HOUSE OF LORDS

# Communications and Digital Committee

## Corrected oral evidence: Large language models

Tuesday 17 October 2023

3.30 pm

Watch the meeting

Members present: Baroness Stowell of Beeston (The Chair); Baroness Featherstone; Lord Foster of Bath; Baroness Fraser of Craigmaddie; Lord Griffiths of Burry Port; Lord Hall of Birkenhead; Baroness Harding of Winscombe; Baroness Healy of Primrose Hill; Lord Kamall; Lord Lipsey; Lord Young of Norwood Green.

Evidence Session No. 5 Heard in Public Questions 37 - 45

### Witnesses

I: Dr Florian Ostmann, Head of AI Governance and Regulatory Innovation, Alan Turing Institute; Michael Birtwistle, Associate Director (Law & Policy), Ada Lovelace Institute; Katherine Holden, Head of Data Analytics, AI and Digital ID, techUK.

# Examination of witnesses

Dr Florian Ostmann, Michael Birtwistle and Katherine Holden.

Q37 **The Chair:** I am very pleased to welcome our second panel. Via this panel, we will look in detail at the Government's White Paper on AI, which was published earlier this year. I will start by asking our three witnesses to introduce themselves. It is very nice to see you again, Mr Birtwistle.

***Michael Birtwistle:*** Good afternoon. Thank you so much for having us. I am the associate director for law and policy at the Ada Lovelace Institute. Ada is an independent research institute with a mission to make AI and data work for people and society.

***Katherine Holden:*** Hello everyone and good afternoon. I lead the work of techUK on a few technologies, data analytics, digital identity and artificial intelligence, so I am having a very busy time. For those of you who have not come across us, techUK is the largest technology trade association in the UK. We have around 1,000 members from the FTSE 100 to a number of start-ups and scale-ups; around two-thirds of our members are SMEs.

***Dr Florian Ostmann:*** Good afternoon. I am the head of AI governance and regulatory innovation in the public policy programme at the Alan Turing institute. That programme focuses on the implications of AI for government through our research and advisory work. I should briefly mention that I also lead Turing's work on the AI Standards Hub, an initiative that is coming out of the national AI strategy; as mentioned earlier, it is a partnership between the Turing institute, the British Standards Institution and the National Physical Laboratory. We might come back to some of the objectives that we are trying to achieve with that initiative later.

**The Chair:** It is very good of all three of you to be here. I should just say that I sat on a panel with Mr Birtwistle recently. However, it was not the panel that raised the point, and he was not the person who did so either. Just in case mention is made of the fact that we were on a panel, let me say that it was a different one. Anyway, we will crack on.

Q38 **Lord Kamall:** Before we start, I should declare an interest. I did some work in co-operation with techUK; I was on the Brexit advisory panel for techUK at the time.

Some of you will have heard the previous session. Clearly, the issue of how to regulate this area came up. Do you think that individual sector regulators address the opportunities and risks of generative AI in large language models better than, say, a single statutory, cross-sector AI regulator? I know that techUK has said that it is happy with the context-specific way, while the Ada Lovelace Institute thinks that this is more challenging. Obviously, you have looked at this internationally. I am not sure what the Alan Turing institute thinks, so perhaps I could start with you, Dr Ostmann, before I turn to the others for their views on this.

Let me also say that generally there is a concern in wider politics—I have heard this from both the main parties—about who regulates the regulators. So if we are going for a sort of super regulator, as it were, who would that be accountable to?

***Dr Florian Ostmann:*** In responding to your question, one thing I would highlight at the beginning is that I interpret the White Paper this way: I think of it as a contrast not between sector-specific regulators and cross-sectoral regulation but between existing regulators and setting up something new.

**Lord Kamall:** That is a fair point.

***Dr Florian Ostmann:*** It is an important distinction, because, if you look at the existing regulators in the UK, we do have cross-sectoral bodies such as the CMA and the ICO. Those bodies have a really important role to play and have done really important work already.

The Turing institute has responded in writing. I share the view that we broadly welcomed the Government's approach, as set out in the White Paper. The premises that led to that approach still hold for the age of LLMs. For one thing, there is existing regulation that applies to AI, including LLMs, even though it might not be regulation that was written for AI.

Secondly, when it comes to thinking about the need for new regulatory measures, existing regulators might be best placed to design those, because risks are often very context-specific. Think about LLMs: the risks of using an LLM to write poetry are different from the risks of using an LLM to make medical treatment recommendations. It is about context-specific risks and the fact that there is existing regulation. The CMA report on foundation models brings that out very strongly; it is an important point.

What raises big questions—this is where the central functions have heightened importance in the age of LLMs—is the fact that a lot of existing regulation is technology-neutral in its approach. Think about the CMA's role in regulating fake customer reviews or the FCA's rules: they are usually defined in terms of outcomes for consumers rather than specifying how technologies are used or developed. AI in general but LLMs in particular have the potential to challenge that approach, specifically when it comes to regulators' ability to understand how technology works and having sufficient transparency around development processes.

