the UK and the companies, on model access, we have what is called logging-free access, where our inputs would not be used in any future version of the model. That is very important. You are quite right.

**Stephen Metcalfe:** Absolutely.

Q666 **Chair:** Thank you, Stephen. Before I come to Dawn Butler, I have one final question on the summit. Which were the areas of the communiqué, or points in the roundtables, on which agreement was most difficult to reach?

**Matt Clifford:** Without, for obvious reasons, going into either the content of negotiations or individual contributions in closed sessions, I would maybe draw three themes where big unanswered questions remain.

One is what we just talked about—open source. There is an increasingly polarised debate about whether open-source models are a good or bad thing. For reference, open source means that the maker of the model makes the full model available, including the code used to generate the training, to anyone who wants to see it.

Advocates of that approach would say that it is better. It is transparent. These are powerful systems and we should know how they are built. We should be able to adopt them and adapt them. We should not let them be used by a small number of American companies.

Opponents say that, if there are powerful capabilities in a model, open sourcing is irreversible proliferation: once it is out there, you cannot take it back, which is a dangerous thing. That is going to be one of the thorniest issues for policymakers internationally. It came up in the communiqué and the roundtables, and there is more to do on it.



There is a big question relevant to your first question, about how broadly we should think about safety. I personally reject the long-term/short-term dichotomy for these things. As I said, 12 months is a long time in AI and we would expect models to be very powerful, next year; but it is completely legitimate to say that when we look at all the potential benefits and harms of AI they go far beyond catastrophic risks. Although the summit was focused on safety and in particular on catastrophic risk—Emran can speak to the huge amount of work that is going on in DSIT on every aspect of that—the big questions that remain unanswered, and on which there will be fierce international debate, are the extent to which things like bias and discrimination, and particular and unequal economic impact, should be considered a safety issue. There is a broad range of views on that.

The final thing I would point to is the question of what I could very crudely put as AI takers and AI makers. Right now, it is true that the number of companies that can afford to train the models at the scale of the systems we have been talking about is very small. If we are talking about models for which the training might, next year, be in the hundreds of millions or billions, the number of companies with that amount of capital, and that have the talent and data to do it, is very small. At the moment they are all based in the US, in London or in China. There are some emerging companies in Europe that want to have a shot at it, but a big topic that came up a lot at the safety summit is how we should think about it as a truly international question. Where does the global south fit into this? If the models are as powerful as the makers claim, is it right that there is this sort of market concentration and, if not, what do we do about it? These are all big questions.

**Q667 Dawn Butler:** I will give you a break from speaking, Matt, and ask Emran: why was the AI Safety Institute necessary, and what is the purpose of it?

**Emran Mian:** The need for the AI Safety Institute goes back to something that Matt was talking about, in developing the capability within the public sector, within government and in public hands, to carry out high-quality, effective evaluation of the capabilities of models, and identify what the safety issues might be.

As Matt was saying, the companies are doing quite a lot of the testing. That is great and positive. It is right that they should do it and that they should contribute to the development of the science on safety; but relying solely on the companies gives us some pause for thought, both because of the commercial imperatives that the companies may have in the rush to market, but also in the classic issue that arises in these kinds of safety conversations, about companies marking their own homework. That is the phrase that the Prime Minister used at his press conference when he was closing the summit. We felt that it was important to begin to develop the capability within government to carry out safety evaluation, for both those reasons.



I would add a third reason, which is that because this is a relatively new technology there is a lot to learn at this stage about evaluation itself. It is that classic stage of developing the science where the case for public intervention and the case for public funding of the research is extremely high.

Q668 **Dawn Butler:** Where are you going to get the experts from, because starting from now it will be too late, won't it?

**Emran Mian:** The good news is that we have already started so we have got a few months under our belt. About six months ago Ian Hogarth came in to chair what we were then calling the foundation model taskforce and to begin to develop the capability here. As chair, Ian has been able to develop a really powerful advisory board around him, which was how we got to know Yoshua Bengio, who, as Matt said, is one of the godfathers of AI. As part of that advisory board we have also, for example, got the head of GCHQ. So we have begun to develop that core of expertise.

The other thing we have been doing over the past few months is hiring experts from universities and the industry, to develop the capability. It has been interesting to see that people from both the academic community and the companies want to come and do this. We worried at times about whether we would be successful in attracting the right capabilities into government and have been pleasantly surprised by the quality and calibre of people we have been able to bring in, precisely because people know this is a significant issue. People who work in the industry recognise that the industry should not be marking its own homework and they want to be part of the public sector capability on this.

None of that is to say that we do not need to keep going at the same pace we have been going at over the past few months. The build is very much work in progress. There is a lot more for us to do in terms of building it. That is why we have been so keen to commit funding to it. We have now committed funding to it beyond this spending review period.

Q669 **Dawn Butler:** How much is the funding that has been committed?

**Emran Mian:** It is funded for £100 million over the rest of this spending review period—this financial year and the next. We have said that we will continue that funding level through to the end of the decade. We wanted to create certainty about its funding.

It is important to stress that we want to do as much of this as possible out in the open. There are clear limitations to how much is out in the open, especially when we are testing some of the most sensitive capabilities—national security capabilities. There is a fine balance between how much information we can provide and providing enough information to create confidence in the evaluations. We also need to continue to be really transparent about our approach.



In the time the taskforce has been going, it has done two progress reports. Alongside the summit, we published a brochure—a Command Paper—laying out the thinking behind the institute and the approach it will take. We think that that commitment to transparency, especially in such an emerging field, will be important in creating and sustaining public confidence, in continuing to bring high-calibre people to the institute and in ensuring collaboration with other countries that are doing this work.

**Q670 Dawn Butler:** When will it be fully up and running? It is not fully up and running yet, is it?

**Emran Mian:** In many ways, it is fully up and running. The summit was an important test for us and it set an important milestone. We wanted to be ready at the summit to present some of the early findings of the taskforce's work. Each of the roundtables that we ran as part of the summit explored different aspects of safety—for example, risks of misuse, and to what extent are AI agents able to exhibit autonomy from the people controlling them?

All those discussions were animated by the taskforce's early work, which we presented to the room. A lot of that work has not been peer-reviewed; it has not gone through the full scientific rigour that we would want it to go through, and institute researchers recognise that, but we are up and running in the sense that we were able to bring that material to the summit and use it to help to drive the conversation.

**Q671 Dawn Butler:** Would you say that it is now fully operational?

**Emran Mian:** I would not say that it is fully operational. There is definitely more for us to do in the build. It is fully operational in the sense that it is already able to help us to drive the conversation at the summit. It is already able to help us to think about the policy environment. It is already fully operational in the sense that it is helping us to think about how we do international collaboration. We notice from that that we are probably a good bit ahead of a lot of peer countries. A lot of people are thinking about how to develop this capability. One thing we in the UK have done successfully is to go early in developing that capability.

In all those senses, it is fully operational, but we need to keep building the team and to build the science around what good evaluation looks like, growing the circle of people who are able to look at the science, peer-review it and continue to improve it.

**Q672 Dawn Butler:** We are always very concerned that we do not have the expertise in government, with things progressing at pace. Is there constant work with outside organisations that are focused on safety such as the Alan Turing Institute? Is there such a relationship, and if so how strong is it?

**Emran Mian:** I give two examples of the relationships we have built and relied on in the run-up to the summit and that we will continue to build and rely on.



## HOUSE OF COMMONS

The first is the set of academic and civil society organisations that have been working on these issues for some time. It is important that we build on what they do. The Alan Turing Institute is absolutely one of those, but in the run-up to the summit we held events and discussions with the Royal Society, the British Academy and techUK. We recognise that there is a broad range of stakeholders and we want to continue to build a conversation with them.

That is one aspect where collaboration is key, but the other is that some of the technical expertise exists elsewhere. In the work prior to the summit, we were working with a set of other organisations—quite a small set of organisations, and some of the organisations themselves are quite small at the moment. All of the ecosystem needs to continue to build. We were not trying to do it all in-house; we were definitely relying on some of the expertise that exists out there.

