House of Commons

Digital, Culture, Media and Sport Committee

Sub-Committee on Online Harms and Disinformation

Oral evidence: Online safety and online harms, HC 620

Tuesday 18 January 2022

Ordered by the House of Commons to be published on 18 January 2022.

Watch the meeting

Members present: Julian Knight (Chair); Steve Brine; Clive Efford; Julie Elliott; Damian Green; Simon Jupp; John Nicolson; Jane Stevenson.

Questions 261 - 298

Witnesses

I: Iain Bundred, Head of Public Policy UK&I, YouTube; Richard Earley, UK Public Policy Manager, Meta; Elizabeth Kanter, Director of Government Relations and Public Policy, TikTok; and Niamh McDade, Deputy Head of UK Policy, Twitter.

Examination of witnesses

Witnesses: Iain Bundred, Richard Earley, Elizabeth Kanter and Niamh McDade.

Q261 **Chair:** This is the Digital, Culture, Media and Sport Committee, and this is our Sub-Committee looking at online harms and disinformation.

We are joined for this session by two familiar faces from our last session: Iain Bundred, head of policy at YouTube, and Elizabeth Kanter, director of government relations and public policy at TikTok. We are also joined by Richard Earley, UK policy manager at Meta, and Niamh McDade, deputy head of UK policy at Twitter. Iain, Richard, Elizabeth and Niamh, thank you very much for joining us this morning.

I will kick off with you, Richard, as you are one of the new joiners. When it comes to what we could refer to as harmful but not illegal content, do you agree with the provisions of the Online Safety Bill? Do you agree that there needs to be effective regulatory oversight of harmful content, rather than simply illegal content?

Richard Earley: Thank you, Chair, and thanks very much for letting me join the session. In short, yes, those provisions of the Bill are effective and will be valuable. The overall structure of the Bill is—as the Government have pointed out and as we have said—quite ambitious in that it seeks to look at the entire sweep of harmful content that can be posted on to social media sites or the internet, from the obviously illegal through to those more nuanced types that are not necessarily breaking laws but can be harmful. That mirrors policies on the platforms that Meta runs: Facebook and Instagram. Policies called community standards prohibit a wide range of content and behaviour being shared on our services and go far beyond the local laws in the countries where we operate.

Chair: We will probably return to that particular issue in a second, but I would like to ask the same question of Niamh McDade in terms of harmful but not illegal content.

Niamh McDade: Thank you for the question and for having me here today. Twitter, on the whole, is extremely supportive of regulation that is fair, forward thinking, and protects the open internet and access for all. As to Richard’s point, we are similar. We welcome the provisions of the Bill and think that those will be valuable.

In terms of legal but harmful, we have previously called for greater clarity in this space. That clarity is necessary not only for us as platforms sitting here today but also more broadly, I suppose, for internet users as a whole. On that, we have welcomed the Joint Committee’s recommendations for more clarity in that space and we look forward to continuing to work with the Committee on that.

Q262 **Chair:** Niamh, how would you define “clarity” in this space?

Niamh McDade: I suppose clear, substantial definitions in terms of what we should be allowing on the platform and what Ofcom expects, not only to allow us to work well with Ofcom, but to ensure that the various terms, conditions and rules, which we as a platform have in place already, are in line with expectations and so work in a collaborative way in that space. To work collaboratively, we would benefit from clarity in that space as well.

Q263 **Chair:** Would "clarity" effectively be a list of what is deemed as harmful content as well as illegal? It is difficult to precisely define the legality issues as well, but it is probably easier to do so than defining harmful content. Would you like to see a checklist of what is harmful?

Niamh McDade: Not necessarily a checklist, no, because of course, as we evolve, as communities evolve and as discussions evolve around the world, what is harmful changes all the time. As a platform, we continue to keep mindful of that to enforce our rules and policies.

More what I am trying to get to here, rather than a specific list, is that collaborative piece: working alongside Ofcom, which we look forward to doing, on the provisions that should be outlined in the Bill and the specifics that we should be focusing on or looking at; and how Ofcom as the regulator believes we are acting as a platform, and how our current terms, conditions, rules and policies that we have in place are working for users and for the system as a whole.

Q264 **Chair:** Okay. Iain, I am turning to you for your view on what is harmful rather than illegal content.

Iain Bundred: Sure. We absolutely support a systems-based approach and the way the Bill has been set out; we agree with this idea that Ofcom, which will be a tough but thoughtful regulator, can hold us to account for our rules and approaches.

