HOUSE OF COMMONS

Culture, Media and Sport Committee

Science, Innovation and Technology Committee

Oral evidence: AI and Copyright, HC 695

Tuesday 4 February 2025

Ordered by the House of Commons to be published on 4 February 2025.

Watch the meeting

Culture, Media and Sport Committee Members present: Dame Caroline Dinenage (Chair); Mr James Frith; Natasha Irons; Liz Jarvis; Paul Waugh.

Science, Innovation and Technology Committee Members present: Chi Onwurah (Chair); Emily Darlington; Dr Allison Gardner; Jon Pearce; Steve Race; Dr Lauren Sullivan; Martin Wrigley.

Questions 1 - 53

Witnesses

I: Vinous Ali, Deputy Executive Director, the Startup Coalition; James Smith, Co-founder and Chief Executive, Human Native AI; and Sebastian Posth, Founder and CEO, Liccium B.V.

II: Sajeeda Merali, Chief Executive, Professional Publishers Association; Max Richter, composer, pianist and producer; and Matt Rogerson, Director of Global Public Policy and Platform Strategy, *Financial Times*.

Examination of witnesses

Witnesses: Vinous Ali, James Smith and Sebastian Posth.

[Chi Onwurah took the Chair]

Q1 **Chair:** I start by welcoming Dame Caroline Dinenage and members of the Culture, Media and Sport Committee to this joint sitting of the two Committees on AI and copyright. I also welcome our witnesses. We are very pleased to have the opportunity to speak to you today about artificial intelligence and copyright in the context of the Government's consultation. We also reached out to some of the larger developers on this issue. We invited Google and OpenAI to join us. They told us that as the Government consultation is still live, they were unable to take part in this public session, and we chose not to press them. However, I want to emphasise that I think having a public understanding of how these decisions are being made within tech companies as well as within Government is very important.

With that, let us go to our guests. I will ask you to introduce yourselves when the question is presented to you. The Government consultation on copyright and AI basically sets out that the preferred option is an opt-out approach, prerequisite on having the appropriate technology for opting out and for increased transparency of inputs and outputs. I will start with Vinous Ali. Please introduce yourself. As a representative of a large tech sector, you are probably still preparing your response to the Government consultation, but can you say a little bit about your preferred outcome?

Vinous Ali: Of course. First, thank you so much for having me here today. I am here on behalf of the Startup Coalition. We have about 4,000 start-ups within our ecosystem and we are here to represent the voice of start-ups, scale-ups and the investors that support them.

I will start by saying the Startup Coalition is very clear that our main endeavour is to make sure that the UK is the best place for start-ups to build their companies and to scale here in the UK rather than moving abroad. We have so many of the raw ingredients here that make the UK a beacon in the world for start-ups and scale-ups. We have deep pools of talent and world-class universities and research institutes. How do we go above and beyond to compete with the likes of the US, which obviously has very deep capital markets and, closer to home, the EU, which has 500 million consumers in a single marketplace? That is the challenge that I set out today. They are the competition. How do we compete?

In the particular realm of AI and the AI ecosystem, there are two critical ingredients. The first is compute and access to compute. We are really pleased to see that the Government will be bringing forward a strategy, and fundamental to it is what was outlined in Rachel Reeves's recent speech: resetting on growth and making sure that barriers to planning and so on are being brought down.

The second piece of the puzzle is access to large-scale data. Again, the Government have made very welcome noises. We are pleased to see the Government looking at opening up their own datasets, whether that is in the Departments themselves or with arm's length bodies such as the British Library or the Natural History Museum with the National Data Library. We are hearing very positive things, and I hope that is encouraging to our start-ups and scale-ups.

Fundamentally, however, publicly available data on the internet is the biggest data source, and to train these models you want to have the largest possible database to access. It means that the models are more diverse and have more inputs. We think that creates better outputs, so much more accurate representations on the other side, whatever comes out of the machine. Our preferred option, if we want to compete with the likes of the United States, which has a fair use policy, is option 2. I say that not to be not constructive with the debate, and in our response and further today we will go into the complexities around that. However, when start-ups are looking at where to build their companies, they are looking not just at London or wherever in the UK; they are looking at the global picture. There are the likes of the US with its fair access and fair use policy, and Japan, Singapore and Israel have all created very permissive regimes for training models. If we want to compete against them, that is sort of the essay question.

Q2 **Chair:** Thank you very much for that. I think you have made the position very clear, but could you just clarify for us: will you be proposing technical solutions to accommodate option 3?

Vinous Ali: A lot of the companies that we speak on are looking at those technical solutions, but they do not necessarily exist in the way they need to today. For example, they are not standardised, and they may not be interoperable or machine readable and so on. Until we have those sorts of technical solutions, it puts start-ups in a very difficult position if essentially we are legislating ahead of what is technically feasible today and, I guess, economically viable for start-ups. We do not want to be in a position where the largest players can pay their way through, and start-ups and scale-ups are locked out.

Q3 **Chair:** Thank you very much for that. We will now hear from James, who is developing technological solutions in this area. James, this is a very complex area, as you all know, but could you briefly describe where you sit in this area, what solutions you are developing and how they address the issues of transparency of inputs and outputs and ease of opt-outs?

James Smith: I will try my best. Hello, everyone. I am the CEO of a start-up called Human Native AI. We are only nine or 10 months old at this point, so it is very early in our journey. We were founded in the UK last year, and we are one of the startups that Vinous represents as part of the Startup Coalition. We are a slightly unusual start-up in that we are trying to help to tackle the issue. In particular, we are looking to try to enable more licensing by AI companies. I come from the product and engineering side, and appearing in front of these Committees is a new experience for me, so bear with me.

Chair: You are doing very well.

James Smith: This is only my second ever day as a professional wearing a suit. My first one was at the House of Lords inquiry in November.

We have been looking at how we build AI models, and at how companies like Google and OpenAI and so on require huge amounts of data to power these models. I think that is all well understood. The internet has generally been already consumed into these models. I think we all have to accept that as truth. Whether you agree that that is correct or not is a difficult issue, and there are some jurisdictions where it may or may not be legal. Our company is trying to help AI companies get access to the content that they need to improve the models beyond that in a fair and equitable way. We run what is basically a data marketplace. We enable AI companies to purchase data legally and fairly from the creative community.

You can think of us as trying to help to get generative AI out of its Napster era and think about a more legitimate era. We have been on an interesting journey over the last nine to 10 months. We have spoken to a lot of people on both sides of this debate, and we have heard interesting things from them. We have heard the creative community express anger and frustration at this issue. I think it is natural that sometimes they feel as if their work is being stolen, and for many of them it is very, very personal. I am sure you will hear from some great witnesses on that later today. From the AI companies' perspective, they also feel frustration. They are trying to compete. They are building new world-class products, and data is essential to build them.

Whenever you hear about AI today, think about it as just a reflection of the data that goes into it. My colleague here made a really great point about consuming the internet. The internet has great diversity of content, but the internet also has incredible bias of content. Imagine if you were building an AI model today and you wanted to counteract that bias. You might look to bring in some really reputable data that helps to balance out, for example, political leaning. You can do that by accessing and licensing appropriate data, and that is what our company helps to facilitate.

On this particular issue for the Government, the challenge is that a lot of the damage on text and data mining has probably already been done. The original sin, if you like, has happened, and I think the question is how we move forward. I would like to see the Government put more effort into supporting licensing as a viable alternative monetisation model for the internet in the age of these new AI agents.

Chair: Thank you very much, James. How could the Government support—

Q4 **Emily Darlington:** I was just going to ask a quick question. What needs to happen technologically to have a licensing system that is still in the hands of the creator?

Chair: That was absolutely the question that I wanted to put to you.

What needs to happen so that the licensing is in the hands of the creators, and is easy to access, simple and comprehensive?

James Smith: That is a fantastic question, and there are lots of great options. This is what our company is trying to tackle. I will lay out a couple of different points.

The first is simply streamlined legal access, and I think this is where the Government could particularly help. There are major challenges for many in the rights holders' sector—particularly things like audiovisual, maybe, for example, television—about clearing the rights for their content. You can imagine that makers of television shows may have writers and producers who they have to clear rights with. Makers also have to consider the talent on screen, may have some music in the background or may even have a Coke can on a table somewhere, and under the current legislation they would need to clear all of those rights before they could offer that content for AI training rights.

Until 18 months ago, no one ever wrote the words "AI training rights" into a contract, and that is a particular challenge that the Government could help with, with some type of more permissive system to enable rights holders to make best-effort attempts to contact people and effectively enable licensing and an extended copyright regime.

The second major challenge is probably technical, and it is access to the data—not only finding the right data, so discovery, but then being able to aggregate it and provide it to AI companies in a way that is sustainable and scalable. For example, let's say that an AI company wanted a UK news package. They would go around and speak to all the news companies in the UK and maybe license from three or four of them, but every time they do a licence it is a bespoke negotiation, a bespoke legal document and a bespoke technical integration to extract the information. Hopefully, services like ours could streamline that process and provide one technical integration to access all that different content. Then, if they think globally and want to access 40 newspapers globally, again you could start to see the benefits of a single aggregation platform for that.

Finally, it is about making sure everybody gets paid, being able to collect all the money in from licensing and being able to distribute it. CMOs may have a large part to play in that, but there are some very interesting challenges there around getting the data together, the legal side, being able to deliver it and then making sure everyone gets paid.

Q5 **Chair:** Thank you very much, James. We will now go to Sebastian Posth, who is joining us virtually. As the co-initiator of the International Standard Content Code, Sebastian is in an area that I am fascinated by from my background in engineering.