**Q673 Dawn Butler:** Some organisations felt left out, didn't they? They did not feel part of the summit, or were not invited. You are the opposite—an institution, the AI Safety Institute—and you bring them in. Am I correct?

**Emran Mian:** Yes, I think that is right. Some of the reporting was probably slightly overdone in the run-up to the summit. A bit of opposition was created between long-term risks and short-term risks. The technology and the power of the models are developing so quickly that some things that are described as long-term risks are not necessarily long-term risks, or at the very least we need to begin to get a scope on them now.

A second dichotomy in some of the reporting in advance of the summit was between a group of civil society organisations that perhaps are focused more on present-day risks versus a set of technical organisations or companies that are more focused on existential risks.

Our experience of the run-up to the summit and of the summit was that that dichotomy was massively oversold. It was really important to us to have civil society organisations as part of the summit. We have published a list of the civil society organisations that came to the summit. I do not think that the list of those organisations fits any characteristic of civil society organisations that are interested in only one or other category of risk.

Equally, when we talk to the companies we find that they too do lots of work on bias and transparency and are focused on it.

Matt talked about the open-source versus closed-source debate. A big part of the advocacy for open source is exactly about creating transparency of these models.

**Q674 Dawn Butler:** The Counter Disinformation Unit is now based in DSIT.

**Emran Mian:** That is right.



## HOUSE OF COMMONS

Q675 **Dawn Butler:** Are you involved in it? What is its purpose?

**Emran Mian:** The Counter Disinformation Unit came to the Department from DCMS as part of the machinery of government change. It is in the part of the department that I lead as director general.

The unit's focus is issues of national security and public safety. It has quite a tight remit.

Q676 **Dawn Butler:** It is focused on national security and—

**Emran Mian:** Public safety.

Q677 **Dawn Butler:** Is it an AI system? Is it a large language model, or is it just somebody looking for things?

**Emran Mian:** We do not have our own large language model; that is not the nature of the work we are doing. The nature of the work we are doing is looking at content within those specific categories on social media platforms and speaking to the platforms about where content breaches their terms and conditions.

Q678 **Dawn Butler:** That is human based. Somebody is going through and looking for disinformation.

**Emran Mian:** It is definitely aided by technology but we do not have our own LLM to do it.

Q679 **Katherine Fletcher:** Whose technology are you using?

**Emran Mian:** We have talked about the approach we take. It is a matter of looking at the technology providers in the market and the ones best placed to be able to support us and the unit's narrow focus, and to be able to identify content so that we can then have a conversation with the social media platforms about whether it breaches their terms and conditions.

Q680 **Dawn Butler:** With that narrow focus, would you expect any MPs to be looked into, reviewed or monitored?

**Emran Mian:** Given the narrow focus that the unit has, that is not something that we are expecting to focus on, but ultimately we will be driven by the focus of the unit on national security and public safety.

Q681 **Dawn Butler:** Matt, Friday is your last day.

**Matt Clifford:** It is.

Q682 **Dawn Butler:** Why are you not going to work for the AI Safety Institute?

**Matt Clifford:** I have patient, but not infinitely patient, shareholders.

Q683 **Dawn Butler:** So you have to go back.

**Matt Clifford:** I do.



## HOUSE OF COMMONS

Q684 **Dawn Butler:** Would you like to work for the AI Safety Institute?

**Matt Clifford:** I will support the AI Safety Institute in whatever way it asks me to, consistent with my responsibilities to my shareholders.

Q685 **Dawn Butler:** Is the Bletchley declaration binding?

**Matt Clifford:** It is a communiqué. It is the sense of the countries. It is not a treaty. It does not create formal obligations on the countries that signed it, beyond it being a public statement.

Q686 **Dawn Butler:** Did the developers who attended the summit agree to provide full access to their models? You spoke of some access.

**Matt Clifford:** If you read the statement from day two, you will see that they agree to provide access consistent with pre-deployment testing for extreme risks for national security.

Q687 **Dawn Butler:** You can control access.

**Matt Clifford:** Again, it is not a law. It is a non-binding agreement between the parties. In theory, they can renege and walk away from it tomorrow.

Our position would be that it is something they are entering into in good faith—that is what we wanted to do—and it is a continuation of work that started over the summer.

To your questions about whether the AI Safety Institute is up and running, you might have seen earlier in the summer that the Prime Minister announced that the three leading frontier labs had already agreed to give model access to what was the taskforce but is now the institute.

We are not asking at this stage for full access to the model weights—the raw model—for the reason that we describe: that is not how they are providing access to the public. Ultimately, to test for risks we need access in the form that the public will have, and that is what we are getting at the moment.

To Emran's point about whether the institute is up and running and operational, what is a real head start, as well as having got talent in over the summer, is having started formally to agree access agreements, which are hugely complex, because there is a tonne of commercial sensitivity, IP sensitivity, security sensitivity—very complex documents. They are being negotiated for those three, but we have not asked at this stage for full model weights because you do not need that to do the testing we have described.

Q688 **Dawn Butler:** I am glad that you mentioned bias and how it should be considered a risk. Do you think that other decision makers and main players understand that?

**Matt Clifford:** You mean the companies.



**Dawn Butler:** Yes.

**Matt Clifford:** I think they are acutely aware of it. They make money by selling services largely to enterprises. Their incentives largely follow the incentives of their customers. It is not my job to speak for them—I would be sitting on the other side of the table—but we have spoken to Microsoft and Google, two of the big players in the space, and they are acutely aware of their responsibilities under existing legislation in the UK and elsewhere and take them seriously, or at least that is what they tell us.

Emran mentioned that for each of the roundtables at the summit we opened with a presentation from the AI Safety Institute showing some of the emerging work. Contrary to reports about us excluding this topic, one of those was on bias and social harms. We have a demo showing how if you change some of the cues given to the model about social class it gives different advice. It gives less ambitious advice if you give cues that the person asking is working class.

Today, that is in the model, as is. It is semantics whether that is a safety question, but most observers in that room certainly felt that it is a question that needs to be dealt with in the next generation of models. Companies broadly accept that.

Q689 **Dawn Butler:** Could those presentations be delivered to the Committee?

**Emran Mian:** May I take that away and come back to you with a response?

**Dawn Butler:** Sure. We can do it in private; it does not have to be in public.

Q690 **Aaron Bell:** Mr Clifford, you have already given us a fair bit of detail, but who will be involved in the writing group for the “State of the Science” report? An expert advisory panel is advising the group. How will the nuts and bolts work?

**Matt Clifford:** We have asked each of the 28 countries that attended to nominate experts. We have asked that they not be people with big commercial interests. There will be academics.

Q691 **Aaron Bell:** Will this be for the panel or the group?

**Matt Clifford:** Effectively, the group is just a small secretariat, given that those on the panel, especially Professor Bengio, are very busy people. I hesitate to make the analogy, but it is a little like the IPCC. A secretariat will physically type the report, but the panel is where the scientific expertise comes from. The secretariat will not be academic experts in AI.

It is envisaged that Professor Bengio, as chair, will convene the panel from countries’ nominations, set an agenda and work closely with the secretariat to shape that into a report based on what is happening with the science over the next six months.





## HOUSE OF COMMONS

Q692 **Aaron Bell:** It says that the report will not produce new material or make recommendations but will look at priority areas for research. Given the speed at which everything is moving, is there not a danger that the process will be instantly outdated, bearing in mind how long it takes 28 people to agree a report?

**Matt Clifford:** It is a great question. One of the things that was a bit of a coup was agreeing to have the next summit not in 12 months but in six months. That push came from exactly that concern.

The UK is fully engaged in all international fora, including the recently announced UN high-level body on AI. One of our concerns about relying on that for the next summit is that these things traditionally take a long time. One of the reasons we wanted Professor Bengio as chair is that he saw the urgency and felt able to commit to having something done by the next summit.

That is a long time in AI terms, but in a way the timing is fortuitous. In six months, we will be on the brink of the next models coming out. Most industry observers suggest that the models will be released next summer. If that is right, I hope that the "State of the Science" report will contain the most accurate synthesis of where the science is on the eve of the next generation.