In terms of specific definitions of harms, we do not want to be slowed by legal but harmful elements and trying to cross-reference them with local definitions. We argue that if Parliament or the Government consider certain harms of such a level of danger, they should be made illegal. That is why we support, for example, the Law Commission approaches.

At the same time, at the core of this is a system Bill that looks at whether we are implementing our terms and conditions. We are fully supportive of that and very much behind it.

Q265 **Chair:** All right. You support a law provided that it works with your terms and conditions. I have to say that is a cart-before-the-horse approach.

Iain Bundred: Apologies if I gave that impression. You asked me whether we wanted clear definitions. We have clear definitions in our community guidelines of what we consider to be harmful. At the core of the Bill, certainly in terms of legal but harmful to adults, is this concept of whether platforms are consistently applying their terms and conditions. We support that.

Q266 **Chair:** Yes, but do you accept the fact that the Bill should effectively push your terms and conditions further and that it needs to be on the face of the Bill? It is not just about making sure you do your job properly or the job that you say you do; it is actually about the fact that we wish to see you become good citizens on the internet, to ensure that you do not promote what would be considered widely across society and Parliament as harmful content. Iain?

Iain Bundred: Absolutely. We have to meet our responsibilities and it is right that Parliament is debating exactly how you want us to do that.

Ultimately, as a platform, we have a responsibility to our users, our creators and our advertisers to make sure that we are actioning that out globally, irrespective of individual local rules. Therefore, we are working hard to make sure that we remove harmful content all the time, that we take the main steps on our platform to both remove clearly harmful stuff and reduce the exposure of borderline content, as well as raise authoritative sources to try to give better experiences for users around sensitive subjects, as well as all the steps around monetisation that I raised earlier and so-called rewards. Those are four steps that we take and there is a lot of information that I am sure we can give you at the end of the session around how we action against that.

In relation to the Bill, the point here is about a systemic regime where Ofcom is empowered to hold platforms to account. We are supportive of that.

Q267 **Chair:** Elizabeth, do you agree with the Bill's provisions in the area of harmful rather than strictly illegal content?

Elizabeth Kanter: It is important to say from our perspective that we have community guidelines, as Richard said. For Facebook, we have 10 different categories. That includes both illegal content as well as legal but harmful content. We do agree with the systems and process-based approach of the Bill. As others have said, we need clear guidance on expectations from Ofcom and how we can meet our regulatory obligations with regard to anything legal but harmful. As Niamh said on legal clarity, we need a Bill that we can implement and translate to our moderators so that when they look at a piece of content that has nuance, they can achieve our regulatory obligation to take content down when it violates our guidelines, and operate in an environment that recognises that there is always nuance in these legal but harmful categories.

Q268 **Chair:** How do you achieve nuance through an algorithm?

Elizabeth Kanter: An area that we look at with regard to legal but harmful content that is very nuanced is eating disorder content. With this kind of content, we need to be mindful of the user's experience.

For example, we have met with Beat, a UK-based NGO, to understand people who have had the experience. We talk with people who have survived eating disorders. They help us understand what kind of content on the platform might be triggering and what kind of content might be useful as other people are recovering from eating disorders. In terms of the way we identify our community guidelines, we try to take into account people with lived experience—for example, of eating disorders—so that we can inform our community guidelines and determine which content will or will not be on our For You feed.

Q269 **Chair:** Thank you. We are focused on eating disorders and that particular area.

Richard, looking at one particular example for you, financial scams have been widely discussed in front of the Treasury Committee and in the parliamentary debate that we had last week. We have a situation at the moment where, effectively, the industry has stopped user-generated content scams. However, other scams are still out there from paid advertising. Do you need to wait for legislation to do the right thing and stop people being robbed blind by a collection of fraudsters?

Richard Earley: I do not accept that we are doing nothing to stop people facing scams. You are absolutely right that the use of the internet to try to carry out fraud or scams is a serious problem that the people who use our services are talking to us about and expressing their concerns about. Ofcom last year found that scams and the hacking of data was the No. 1 concern of people in the UK when it came to using social media.

You mentioned both what we call organic content, which means posts put up by ordinary people, and advertising content by paid advertisers. For both of those types of posts, we already take a number of different steps to try to identify and remove fraudulent or scam activity. A lot of that can be to do with the use of technology to find and remove fake accounts because often people who are involved in this sort of activity use fake accounts to do it. Our technology is effective at finding and removing fake accounts, often even before they have been fully created.

On the advertising side, my colleague Tom was speaking earlier to you about how whenever someone wants to place an advert on one of our services, they have to go through quite a lengthy process, which includes of course complying with all the policies we have for that normal user-generated content—

Chair: Just to clarify, you mean your policies?