I would think that standardisation has a huge part to play here. We have heard a little bit about all the different ways in which rights can be asserted and content can or cannot be aggregated. Sebastian, as well as introducing yourself, would you briefly outline some of the key technical requirements for effective rights reservation methods, such as the ones you have co-developed? We have heard from James about the need for

AI developers to go out and find the content and negotiate. I am also thinking about individual authors or singers and how they can assert or reserve their rights easily and accessibly. Perhaps you could talk us briefly through that.

Sebastian Posth: I am a German citizen currently residing in the Netherlands, so I bring some of the European background, and not only in the discussion that is taking place currently in the context of the code of conduct of the European AI Office. For the last 20 years, I have been working in publishing, together with media organisations, in the context of digital distribution. From this background, I have been one of the co-initiators of the ISCC, which you have already mentioned—the International Standard Content Code. I served as a convener of this working group and its successor, which is a technical report at ISO level on what is called soft binding.

Chair: ISO is the International Standards Organisation, so that standard becomes comprehensive and universal.

Sebastian Posth: Correct. The ISCC has been developed in a working group with experts from 29 countries over the term of five years. The ISCC was published as ISO 24138 in May 2024. It is a rather new standard, and so it is important to talk about it and to introduce its capabilities. Liccium, of which I am the founder and CEO, is a rights management platform that allows individual creators such as you have already mentioned but also larger rights holders to make what we call asset-based declarations of their rights and preferences, and other metadata. In this context, it is relevant that these preferences are about the use of the content for training models and applications of generative AI.

I will also note that I am not a lawyer but a technologist, like my colleague, developing software that can be used by creators and rights holders to simplify workflows and processes. I am happy to introduce you to the technical possibility of making asset-based opt-out declarations. That is obviously not the first choice of creators and rights holders who, for good reasons, have a preference on opt-in or content licensing when it comes to training AI models.

Looking at the technical solutions and the key requirements, we have to observe the special situation that we are in, where in most cases the AI model providers already have access to the content. That is of interest because the question basically is how two parties—the rights holder on one hand and the model provider on the other hand, who are not in a contractual relationship—can both have access to the same content and exchange information about it.

There are three components to this. One is content identification. You have to be able to unambiguously identify the media content that you are talking about. The second is what I call metadata binding. You have to find a way to inseparably bind metadata and rights information like training preferences to the media asset itself. Then you have to develop a vocabulary that makes sure that everyone knows what to do when you

are talking about opt-out—so, opt out of what? Opt out of text and data mining, or opt out of training models of generative AI? These are the questions that have been raised in the context of the European discussion.

Q6 **Chair:** These are the questions. Do you feel that your standards address those questions?

Sebastian Posth: Yes.

Q7 **Chair:** James raised the issue of the retrospective, and he said that the original sin had already been committed, as it were. Could your standards also be implemented retrospectively?

Sebastian Posth: Yes. Let me briefly explain how the soft binding approach would work. Soft binding is in comparison to hard binding, where you try to include rights and metadata into the media file itself, so in embedded metadata. This is a problem, because metadata gets stripped when content is being shared online. It is very easy to remove the metadata. We provided the standard that allows anyone who has access to the content to apply an ISO standardised algorithm and generate a short identifier string from the media file. That means that anyone who has access to the same or similar content can generate the same or similar ISCC code from the media file itself.

The second step is that the rights holder would bind metadata to the content-derived identifier that has been generated and make this dataset, the ISCC code and the rights holder's preferences publicly accessible so that anyone with access to the content could resolve.

You mentioned binding metadata retrospectively. This is possible because you can update these declarations. Let us say, the AI model provider uses specific media assets for training. In this process, they would generate the code and then get access to the latest update of this declaration that has been, in our case, digitally signed by the creator and rights holder. This is also a way to update licensing terms or rights reservations to the content.

Q8 **Chair:** Thank you very much. There seems to be a combination of identifying the creative content plus easily and universally accessible databases of what that IP means. Vinous, do you want to say something quickly?

Vinous Ali: Yes, very quickly. I think these models are incredibly interesting, but I would make two arguments. The first is that the US does not require any of this, so how do we keep UK start-ups here—their talent their tax receipts and so on? That is my first point.

My second point is that is all great, but is it accessible to start-ups? What if someone decides to opt out in the future? That sort of continual checking and updating is a compliance cost that we should consider.

Chair: I think the Government will do that in the consultation.

Q9 **Natasha Irons:** I am very interested in the perspective of start-ups, and how they scale up their business models and things of that sort. On the

other side of that, how do you envisage creatives being paid for their content? How do you envisage that industry still thriving if, as you stated, you prefer option 2, which is to do what America is doing? How do you expect UK creatives to protect their incomes if we move to option 2?

Vinous Ali: Respectfully, I would say that the US has a thriving tech sector and it also has a thriving creative industry. Japan, Singapore and Israel all have creative industries that are thriving.

Secondly, I think this binary is misleading in some ways. This is not about a right to copy, paste and regurgitate work. The models are being trained on vast swathes of data. I am talking about millions or perhaps billions of datapoints. They are not reading, copying and pasting. They are learning. You can think about a student who goes to university where there are very clear rules around plagiarism. They cannot simply copy and paste previous work.

Chair: I do not want to rerun the argument about the creative industries versus the tech industries, because that is an argument that is being talked about a lot. We want to understand what the technical options are, and we do not have very much time.

Q10 **Dame Caroline Dinenage:** Vinous, you talked about the situation in the US, but fair use is not settled in the US, is it? That is why we are seeing so many lawsuits. To what extent does that come into play?

Vinous Ali: I think we will see how everything plays out in the US. I have absolutely no control over what happens there, but I will say that the Administration today is unlikely to move to a much more restrictive position. That is just one market. There are other markets that are seeing this as an opportunity. The creative sector benefits from artificial intelligence.

I completely agree with you. This is not a binary, zero-sum game; it is additive. It is about how you put these tools into the hands of creatives, whether they use them in post-production or to create new forms of art. It is certainly not a zero-sum game at all, and I do not want to suggest that for a second.

Q11 **Dame Caroline Dinenage:** Last week we saw the launch of the China-based AI model DeepSeek knock a combined trillion dollars off tech companies' market caps. OpenAI told the *Financial Times* that it has evidence that its intellectual property was used to build DeepSeek. We have talked a lot about competition, and you all talked a lot about competition in your opening words. However, is it not just plain common sense that when the door is open to widespread IP infringement, it will come back to bite you? To what extent do you think that is karma, James?

James Smith: That is a difficult point. There are lots of things at play here and I mean geopolitically as well. There are a lot of interesting things happening.

People will use whatever tools are available to make products and make things better. The open web was there, and there are cases that are quite

rightly being litigated to find out whether that was open to be used. There is a lot of mindset in tech companies, because of the competitive pressures that they are under, to move very quickly and to build with what they have available to them. It looks as though that is what this Chinese company did, although obviously that is still to be finalised. If it did, there could be an argument that it is an interesting use of other tools and other technologies.

Perhaps a more interesting argument for the creative sector here is that this presents more opportunity for them. I go back to one of the earlier questions about how they will make money? The more AI companies there are that are looking for data to improve their models—I think that that represents opportunity for the creative community, and new and interesting monetisation models. There will always be a great creative community and a great creative sector because people will look for human connection when they look at pieces of art, music or other things. I do not think that that is going away to be replaced by AI music any time soon. However, I think that AI tools can help improve and create new opportunities that are then trained on great creative works. This can be a win-win if we can engage with it in that way.

Q12 **Dame Caroline Dinenage:** Vinous, Sam Altman said on Reddit on Friday that OpenAI has been on the wrong side of history when it comes to open sourcing its technologies. Do you think it would have been more apt to say that OpenAI has been on the wrong side of history when it comes to IP infringement? How does that bode for the future?

Vinous Ali: The Chair said right at the beginning that OpenAI was asked to be here and could not come. I am really not here to speak for OpenAI. That is not the constituency that I represent at all.

The DeepSeek explosion is really exciting for us sitting here in the UK. I have heard a million times that we are not going to create these types of models anyway, that we are already out of the game, that that ship has sailed, that it is far too expensive, and so on. What DeepSeek shows us is that it can be done a lot more cheaply. We can expect that efficiencies will continue to be made, opening up the competition again. I want UK founders to be leading that charge rather than going to the US and, as James says, moving quickly there because they are enabled to.

Chair: I am keen for us to continue on the technology and not to rehearse arguments.

Q13 **Steve Race:** I note for the record that the Startup Coalition provides the secretariat for the ClimateTech APPG, which I chair.

I want to move on to the transparency of inputs and outputs. The UK wants to see ever more start-ups setting up and then scaling up, leading on AI but also utilising AI for new and exciting businesses. Vinous, what might be the impacts or the opportunities for you and your members around the requirements for transparency, in and out, of that data?

Vinous Ali: What we hold it up to is whether it is something that start-ups can comply with. That is always the biggest question. We are very

excited to look at the technologies that James and Sebastian are putting forward, because those technologies are all about making these arrangements much more streamlined, more accessible and so on. We are excited to see where that technology goes.

We look at what is happening in the EU at the moment. So far, we see that these transparency requirements could be put at a threshold that essentially locks smaller players out of the market because perhaps they cannot comply, or they cannot continue to go back and make sure that they are updating it, and it just becomes a further compliance cost for them. We want to make sure that wherever we draw that line, it is accessible to start-ups. That is the same for whether we reopen the debate around robots.txt, and so on. It is about making sure that start-ups can compete to upend the market once again.

Q14 **Steve Race:** James do you have anything to add on that?