Q693 **Aaron Bell:** You talk about future summits being in Korea in May and then in France.

**Matt Clifford:** We have not confirmed the dates, but it is Korea in the spring/summer and France in roughly a year.

Q694 **Aaron Bell:** Will they both be safety-focused summits, or will they be wider AI summits?

**Matt Clifford:** It will be for the hosts. It has been pitched to them and accepted by them on the basis of it being the AI safety summit.

Q695 **Aaron Bell:** Can you confirm whether Jonathan Black will remain as the PM's representative at the summits? I appreciate you cannot speak for him.

**Matt Clifford:** It was an unusual summit. Both the sherpas, or PM's representatives, were hauled in from outside. Jonathan is currently on leave from the civil service as Heywood Fellow at the Blavatnik School. I believe he is returning to do that work, but you would have to speak to him.

Q696 **Aaron Bell:** If we are to have a global summit every six months, it might make sense for someone to take on the role permanently. Do you agree?

**Matt Clifford:** That is a matter for the Department. Emran has my phone number if he needs any advice!

Q697 **Aaron Bell:** Mr Mian, if we are to have regular summits should someone like Mr Clifford or Mr Black be appointed permanently?



## HOUSE OF COMMONS

**Emran Mian:** We were the hosts of the summit and of a new stream of work; it was the first time an international conversation had been convened around this set of issues. We did feel that we needed to lean into it very hard, and it was great to be able to rely on the expertise of Matt and Jonathan.

The task for the next two summits is slightly different in nature. We are not hosting. We set a lot of the tram lines by hosting the first one, but that is not to underestimate the amount of work the UK would want to do, having, not least, in some way contributed to the setting up of this stream of conversations.

We are thinking about what capability we need in the Department to be able to do that. We certainly need to be able to grow our international expertise, including on AI specifically. We have a head start, thanks to the summit, but we have not made decisions about leadership going forward.

Q698 **Aaron Bell:** A decision will be made at the time, but you envisage sending the Secretary of State to future summits, unless the Prime Minister himself wanted to go.

**Emran Mian:** That would be a decision for Ministers to take, but both the Secretary of State and Prime Minister dedicated a massive amount of time to this, in advance of the summit and at the summit itself. They continue to do so, including engaging closely with many of the key AI countries. The head of state of Korea will visit the UK later this month, so engagement remains very close, and the level of engagement among senior Ministers, including the Prime Minister, remains high.

**Matt Clifford:** May I add briefly—it is nice to have the opportunity to put it on the record—that in landing in the Department 10 weeks ago one of the things that was immediately clear was the extraordinary talent base in DSIT focused on AI? I think that, with a tiny bit of bias, 10 weeks of that group engaging with the world on AI has probably created the best qualified group of civil servants globally on thinking about the international aspects of AI. It just has not been done before, so the summit work winds down but the AI policy directorate within DSIT and the international elements of that give us, as a country, a really strong base for continued leadership.

Q699 **Chair:** That is very good to hear.

On the day two declaration, have all the developers who attended the summit signed up to give access to their models?

**Matt Clifford:** Day two had a separate invite list. We invited nine developers, seven of whom we call, as a shorthand, the White House seven. Just before the summit, seven of the biggest companies most at the frontier of AI agreed voluntary commitments on AI safety.



## HOUSE OF COMMONS

We invited those seven. We added Mistral, a relatively new French company that has raised a lot of capital. It is, crucially, focused on the open-source side, so it was an important voice to have at the table.

We invited X. A little-known entrepreneur, Elon Musk, is the CEO.

All nine of the companies signed up.

Q700 **Chair:** Do you expect the list to grow, a bit like the White House list?

**Matt Clifford:** That is right. Our aspiration, although it is now work for other people, would be that anyone who is building sufficiently close to the frontier that their capabilities and systems are unknown at the time of training ought to be party to that agreement or something like it.

Q701 **Rebecca Long Bailey:** Mr Mian, how will DSIT ensure that the UK's international AI safety work complements rather than duplicates other international efforts such as the G7 Hiroshima process and the US Artificial Intelligence Safety Institute?

**Emran Mian:** Work is being led from the same parts of the Department. We are closely linked. The same civil servants and Ministers who were feeding into the G7 process were the same officials and Ministers who set up the AI safety summit. I am confident that we had that co-ordination and collaboration in advance of the summit and will continue to organise it on that basis.

The exciting area is collaboration on safety, partly because the UK has made a good head start and has a lot to bring to the conversation. We have already brought a lot of expertise into Government to form the basis of the safety institute, and we have started work, including that which we were able to present at the summit. That has given us a head start in forming collaborations with others around the world who are going to be doing the same work.

It has helped the UK to position itself as a centre where people who are not British may want to come to do research, and we are finding there is interest in coming to work in the safety institute.

The US confirmed shortly before the summit that it was going to create a safety institute. We are already picking up the conversation with the US about what collaboration between the safety institution that they may create and what we are creating looks like.

Q702 **Rebecca Long Bailey:** On the White House, what assessment have you made of the Executive Order that was made last week on safe, secure and trustworthy AI?

**Emran Mian:** We are studying the order in quite a lot of detail. We had good conversations with US Administration officials in the run-up to it.

Our initial assessment is that it covers a lot of the same areas as we were covering. If you take the combination of the AI safety summit and the



## HOUSE OF COMMONS

White Paper on AI, we think that it covers a lot of the same ground. Inevitably, because their legal system and Government apparatus are different, some of it has come out differently in terms of how they will look at issues. Given the scope of issues that they have identified in the Executive Order and the scope of issues on which we are focused, it feels that there is a high level of commonality.

**Q703 Rebecca Long Bailey:** I put the same question to you, Mr Clifford.

**Matt Clifford:** I don't have a lot to add to what Emran said. The US, from a western, non-China perspective, is the key player. In the run-up to the summit—weekly, then daily—we had close engagement with the White House and Department of Commerce, where its safety institute will fit. What Secretary of Commerce Gina Raimondo said on stage at Bletchley last week was a full-throated endorsement not only of the UK's leadership but of the need for very close collaboration.

The EO is extremely broad in its reach. It goes way beyond safety into a lot of other questions, which is necessary and makes sense, but on the safety issue, if you look at the categories we discussed earlier—misuse and issues of control—you see not only a very similar vocabulary or taxonomy of risks but a similar approach.

You also see a lot in the EO on the emerging consensus on the role of the public sector, building state capacity and being able to go toe to toe with model developers to mark their homework, to use the Prime Minister's phrase.

**Q704 Rebecca Long Bailey:** The EU has taken a slightly different approach. What assessment have you made of its draft AI Act?

**Matt Clifford:** It is obviously still in trial. We have yet to see the final text, and I understand that there are still bits to be debated, particularly on how, if at all, to regulate the most powerful general purpose models.

We spent a lot of time with the Commission in the run-up to the summit. Emran can probably better speak to the broader policy piece, but I can speak in the context of what we agreed at Bletchley.

President von der Leyen's state of the union address had one segment on AI. The language she uses praises the EU AI Act—it would be strange if it didn't—and it is very similar to that used by the Prime Minister about the risks she worries about and the international collaboration needed to mitigate it. In her public remarks last week, we were encouraged to see that she welcomed the "State of the Science" report as it aligns with the IPCC-style power that she called for.

The EU will obviously take a different approach. I will not get into Emran's territory, but relative to what you have seen in the White Paper around a sector-led approach to narrow AI, that is clearly not the approach they have gone for. I want to draw a distinction between the broad regulation of AI and AI safety. On AI safety, what we saw in the



## HOUSE OF COMMONS

run-up to the summit and at the summit was tight alignment on approach.

Q705 **Rebecca Long Bailey:** What assessment have you made of the EU's draft AI Act?

**Emran Mian:** Matt focused on the safety elements, where there is high alignment, as shown through the Commission President's engagement and that of Commission officials in the summit and in its run-up.

The other part for which the EU is regulating is the sector-based harms that could be driven by AI—for example, where data protection issues are created or exacerbated through the training of AI or through its deployment.