The technology-neutral aspect creates challenges. That is where the central functions described in the White Paper play the really important role of supporting regulators to deal with that. There might be a need for cross-cutting measures that enable regulators to better understand, for example, the capabilities of LLMs and the measures for increased transparency that all regulators can draw on.

**Lord Kamall:** Thank you for that, and thank you for your clarification earlier, too. May I turn to you, Mrs Holden?

***Katherine Holden:*** Yes, absolutely. Broadly, techUK is supportive of the UK's approach, as set out in the AI White Paper. We think it is proportionate, risk-based and outcomes-focused, which we obviously think is incredibly important. As a result, therefore, we think that regulation should be determined at the sector regulatory level.

There are a few reasons for that. We think that regulators have a good understanding of the regulation that already exists within their own remits and how the technology is being applied specifically in their area. They are obviously well placed to determine where gaps might exist. We recognise that some regulators are particularly advanced in this area—our members work in particular with the FCA, the CMA, the ICO and Ofcom—and that there are, I think, more than 95 existing regulators in the UK. At the moment, there are varying degrees of expertise and capacity when it comes to AI.

We would like to stress that, although we think that the UK's approach is the right one, for it to work and be successful, regulators need to have the capacity and capability to fulfil their new roles and requirements. Ultimately, this requires additional funding, which we hope to see secured in the Autumn Budget in November.

The central support function will also be key to ensuring a coherent, consistent approach to AI governance across different regulators. I agree with Florian's point: we will definitely see that we have to look at this issue from a horizontal angle but also on a vertical level. Areas where we may need to take more of a horizontal approach include guidance on how different regulators should adopt the AI principles and the implementation of those principles into practice. We will need some concrete guidelines there for this approach to work. How we define high-risk AI applications is equally key. We cannot have different regulators defining high risk in different ways. That would cause considerable confusion.

In summary, we very much agree that looking at these things from an individual regulator perspective, bringing in that central co-ordinating function, will be a successful mechanism if it is implemented successfully.

On your point about what that looks like in practice, we would say—that is, our members believe—that this central function should be meaningfully independent of the political cycle, the parliamentary process and ministerial briefs and should be a body that reports almost directly to Parliament.

**Lord Kamall:** Thank you. To understand the spectrum of views on this, I turn to Mr Birtwistle.

***Michael Birtwistle:*** Thank you. The short answer is that there are a lot of advantages to regulating AI as close to the point of use as possible, because the impacts are highly context-specific, as you have heard, but doing this is very challenging from a design and implementation perspective, because it has some pretty high conditions for success. Those conditions look like a regulatory ecosystem that is well resourced

and empowered to address the challenges of new technology, and has coverage across all areas that you would expect AI to be impactful, and this sort of co-ordination question between the regulators.

The research we published earlier in the summer strongly indicates that those conditions are not currently in place in the UK. I am happy to go into more detail on what those gaps are, if helpful, but on top of this, as was alluded to in the last session, we have the challenge of foundation models. They create novel governance challenges, because they hold the prospect of a core part of our digital economy being built on top of a platform of highly capable general-purpose systems, probably developed by a relatively small number of companies based outside the UK. There is also evidence of significant use of foundation models increasingly in public services as well: we have just published an evidence review that covers some of that early use across the UK public sector.

Many of the impacts and risks of those general systems will flow from decisions that are made upstream by developers. So, coming back to your question, we have not come to any firm conclusions on how or where you house this, but we think about it in terms of regulatory capability. We need regulatory capability or an accountability framework that looks holistically at the top of the AI life-cycle—developers and also potentially hosts of foundation models. Outside of safety case-based regulators, the CMA and a couple of others, there are very few regulators that have the scope and powers to look up that value chain at developers, because of the context-specific approach. The White Paper targets primarily the use of AI rather than its development, so there is very little protection currently in managing risks that occur upstream.

Q39 **Lord Kamall:** Evidence from the Ada Lovelace Institute says that the Government are setting themselves a harder regulatory challenge than other international legislators. Which international legislators did you have in mind, and which ones do you suggest we should be looking at?

***Michael Birtwistle:*** I suppose the comparison is with the EU, where the bulk of our analysis outside the UK regulatory proposals has been.

**Lord Kamall:** Okay, fine, I shall stop there for now.

**The Chair:** Thank you. We will come on in some of the next questions to the central support function and the ability of regulators later.

Q40 **Lord Young of Norwood Green:** I must admit that when I listened to you say there were 93 or 95 regulators and they have a coherent strategy, I was impressed. Florian, you mentioned the ethical implications of emerging technologies. Will you expand on that? Michael, I noticed that you led the Government's Centre for Data Ethics and Innovation. I would like to hear more about that concept of the ethical implications of emerging technologies. What exactly do you mean by that?