James Smith: Yes, I echo Vinous's points. If the Government want to encourage growth, I think they will want to find solutions that enable SMBs to compete effectively on both the AI and the creative sides. There are real opportunities. For example, we are working with a UK-based start-up today, a very small operation that was only founded last year. It is licensing data to improve its models, and it is doing so in a fair and legitimate way. The founders are very ethically minded and want to do this. We should be encouraging that. We should make it easy for them to access amazing content that the UK produces if the Government want to go after growth.

Likewise, we should make it easy for small creative communities that have amazing content to participate in a market where there is demand for their content. By doing so, by helping them to clear rights and other things like that, that the opportunity is there for them.

Q15 **Steve Race:** If I am right, your company, James, might be one of the tools that would help start-ups and scale-ups to comply with any new transparency requirements. Is that right?

James Smith: That is right. We would help them to license the data and, through licensing data through us, they could produce a compliance report, which they could then submit to any transparency requirements in any jurisdiction. If they were to acquire only licensed content through us, they would be easily able to comply with transparency requirements.

Q16 **Steve Race:** The consultation says that some R&D funding might be made available to support new tools to encourage greater transparency. Do you think that is necessary? Do you think that is a necessary part of what might be a new ecosystem, or do you feel like you are doing it already and that that is unnecessary, and that there is plenty of competition out there to help support companies to do this better?

James Smith: There is a very interesting opportunity here for the Government to get involved and produce some new paths. There is a lot of great content and great data within the Government's realm of influence, shall we say, where they could help to move things forward. For example, you will notice that the action plan lays out a couple of

British institutions. I have been fortunate enough to have spoken to some of those institutions, and what they are concerned about is whether we can get 10 data privacy officers involved before we can even move forward in understanding what this is. The current regime around data in the country is very restrictive, so if we wanted to do it, I think that the Government getting involved in trying to find ways to move it forward would be really interesting.

Q17 **Steve Race:** That is a good segue into my next question. What do you think might need to be changed in the regulatory underpinning to encourage this?

James Smith: I will preface my answer by saying that I am not a lawyer. We have some experts behind us. From my perspective, changes might include making it easy for companies to be able to say, "We have done our homework, we have contacted as many rights holders as we can and we have cleared the majority of rights". There are similar regimes available in situations where the copyright holder may have passed or is uncontactable. There are similar provisions that allow for works to be generated and sold on. Those types of things would be very interesting. Anything that the Government could do to encourage that type of work in licensing would be very interesting.

Q18 **Steve Race:** A final question for me is on consistency in transparency. Vinous, consistency is probably key to start-ups and scale-ups understanding what they are supposed to be doing. Where do you think that the Government can come in to avoid unnecessary complexity and to give people a clear steer on what they are supposed to be doing?

Vinous Ali: We are probably going to cite this report for many, many years to come, but the AI action plan is very clear. It says that the UK should not go beyond what the EU has proposed. Unfortunately, I think that if we are not looking at option 2, which is aligning us to more permissive regimes, we certainly should not be going beyond what the EU has already set out, because we are asking our start-ups to try to create two different regimes that they need to comply with.

In many respects, it is a shame that it has taken us this long to reach a position of consultation. It is very bizarre to me that the EU is ahead of us in this respect. However, ambiguity in itself is to the detriment of start-ups and scale-ups. Given all the hype around generative AI, you would expect to see a thriving generative AI ecosystem here in the UK already. Actually, the data shows that numbers are declining, so the trends are not looking good.

The House of Lords Communications and Digital Committee published a report just yesterday, and even the likes of the creative sector said that the ambiguity of the UK's current position on the application of copyright in relation to AI was consistently flagged as a factor limiting growth and innovation. There is an element of getting it right, but others have come before us, and we should look to them if all else fails. There is also the point about doing it quickly.

Q19 **Chair:** Before we move on to some more details of rights reservation, do

you think—yes or no—that it is possible for the rights environment to be entirely transparent so that a large language model can identify every single part of creative content that is input? Do you think that is possible technically?

Vinous Ali: No.

James Smith: Yes

Sebastian Posth: I would say that this is possible. I would say yes.

Chair: It is a yes, so we have one no and two yeses.

Q20 **Martin Wrigley:** It is a very interesting dilemma. I have spent a number of years working with start-ups. I was in that sort of industry pre-pandemic, so I have seen all this under a number of things. I am also old enough to remember Napster and digital rights management going through a whole range of this stuff before, and I am wondering how much of this we are recreating.

I am very confused on this issue, because I have met a number of people now who have described their job title as the manager of responsible AI. It is not clear to me whether there is somebody who is managing the irresponsible AI, or if it is just being unmanaged. How do we tell the difference between good-faith actors and bad actors? All the things that we are talking about and all the stuff that I have seen from you so far is relying on everyone being a good-faith actor. How do we enforce? How do we determine?

James, you were talking about your compliance reports. That is fantastic and great, and it sounds like a really good initiative, but it relies on those people running compliance reports to identify and acknowledge all the input that is already in the memory of their neural net. How can we enforce any of this? How do we stop this from turning into the next George Harrison being sued for "My Sweet Lord", because what has popped out of an AI system sounds remarkably like somebody's previous work?

Vinous Ali: As with everything, first, the UK Government can take a lead here in identifying good actors and bringing them to the fore. AI presents a huge opportunity, as many of you will know, to the public sector, whether that is in healthcare, in supporting decisions that DWP or the MOJ have to make or in enabling teachers to take some of the bureaucracy off their plates so that they can concentrate on their students. It would be a fantastic start if we could get the Government to procure this stuff and bring it in, and have systems in place there to bring the good actors to the fore.

The point about the outputs is interesting. Going back to my earlier point, this is not about models regurgitating things. Frankly, if we are seeing things on the output side that look like copies, there should clearly be stringent enforcement mechanisms to allow rights holders to take issue with it and to have their day in court. There is that output side and capturing bad actors, and then there is what the Government can do to bring good actors to the fore.

James Smith: From my perspective, our company is very much focused on the input side of the problem. We are not working on, for example, output detection or things like that. I should preface my answer with that.

We are talking to and working with some of the world's best labs in AI. They are coming to us and asking for legitimate access to data. They are at the forefront of their field. They already have the world's most capable models, and they are looking to improve them by getting legitimate access to great content. By doing that, I think that already is the natural world of competition among these products. The best products will be those created by the best actors who are legitimately accessing content, because that content improves their models and their capabilities.

You can imagine a world in a few years' time where we have these very powerful models in all these different areas. I think the most powerful models will come from the companies that we consider to be the best actors. The companies that are not behaving legitimately are therefore maybe just getting the open web content, which effectively gives them the capabilities of today's models. Tomorrow's models will be built on open web data plus access to licensed content that is not publicly available today.

Sebastian Posth: I believe that the question of opt-out methods—the way to make opt-out declarations—is related to input transparency. Let's say AI model providers generate identifiers and look for preferences. They already generate codes for the content that they intend to use for training. In the case of James's model, it is pretty clear, because content is actively provided and distributed to the model provider. But in the case of opt-out, there is still the question, "Did you really remove our content? What have you trained on, and were our terms accepted and respected?"

Generating codes for content used for training allows monitoring of whether opt-out declarations have been respected or not. The same goes for output transparency, where it is possible to generate an identifier for each asset that is synthetically generated. We already see with US-based model providers that they identify the content that they are generating cross-platform, so that is also possible. I would say transparency can be achieved at a fairly detailed level.

Chair: Thank you very much. That is very interesting. I am going to ask Jon Pearce now to talk a little bit more about the technology.

Q21 **Jon Pearce:** I particularly want to address some questions to you, Sebastian. You have already started to go into some of the detail but I am very interested in how your protocol, TDM·AI, works. How do you think it helps us with the toolbox of dealing with these issues, and where do you think any gaps might lie?

Sebastian Posth: The TDM·AI protocol is more or less a description of a method that can be implemented by organisations, because most of the technology that is being used is openly accessible under a permissive licence—for instance, the ISCC code. Anyone can use the ISCC generator

and generate the codes, be it a model provider, a rights holder or a service provider.

The three components of the protocol are generating the ISCC code, binding rights and metadata to the ISCC code and digitally signing this declaration. One thing that is often overlooked is that you have to understand who actually made a declaration—who was entitled to provide rights information to a specific media asset. Things such as certificates and verifiable credentials are part of that declaration process to make publicly accessible the rights preferences that are digitally signed.

Q22 **Jon Pearce:** Great, thank you. That is helpful. Can I just check, is the protocol available for rights holders in the UK?

Sebastian Posth: For sure. The components are described, and start-ups and media organisations can make use of that. In practice, we already see implementation by international book publishers, generating these codes on premise, so you do not need to send any content to a service provider. They generate the ISCC code—the identifiers—on premise, then they attach their licence and put it in a publicly accessible registry.

Maybe I will make a comment on the registry. It is important to understand that the registry technology should be federated, because it is hard to find or implement centralised approaches in an international setting. Imagine UK publishers, Italian book publishers, German book publishers and US-based record labels. They would need to agree on one registry where you could access the ISCC codes and the preferences. That is unlikely, so we are trying to achieve a system of federated registries where each media sector can make these declarations locally, and the ISCC codes serve as a means for the interoperability and aggregation of these statements.

Q23 **Jon Pearce:** I have one more question on this for James. What are your thoughts on the protocol, how it works and how useful it is?

James Smith: It is interesting. I think the challenge will be whether all the creative community can participate in this. I think we often forget, as technologists, that not every company we speak to is a tech company, and they may not have the capability to do this. The size of some of these archives is interesting. We are speaking to rights holders who literally have archives that go back centuries. If they make those archives available at scale, they need a commercial opportunity to do so. One of the challenges is that there is always value in defending things, but often what motivates is the opportunity to create a new opportunity.