Our approach is set out in the White Paper. We begin by looking to our existing sector-expert regulators to use their powers within the regulatory framework. We are in regular conversations with them about them doing that.

To continue with the example of data protection, I saw the Information Commissioner yesterday. The Information Commissioner has given evidence to this Committee about the approach that it is taking. It has laid out a set of principles in terms of how it sees the data protection framework applying to those who are training and deploying AI models. It has already taken at least a couple of significant enforcement actions where it thinks there is either a breach or a potential breach of data protection. It is already being active in using its sector-based powers, and they are very broad powers, of course, because data protection spans very many sectors of the economy.

We are seeing that from the other regulators as well. The medicines regulator is doing a lot of work on the deployment of AI in healthcare. The Competition and Markets Authority is already thinking quite deeply about some of the competition issues that may be caused here. It has published an initial report on what it sees as some of the issues that may come through. It is doing that using its existing powers under the competition framework. Of course, we are in the process of strengthening or adapting the powers that the Competition and Markets Authority has for digital markets specifically through legislation. That is before the House at the moment and will be carried over from the previous session.

We are continuing with the approach that we laid out in the White Paper. We have a set of expert regulators here who we think can deal with some of the specific and sector-based harms in the same way that the EU is trying to deal with them through the EU Act. We think they can do so using their existing regulatory frameworks. As they do so, if they are finding lacunae or issues where they are not able to take the enforcement action that they think they may need to take to protect citizens, consumers and other businesses, we are talking to them regularly to understand what the issues are.



Q706 **Rebecca Long Bailey:** Clearly, what you have just set out is slightly different from the approach that the EU has taken. What are the reasons that we have chosen to diverge from the EU's approach? What deficiencies did it have in its proposals?

**Emran Mian:** I am not sure I characterise it as deficiencies in its proposals. It is more that we are able to draw on quite a wide set of sector-based, and broad-based regulators, as in the data protection area, that we think are already able to take action using their existing powers around some of the specific harms or risks that may be associated with AI.

The gap has been on the safety side, which is why we are investing so hard in the capability on safety, focused so hard in the summit, and making progress on the policy in the international collaboration for safety, but when it comes to non-safety issues, when it comes to issues associated with the deployment of AI, we are definitely continuing with the approach that we have in the White Paper of using our sector-based regulators.

Q707 **Chair:** On the US's approach, Mr Mian, when Vice-President Harris was in London last week, on 1 November, she talked about the US AI Safety Institute within the Department of Commerce, and she was very clear that its role would include issuing technical guidance to sector regulators in the US and she gave some examples of rule-making and enforcement on issues such as authenticating content created by humans, watermarking, AI-generated content, identifying and mitigating against harmful algorithmic discrimination, ensuring transparency, and many more things.

Is the UK AI Safety Institute intending and expecting to give similar direction to regulators?

**Emran Mian:** It is probably slightly too early in our journey with the safety institute to give you a completely definitive answer to that. The kind of thing that we absolutely need to think about is that there has been some commentary and some stuff written about how the deployment of AI in financial decision making may add to systemic financial risk or may change the character of systemic financial risk.

It is an important research question for the safety institute to think about whether, as part of its evaluation of models, it is able to get some grip on the extent to which models might create that risk. If it is able to do so—and that is an open scientific question at the moment; I do not think anyone has done that evaluation work yet—the safety institute may very well be best placed to do that work and then provide the sector regulators with the findings of the work that it is able to do.

I am just trying to think ahead here rather than trying to give you a definitive answer. That is probably the kind of thing that the Vice-President was also indicating.



Q708 **Chair:** This is a crucial point, and the Committee is worried about it. In looking to the long term, I accept Mr Clifford's view that you can separate the long term and the short term and some of these things apply in both, but there are some here-and-now questions for regulators. Kamala Harris has said very explicitly that their safety institute will be issuing technical guidance to regulators in all these respects. You said that the UK institute may look at this, but it is not at the point of doing that yet. Meanwhile, the US is getting on with it. As Rebecca Long Bailey said, the EU is proceeding with its Act.

I noticed that the Secretary of State, in talking about the American institute, said that, yes, they have plans for institutes, but we can do it a lot quicker because we already have that initial organisation. What you have just told the Committee, actually, is that they are proceeding with giving direction to regulators but that we are, in fact, behind. Is that a reasonable characterisation?

**Emran Mian:** Some of this may come down to the differences between the two systems and where we would do certain things. Another example of what the US safety institute may look at is algorithmic transparency, which is one of the things that the Vice-President has identified that their safety institute may look at.

Through a group of officials in the Department working closely with outside experts—the Centre for Data Ethics and Innovation—we have already created an algorithmic transparency record, and there are a number of deployments of algorithms that have already gone through that process to recording algorithmic transparency.

That is a functionality that we have not waited for the creation of a safety institute to have. It is a functionality that we have already done through a part of the Department. That is an area where we need to continue to make progress.

Q709 **Chair:** On that particular point, because obviously algorithmic transparency is important, that is a voluntary standard, is it not, for public sector organisations?

**Emran Mian:** It is a voluntary standard; that is right.

Q710 **Chair:** What we are talking about in the US is directions to regulators. Correct me if I am wrong, but since this was initiated in 2022, over a year ago, only six reports have been issued. The reporting standard was not even mentioned in the AI White Paper.

**Emran Mian:** Yes, this is why I say that we need to do further work on algorithm transparency. I was drawing out the question of whether we would do it within the safety institute or we would do it differently. That may be a question that we resolve in the UK perhaps in a different way from the way the US does it.



## HOUSE OF COMMONS

Creating the transparency around this and the focus on ensuring fairness and dealing with bias is very much recognised in the White Paper, is very much recognised as part of the Bletchley declaration, and is recognised in terms of the work that we need to do following on from the summit. As Matt said, the companies are quite focused on it as well, not least because it is important to their commercial deployment of the products and services that they are creating.

**Q711 Chair:** The algorithmic transparency recording standard was proposed in 2021, yet in 2023 it is not even mentioned in the White Paper. Obviously, it is a crucial aspect when it comes to bias to be able to see what algorithms are recommending. It is not mentioned in the White Paper, there is a voluntary standard, and there is very low compliance in terms of reporting back on it.

To go back to the US in terms of its specificity, I looked at a draft that has come out of the Executive Office of the President of the United States going into some detail. It is a draft prior to rapid adoption. All Government agencies in the US must designate a chief AI officer within 60 days of the issuance of the memorandum that is being consulted on. There is a very long list of requirements of those officers. It is to report on the use of these agencies in law enforcement, of surveillance-related risk assessments about individuals; criminal recidivism prediction; offender prediction; predicting perpetrators' identities; facial matching; deciding immigration or asylum or detention status; detecting student cheating or plagiarism in education; and tenant screening and controls in the field of housing.

This is a very detailed set of requirements that are being placed by the US Administration on Government agencies—very specific now. The contention is that we are ahead of the game here, but it seems that we are very much behind in implementing the here-and-now regulation that the US is. Is that not the case, Mr Mian?

**Emran Mian:** No, I do not think that is our assessment. I say all that with the caveat that we need to continue to study the Executive Order and we need to continue to talk to the US Administration about the advances that they are making.

Clearly, the US Executive Order is a significant and substantial measure. I do not want to suggest in any way that there is nothing for us to learn from it. We need to continue to learn from the US and we need to continue to learn from other countries that are moving forward on this.

**Q712 Chair:** The Secretary of State said we can do it a lot quicker, but we are already behind.

**Emran Mian:** On safety, we are moving very quickly.

**Q713 Chair:** It goes back to my first question to Mr Clifford: are issues like bias in algorithm use and the deployment of AI in tenant screening and controls not safety?





## HOUSE OF COMMONS

**Emran Mian:** As we went through a few minutes ago, we are not characterising safety as focused only on long-term issues. Safety includes issues that are presenting now.

Q714 **Chair:** Okay, so where are the directions to regulators from the Department or from the Government that are the equivalent of what the US is already doing?