***Dr Florian Ostmann:*** I know there are different frameworks to think about these things in a different way. We heard earlier about the current focus on safety of AI systems, and there is a long-standing debate about

how to distinguish between safety issues and ethical issues. At the Turing, we think of ethical AI as a very broad category that includes thinking about issues that can arise in the design and development of systems but also in the way systems are used. So we might have a system that is designed and developed to a point where it works reliably if it is employed in the way it is intended to be deployed, but it could be misused for malicious purposes. LLMs are a great example: they can be misused for fraud purposes, for example. Also, we think more broadly about downstream societal impacts, which would include thinking about the impact on employment, for example, or the impact on democratic processes that might result from misinformation spurred by LLMs. Michael may have things to add.

***Michael Birtwistle:*** Sure. To clarify, I did work at the Centre for Data Ethics and Innovation, but I do not speak for it in any capacity today. I very much agree with that analysis of risks. When we talk about ethics in this context, we are talking about a holistic view of all the risks that flow from these models. If we are talking about foundation models or large language models in particular, I agree that reliability, discrimination, misuse, systemic risks, as Florian described, are absolutely some of the major concerns around foundation models.

There are also features about the way they operate within the value chain that amplify those potential harms. So there is their generality and ability to achieve a range of tasks, which means that they dealt built upon and create a sort of upstream risk that can flow into many places and sectors. There is the fact that they have fast and sometimes unpredictable jumps in capability. There is the fact that they can be fine-tuned or retrained, which disrupts how developers up the chain can see the risks that might occur lower down. There is also their wide-scale accessibility, which puts powerful AI capabilities in the hands of a much larger number of people. When we think about ethics, we are thinking about all those risks in the round and making sure that they are properly managed.

**Lord Young of Norwood Green:** Finally, Katherine, do you think all those regulators take that same approach?

***Katherine Holden:*** Probably not at this stage, no. With 95 regulators, there are obviously some that do not have the bandwidth and capacity or just have not even thought about AI and what it means within their own regulatory remits. As I said, I think that we have, within the UK, a really strong and robust regulatory regime; some of the regulators are absolutely world leading in this area and are really leading from the front on a global level. Saying that, there is a lot of work to do to get some of those other ones up to speed. Not everyone may need the same level of expertise and capacity on this particular issue, but there is probably a baseline understanding that everyone, in government and outside government, needs when it comes to AI and what it means for them.

**Lord Griffiths of Burry Port:** We, as a committee, put a report together called *Regulating the Regulators*, and it may be good, in view of what has been happening in this fast-developing field, to measure some

of the conclusions we reached there against some of the questions that are emerging here. We did have that report and it was rather well received.

**The Chair:** Noted. Thank you very much, Lord Griffiths.

Q41 **Baroness Featherstone:** The White Paper's holy grail, as you said, Katherine, is this central support function, so what will that function have to do to make sure that regulators are properly and adequately supported to deliver on the Government's objectives? Can you identify any weaknesses or omissions? I have a list of what they are meant to do, but where are the real problems?

***Katherine Holden:*** Goodness. Where to begin? Alongside building the individual regulators' capacity and expertise—we cannot shy away from that; it also needs to happen at the individual regulator level—the central function really is key to the success of the White Paper approach. As you said, if you read the White Paper, from reading it many times I think it is at point 3.1, a considerable list of roles and expectations have been given and placed on the central function. It is, first and foremost, critical that that central function is well resourced, with a diverse pool of experts.

In our response, you may have seen that we spoke about the need for skills and experience in areas such as auditing, risk management, privacy, security, law, AI ethics, stakeholder engagement, standards and understanding in civil society issues—and, of course, we cannot forget practical experience of implementing AI solutions. We recognise that there is obviously a resourcing issue when it comes to AI, so we will have to think of innovative ways of increasing capacity. That may be secondments, placement schemes, advisory committees or rotating seats; we would be open to a number of those ideas. It is important that we look at the very long list of roles and responsibilities, set the central function and determine what the most critical tasks are that it needs to focus on first.

For me, and I think for our members, that is probably about ensuring a coherent and compatible approach to AI governance across the different regulators. So the most critical tasks are implementing the AI principles, as I said before, and producing a common framework for determining high-risk AI. That is essential to ensure that we do not have, or that we at least mitigate the risk of, regulatory overlap, duplication or issues that may fall through the gaps.

Another key piece of work that may be worth doing to reduce the burden of workload for the central function is to conduct a review of the tasks that individual regulators may already be doing themselves. For example, nearly every regulator will have some kind of horizon scanning capabilities: how do we build on that rather than duplicating efforts? Equally, and this is my final point, it would be useful to determine how this new function would work alongside existing bodies: as we heard in the previous session, the Office for AI, the Centre for Data Ethics and Innovation and the Frontier AI Taskforce. We can do a tremendous

amount as long as we know who is doing what and how they work together in harmony.

**Baroness Featherstone:** I can see that you have thought about this.

***Katherine Holden:*** I have thought about it a lot, yes.

**Baroness Featherstone:** Do either of you two want to add to that list of functions? I have many supplementaries about them.