I am really supportive of these types of protocols. I think there are loads of opportunity there. I really like the idea of them all being federated, a bit like, for example, how DNS systems work. When you type in "Google.com" on a web browser, it finds where that server is using a federated set of computers around the world, so it is not one single point of failure. That is a great way to do this, so I am very excited about this approach. I think it will lead to a lot of interesting challenges around how

creative communities engage with it, but over time, if they have the incentive to do so, we will see more and more engagement.

Chair: This question of how easy it is for individual creators to access it is a key one, and perhaps we can return to that in the next session. Allison, the labelling of the output of AI models is also an important topic.

Q24 **Dr Allison Gardner:** James, I know you mentioned that you do not look at output labelling, so I might broaden this and have Sebastian's input. I am interested in whether the soft binding and hard binding protocols that you are doing for input could be expanded to output labelling—with generative AI, understanding that something has been produced by generative AI—and how we could do that. I am quite interested in your thoughts on what information should be in output labelling, and how that information would be presented or made accessible to people so that they could understand that something had been generated by AI.

James Smith: I will take a quick swing at it, and I am sure Sebastian has great thoughts on this. This is definitely not my area of expertise, but I am excited about the idea of lots of new types of content and output being created by generative AI. I should say that I am definitely an optimist about the future of AI. I realise that might be a controversial opinion, but I think it will generate a lot of interesting things.

It is on the AI companies themselves to ensure that their outputs are correctly labelled and that they are correctly machine readable, or otherwise, to say, "This was created by our tools" and provide some sort of information. That will be key to a lot of opportunity, and to ensuring that we understand the difference between what is AI-generated and what is human-generated, because that line will start to blur considerably. We will see more and more efforts, particularly among the best actors in this space, to make it much clearer what has been output by their systems. The challenge will come when the bad actors have systems that are good enough to produce convincing output. Hopefully, that is not the case yet, but it will be soon.

Martin Wrigley: They already do. This was the announcement last Sunday, when they were talking about pornography produced by AI now being so good. They already do. We are already there, so I am sorry; it is too late.

Q25 **Dr Allison Gardner:** Yes, and it is one of the big problems that we see on social media in videos and so on. That is in a different inquiry that we are doing. Sebastian, can I go to you on the same question, but particularly about the technical aspects of what would be involved in such labelling and how it would be presented?

Sebastian Posth: I would like to highlight that it is not only simply a flag for synthetic content, but you may also want to consider the downstream use of synthetic content. As just one example, with images in the news industry we have the IPTC metadata standard. It has what they call the digital source type, where you have something that says, "This image is 100% synthetic and AI-generated," and then we have edited versions that have been supported by the use of AI, and human-generated

content. You have flags for different degrees of AI involvement that bring transparency to its use. The IPTC standard is broadly used and broadly accepted as a way to flag content that has been used. It is basically adding this information to the ISCC code and having a declaration requirement for the model provider or the system provider that generates the generative synthetic content to generate the code and add this flag.

Q26 **Dr Allison Gardner:** I do like the sound of a flag. I am still trying to understand how your bog standard user scrolling through on the internet and looking at images and video would know whether it was AI-generated. It is great that they see those flags in the industry, but how would it be technically presented?

Sebastian Posth: If, say, a social media platform has user-generated or user-uploaded content that is being shared on these platforms, the platform provider can easily generate the code and have a look at whether some flags are associated with this particular ISCC code or a similar ISCC code. That is a persistent way of binding this flag even if it is distributed online.

Dr Allison Gardner: That is really interesting; thank you. I want to ask about consistency, but do you want to jump in, Chair?

Chair: I think we are going to have to move on, because I do not have much time. On the point that Allison was making—the platforms know that it is synthetic, but how does the person using it know?—I guess there will be an opportunity to click on something, somewhere, and it will let you know. That is part of the discussion that the sector needs to have with the Government about what we want to see.

I will move on now, because we have only seven minutes left. Lauren wants to pose a general question to everyone, and Paul wants to come in as well. We have seven minutes and three witnesses.

Q27 **Dr Lauren Sullivan:** I will be nice and quick. My question was going to be about the hurdles, whether technical or non-technical, as we go ahead. But to pick up on some of the comments and questions, we have issues with trust. It relies very much on honesty from good companies, and whether you are licensing and whether you are showing that. You mentioned that monetisation might be an option for how creatives can share in the growth. What kind of sanctions do you think we need to instil or be aware of so that there is that trust, and so that the good actors are rewarded and there is enough weight to hold to account those that are not so good—as well as the hurdles we are facing?

James Smith: Again, I preface this by saying that I am not a lawyer, but my understanding is that in the United States, for example, they have statutory damages for copyright infringement. I believe we do not have those in this country; I am just checking with my lawyer. I think if we wanted to put a sanction-type system in place, something along those lines might be interesting.

Vinous Ali: I will take the question on hurdles, if I may. It was interesting to watch the debate last week in the other place on the Data (Use and Access) Bill and the amendments that have passed, led by

Baroness Kidron there, and I just want to put on record the big hurdle that that puts in front of the UK's tech sector. You had the Prime Minister talking about how he wanted the UK to be a maker of these technologies rather than a taker. The Kidron amendment, which puts extra-territoriality into these discussions where it currently does not exist, essentially would mean that these models cannot even deploy into the UK. It is not just that they cannot be made here, but they cannot then be deployed here. I think that puts us, as a country, at a huge disadvantage.

There are many benefits to artificial intelligence being diffused right across our economy. I touched on the benefits to the public sector earlier, but there are huge benefits to the creative industry. I think the University for the Creative Arts estimated that 40% of employment in the creative sector would be in "createch"—the intersection between creativity and tech—by 2030. I hope that when this place comes to consider the amendments in the data Bill, we keep in mind that very broad picture and do not lose the opportunity entirely. I will concentrate on the hurdles.

Sebastian Posth: When it comes to the hurdles facing the technology, I understand concerns from the rights holder side or creator side about the use of such technology, and then obviously costs and effort are involved. We need not to ignore that; we must take that into account. On the other hand, the experience with rights holders that are using the technology is much simpler than we anticipated. It will become standard practice in dealing with this kind of situation where the preferences are published in such a way.

Chair: Thank you. I am going to ask Paul and James to come in very quickly—one minute each.

Q28 **Paul Waugh:** We had the Minister, Chris Bryant, before our Committee last week, and he made it absolutely clear that one of the central requirements in this consultation is that if the technical solution cannot be found, the Government will not go ahead. That is why I am interested in what you said, Vinous—you think those technical solutions are not possible, which is why you are backing option 2. James and Sebastian, you are basically both suggesting that it is possible. I want to ask you about timeframes. We deal in timelines in this place. This consultation is time-limited, as this Committee's report will be. How realistic is it that those technical solutions could be provided and made robust enough to last within the horizon of this policy?

James Smith: I think that is the question. It comes down to the requirements of the technical solutions. If you just want a simple opt-out mechanism that enables, for example, an entire organisation to opt out, you could argue that the robots exclusion protocol does that today, or you could argue that it does not. I think it depends on the requirements before we can put a timeline on it.

Sebastian Posth: We must be clear that ISCC and this situation is new, and that people need to get familiar with the technology. Software providers and service providers, need to start getting familiar with the opportunities that come with this technology. It is at an early stage, to be

honest, and the feasibility of the registry technology is currently being tested, supported by the European Commission. A tender for a feasibility study for ISCC-based opt-out declarations was published two weeks ago. There are a lot of things going on.

Q29 **Mr James Frith:** James, you talked earlier about the original sin having already happened. That talks to a fixed point, when some would argue that the original sin is continually repeating itself in new forms and new guises. Have you heard of this idea of grounding? You no doubt have. I have read your biography. To those of us who are concerned by the apparent dismissal of creatives' concerns, will you explain the threat posed by the advance of this tech to the point where the original sin, to use your phrase, is repeated every day on other people's work?

James Smith: There is a period where these models are being trained. If you use any of these models, sometimes they will say to you things such as, "My knowledge cut-off was at a certain point". Ongoing access to information through grounding is a real opportunity for organisations such as those in the news sector to provide up-to-date information. Those licensing deals are happening today. I think that there will be greater opportunity to continually move forward. If rights-holders are taking control of their content, expressing opt-out by whichever mechanism exists and then engaging in licensing, there is an opportunity for the creative sector to really participate in the AI economy.

Chair: Thank you very much, James. That is a really good message to end the session on. I am afraid that we will have to leave it there. We have been able to take not a deep dive, but a shallow dive into the technical options to address the Government's preferred option in their consultation. We have heard that there are technical options with different levels of implementability. We have heard from one of our witnesses who does not think that it is possible to have a totally comprehensive opt-out regime, which I think is essential, and from two of our witnesses that it is.

It is also clear for the Committee that the technical solutions are technical and complex, and therefore difficult to understand. We need to see how creatives, from the individual artist to the huge news conglomerate, can or cannot accept those technical solutions and continue to make a living, which I think is what we all want. We will be hearing more on that in the second part of the session, but for the moment let me thank you very much for what has been really interesting evidence.

Examination of witnesses

Witnesses: Sajeeda Merali, Max Richter and Matt Rogerson.

[Dame Caroline Dinenage took the Chair]

Q30 **Chair:** For our second panel, you may notice the Chair has changed seamlessly from the brilliant Chi Onwurah to myself, because we are holding this session jointly with the Science, Innovation and Technology

Committee. On behalf of the Culture, Media and Sport Committee, we are very keen to welcome our second panel today: Sajeeda Merali, who is the Chief Executive for the Professional Publishers Association; Max Richter, the composer, pianist and producer; and Matt Rogerson, who is the Director of Global Public Policy and Platform Strategy at the *Financial Times.* You are all very welcome.