**Emran Mian:** As I was trying to articulate, we are not certain that we need to issue equivalent directions to our regulators. The example that I used was of the Information Commissioner. The Information Commissioner has taken—and, of course, we have been having conversations with it—a proactive approach to laying out the application of the data protection framework to the training and deployment of AI. We have not felt that we need to issue any additional direction to the Information Commissioner around that because what we have seen both in conversation and in what the Information Commissioner’s Office has done publicly is that it has taken a proactive approach to this.

If there are areas where we do not see that same proactive approach, it may be things that we need to revisit, and it may be then that we have to get into a different mode with the regulator. I was using the Information Commissioner’s Office as an example of where we have seen that proactive approach.

The other aspect of what you were talking about was some of the measures that the US is taking to ensure the adoption of AI across public service and charging various officials and various agencies to bring forward use cases of adoption of AI and to do so in a safe and fair way.

Having the algorithmic transparency work and having the focus on fairness alongside, through the Central Digital and Data Office, having a focus on the deployment of AI through the Government’s digital profession, which we have had now for a couple of years, is our approach to ensuring that we have good adoption of AI in products and services where Government are involved.

Q715 **Chair:** Does the Information Commissioner’s Office have all the powers it needs to do everything that might be required for current applications of AI risk?

**Emran Mian:** As the Information Commissioner’s Office has said, it has felt able to lay out how a data protection framework applies to both the training and the deployment of AI tools. It has done that publicly.

Q716 **Chair:** Does the Department have all the powers that are needed?

**Emran Mian:** Yes, our assessment at the moment is that we do not need to provide any additional powers.

Q717 **Chair:** To the Information Commissioner’s Office.



## HOUSE OF COMMONS

**Emran Mian:** But we will keep that under review, and we will keep it under review with other regulators as well.

Q718 **Chair:** Data privacy is an important aspect. Would you share with the Committee your assessment of the adequacy of the ICO's powers?

**Emran Mian:** Can I take that away, Chair, and look at it? I just want to be careful of not duplicating something that the Information Commissioner's Office has already done. It has already talked about and already published how the data protection framework applies.

Q719 **Chair:** I am sure it has. I am interested in the Government's view and the Department's view. You have just said that you are satisfied that it has the powers that it needs in the face of the new risks that come with AI. If that is your view, that it does not need any new powers—this is a very important matter for the Committee—we would be grateful for your assessment. I cannot see any reason that should not be provided.

**Emran Mian:** As you know, in parallel, we had the White Paper on AI. We have had the consultation out on that, and we need to take and publish the next steps. Either through that forum or through a direct letter to the Committee, we can cover that point.

Q720 **Chair:** If you have made an assessment that the ICO has enough powers, we would appreciate seeing that assessment.

**Emran Mian:** Okay.

Q721 **Tracey Crouch:** It is not just the ICO, though. In a previous answer, you talked about other regulators and their use of AI. Are you certain—you used the phrase "certain" in a previous reply—that they have the legal powers, as set up through the legislation that establishes them to deal with some of this? We would not want them to be acting ultra vires on some of these issues.

**Emran Mian:** I think that is right. If I may take a step back, the thing that is really important for us all to notice about AI is that it will be deployed across very many different parts of the economy. It is not going to be deployed in only one single sector, which is why looking at whether different regulators in different sectors and the different legal frameworks in different sectors can deal with some of the risks or some of the potential harms is important.

That inevitably means that any time you ask me or you ask the Minister about whether we feel confident that we have the right approach to regulating AI, it is going to be quite a long answer because the deployment of AI is going to happen across a whole range of different sectors.

The one area where it is a simpler and a more straightforward answer is when it comes to frontier safety, where we can invest up front in a set of capabilities that allow us to assess AI safety, and to do that pre-



## HOUSE OF COMMONS

deployment. That is the thing, as part of the safety summit, that we were really focused on making progress.

The key to deploying AI products and services into the economy is to rely on individual regulators to see where, using their existing powers, they can do what they need to do to be able to deal with those risks and harms.

What we have said in the White Paper is that we will keep that under review. In fact, we used the White Paper to seek wider views and to get more evidence about how people see this playing out. We remain part of that conversation with each of the major regulators.

**Q722 Tracey Crouch:** One of the things that arguably could have helped to regulate the regulators and provide the required guidance and direction would have been to have a narrowly focused AI Bill in the King's Speech, which the Committee recommended. I cannot help but notice that it was not in. We seem to be lagging behind other jurisdictions that are setting regulatory standards. We are just relying on quite vague, I am sorry to say, discussion or guidance around what other regulators might be doing. Could you explain why you do not think that an AI Bill is necessary?

**Emran Mian:** The view that we have taken at this point is that we do not think an AI Bill is necessary because we have not yet seen evidence that regulators lack the powers to be able to deal with the risks and harms that they are being presented with.

The Information Commissioner is already taking active enforcement action. It is using the powers that it has to take those actions. The Competition and Markets Authority, within the framework it operates in—a framework that we are in the process of legislating to adapt for digital markets—has already been doing work on foundation models and to what extent foundation models may create competition issues that in time the Competition and Markets Authority may need to deal with.

The approach we are taking is very much evidence-led. If the regulators were flagging to us, or flag to us at any point, that they feel there are issues arising that go beyond the powers that they have and significant harms are being created or significant harms are at risk of being created that need dealing with, we would want to reflect on that.

**Q723 Tracey Crouch:** You are specifically ruling out an AI Act here in the British Parliament, basically.

**Emran Mian:** I do not think it is my place as an official to rule it out. What I am playing back is the current Government policy that Ministers have sent me here today to give evidence to the Committee about, which is that we do not have an AI Bill in this Session of Parliament, and that the approach that we are taking, as laid out in the Government White Paper, is to rely on the existing powers of the regulators.

**Q724 Tracey Crouch:** You both said in response to earlier questions that this



## HOUSE OF COMMONS

area of policy is developing so quickly that we are now having six-monthly summits. Do you think, therefore, that there is a real possibility that we will have missed the boat on any substantial regulation as a consequence of the guidance that you are giving Ministers—to rely on the existing powers of regulators in what is probably an area of policy that really requires some defined regulation?

**Emran Mian:** Our assessment has been that the area that is moving the fastest is the development and the deployment of the foundation models themselves. What we need to do most is to build the capability to be able to assess those models before deployment. That is why we are so focused on building the capability to be able to do that work effectively, including building that capability within government, but also to work really closely with the developers that they—

Q725 **Tracey Crouch:** You are relying on the regulators as they exist. You are talking about guidance for the medicines authority, for example. Do you think it actually has the tools, the capability and the capacity to be able to apply that regulation in the way it ought to be done?

**Emran Mian:** Skills, capabilities and expertise is a slightly different issue that arises regardless of whether we have new legislation or not. In the world where we do not have new legislation, which is the current Government policy, we still need to work really hard on the skills, capabilities and expertise of the regulators. That is absolutely clear. There is a set of steps that we are taking with each of the regulators to understand what their current skills and capabilities are and to identify how we can support them in improving those skills and capabilities.

Q726 **Tracey Crouch:** I am one of the members of this Committee who is quite new to this area of policy. I have to say we have heard some witnesses talk about the potential harms of AI, and I find it, like some members of the public, quite frightening. It is not really filling me with much confidence that this is going to be an area that is properly regulated and controlled.

Is that an unfair assessment of where Government are in terms of safety, regulation and guidance given that we do not seem to be as far ahead as other jurisdictions?

**Emran Mian:** I do not think that that is the assessment that we have made. The assessment that we have made is that the UK is taking some really strong first steps in building the capability to assess the safety issues that arise from frontier models. That is a capability we have done a lot to build over the last few months and we will continue to build over the next few months. That is a scientific enterprise that we are really well advanced on.

Alongside the work that we are doing, we need to stay really close to the companies that are improving their own capabilities to do this work. We need to help to keep the pressure on them so that they continue to



## HOUSE OF COMMONS

improve their capabilities. I feel that on that front the UK is in a really good place.

Alongside that, there is the issue of where the deployment of AI products and services creates specific risks or specific harms. What I am saying to the Committee is that we are seeing really strong evidence already of individual sector regulators and, indeed, broad-based regulators taking early action to get on top of those issues. That should give members of the Committee and members of the public confidence that those regulators are taking those steps.