***Dr Florian Ostmann:*** First, I very much agree with both my fellow panellists that the existing approach is demanding in its conditions for success; I did not mean to downplay that in my earlier answer. I agree with Katherine. The way I would describe it is that, at the high level, there are co-ordination and capacity building needs. When it comes to co-ordination, it is about such things as making sure that regulators interpret principles in a similar and coherent way but also that there is appropriate risk ownership in mapping out regulatory risks and then assessing whether those risks are covered by existing remits; that there are not duplicative efforts to regulate the same risk by more than one regulator; and that no risks remain unaddressed because one body assumes that it is being taken care of by another body.

**Baroness Featherstone:** The gaps.

***Dr Florian Ostmann:*** Yes, so that things do not fall through the gaps. On the capability building side, in developing skills and expertise, financial resourcing is key. We consulted with regulators and facilitators at a round table as part of the White Paper consultation, and that came out very strongly.

Thinking about regulatory powers is also really important. It is easy to forget that the powers of individual regulators vary significantly. This includes rule-making: some regulators have their own rule-making powers while others do not. It also includes information-gathering powers. Some regulators, for example DRCF members, are quite well positioned, but other regulators have limited information-gathering powers. Then, of course, there are enforcement powers, which can be quite limited. Some regulators have to rely on the courts for enforcement action, which of course limits their ability to take swift action.

**Baroness Featherstone:** Michael, do you have anything to add?

***Michael Birtwistle:*** Yes. My fellow panellists have given very comprehensive accounts, and I agree with almost everything that they said. The connection back into policy-making will be critical. Governance tends to break with transformative technology when it does not contain the concepts to adapt. A good example of this is self-driving cars, where we license human drivers. Obviously, that breaks as a model when you introduce self-driving cars. The Government are taking action on that through the connected and automated mobility strategy. There will be other areas; IP law is another one that I know this committee has taken evidence on.

Understanding those gaps will be real; it is not going to be an action that regulators can do anything about, because it will be outside their existing

scopes and powers. Having that connection back into government policy-making as a sort of priority line, ideally as a formalised process, will be really important.

The integration of those affected by technology, such as civil society, into policy-making on AI will be important. The organisations that represent those groups will often be at the forefront of understanding how AI is affecting people. Being able to feed those thoughts into the central AI risk function, for example, will be crucial.

I really associate myself with the comments that were made on resourcing. I was hoping to say a bit about the scale that we are thinking of. AI, in particular foundation models, is rapidly becoming a core part of our digital infrastructure. There are many emerging similarities between its role there and how we think of things such as search engines, social media and major online retailers as digital platforms. When we think about how we should govern technologies or industries that are consequential—this term has been used in previous evidence sessions—and that form part of our critical national infrastructure, or where safety and trust are critical or it is really important for people or a market to have access to a particular product or service, we tend to ensure that there are well-equipped institutions backed by effective regulation, often with a safety case-based approach to that.

The objectives of those regimes tend to be not just the management of extreme risk but the creation of whole-economy or whole-society trust in those technologies or industries. We trust that planes are not going to fall out of the sky, because we equip and empower the Civil Aviation Authority properly. The same is true for medicines and medical devices, life sciences, food, rail, automotive, finance and cybersecurity. Those regulators are funded in the region of tens of millions of pounds per year. Regardless of whether AI gets governed on a distributed or centralised basis, we think that is the right frame of reference for planning the regulatory response.

**Baroness Featherstone:** Okay. Katherine, you said that the central support function should be independent of government and Ministers—basically, that it should just be independent. Would there be a trade-off if it were in the department? I know what you think. What do the other two think?

***Katherine Holden:*** Do you want me to add something on the trade-off part first?

**Baroness Featherstone:** Yes.

***Katherine Holden:*** There are always trade-offs with all decisions. I guess that making it independent would take more resource and parliamentary time to put it on to a statutory footing. I guess our members think it is really important for it to be independent of the parliamentary process so that it can focus on the task at hand here and so that it is not flexible depending on ministerial brief—particularly ahead of another general election, where anything could happen. We would want to make sure that this body, this central function, was able to and

had the mandate to carry on with the task at hand, as I said. Where our members are looking for certainty and stability in this area, that would provide business confidence and clarity so that they can invest in the UK.

**Baroness Featherstone:** Do you two agree or disagree?

***Michael Birtwistle:*** I cannot see many disadvantages to making it independent.

**Baroness Featherstone:** You have all referred to the need for resource. Can you put a figure on that?

***Katherine Holden:*** Oh, goodness. Hearing what Michael said, that seems to be the ballpark in the areas that we have been talking about.

**Baroness Featherstone:** So tens of millions. The White Paper outlines a series of related deliverables. There is a kind of shopping list of things that are given a timeframe. One example is: "issue the cross-sectoral principles to regulators, together with the initial guidance covering implementation, within six months". There are times against all these issues. Do any of you know when that time started? Is it now or before?

***Michael Birtwistle:*** To be generous to the officials working on this, the advent of foundation models has slightly disrupted that. I expect a fresh road map; I think it is one of the next projects to be announced. It is something that we expect in the near future.

**Baroness Featherstone:** You are all nodding.

***Katherine Holden:*** Agreed. We had an event with the Secretary of State this morning. She mentioned that further details will follow by the end of the year, so that is the timeframe that we are working to.