The Culture, Media and Sport Select Committee, in this Parliament and the last, has been taking a close interest in fair remuneration for creatives. We have been looking at the opportunities of AI, but also the risks. Sometimes, there is a tendency to pit the two against each other—to pit the creative industries, publishing and others against AI and innovation. We are trying to navigate our way through it, and that is why bringing together the individual Select Committees is really useful.

Before we kick off, I want to come to each of you very briefly and get your reflections. You were all here for the previous part of this session, and I want to get very brief reflections on what you have heard so far. Sajeeda, I will start with you.

Sajeeda Merali: Thank you for inviting me along today. I will let you know who I represent, because I think it is quite relevant to the conversations that we are having today. The PPA represents around 200 publisher members from the magazine media sector. We are part of the broader creative economy, which is worth £4.4 billion to the UK economy and employs 55,000 people across a range of skills—journalism, videographers, data analysts and so on.

We have a really broad variety of members. Among the 200-plus members, there are a lot of large multinational consumer magazine businesses: UK household names, *Good Food*, *Good Housekeeping*, *British Vogue, Grazia, The Week, New Scientist*—you will recognise many of them—but also business media. This is for decision makers who are helping leaders—people who lead some of the biggest industries in our world—but also independent publishers, so SMEs all the way up to entrepreneurs. Some of them deliver important information to diverse groups: *Inclusive Design,* an interior design magazine for the disabled community, or *Black Business,* for example. I think we are thinking about it absolutely to your point—large conglomerates, but also those very small creatives that are trying to reach niche audiences.

It was interesting to hear such a diverse mix in the earlier session. We face a lot of challenges that hopefully we will get a chance to share. I am definitely on the side of what Human Native AI said—that there should be an opportunity to protect copyright, to create a licensing economy and to think about how we can ensure that AI is part of the growth economy that we are all looking for.

Max Richter: My initial response to what we have heard so far is that AI is obviously a technical subject—the conversation was very technical—but I think the technical impacts different areas of human culture in different ways. It is obvious, for example, that there is a tremendous benefit to the use of AI in medical research, scientific research and all kinds of technologies. It is not obvious how AI can "improve" creative things. I do

not think we have heard people say that the human race has run out of ideas for music, and we need to automate that. The argument is much less persuasive when it comes to the creative field. That is my initial takeaway. I am not totally sceptical; I am just not as persuaded as I am in technical fields.

Matt Rogerson: Thanks for the invitation. I think you are probably aware, but the *FT* group is headquartered in London pays its taxes in the UK, which is quite rare these days. It comprises 18 different companies, including the 135-year-old pink *FT,* the FT Live events business, an investment data business called FT Locations, a publication called *Sifted* that looks at EU start-ups*,* and *Endpoints*, which looks at biopharma. Those are just five other companies within the group.

As a business, we are embracing AI. We were the first UK publisher to sign contracts with OpenAI for use of our journalism in training and also in display within their services. We are using AI in the business and are training staff to be AI literate or AI fluent. In the newsroom, we are looking at AI investigations and investigating datasets. Our business has no cost to the taxpayer; we are completely commercial. We create overwhelming positive externalities for wider society, but we do need to be able to make money, and I think that is where we are worried about AI and this consultation in particular.

On what we heard earlier, I think there is some misdirection that we have heard throughout this debate. One is the concept of publicly available data. When you hear that phrase you just think that they want to scrape everything, because that is what has happened for the past few years—everything you have ever written on the internet, whether you are a consumer or a business, has been scraped by these businesses for use in training. When they talk about more diverse and better outputs, it just means they want everything. There is no selectivity about this; they have scraped everything and used it for training these models.

The other misdirection is that the case law is settled in the US. It really is not. Fair use has not been proven to cover either what has happened in training or what is happening with output—I will come to that—and that is the real concern for us. It is both the training and the substitutionality of the outputs. The worst thing for a news publisher is to have the relationship between you and your reader mediated by an aggregated platform. We have had this for 20 years. Google will always put bigger blocks of text in from other companies so that the user does not click through. This is the big concern with the AI world—that they will extract the value of all the journalism for training and extract the value for outputs, and no value will flow back to our industries. The worry for us is that that will lead to a hollowing out of our creative and information industries.

Chair: Brilliant. You have all set the scene beautifully. Thank you very much. We have just under an hour—about 55 minutes—to get through a lot of burning questions. I will start with Liz Jarvis.

Q31 **Liz Jarvis:** Good afternoon to you. I will start by asking Sajeeda and Matt about the challenges that AI development poses for rights holders in the creative industries and media sectors.

Sajeeda Merali: I want to start by saying that as an industry there is a huge amount of AI innovation across the publishing sector already. AI is not new for the publishing sector. We are using it to drive efficiencies, to understand our customers better, for new product development and also for creative solutions. We are an industry that has innovated throughout the digital era. When the internet came about, we had to pivot from print to digital. When smartphones came about, we had to think about new ways that consumers wanted to access the information, and the same with social media. We think about this as digital 4.0, and I think this is another one of those big moments where the industry is looking to gain a level playing field to continue its sustainability.

The thing to remember is that our industry's bedrock is fact-checked original content that is produced by humans. There is a human accountable, and a trusted brand accountable, for what gets out there. In the digital economy, the investment that goes into that fact-checked journalism currently comes from traffic to publisher websites. To unpack that further, traffic then turns to revenue largely through advertising, but also through conversion to subscriptions, selling tickets to events or e-commerce.

Over the years, members have consistently told us that the majority of their traffic comes from Google Search. People who previously relied on Search for day-to-day queries are now increasingly turning to large language models to have those queries answered, where publishers do not get any traffic. People get the answer and do not have to go on to the publisher's website, and therefore there is no revenue. AI might represent the next digital evolution, but in the absence of copyright protection, this is a huge threat to the commercial structures that are needed for fact-checked journalism—trusted editorial—to be sustainable.

These AI models are being used by our readers essentially as a competitor for time, engagement, attention and source of information. You cannot have a competitor that has been granted free access to use your assets and your content to build a product that they are potentially creating billions from, and not have any protection from copyright law.

Matt Rogerson: I agree with all that, but more. The Government's consultation focuses a lot on model training. For the news industry, the real concern is what is happening today. It is about substitutional products today. There is a search engine called Perplexity, which consumers pay for, and none of the money comes through to the content that is featured in the responses that come from that search engine. Perplexity is a good example. Last week they changed their terms of service, and I think I know why. Previously their terms of service said, "We abide by robots.txt. However, we reserve the right to come on to your website and create a short summary of the stories that are on your website and provide them to a user." What that means is that they are not abiding by robots.txt. They are scraping the article and then

providing a summary for the user. Nobody knows how they are getting access to that information, because lots of publishers block them. There are views that they may be using a spoof server to get in. Now they have changed to a position where they say they use third party crawlers to get access to that data.

We know that Google and Microsoft both sell access to data that they get from the search index to large language models. I think Microsoft sells 14 different large language models, so that extracts value from news publisher sites and sells that value to LLMs. They query the search data to provide an outcome. The problem with large language models is that they are a bit like an expensive car. You can have it in your driveway, but you need fuel to get it off the driveway and go for a spin. You need data from publishers like the *FT*, *The Guardian* and others to make the large language model make sense. At the moment, there is no payment flow whatsoever from the people providing that service to large language models back to the source of the originating content.

We have a consumer business, and we also have a B2B business and B2B licensing. You have products coming to the market where a business can query the search index to provide things such as long research reports. You can quite quickly see that that could become substitutional for the sort of products that we provide, and that other companies provide to businesses. At that point, you see revenue streams at serious risk and there is no way at the moment of getting a replacement for that.

Q32 **Liz Jarvis:** Can I ask you both what is your preferred outcome from this Government consultation?

Sajeeda Merali: We still have until 25 February, and we are consulting with our publishers about that. Summarising some of the topline themes that are emerging, we want a recognition that content has already been used and that is in breach of copyright law. We want recognition that copyright law exists and is applicable. We want recognition that the original creative content that is being produced by so many of our publishers is a key part of the growth that we are looking for in this new AI economy, but we also need a practical solution for creative industries.

If I reflect back on the scale of publishers that I spoke about, not everybody's pockets are deep enough to try to enforce that this works. We need to think about frameworks that allow all sorts of publishers to come to the party and play within this. When we say "practical", we mean that the burden of implementation sits with the platforms, and the burden of enforcement needs to sit with an independent regulator.

Matt Rogerson: It is interesting that the option is not there, which it should be, to keep UK copyright law as it is and introduce transparency measures. There is nothing wrong with UK copyright law. It is completely clear that what has been going on is a breach of UK copyright law. I think Baroness Butler-Sloss, who is more learned than I am, said in the House of Lords last week that there is nothing wrong with the UK copyright law. She can discern this as a breach of it.

The transparency measures that we worked on with officials under the last Government are pretty positive if you can get them right, and if UK businesses can start to be able to realise the value that has already been extracted through the training using our data and start to build a market in the UK that functions. We have huge advantages with the creative industries in the UK. It is a £126 billion industry, with a huge legacy and archive of data that can be used to train models. Keep existing copyright law as it is and introduce transparency measures. The question then is around enforcement. I think we may get on to that in a different question, so I will leave that for now.

Q33 **Liz Jarvis:** Max, you signed the statement objecting to unlicensed use of copyright works for AI development. How prevalent do you think this practice is?