If we identify examples of harms or potential harms arising where we do not think regulators are taking the right steps or where we do not think regulators have the right powers, we would obviously need to consider and come back to that.

**Q727 Tracey Crouch:** You mentioned, in response to the Chair, the White Paper and the Government response. I appreciate that you cannot give the grid slot date, but could you say whether we are expecting it before Christmas?

**Emran Mian:** I think the Secretary of State said that we were hoping to publish it this year.

**Q728 Tracey Crouch:** Okay, thank you. Will it be accompanied by an AI regulation road map, as promised in the White Paper?

**Emran Mian:** We will take the next steps that we said that we would take as part of the White Paper.

**Q729 Tracey Crouch:** We saw in the news that Ministers are using AI as part of their submissions and other areas of policy. They have subjected themselves to the experiment of AI being used on Government documents. Out of interest, have you used any AI to write the Government's response on AI?

**Emran Mian:** The use case that we are looking at within the Department is on correspondence, where there is a lot of correspondence and where the responses are quite similar across different pieces of correspondence. There, we think there is some potential application of AI in generating really timely responses to correspondence. That is a specific example of where we are looking at using AI. We have not used AI to generate any of the products for the summit or the White Paper.

**Q730 Chair:** We are grateful to you for coming. You are an official in the Department. You are not a Minister. We are pushing because we are interested in making sure that we can translate the impressive ambitions and sophistication of policymakers into practice. None of us doubts that the work being done on frontier models is very valuable, but there are these here-and-now risks, and it is important to do both at the same time.

In the White Paper, it was anticipated that legislation might be needed to



require the regulators to have due regard to the cross-government principles and cross-economy principles set out. Why should that no longer be the case?

**Emran Mian:** It is something that we are consulting on, as you say, as part of the White Paper. It is not something that we are taking forward as a consequence of the King's Speech. Instead—it is the reason I wanted to go into it in detail—we are already seeing the regulators take action and have due regard to the issues that are presenting here and now. We are absolutely focused on the issues that are presenting here and now. There is no way in which we are only focused on issues that are going to arise only in the future. We are absolutely looking at the issues that are presenting here and now, and, indeed, so are the regulators.

Q731 **Chair:** Let us give an example. The Prime Minister raised in his discussion with Elon Musk on Thursday night what we call in our interim report the "misrepresentation challenge"—what people call "deepfakes"—which is obviously very important for elections, of which there are some very important ones next year.

In the UK, which regulator is responsible for making sure that deepfakes do not disrupt our elections?

**Emran Mian:** Ultimately, the responsibility for electoral integrity rests with the Electoral Commission. The Electoral Commission is one of the regulators that we are talking to in the context of the White Paper.

Q732 **Chair:** Does it have the powers it needs to make sure that deepfakes are not a problem in the elections now and next year?

**Emran Mian:** I am afraid I would have to come back to the Committee on detail in relation to the Electoral Commission. Specifically, I think the question you are asking is about deepfakes. We have not to date seen widespread use of deepfakes in political information. There have been some instances—

Q733 **Chair:** With respect, that is not the question. The question is a different one, which is germane to the Department, of which you are director general. Given the approach, which in our interim report the Committee welcomed, of working through existing regulators, concomitant with that is the need to satisfy yourself as the lead Department of Government that regulators have the powers that they need to operate in an AI-enabled world. If not, they may need to have some powers taken through legislation in the King's Speech, which will be a DSIT Bill.

It is reasonable to know whether you have made an assessment as a Department of whether the Electoral Commission has the powers that it needs to make sure that it has the powers to pounce on any use of deepfakes, which could be absolutely crucial in this country, and stop them disrupting an election campaign.



## HOUSE OF COMMONS

**Emran Mian:** Forgive me, Chair, I thought you were asking a different question. I thought you were asking about what information the Electoral Commission holds about the use of deepfakes.

Q734 **Chair:** Does it have the powers that it needs?

**Emran Mian:** Yes, that is a question that flows from the White Paper. That is a question that we will address as part of the response to the White Paper. I do not have the answer to it here and now for you today, but we are considering it as part of the response to the White Paper.

Q735 **Chair:** To Tracey Crouch's question, this comes into whether we risk falling behind other jurisdictions on that. I have read out some of the terms of the Executive Order that has come from the White House to US regulators. The Committee is going to Brussels next week to talk to the Commission and others. It is proceeding. Yesterday's King's Speech was the last King's Speech, the last legislative programme of this Parliament, beyond which it will be 2025 before it is possible to legislate.

Were it to be the case that some of the regulators, even one of the regulators, needed their powers tweaking slightly, that has to be done through Parliament. If you have come to a view that it was not necessary to include even a small, focused Bill in the King's Speech, surely that is on the basis of an assessment that that is not needed.

**Emran Mian:** That is the view that we have taken at the moment—that we do not need to legislate. Of course, we have not yet published the response to the White Paper, so it is a view that is—

Q736 **Chair:** It is possible that there might be a Bill, albeit not one mentioned in the King's Speech. The King's Speech says, "Other measures will be laid before you." This may be one of them.

**Emran Mian:** No, we have no current plans for legislation.

Q737 **Chair:** Right.

We have talked here about the ICO. We have talked about the MHRA, the medicines regulator. We have just talked about the Electoral Commission. You can go on and on. Surely, if every one of those regulators was set up using statutory powers, and those powers were optimised for a world of AI that, as Mr Clifford has said, has changed so rapidly, it would be the most bizarre remarkable coincidence that there was not a single thing that needed to be tweaked in those powers across any regulator, would it not?

**Emran Mian:** I have just two quick observations. The first is the thing that is moving very quickly is the computational power of the models. The ways in which the models are deployed may be things that are very familiar. If a model is used to generate a cyber-attack, that is very familiar.

Q738 **Chair:** You do not think there are changes and developments in the application of AI technologies that may not be frontier but may be being



## HOUSE OF COMMONS

deployed more widely than in the past.

**Emran Mian:** Chair, there may be. What I was indicating is that lots of the deployments are familiar. Lots of the deployments are things that are already done. Indeed, many of them are already done using technology.

The second observation I wanted to make is that some of the legislative framework that we are talking about here is either very recent in itself or is in the process of being renewed.

An example of something that is very recent is, of course, the Online Safety Act, which received Royal Assent only a couple of weeks ago. We think that as a consequence of that Ofcom has a lot of the powers that it would need for harms that are generated. That is very recent legislation.

Q739 **Chair:** Exactly. So—

**Emran Mian:** Chair, may I finish my point? There is legislation in the current Session that continues to adapt the framework. On digital markets and the regulation of competition in digital markets, that is legislation that is currently going through Parliament.

Q740 **Chair:** Does that not precisely make my point? In both those respects, Ofcom is acquiring new powers through a recent Bill and the CMA is acquiring new powers through a current Bill. Therefore, for all the other regulators, if it is proved necessary to update their powers, how extraordinary would it be that there was nothing that needed to be updated in the powers of other regulators that do not have the coincidence of a Bill being before Parliament just now?

**Emran Mian:** I see the question, Chair. I suppose the point I was trying to make was that it would be extraordinary if the legislative framework had been static for the past 10 or 20 years. It has not been static for the past 10 or 20 years. Indeed, the Online Safety Act only came into law a couple of weeks ago. We have digital markets legislation in the current Session of Parliament. Indeed, we also have data protection legislation in the current Session of Parliament. The legislative framework is not static, and the legislative framework has in itself been adapting to technologies. Of course, AI is not a completely new issue. AI is already being used in all kinds of technology deployments.

The point you make would have absolute force if the legislative framework itself had not been changing over the past period, but it has been changing quite significantly, which is why we think the judgment about whether to legislate further is a really finely balanced one, because we think it is really important first to test whether the legislative framework that we already have and the legislative framework that we are still in the process of adapting gives us the powers and gives regulators the powers that they need to deal with the risks.