**Baroness Featherstone:** It will be delivered shortly.

***Dr Florian Ostmann:*** I want briefly to add to your previous question on independence. I thought it might be worth sharing this. We did a lot of research and wrote a report on common capacity for the regulation of AI. It was largely based on interviews with the regulators themselves. The independence point came out there very strongly, so it is a view not just from industry but from the regulators themselves.

**Baroness Featherstone:** Thank you. I think that we got that message.

Q42 **Lord Hall of Birkenhead:** In terms of this body and the complexity of the long list of things that you are asking it to do—rightly—I wonder what importance you would attach to horizon scanning and being the place where you would go to say, "This is something you need to be thinking hard about". All the things that you have described represent a very weighty brief. You would hope—at least, I would hope—that horizon scanning would be up there as something about which you would say, "This is a real set of expertise". I do not know whether you think that is important. What weight would you attach to that?

***Katherine Holden:*** Horizon scanning is fundamental. It is incredibly important, particularly for anticipating some of the emerging and most significant risks, as well as opportunities, in the field. From techUK's

perspective, we are not too prescriptive about exactly where that sits as long as it is happening and doing so effectively. We have a number of other bodies. Obviously, we have the Centre for Data Ethics and Innovation as well as the Office for AI. The Frontier AI Taskforce may have a role to play when it comes to horizon scanning related to generative AI, for example. We do not mind exactly where it sits so long as it is clear who is doing it and how they are disseminating that knowledge to others in the field.

**Lord Hall of Birkenhead:** The trouble is that, if it is happening in lots of different places, that can be a "thousand flowers bloom" situation or it does not get through to the people who need to be included by it.

***Katherine Holden:*** Exactly. So we need clear leadership from government on where that responsibility sits, and there need to be clear communicators to other parts of government, as well as to outside of government, on what the real opportunities and threats are.

**Baroness Harding of Winscombe:** Can I ask one small follow-up question? There is a link. Mr Birtwistle, you listed a number of different sector regulators as examples, such as the CAA or the medical ones, yet all three of you stated quite clearly the value in having AI regulation bedded into individual sector regulators rather than creating an AI regulator. Are there any parallel regulators dealing with this? The one that was in my mind is the National Cyber Security Centre, which is not trying to do detailed regulation in every sector but is crossing every sector. Are there other existing organisations that meet this central function but do not overarch the regulation of everything?

***Michael Birtwistle:*** The NCSC is certainly an interesting model. As I said earlier, our concern is about addressing gaps in the model proposed in the White Paper. Speaking of that, we do not have regulatory functionality in looking at the value chain. There are also gaps in our sectoral patchwork; for example, we do not have a recruitment or employment regulator, which would apply the principles in that domain and look at some high-risk use cases.

The White Paper is also not clear on how the principles would be applied or enforced in the public sector, some of which is directly regulated and some of which—for example, benefits or tax administration in central government—does not have a regulator. So there is an argument for making sure that those functions are delivered in order that the systems are credible. We have not taken a position on whether that should sit within a new body, should somehow be integrated into the central functions or should be distributed among existing regulators.

**Baroness Harding of Winscombe:** That is helpful. Would anybody like to add anything on that before we move on?

***Katherine Holden:*** Not particularly. I just want to say that I am not aware of any similar organisation in the AI field. For me, that is why this central function is so important: because nothing like that currently exists.

Q43 **Baroness Harding of Winscombe:** If we can flip from looking at the central function to looking at what the individual sectoral regulators need to do in order to rise to the challenge of regulating AI in their sectors, what options do you think there are for improving the ability of those regulators to oversee their use of LLMs?

***Dr Florian Ostmann:*** I will start with one small point. It is a really important point that may easily be overlooked and it came out in our consultation with regulators for the White Paper: the ability to share information between regulatory bodies, which is perceived as very difficult at the moment. Obviously, there is a lot to be gained from regulators sharing the results of their investigations and sharing that information with an adjacent regulator whose remit touches on the same use. That could probably be addressed relatively easily through legal action—that is, legislative action—and could make a big difference.

On LLMs specifically, as I mentioned earlier, the challenges that are heightened for LLMs compared with more traditional uses of AI related to transparency and regulators' ability to gain access to documentation about systems. The fact is that a lot of LLM systems rely on API access, which creates transparency challenges. There would be a lot of inefficiency if we ended up with a scenario where every regulator had to knock on companies' doors to do its own information gathering, potentially using its own framework for what information it expects to be shared. Some sort of agreed framework for information gathering, information sharing and reporting across regulators could make a big difference to the resource burden on individual bodies.

Recently, in the context of generative AI more broadly and frontier AI, on specific measures that might be useful, there was a proposal for the registration of models so that, if a company develops an LLM, it has to register that with a central database; that might be administered by government, for example. Then there is the implementation of certain forms of reporting on a model's characteristics and capabilities, as well as reporting on incidents, so there is a lot that could be done on shared frameworks and making information available.