Max Richter: It is universally prevalent. I think all music has already been hoovered up, so what we are trying to do at this point is to establish a situation where that can in some way be remedied and protections put in place for the future.

Q34 **Natasha Irons:** Thank you for coming in this afternoon. I am thinking about where we have already been in this journey and the previous consultations and discussions. You mentioned that you have met with a group of officials previously. Matt, given your sector's response in the last Parliament, are you surprised that the Government have renewed plans for broad exceptions for text and data mining?

Matt Rogerson: Yes, I think it is very surprising. There has definitely been a change from Opposition to Government, and that is quite troubling. I think the Prime Minister and the Secretary of State for Science, Innovation and Technology have spent a lot of time with the biggest US tech companies, and it is understandable to me that they would want access to all this content for free. What business would not want free things? They have a huge liability on their hands if we can enforce copyright, because the scale of the processing of data and the scale of use of copyright content is so vast.

What is equally scary is the move away from thinking that competitive markets are a good thing. If AI technology is as powerful as people suggest, the very worst thing you could do as a Government is have three to five very large companies from outside of the UK dominating that technology and being integrated into things such as public services because, to put it bluntly, they have got you over a barrel with cost. We may get on to DeepSeek versus US tech, but if you look at the cost of building these products in the US compared to other countries where they are genuinely innovating, there is no magic money tree. We must pay for that. As consumers and as business, we have to pay for that. I am very surprised about the change on copyright, and I am very surprised about the attack on competition as well.

Q35 **Natasha Irons:** Sajeeda, are the current proposals at least a slight improvement to the previously proposed TDM exemptions?

Sajeeda Merali: There is a whole timeline around the TDM piece. It is worth remembering that back in 2022, when I think it was kicking off, AI was something on the horizon that needed to be addressed. I remember coming back from Christmas in January 2023 and ChatGPT was literally the only thing anyone was talking about. AI is definitely the word of the year. The fundamental legal concepts have not changed. I have an example that Matt might be familiar with. When the iPad first came out, it was argued by *The Times* that the digital edition should be treated the same as the print one and should be zero-rated for VAT. The case was won on the principle that when the law was drafted you could never have imagined that an iPad would have existed. What they did know was that there would be a daily edition—a collection of news distributed to the public that was of social value—and they won on that basis.

I think this is a similar scenario. The whole AI piece has exploded, the investment into technology has gone up, as has the number of players and the geopolitical implications. That has changed, but none of the issues around copyright protection or the need for a practical regime to enforce it has changed.

Q36 **Natasha Irons:** Given some of the evidence we heard in the first session, what is your response to the fear of the smaller start-ups in the UK that the bigger boys in the race could hold back their development and stop them moving forward?

Matt Rogerson: I think we have all got our favourite Keir, haven't we? My favourite was when he talked about Britain being on a mission. If the Government had come to the content industries and said, "There is a mission here. We want to get smaller AI companies working with the content industries to develop models in a way that puts Britain at the forefront", the content industries would have been up for that. Instead, it feels as though what is happening is that this is all being done to the UK's content industries. None of us has met the Secretary of State for Science, innovation and Technology. I think we have a meeting in April with him, but he has not heard our concerns about the kind of proposals that they have in the paper.

I think that is a missed opportunity, because the DeepSeek development shows you that the field is not settled. If you got a collaboration between smaller AI companies and the UK's content industries, with the creative history that we have, you could do something really interesting.

I am quite surprised to hear start-ups arguing for blanket commercial TDM exceptions. It seems to me that does not advantage them at all. It does not enable any leverage for the British Government to say, "You are a UK start-up, you pay taxes here and you employ people here, so you can get access to this data". All it would do is to create a free-for-all for large tech companies from anywhere in the world, and I think that would be a huge mistake.

Q37 **Mr James Frith:** Good afternoon. Sajeeda, has the European Union's opt-out approach been working for the creative industries?

Sajeeda Merali: Much like the EPC—the European Publishers Council—and EMMA—the European Magazine Media Association—the PPA is definitely supportive of the advances made around the use of the AI Act, but we echo their concerns about having a law that enforces it meaningfully. From our perspective, we have the same philosophy, but they have struggled to create a practical solution. All platforms are lobbying to say it is too difficult, so they are not incentivised to solve the problem. That is our response on that one.

Q38 **Mr James Frith:** What do we need to do differently to the EU?

Sajeeda Merali: This is about having a legal framework that exists in our copyright law. It is about reinforcing this and fining for misuse, but the key is transparency. Transparency is a big part of that, making sure that there is a requirement to report on usage by the LLMs and the platforms; this can be done, despite the claims to the contrary. Then there is paying content creators fairly. Models exist now—licensing companies, Apple News. LLMs need to have collective or individual agreements to make that work.

Q39 **Mr James Frith:** Matt, what is your view on the EU now considering whether central registry opt-out is feasible? It sounds like a limited library model, almost. How do you take it?

Matt Rogerson: I think there is heavy pushback from the tech community in Brussels. What they dislike, and what is absent from the UK proposals, is that if you have not complied with the law in the EU, you cannot release a product in the EU. That is the biggest disincentive to piracy that there is. You could not contract with government if you had not complied with copyright law in the EU. We went through the IPO roundtables, and it was collapsed by the technology companies. The next thing we know, Microsoft has a massive deal with the UK Government. That is not sending the right message for how you build these models.

I think we are missing bits in our proposal. In the EU, from our perspective, we think that greater transparency is key. We are not saying that we want to know all of the model weighting that the EU use, or any kind of secret source, but knowing the data that goes in is fairly basic stuff.

Q40 **Jon Pearce:** A quick question to all of you. I think I know what you are going to say, but it would be interesting to know if there are any positives to the fair use base regimes in America, Japan and Singapore? Are there any positives that you think we could draw on or lessons we could learn?

Max Richter: No, not from an artist's perspective. Speaking as an artist, not as a developer or part of a tech company, my interests really are just to do with my rights in the work I have made. I personally think, and I probably many artists would agree, that it should be an opt-in model. I think opt-out puts the onus on individual artists to police these gigantic multibillion dollar tech companies, which is a constantly shifting landscape. How can an individual artist possibly be expected to do that?

From my standpoint, there needs to be a verifiable, transparent value chain from the training data all the way through to the outputs. Any

monetisation of a work that is derived from training data by an artist should pay through to that artist. They are derivative works. I think there should be no obligation to opt in to that process. Artists should be free to either participate in it or not.

Matt Rogerson: To give you a sense of the scale of what individual web publishers are up against, we are working with a responsible AI company. They are called MESA.AI, and they work with publishers to build answer engines using AI. They are brilliant, because they are positive about its use. They found recently that there are 1,000 unique bots scraping data from publisher websites; that is across 3,000 publishers that they found. We do not know who those bots work with, but we know that they are working with AI companies. On average, publishers are being targeted by 15 bots each for the purpose of extracting data for AI models, and they are reselling that data to AI platforms for money. That is where the money is. Do not go into journalism; go into scraping, because that is where the money is. Then there are violations of robots.txt by companies that are mixed up with that. Hopefully, they will submit to this Committee.

There are a lot of economic reports out there. A lot of them are written by ChatGPT itself, so I would disregard those, but Daron Acemoglu and Simon Johnson, who are both in IT—Simon Johnson used to be the IMF's chief economist—say there are five key things you need to do to rebalance AI towards the consumer and citizens. The first of those is that the government need to establish clear ownership rights of all consumers over their data and should tax digital ads. On that one, this consultation would pass the ownership away and say that we can only assert rights after the event of the data being taken. On digital ads, I think this week it was suggested that we are going to drop the digital sales tax that was brought in under a previous Government.

The second thing is that they talk about rebalancing the tax system so that you have a more symmetric tax structure equalising marginal rates for hiring and training labour, and for investing in equipment and software. That is so the incentive for AI is not that you basically fire people and put AI in their place. If you want a complementary AI plus human approach, you need to think quite early about how you reframe the economy. It seems to me on those two measures alone, the UK is going in the wrong direction.

Q41 **Dr Lauren Sullivan:** The Government consultation highlights collective licensing as convenient for both creators and AI developers. Do you agree?

Sajeeda Merali: Absolutely. By retaining UK's gold standard copyright protections and by introducing robust transparency provisions, I think the Government can lay a really interesting dynamic licensing marketplace. We want the right to have it, but we would not necessarily want to mandate collective licensing. A lot of the licensing that has happened already has been direct deals between publishers and the platforms, but voluntary collective licensing should be a tool available to publishers, particularly the long tail of publishers. If you think about the utility curve

and you have 10 of the biggest publishers doing a deal directly with the LLMs, they probably have 85% of the content. That is the nature of the market. Collective licensing will allow some of the smaller publishers to play and be part of that.

It is worth saying that collective licensing already exists in the UK. From some statistics sent over to me by Publishers' Licensing Services, it currently facilitates the reuse of published content by 12.7 million students, over 9.5 million employees and over 40,000 businesses, schools and universities. It also provides access to over 8 million print and online publications, and over the last year it has distributed over £43 million in collective licensing back into publishers' businesses. There is a route to get there.

Max Richter: We are used to it, in a sense, already in music with the blanket licence, for example, like the BBC has for the use of music in broadcasts. In principle I am not opposed, but it is to do with how a thing like that is implemented and how that work is then valued and monetised.

Matt Rogerson: I think that collective licensing is good, especially for smaller publishers. This ground has already been looked at through the CMA and Ofcom, who did some work on a news media code and how that would look for the UK. It was not specific to a particular use case for journalism, but they said that there should be the ability to do both bilateral arrangements and collective arrangements. I think it would be good in this context as well. The challenge, more than anything, is knowing where your content has been used and who is using it. That, again, comes back to transparency of trained data, but also RAG.