**Matt Clifford:** I might make just one very small addition. Outside the frontier general purpose AI, I would encourage you to think of the impact





## HOUSE OF COMMONS

of AI as one of scale rather than one of kind. I know nothing about the Electoral Commission, but I assume that, if someone creates a photoshop of Greg Clark MP doing something that he did not do that might put off your voters, it has powers to intervene.

AI makes that significantly easier and scales it, but these are questions of capacity and capability rather than of powers in general. This is probably true across the board. We want our medicines to be safe and effective, whether they are designed by a human or a machine. We do not need new powers to do that. We may need new capabilities, new skills and new people. The one exception is general purpose AI, which is where the AI Safety Institute can make a unique contribution.

**Q741 Chair:** You make a fair point that the regulators may have the powers they need. The view of the Committee is that it is worth checking that rather than assuming it.

Secondly, we noticed that other jurisdictions, including the US, which is not known for heavy-handed regulation—the Committee visited Washington and spoke to the White House and Members of Congress and others—felt the need to date their regulations in the way we have talked about.

Let me go to my colleagues who have not asked a question yet.

**Q742 Katherine Fletcher:** I want to pivot us slightly. The capacity AI is generating is bringing enormous opportunity. Part of what we deal with is not only, to the Chair's point, regulation not being in place, because, to take your example, you had to be really quite good at photoshop to produce a decent fake and it took ages, so the volume is not there. There is that side of it.

I will start with you, Mr Clifford. We have to balance this safety and our ability to check it is safe with an ability to create an ecosystem that does not just sit with a small number of well-financed developers and a large consumer pool, aside from national security. I have had a couple of conversations with service providers, primarily B2B. We are quite good as legislators at doing B2C, business to consumer, very well, and we are creating this new safety institute to look at the developer model to make sure that they are not building something to be sold that is bad, but what about the B2B community? How do we make sure that we are not, effectively, creating the world of the "Hitchhiker's Guide to the Galaxy" by Douglas Adams where it is the only company and nobody else exists?

Where do we need to go to make sure there is room for B2B companies to take a model, put their own dataset through it, innovate and sell it as a service?

**Matt Clifford:** I almost feel that I can answer that more with my normal private sector hat on. My day job is helping entrepreneurs build technology companies, primarily in AI, which is, to the extent I have expertise in it, where it comes from.



## HOUSE OF COMMONS

The first thing I would say is that is happening. The UK is already a world leader in AI after the US and China in that order. We certainly have the third-largest AI ecosystem at the start-up and scale-up level, and there is every reason to think that will continue.

If you think about what the underlying ingredients of a vibrant ecosystem are, it is really the ones you would expect: talent and capital. The new ones for AI are the computational power to actually do something and data. Do we have the data that we need to train it?

On talent and capital, I have been fortunate to be a part of building the UK's tech ecosystem over the last 12 years. The growth has been extraordinary. It is only the third trillion-dollar tech ecosystem in the world, built here in the UK.

You see that across the board in AI. As I said, the only country that actually has a world-leading frontier AI company outside the US is the UK. DeepMind was founded up the road and remains there, and, in fact—hopefully, they are not watching in Mountain View—engineered a reverse takeover of Google AI earlier in the year. All the research is happening here in London.

**Q743 Katherine Fletcher:** That is exciting, but is there a risk that in pushing for safety a bunch of us that are a bit more like potato print computer people regulate something that we do not understand and stifle the innovation?

**Matt Clifford:** I do not think that is a risk. One of the things that I was asked a lot when I took on this role was, "You are an entrepreneur. You are an investor. Why emphasise the safety angle when we could be emphasising the innovation angle?" My response was that only AI that will be adopted is safe AI. Tracey Crouch made the point that this is a topic on which the public—

**Q744 Katherine Fletcher:** I agree, but here is my question. If making it safe means that a regulator demands that a smaller SME has to put such checks and balances in place that it could never hope to achieve the capital to start up again, we end up with big developers only being able to do it.

**Matt Clifford:** That is a huge risk, and that is why we have taken the approach we have of saying safety has to happen at the model layer. An SME is not going to build their own models; they cost hundreds of millions of dollars to build. Actually, the SME needs to know that when it uses an OpenAI model, a Google model or an Anthropic model the developer has invested, as you say, lots of money and time in making it safe. That is actually the right approach.

There are lots of reasonable critiques about the necessary relationship between Government and the big corporate actors, but I disagree with some commentators that this is regulatory capture. In fact, the approach that we took in the summit is that if you are building something unusually



## HOUSE OF COMMONS

powerful you should face unusual scrutiny. It is not SMEs that are having to give access to their models for safety testing; it is seven very large corporations. The way to get safe adoption of AI by application builders is to reassure the public that the underlying models are safe.

Although I do not work in the Department after Friday, one thing that I would praise about the White Paper is that what it broadly says is, "If you are not building one of these extremely capable general-purpose models, go to the races. Government do not want to get in the way of what you are doing, and they do not want to create a new regulatory burden."

That is the potential critique of the EU AI Act and of the EO. They basically try to anticipate all the things that might go wrong in future at every level right down to the two-person SME. Instead, the White Paper says, "Let us let them run. They are not doing something unusually dangerous. Let us preserve our focus and let us preserve the onerous burden on the companies that are best equipped to shoulder that," and that is exactly the approach we took in the summit.

**Q745 Katherine Fletcher:** Mr Mian, to build on that—I do not know whether you want to make any broader comments—say that a B2B company has thrived thanks to the UK's watch-and-see regulatory environment. Say the model is proved safe, but because we are potato printing with regulation it cannot cope with the data protection provisions. The way it has been explained to me as an example is that it is like asking us to set up and sell a post office but then telling us we have to read everybody's letters to make sure it is safe, and we cannot do that because then we violate data protection.

Is the Department minded to make sure that commercial interests need to be protected as well as individuals? I do not push back from it needing to be safe. Perhaps on data and personal information, is that something you are looking at?

**Emran Mian:** This is something we are looking at, and it is something we said in the White Paper we would look at. You give a really good example of the issues that might arise where there is not sufficient transparency between the developer of the foundation layer and the business, and it may be a much smaller business that is then deploying that foundation model for a specific purpose. If the organisation that is deploying it for a specific purpose does not have sufficient transparency around the data used in training the model or whether the model displays any particular biases or—

**Q746 Katherine Fletcher:** Or checking its outputs for problems.

**Emran Mian:** Exactly. That can cause an issue and liability can end up in the wrong place.

For the most part, that is not a new problem; it is a problem that arises in all kinds of industries as well where large developers of foundations are providing a service to other businesses. Usually, the approach to dealing



with that is through the contractual relationships between the two firms. We said in the White Paper that that is an issue that we want to continue to look at. Given the imbalance between the size, the scope and the financial power of the two different kinds of organisations, you might end up with some challenges in transparency.

**Q747 Katherine Fletcher:** I grant that commercial terms can prevent liability. The small user of the foundation model can say, "If this spits out a load of racist something, I'm going to sue you, and it's not my fault." I am interested as a representative of the people of Lancashire in making sure that it does not spit the racist stuff out in the first place and that it does not mean that you are losing your data protection because the SME that is using the model has to be so in the detail continually to check and make it safe. How are you going about squaring that circle?

**Emran Mian:** It is so important that the safety work and the testing happens at the foundation model so that we create the confidence not only for citizens but for small businesses or other businesses. It is important that these issues have been looked at in the foundation layer itself so that they do not have to do that work necessarily each time that they are making a commercial contract, because they know that it has been done through the public sector.

**Katherine Fletcher:** It is base-layer safety regulation so that we do not stifle application and use in the organisation. Great. Thanks very much.

**Q748 Stephen Metcalfe:** This has been a long inquiry, and, as we know, a week in AI is a very long time. I do not know where I have heard that before. Technology is developing. I recognise the issue of legislation is very tricky, and I am not sure there is the complete agreement on the requirement for legislation and regulation that there was perhaps at the start of the inquiry.

The Chair mentioned that we visited the US in September and went to Washington. We also went to Boston, where we heard from a quite senior lecturer at MIT, Randall Davis, who told us—and I think this is where the Government are coming from—that we have almost all the regulation we already need. It is about using existing frameworks and remembering that AI is a tool. It is a means to an end; it is not the end in and of itself. We need to try to remind ourselves of that when looking at legislation.