There is a lot to be gained by looking at the role of standards and assurance ecosystems. The Centre for Data Ethics and Innovation has done important work on AI assurance, the role of third-party assurance providers and the role of certification for AI systems. A well-working system of standards and certification would take a lot of the burden off regulators, because they could then rely on the signalling that is implicit in certification as well as speculation on the use of standards.

The challenge is that the space of AI standards is already very complex. That is partly why we set up the AI Standards Hub. One of the challenges that we are trying to address with the standards—this goes back to the capability/capacity-building point that we discussed earlier—is enabling regulators to understand what AI standards may relate to their remits or regulatory objectives. It is also about developing a sense of the adequacy of different standards in order to be able to assess whether a certification scheme, for example, sends a sufficient signal.

***Katherine Holden:*** I would like to put on the record the fact that I agree completely with Florian's points about the importance of things such as registries, auditing, standards and assurance. It is all really important; there is a healthy mix of ideas that we need to progress further.

Broadly, having consulted our members, I would say that many of the regulators currently have sufficient powers to oversee LLMs although, as we have said, the capacity to address some of these issues is perhaps another thing. I know that the ICO highlighted something in an event over the summer. This is a direct quote: "Laws already exist to protect people's rights, including privacy, and apply to generative AI as an emerging technology". We have seen this recently with the ICO's preliminary enforcement notice against Snap and its gen AI chatbot. A lot of regulators—particularly the likes of the CMA, with its work on competition, and the ICO—are using their powers to try to address some of the risks associated with this technology.

Of course, we cannot be complacent in this space. We need to monitor the technology as it continues to evolve. We would recommend, first and foremost, conducting an in-depth review of the risks that are related to this technology to determine exactly what is new here and how any novel risks that are presented are being addressed by existing regulation. We would say that the Frontier AI Taskforce could play a significant role in conducting this research, although the task force will also need additional resourcing in working with industry, civil society and others. Before we leap into where regulators perhaps need additional powers, which they may, and have a knee-jerk response, some in-depth research, consideration and debate need to happen in that space.

**Baroness Harding of Winscombe:** Just so I understand, are you suggesting that that review needs to be done centrally, rather than individual regulators doing their own reviews?

***Katherine Holden:*** That is a good point. Yes, it probably needs to happen centrally.

***Michael Birtwistle:*** I very much agree with Florian's points about powers. What Katherine said is true of perhaps five or six regulators, in particular digital regulators, that are in a better position to cover those bases through seeking information and so on. However, even if you take an example like Ofcom, that is at least half a dozen regulators within a single body. The remit, powers and scope that it has within each of those domains are very specific. The way it can use those powers is specific to each of those domains, so there is a huge variance even within regulators. Outside the five or six digital regulators, there is a very long tail of dozens of other regulators that do not have access to those powers. Some of that work of assessing those risks is being picked up by the central AI risk function, which is being set up in the Office for AI, the AI team in DSIT at the moment.

I very much agree that assurance and audit will be very important, not just for the general ecosystem but for regulators specifically. Some good work came out of the DRCF on the importance of audit, and there was

some really good work from CDI on how the UK could develop an assurance ecosystem.

Another thing that could be done to equip regulators better concerns the way in which they are asked to comply with the principles. The duty to regard proposed in the White Paper is very minimalistic, and some regulators will tell you in private that they are sceptical of the impact it would have on their decision-making to comply with the principles. You could do a lot more to beef that up—for example, by creating reporting obligations to Parliament and some actual objectives or measurements for how those principles are seen as having been lived.

Finally, on your earlier question about other regulatory models, we will shortly release a report—I hope to share an advance copy of it with the committee—looking at the Food and Drug Administration in the US as a potential model, and at the pre-release requirements to market and so on that are used there for comparably complex technologies with unknown onward impacts. There is a lot that we can learn from those approaches, so I am happy to share that with the committee in due course.

**Baroness Harding of Winscombe:** That would be very interesting to see. You alluded to this in earlier discussions, but could be clearer on whether you think existing regulators will be able to address the upstream challenge of regulating the foundation models? Do they have the powers to do that? If they do not, what should be thinking about to address that?

***Dr Florian Ostmann:*** It depends on the regulator. If one looks at the landscape of regulators as a whole, starting with sector-specific regulators such as the Financial Conduct Authority, it usually has a perimeter that is defined by certain entities falling within the perimeter and outside the perimeter. In many cases, not in all, the perimeter will include companies that tend to use systems rather than develop them themselves. The customer service chatbot used by financial services companies might in many cases rely on something like ChatGPT, which is developed elsewhere.

So the powers in many example cases will be limited to a user that is not a developer. That does not mean, of course, that the responsibility can be outsourced. It is an important point of regulatory policy that the company that is using the system is responsible for making sure that it meets regulatory expectations. However, there is a big question about the ability of those companies to carry out the necessary due diligence if they are not supported. It might, for example, be very difficult to get access to information about how the system was developed or how the data was used if there is no regulatory or legislative support to encourage developers to make that information available. We see that in the current draft of the EU AI Act, which has a proposal that is aimed at requiring developers to make that information available so that downstream users can carry out that task.