Q42 **Dr Allison Gardner:** Going on to data transparency, Ed Newton-Rex has said, "What is the use of knowing your data is being used if you cannot do anything about it?" Sajeeda, is data transparency a compromise to allow for unrestricted training?

Sajeeda Merali: No, I think transparency is a part of the solution. You need licensing, you need copyright law to be applicable and then you need transparency to enable all of that. When we think about transparency, we probably think about it in three areas. You have transparency for the consumer, and I am just thinking about what has been said recently: AI platforms are big commercial organisations, and they are saying that transparency would be commercially sensitive. To draw a comparison, if you look at the fashion industry, I want to know what material a jumper has been made out of. If it is a lovely cashmere jumper, I know that it will have a particular value, and if it is polyester I do not expect to pay the same amount. An element of it is the consumer's right to know what they are digesting. It is the same as when you go to buy food, and you want to know what ingredients are in there. That is the consumer side.

When it comes to publishers and transparency, and understanding how they get fair value, if the AI companies are saying that the data used to train the models is commercially sensitive, I suppose the commercial

sensitivity that they are worried about is that publishers will ask for fair value. To be honest, we would like transparency on that one.

The final point on transparency is about the implications of any decisions that are made. We had a digital collective meeting where a lot of our digital product leaders discussed this today. One of the things that came up was that they are not sure what the implication will be if they opt out of things. If they opt out of Google for their large language model, does that mean they will be penalised from Search and therefore effectively worsen their commercial situation? It is a big topic.

Q43 **Dr Allison Gardner:** There are a lot of unintended consequences. I want to come to Max in a moment but, Matt, linked to some of the comments there, what would be the commercial trade-offs for AI developers if the Government were to introduce data transparency? Regardless of the approach, but in light of these unintended consequences, if we had an opt-out approach, it might skew the available data for their training. As well as the impact on artists, it might impact on the quality of data available to AI developers. What would be the trade-offs for AI developers?

Matt Rogerson: Ultimately, that is what you want to create—that the material that goes into an LLM is what comes out of it. If you train on rubbish—surprise, surprise—you get rubbish. The same is true when they are scraping on websites for live information. If they are scraping really poor websites, they are useless and they will go out of business. This is the value that we bring. The problem is that we cannot see where the value is, because we cannot see who has stolen our content, and we cannot even see who is taking our content in real time.

The two sides of this debate can work together. We are just at the stage now where these very large companies that usually make margins of 90% might have to take a smaller margin. That clearly will be upsetting for their investors, but that does not mean they should not. It is just a question of right and wrong and where we pitch this debate.

Unfortunately, I think the Government have pitched it in thinking that you cannot reduce the margin of these big tech companies otherwise they will not build a data centre. I think they are just wrong. Morally, and from the standpoint of improving investment in what we do, which I think is societally important otherwise I would not have worked in the industry for 10-plus years, it is possible to get this right.

Q44 **Dr Allison Gardner:** It is good that it does not need to be a binary debate. Max, as an artist currently, what do you do to work out whether bad-faith actors have used your work legitimately?

Max Richter: There is really nothing I can do. There is literally no way to do it. There are a couple of music AI models out there, and it is perfectly easy to make them generate a piece of music that sounds uncannily like me. That would not be possible unless it had hoovered up my stuff without asking me and without paying for it. I think that is happening on a huge scale. It has obviously happened to basically every artist whose work is on the internet.

In the longer term, that will lead to societal consequences. You will get a vanilla-isation of music culture as this automated material starts to edge out human creators. I think you are also going to get an impoverishing of human creators. It is worth remembering that the music business in the UK is a real success story. Its income was £7.6 billion last year, with over 200,000 people employed, and that is a big impact. It is the second biggest music exporter after the US. This is a big thing. If we allow the erosion of copyright, which is really how value is created in the music sector, we will be in a position where there will not be artists in the future.

Chair: That is a good point. Talking of which—

Q45 **Mr James Frith:** Thank you, Max. Matt said something before about acknowledging copyright as is but adding transparency to it. With that in mind, I will ask you, Max, for your views on copyright protection for AI-generated work. Do you think that the status quo is correct that such work should have copyright protection, or do you agree with the Government's instinct that these protections should be removed?

Max Richter: That is a very complex question. For there to be a verifiable situation regarding outputs, there must be a way to register and measure them and log that information. Whether that is an additional new right—these are really technical questions, and they would have to stand on some kind of legal platform. I am a composer, so I do not know quite how that would be done.

Q46 **Mr James Frith:** You referred to the distinctions between sectors. Do you think that the Government should consider sectoral opt-outs—that is unfortunate phrasing—or sectoral alternatives to opting out? Should we have a distinction between sectors because the creatives have not run out of ideas?

Max Richter: Exactly. I think that science, technology and the creative arts are different things. We value them differently and they have completely different economic models. They need to be treated differently.

Matt Rogerson: We should also have looked, through this consultation, at whether the non-commercial TDM exception was working as it is meant to. There has been evidence that academic research has then been repurposed for commercial businesses. Andres Guadamuz, who I think has probably given evidence to one or other of your Committees, talks about where web videos are being used for academic purposes and then suddenly they appear as training aids for Llama. That laundering of IP, which has been going on for many years, should have been looked into.

The second question is bigger: why are we here, and why are we making these changes? I think we are told that it is to advance humanity in medicine, and often the AlphaFold discovery is used as the example; we are told that AlphaFold is going to solve everything. As far as I am aware—I have done a little bit of reading into it—it is based on the protein databank. The protein databank is a public investment since

1971. It has paid public citizens and clinicians to code data. That has then formed this bank, which DeepMind, with AlphaFold, has admittedly done a good job of coding, but it does not solve everything because AlphaFold does not know structural biology. It exists in a very narrow context. It gives us a false sense that AI will solve everything and drive humanity forward, when in fact it is a probability machine, and it comes up with outcomes based on probabilities. It is not going to solve every problem.

Why does it need the full archive of *FT* content to solve cancer? It just doesn't. It feels to me as if this is all about how we water down copyright so that general-purpose AI can be used to fetch stuff from the internet and not give any remuneration to the website owner or the IP owner, so that it enriches the companies that sit in the middle. Just for context, Google invests about £1.5 billion a year on DeepMind. It invests £20 billion a year being the default search provider on Apple phones. If this was all about advancing medical technology in humanity, it would probably be investing more than that.

Q47 **Chi Onwurah:** You have each made very powerful arguments for the importance of copyright and the creative industries, the importance of the creative industries and the uniqueness of human creativity, and that relationship to copyright. Max, you say that you are an artist, not a technologist, and that technology and science and creativity and artistry are very different. They are integrated, and they are integrated in this question. In fact, they are integrated in all our lives. Our Committee emphasises that science and technology are part of all our lives, and that we need to be able to ask questions of them.

I want to try to summarise the implications of what you are saying for technology. I think you are saying that you want option zero, which is the status quo, with greater enforcement of copyright. That would mean, as I understand it, that the bots and the crawlers and so on would only take data from websites where it had been identified as being publicly accessible through the existing mechanism of robots.txt. Therefore, it is robots.txt plus some increased collective agreements where possible, and the AI companies should negotiate all those collective agreements individually. Is that what you are saying?

Matt Rogerson: These companies can go to Mars. They can read. Their scrapers can read machine-readable notices on websites, so that is not beyond the—

Chi Onwurah: That is robots.txt.

Matt Rogerson: There is a follow-on to robots.txt in the making. The IETF—

Chi Onwurah: TDM.AI?

Matt Rogerson: No, there was a working group called control.ai, which was changed because everyone realised that it does not control scraping; it just sets a preference, so it is called preferences.ai. That is in

development. There was a question earlier about the timeframe for development. It requires international collaboration, so it will be probably two or three years before that standard is determined.

With collective licensing, AI companies should come to businesses to license that collectively or directly. I can send you a list tomorrow of the licensing officers who work through all of the UK press—it is not hard. The NLA, the Newspaper Licensing Agency, licenses on behalf of the whole of the sector. You are asking if it adds friction or not. The Newspaper Licensing Agency does that already. It will license to media monitoring services and others. It is not about the volume of people that they have to talk to. It is because they do not want to pay.

Q48 **Paul Waugh:** I declare an interest by saying that I am a member of the National Union of Journalists. Matt, can I ask you about inferencing and artificial or synthetic data, which indirectly affects copyrighted works? How much of a threat or an opportunity is that?

Matt Rogerson: Inference is given very little airtime in the consultation, but every publisher I speak to is super concerned about it. The reason, especially with the two big ones—the two biggies, they are not start-ups any more—is that you have to give all of your data to Google and Microsoft to appear in Search, because they have to search your pages to get terms from it to give you a ranking in Search, so they have everything.

Unbeknown to website owners, they have been selling that as a commercial index to AI developers. That means that you can get the car off the drive because you have put some good fuel in it, and you can drive around for a bit. None of that money comes back to the publishers whose data is used in that inferencing. We also believe that there are a number of other providers of that service, although their identity is not known at the moment. What is disturbing is that when we have asked for the identity of who those services are being provided to, we are given no transparency whatsoever.

Q49 **Paul Waugh:** DeepSeek has put on the agenda the fact that OpenAI is, ironically, worried about its own IP being lifted and stolen. Beyond the irony, Matt talked about a Keir moment. Max, there was a Keir moment when he made an arts speech before the general election, where he said that he had studied flute at the Guildhall and valued the arts for the economy and for wellbeing. This is someone is unashamedly passionate about classical music. Is there a Keir Starmer who you want to appeal to in this debate?