Bearing in mind how quickly things move and develop, you talked about "into 2024", but there is a whole decade ahead of that and beyond. What future work are the Government doing to make sure that we are fleet of foot enough not just to deal with what happens next year but beyond that into the future? I understand that the UK is setting up an AI research resource. Is that going to take on the role of looking into the future and deciding what direction we should be taking? Can you update us, Emran, on the work that that is doing and how it will continue and keep relevant?

**Emran Mian:** I will start with the research resource itself, the supercomputers, essentially, as they are often referred to. We think this



is another really important part of developing the public capability of the UK. In addition to some of the other public capabilities that we have talked about, which will also fundamentally position the UK competitively on AI and its uses, having a public computing resource that is accessible to the scientific community for the purposes of both safety research and other scientific research, and that can be enabled by and indeed sped up by the use of AI, is really important.

We think it is also going to be important to provide access to businesses to some of that computing resource. UKRI already does that on some of the facilities that it runs, including some of the supercomputing facilities. We think that is important. Indeed, the EU has also said that it will be providing access to its supercomputers to smaller businesses. That is an important part of our thinking, too. You are absolutely right. Continuing to develop where we are on computing resource is an important part of being ready for the future.

The other part—it goes back to the conversation we have been having about regulation—is not assuming that we have the final word on regulation but keeping these issues under review. Our current assessment is that we have the regulatory framework that we need, but that is not an assessment that is now closed and we have shut the folder.

**Q749 Stephen Metcalfe:** Thank you for that. I am very conscious of time, so I will try to wrap up my last couple of questions as briefly as I can.

This is about the use of AI across Government. The Treasury has taken quite an interest in how AI might be used to streamline some roles in Government. What role has DSIT taken in talking to other Government Departments about their use of AI and how it might be deployed?

**Emran Mian:** We have been focused on building the UK capabilities, including the ones that we have just been talking about, the policy landscape and the regulatory landscape, and taking key actions to show what the potential of AI is. That is something that risks getting lost in some of the conversation about safety and risk; that we fail to focus sufficiently on what some of the potential is and what some of the benefits are. A lot of those accrue to the public sector and a lot of them accrue to citizens as well. They are not purely private benefits.

That is why in the run-up to the summit we did quite a lot to show what some of the examples are and, indeed, to take some new steps on them. In the run-up to the summit, we announced further funding through UKRI for deploying AI in healthcare and using it to do really cutting-edge work in scanning.

We have also been talking to Department for Education leads about the use of AI to help to reduce the burden on teachers associated with some of the activities that eat into their workday and whether there are ways in which the deployment of AI can give them a better balance in workload and allow them to focus more on the classroom environment.



Q750 **Stephen Metcalfe:** Where that AI is being used, do you think there is an obligation on us as recipients of the AI to be informed of that, or is that a bit like saying that you need to be informed that someone has used a tool? Should being informed be mandatory or voluntary?

**Emran Mian:** That is a good question. It is not something that we have thought deeply about. The challenge is that AI is already used in a massive number of products and services and is already being used to generate benefits in a massive number of products and services. It is often the hidden technology layer, not hidden in any sinister sense, but simply hidden in the sense that the consumer or the customer does not necessarily need to know what is happening inside the machine. What they need to know is that the outputs of that machine can be trusted and the way in which it has been developed is safe and manages risks for them.

It seems to me that we should continue to prioritise those elements rather than prioritising making it very public exactly where AI has been used in the deployment of a product or service. Matt may want to come in on this one as well.

**Matt Clifford:** That is right.

Q751 **Stephen Metcalfe:** Okay, thank you.

My final question is about the regulatory landscape of those advising the Government on AI. We have the Centre for Data Ethics and Innovation and the AI advisory board. Some of those are being disbanded. Why are those decisions being taken, and what do you think the landscape will look like in the future?

**Emran Mian:** The key thing for us is to continue to profit from and continue to draw in a really wide range of expertise. We have done that through a number of board arrangements in the past, including the board for CDEI, and that continues to be a big area of focus for us. That is why we brought in Ian Hogarth from outside government to chair the taskforce. It is why we have an advisory board for the taskforce that draws in lots of key experts from both academia and industry. It is why we have been drawing on the expertise of people like Matt. That is something that we need to continue to do.

It also informed our judgments about whom to invite to the summit. The judgment about whom to have at the summit was a judgment about which countries are part of the conversation and about which civil society organisations, which academics and which universities should be part of this conversation as well as the companies. What we are trying to demonstrate throughout is a really open approach to bringing in expertise from all the different parts of the industry.

Q752 **Stephen Metcalfe:** Thank you. I might have said that was my final question, but I have one final comment that I would like to challenge you both with. It has been said over the years that there is nothing artificial



## HOUSE OF COMMONS

and nothing intelligent about AI. Do you agree that that statement is still true, and do you see a time within the foreseeable future where neither of those statements is true and that we are approaching a technology that potentially has sentience?

**Matt Clifford:** A nice easy one. I will say, apologies, Chair, that I need to go to fulfil my final diplomatic duty and get on the Eurostar and do the handover to the French negotiating team very imminently.

I suppose, in my final three days in public life, to give a politician's answer, I sort of think that in a way it does not matter, in the sense that sentience is one of those things that is of huge interest to science-fiction writers, but, regarding the harms, outcomes and behaviours that exist, to give a slightly facetious example, if you are killed by a killer robot, you do not really care whether it is sentient or not. If you are discriminated against in a job application process, you do not really care whether the machine is sentient or not.

My view as someone who has followed the field for a long time is that we have no idea at all how to make machines conscious or sentient, but we do know how to make them intelligent. If we take the broadly agreed definition of intelligence as the ability to achieve goals in an uncertain environment, absolutely, we have systems that are starting to be able to do that. They are limited by context. They are limited by complexity. Those limitations are falling away at an extraordinary pace.

I do not think that means that we will have artificial general intelligence, human-level intelligence, across all domains next year or the year after, but there are very intelligent people operating at the forefront of these domains who think it is totally possible we will have that within the decade, and that will be a shock to our entire way of thinking about life. Whether or not that is coupled with sentience, which I think is extremely unlikely, we need to prepare for it, and the summit was a small step in that direction. As a final thing, the interesting thing about this technology is that, even if it falls far short of that, it still can be absolutely transformational.

Q753 **Stephen Metcalfe:** Absolutely. Thank you. Emran, do you want to add anything to that?

**Emran Mian:** All I can add to that is perhaps something even more frustrating. There is an active debate in philosophy about what constitutes sentience for human beings, never mind what constitutes sentience for machines. The key point is what Matt said: ultimately, it does not matter whether they are sentient or not. The thing we need to look for is protecting our agency, protecting our autonomy, protecting our interests and ensuring that we minimise harms.

Q754 **Chair:** I have just one point of clarification. In answering Stephen's questions, Mr Mian, you mentioned the advisory board of the Centre for Data Ethics and Innovation. The Committee's understanding was that



## HOUSE OF COMMONS

that advisory board has been abolished. Has it not?

**Emran Mian:** Yes.

Q755 **Chair:** Right. Is there a reason?

**Emran Mian:** I think it was simply to do with the fact that members of the advisory board had reached the end of their terms. The key thing for us, as I said in response to Mr Metcalfe, was to continue to ensure that we are getting expertise across the broad range of AI issues and related ethics issues. We are confident that we are doing that, but that is something that we will continue to look to find ways to do.

**Chair:** When people's appointments come to an end, it is normal to replace them, but I will reserve that, conscious of Mr Clifford's time and conscious of the fact that you have both been extremely generous in coming to give evidence today. I am very conscious that you are an official at the Department and you have to explain current policy. We look forward to having the Secretary of State, who has been very generous to commit to come to give evidence to the Committee. In terms of future policy, of course, this is a matter for her. We are very grateful to you, Mr Mian, for coming from the Department, and Mr Clifford for coming in your final days—final hours, indeed—of service to the public, and we thank you for that and thank you for appearing today.