For cross-sectoral regulators, it is a different picture. The CMA, depending on the concern, will be able to investigate or issue measures

that are addressed to developers. The important thing there is that the powers and remit of the CMA or the ICO as cross-sectoral regulators focus on particular issues, such as data protection in the case of the ICO.

It is a patchwork, but what emerges, if you think about it that way, is that there are potential gaps. Think about regulators that have powers to access only users, and the cross-cutting regulators whose remit is limited to particular cross-cutting issues at the development level that might be unaddressed and where there might be value in putting in place, at the very least, transparency or reporting requirements to support sufficient flow of information.

**Baroness Harding of Winscombe:** That is extremely clear and helpful. Thank you. Does anybody else want to add anything?

***Katherine Holden:*** Again, there is a role for the central function in being able to provide a holistic perspective on the AI supply chain, which as we know can be very complex in a lot of cases. There is a role for them in ensuring a co-ordinated approach across the different regulators that may be operating across that supply chain to clarify where the issue is and therefore which regulator it falls within.

We have spoken a lot already today about the importance of AI assurance. In the AI White Paper, the UK is seen as really wanting to lead in this area, so we need to do more work to promote AI assurance across the supply chain in order to be able to identify as quickly as possible where the problem is and to address that. We need to look at that area in more detail.

***Michael Birtwistle:*** I have talked a bit about this already, but it does feel that the big gap currently in the Government's approach is having the regulatory capability to look upstream. That is understandable, given how quickly things are moving, but it needs to be addressed. The MHRA, for example, is a safety case-based regulator that can look all the way up the chain. Most regulators cannot.

Foundation models have a downstream proliferation of risk, and regulators will be constrained by law to dealing with those only within their domain. If an upstream foundation model has a medical device product based off it, the MHRA will only be able to talk to the developer and say, "This needs to be non-discriminatory when it's going to be used for medical devices", but it will not be able to effect those changes across all other users of that foundation model.

Looking at where the solutions lie, around the AI summit we will see lots of talk about voluntary commitments and safety policies such as the transparency requirements, the information sharing, the evaluation and the benchmarking of these models. Those need to form the basis of a legal accountability framework in due course.

**The Chair:** Thank you. Lord Foster, you have a supplementary, or are you just scratching your nose?

**Lord Foster of Bath:** I am pondering, but no.

Q44 **Lord Griffiths of Burry Port:** I am just amazed at the speed at which all these things come to pass. I have been given a question to ask that says "as LLM technology proliferates". I have read enough to know that the proliferation may take us into places that, at this minute, we cannot even anticipate or know much about, and it becomes more widely available. That is the thing about this technology: its findings and its usefulness will be available to people who know how to use it and to people who do not know how to use it, and to people who have an interest in using is in ways that all our regulation would seek to obviate. It all feels worrisome.

Here we are, sitting as a body from the House of Lords and the British Parliament. It takes us however many months just to have this report; other things are going on, of course. We will then have a debate in Parliament and produce a law. By the time the law is on the statute book the debate will no longer be there, it will be somewhere quite different, and we cannot always anticipate that well.

I have always been concerned about nuclear physics, and people will have heard me say this before. I knew a very famous man who was in the Manhattan Project, and he was fired up by his science. His science was driving him forward; others were driven by the commercial things that they were going to have at their disposal. But he saw what the military and the politicians wanted to do with his science, and he stood out and founded the Pugwash conferences, which went on reviewing nuclear physics thereafter.

I am anxious to know whether, within the field of technology where people are at the cutting edge, there are voices with sufficient room to raise the questions that we are worried about, rather than their being driven by their research and by the commercial factors that inevitably operate in these fields.

How would you help? There are clever people around this table, and there are people for whom this has been just a seminar; we are recipients, blotting paper. How would you reassure people like me that that voice from within the field, where people who know are at the cutting edge, is having a chance to be heard?

***Michael Birtwistle:*** Part of the solution to that challenge is well-equipped institutions. I have already talked at length about ensuring that we have well-empowered and well-resourced institutions, but ensuring that we have them and that they are on a legislative footing and are well equipped ensures their capacity both to react at speed and to hire the expertise that will be needed to evaluate and understand these models properly.

I would also say something about urgency. It is much harder to regulate something once it is already in mass usage because of the impacts on society and an economy that has developed dependencies on those technologies. In the past year, we have seen LLMs go from impressive tech demos to being integrated into our search and productivity software; there are now celebrity-fronted personal assistants. The very nature of foundation models is that they act as a general foundation for a

huge number of downstream products, so we can expect those dependencies to develop.

To address your point about the speed and ability of Parliament to react to this, even if the Government were to publish a Bill tomorrow, it would likely be a matter of months, if not years, before that would be in place. We can expect the integration of foundation models into the digital economy to continue in the meantime, so we need to recognise that there will be an intervening period where systems will continue to be deployed without formal regulation. It will be critical for Governments at the summit to secure meaningful, detailed commitments from industry.