Max Richter: Yes. I am biased, of course, but I think that music is one of the most valuable cultural expressions of human life. It has been with us for hundreds of thousands of years. It is an almost unique way to experience how somebody else sees the world. If somebody writes a love song, it touches us and moves us because that person has been in love and that has informed how they wrote that song. This is what music does. It allows a special empathy, a special communication between

human beings, so it is incredibly valuable. You do not need to be able to understand what somebody is saying, but a piece of music will speak to you. Yes, I think that it has a special value in human culture.

Q50 **Paul Waugh:** Matt, we heard from the tech representatives earlier that the DeepSeek example shows that there can be a viable British AI start-up that can outdo the big boys in America, because it can operate cheaper. Can you also flip that and say, "Yes, you can have a British start-up like DeepSeek, but it does not have to rely on breaching copyright?" In other words, it can be a world leader?

Matt Rogerson: Yes. Everyone is taking their own lessons from DeepSeek, aren't they?

Chi Onwurah: It was trained on an existing LLM and it breached copyright.

Matt Rogerson: That is the other thing. Let's kibosh the idea that DeepSeek and Llama are open source. They are not. Llama is trained on Russian pirated datasets of bestselling books. It did that because OpenAI did that. Then it tried to clean all the copyright markers off them so that nobody would be able to spot that it had done that. Neither are open source. These models are being put out into the market so that other, cleverer people can find new paradigms and change the way that you develop AI.

It felt like a wake-up moment when the week before you had the four or five men stood in the White House with their £500 billion project, and then genuine innovation comes at you from people who had resources restricted. In that context you do wonder how we can bring together the different parts of the supply chain in the UK to think about a mission for a UK version of DeepSeek. I think that content owners will be really willing to have that conversation, because we all believe in the value of the country that we live in.

It is just that this is all being done to us, and it feels as though the game has been given away too early. It is believed that Oracle, Microsoft and Google are the only people who can do this. There is no technological moat to somebody creating a new LLM that is better than theirs, and doing so in an ethical and responsible way.

Q51 **Steve Race:** I will ask a question of you all, on the basis that no question is a stupid question around this table. To what extent does it matter whether we get it right in the UK, given what is happening in China with DeepSeek and with other models with the US? The UK Government are trying to balance the third biggest global creative industries market versus the third or fourth biggest AI market. On the AI market side, there is the offer of huge investment and lots of jobs. On the other side, on a more restrictive model, there is protection, rightly, for creative industries. If we go for a much more restrictive system, will we lose all the jobs and it will happen elsewhere in the world? What is your view on the global dynamics around this?

Max Richter: From my standpoint, speaking personally as an artist, I do not think I would be massively opposed to opting into a system that valued the work properly; that is all. Copyright is intact, and there is transparency in what is being used and how the output is being monetised. I do not think that people would argue with that. It is not as if I am a luddite who wants to stop clocks. If something fair and equitable can be found, I do not think that people would resist that. That would lead to a high-quality situation with high-quality systems coming out of the UK.

Matt Rogerson: In every Parliament around the world, there are lobbyists from the tech companies and trade associations that work for tech companies, saying, "Japan, Singapore and the US have changed their laws to enable AI, so you need to do it too". However, when you look into it, you find that the US case is not settled, and Singapore's creative industry—I do not know how big it is, but it is not massive. In Japan the law is not as permissive as people suggest. If the resulting AI model competes with the original source material, that is in breach of copyright. A lot of the stuff that we are seeing and a lot of the stuff that Max has described would be competing with original source material.

There is a game being played at the moment, and the UK's position in that game is incredibly important. We cannot do anything about it as businesses. It is up to the Government, but it is also up to you. There is a vote next week in the House of Commons on the amendments that Baroness Kidron has brought forward. We could say quite early on that we believe in creativity and copyright and that we believe in transparency. If you are in the Chamber and you feel like standing up for what is right, I would urge you to vote for them.

Sajeeda Merali: From our perspective, we want to create a framework that allows humans and machines to co-exist in a way that is positive for humanity, society and the economy. However, this is about balancing a short-term growth agenda with long-term sustainability for the ecosystem. Many of us believe that the creative sector is a big part of that long-term sustainability. It is about finding a framework for both to go hand in hand.

Q52 **Martin Wrigley:** This has been very interesting, and I was fascinated that you used the iPad as an example. A lot of people say that that was stolen from the film "2001: A Space Odyssey" in 1968, which had Newspads in its early scenes. That in itself should probably still be in copyright, but there we go.

You talked a lot about transparency, licensing and copyright. We have talked a lot about the current state of AI, where we are in a mass of newly working, generative pre-trained models—GPT. We have yet to see real AI; we have yet to see actual continuous learning AI. Your suggestion, from what I hear, is that although the consultation talks of three scenarios of loosening copyright, you are talking about whether it should be looking at tightening copyright, and if so in what case, to cope with a changing AI technology. What are your suggestions?

Matt Rogerson: Explain the last bit again.

Martin Wrigley: The AI that we see right now is fleeting. In two, three, five or 10 years' time, it will be continuous learning AI. It will not be trained on, "Here's a language model, this is what you have learned and therefore this is what you can talk about". It will be continuously learning. It will be always fed data. It will not make the same decision twice because it is always learning from its past mistakes. We are in the realms of Voight-Kampff tests and trying to figure out if it is really AI. However, what we have is not the endgame. How do we look at copyright in AI in a changing world? Should we be thinking about adding to copyright law rather than reducing it, as perhaps the consultation is suggesting?

Matt Rogerson: I will be a Gary Marcus for a minute and say that promises of AGI have been made. We will see whether we get there. We can only deal with what we see in front of us, and what we see in front of us is people taking our content, using it for training and using it in substitutional ways. From our perspective, we will prosecute the same argument in every country where we operate and where we see our content being stolen. That is my response on that.

Max Richter: I think that it is impossible to tell. It is impossible to predict how this will develop. Copyright law works and it is pretty settled, especially in the UK. The UK has excellent copyright legislation. It has led to a healthy creative ecosystem, a healthy population of music creators, writers, publishers, recording studios, TV channels, festivals, you name it—an enormous, interconnected web that depends on that existing. It works. It might be necessary to make additional categories of works, which would be works that are derived from AI-generated material, but it is difficult to predict. It is a moving target.

Sajeeda Merali: I echo those comments. There are so many use cases as to how publisher content or creative content is used. It could be for training, it could be to display the results or it could be for fact checking. In the world of continuous AI, it might be something else entirely. With any licensing we need to be able to evolve it as well, but I echo all the comments that have been made already.

Q53 **Mr James Frith:** There is a big ticking clock on this, which I think has been started unintentionally. We risk backing ourselves into a corner—a tight spot—on an issue that is well intended for growth and embracing new technologies and new opportunities, but that starts by undermining one of our biggest strengths as a culture. If you were in a parliamentarian's shoes, what would you be doing?

Matt Rogerson: I would probably ask for a more detailed impact assessment than the Government have made so far of their proposals. The numbers that have been put out into the ether—the £400 billion—were mainly from a ChatGPT calculation that was done in a lobbying report for a tech company. If you look at the 1.5% growth that the IMF has suggested, it also suggested that the labour market would reduce by

5.5% as a result of that 1.5% growth. What is the cost of that, and who does it impact most? Again, if you look at the MIT work that is done by humans, it suggests that the people it will affect most are women in menial jobs. Therefore, who does it affect and how does it affect them? Also how does it impact on all of our industries?

More broadly, because our industries are not widget factories—although I love widget factories—what is the societal impact of stressing our sectors that are already pretty fragile? We are a successful news organisation and we are growing, but we are quite rare in the context of a market that is severely stressed. If substitutional products are coming through because the Government feel like they have to build a new data centre, what is the impact of that on holding power to account? I do not think that any of that has been considered. What has been considered is that the biggest tech companies have asked for a watering down of copyright law and the Government thought that that was an easy win.

Max Richter: The societal impact is really important. Thinking only about the music sector, because that is where I am from, it has already weathered an enormous earthquake with the onset of streaming, which has completely changed all the financials of music. The average income for a professional musician in the UK is £30,000. Over half of professional musicians in the UK earn £14,000 or less. There is no room for less money to happen. The sector is already fragile and basically exists on people's dedication and passion. I would be very cautious about adding additional stresses to that.

Sajeeda Merali: It boils down the fact that the fundamental principles of copyright exist. It grants the artists rights and tries to move the burden of those rights to the people who seek to reuse their work. All we are asking for is that the intent and benefits of copyright are preserved—we are doing that for the next era—and that the Government do not accidently change a gold standard UK copyright law that has existed since 1709 in the face of the seduction of new technology and the desire to seek investment into the UK. UK copyright is clear.

Matt Rogerson: Sometimes people think that these models are magic and that they can create anything. They are probability machines, so they just create material based on what they have consumed before. They cannot produce journalism or replace journalists. If you believe that journalism and journalists are valuable, AI cannot replace that. That is something to think about. Sustainability of news was a thing that we considered a few years ago, and which I care about passionately, but AI cannot be a journalist.

Chair: Very good—thank you very much. That brings us to our conclusion. I thank you all for giving us your thoughts today. Thank you also to our friends from the Science, Innovation and Technology Committee for helping us to broaden our horizons. We have to conclude that this does not need to be a zero-sum game. The creative industry is worth something like £125 billion to the UK economy—10 times what the AI industry is worth right now. We want to be able to somehow embrace

the benefits of AI, but balance that against the real benefits of creativity and copyright. We are grateful for your input today to help us build a better understanding, and both our Committees want to carry on talking about how we can best influence that discussion. Thank you so much.