Again, on having voices in the right places, I go back to the points I made about civil society and effective people's representation in policy-making processes. It is hard to do that, but it is important to do it. There are a number of opportunities in the Government's policy-making process and in the establishment of these new institutions to build that voice in.

***Katherine Holden:*** I agree with everything that Michael said. There are three areas here that will, I hope, address your question. We need to make sure that there are clear lines of escalation for people working on the technologies within their organisation, within the regulators and as part of the central function, so that when they flag risks or areas that they feel their technology is navigating and moving into and with which they are uncomfortable, those things can be escalated and highlighted as soon as possible. That will be particularly important for things like some of the potential emerging cybersecurity risks that could be associated with this technology.

As Michael said, particularly with our upcoming AI safety summit at Bletchley Park on 1 and 2 November, establishing global standards in this area is important—indeed, it is fundamental. It was described to me this morning like this: it is the same as our efforts on climate change in that, unless we all work on it together in a globally co-ordinated fashion, it will not matter in essence because someone else somewhere else in the world could be doing something else that is detrimental. Global standards and co-ordination in this area are really important.

Finally—I know that we have touched on this issue throughout this afternoon's discussion—this is why having an agile, pro-innovation, flexible approach to regulation in the form of the AI White Paper and how it is taken is the effective approach. It can flex and move with the times as the technology evolves. My concern with the EU's approach, for example, is that, when we saw the advent of generative AI, it suddenly meant that there was a lot of scrambling around and having to update the work there in order to make sure that it was in accordance with the technology as it was evolving. In the UK, we are future-proofing our approach in this area by taking the approach set out in the White Paper.

***Dr Florian Ostmann:*** I very much agree with all the points that have been made, especially on the importance of open and inclusive processes and thinking about all these things in the context of the AI summit, for example; the rest of the AI summit is upcoming. I think we are all also

involved in the AI Fringe as an effort to contribute to creating spaces for inclusive discussions on these matters.

I also agree with the point about international alignment and the importance of thinking about this at an international level. Some of the processes at the international level are quite fast-moving, such as the G7 Hiroshima process, which addresses LLMs specifically; it builds on the OECD's work but is very much focused on LLMs and related discussions.

I want to add one new point. Going back to what we have covered so far, which focused on the White Paper, and thinking about the work of regulators, when it comes to the wide availability of this technology for end-users, there is an important point to consider about the range of users, including users who might not occur in regulated entities. We have been thinking about regulatory bodies. They all have their regulatory entities, usually organisations, but one aspect of LLMs is that they are often widely available. Something like ChatGPT is available to an end-user sitting at home. There are a lot of open-source models that can be accessed widely and freely. So, to some extent, we then get to the limits of what regulators can do. There is a need to think about different forms of law to address other potential measures, especially when it comes to malicious use or misuse, that lie outside the scope of regulation.

Q45 **The Chair:** Thank you. Are there any new challenges that you see emerging for regulators as a result of the proliferation of LLMs that you have not already touched on? Do not worry if you feel that you have covered everything. I just want to give you a chance to highlight anything that we have not covered but you are keen to ensure we understand.

***Michael Birtwistle:*** One thing to keep an eye on is what you might call tool-assisted foundation models, where foundation models have access to external data sources or have text-to-action capability. The increased autonomy of those systems in being given tasks may alter the use cases and introduce new risks.

There is also an interesting interrelated set of questions about future technological capabilities in inferential biometrics, specifically systems that are designed to predict internal states such as human emotion. There are also the implications of the next generation of AI-driven personal assistants. The capacity of these systems to provoke or generate emotional responses from users, as well as the capacity for those users to develop relationships with the assistants, will have a really interesting set of implications in the medium to long term.

**The Chair:** That does sound quite scary. Mrs Holden, Dr Ostmann, is there anything you want to add? As I say, do not feel that you need to.

***Katherine Holden:*** There is one risk that I would highlight just because it came up quite a lot this morning at our pre-summit round table. We know that it is a risk and one that is very much live at the moment, but we need to address it further. It concerns the proliferation of deepfakes, misinformation, et cetera. This is an area where we are only just starting to see what the technology is capable of. Actually, it is an area where

further measures probably need to be put in place to make sure, particularly ahead of a general election, that we are able to address that risk, which is very live at the moment, effectively.

***Dr Florian Ostmann:*** One tool or measure that we see emerging in lots of jurisdictions—at the EU level, in other countries and in the G7 process—is the requirement to declare or make it visible when an output is generated by an LLM or another form of generative AI. Of course, that is particularly difficult for text; it is easier for images and video material. That is a really important tool to consider. There is a question here: where would that fit in the UK context? Who would take care of making it happen in the White Paper process? Again, perhaps that is something for the central function; a legislative measure might also be needed.

**The Chair:** Thank you, all three of you, very much for your testimonies today. They have been incredibly helpful. It has also fascinating for us that Parliament has suddenly found out how to work the air conditioning system just as we are going into colder climes. It has been quite cold in this room, but that has in no way deterred you or undermined the quality of your testimonies. We are very grateful.