Richard Earley: Yes. I mentioned the community standards, which are our policies—

Q270 **Chair:** Sorry. Why is someone like Martin Lewis, the money saving expert, who does not have a great deal of hair but does have some hair, pulling it out with frustration at the number of times he sees his face popping up on your platform advertising a scam? This is paid-for advertising, not user-generated content. The idea that you are systematically over a period of years making money from our constituents' misery over being defrauded seems incredible to me and it probably seems incredible to the public. You are making money from that and you continue to make money. You are still waiting for legislation to come forward before you appropriately act to exclude these scams from your platforms permanently. We see the evidence all the time. What do you say to that, Richard?

Richard Earley: Again, I do not agree that we are waiting for legislation to act on the side of advertising any more than on the organic side. Our system ensures that when someone wants to place an advertisement on one of our platforms, the advert goes through a review for compliance with both our ordinary policies and our higher advertising policies and—

Q271 **Chair:** Sorry. Does that include being FCA-authorised? Do they have to be 100% FCA-authorised before they can advertise on your platform?

Richard Earley: That is exactly a step that we announced at the end of last year that we will start taking to ensure that when it comes to being—

Q272 **Chair:** Sorry, Richard. Let us get this straight. How long have you been carrying adverts in the UK? Many years. During that time, people have lost thousands of pounds. Some people have lost their entire livelihoods. Frankly, people have committed suicide as a result of the scams. You have continued to take advertising throughout the entire time from organisations that are not FCA-authorised. Only now are you pulling your finger out and bringing in that crucial change. Richard, you have not done enough.

Richard Earley: I completely agree with you about the serious harm that scams and frauds can cause to people in the UK. We have been working to try to tackle these uses of our platform ever since we first began to run adverts on our platforms many years ago.

Certainly, a big challenge we face is that fraud and advertising scams are, by their nature, designed to be difficult to tell apart from authentic advertising. That is why we have invested so much money in not just the pre-upload review that I mentioned, and our ability to see how scams are performing and use technology to look at any inauthentic signals from the producers of the adverts themselves.

The announcement we made last year is another step in our work towards tackling this problem. As I mentioned before, this continues to be an area where we continue to work, but it is one where there is a serious challenge not just for social media, but for parts of the economy.

Chair: I will be honest with you. This goes to all social media platforms, frankly. You ought to pay back the money that has been defrauded off the British public over many years while you have been taking adverts that are not FCA authorised. In the dim and distant past, I used to work in the newspaper industry. At that stage, we used to have mail order—

MOPS, it was called. You used to be able to see precisely whether or not a company had a deposit; whether or not it was, therefore, authorised to trade; and whether or not it could effectively advertise in your newspaper. If you found that there was any sort of linkage or the idea that these companies could not do so, you phoned each other up. You got on the phone and you told each other, "This company is dodgy. Do not take their adverts".

I do not see anything in what you have just said there—in fact, I know across the piece, across all your companies, none of you has effectively even shared this data and this information with each other. You do share it to the police—they are already overworked—but you do not do it with each other. Secondly, as a result, you have for many years taken money from these scam artists and then not made sure of the one thing you could do to stop it, which is to prevent anyone advertising with you who was not FCA authorised. Personally, I think that is a disgrace and it has been going on for far too long.

Each and every one of your companies should refund the British public any scam money they have had removed. That is not a question. I have to say that that is an observation at this point. Anything you say in that regard—I just think you have not done enough over a long period of time.

Q273 **Julie Elliott:** Hello to the new panellists who have joined us today. I want to move on to something that happened last year. A coalition of 60 children's charities around the world wrote to what was then Facebook stating that disclosures from *The Wall Street Journal* raise substantive questions about how Facebook identified and responded to harms on its platforms. What are you doing about this to reassure the sector? This is a question for Richard.

Richard Earley: Thank you very much for the question. On that letter specifically, we have written back in extensive detail to those organisations that wrote to us. They include some of the organisations that we work most closely with throughout Europe and elsewhere in the world.

Julie Elliott: Are you prepared to share that letter with this Committee?

Richard Earley: Yes, indeed, I would be very happy to share it with the Committee. As I was saying, we have already worked with a lot of those organisations over the last 10 or 15 years, when we are discussing the ways that our policies and our tools should be built to protect young people and others on our platform.

The challenge we face is that constantly the way that people use technology is changing and the technology itself is also changing. We do have to continually be reviewing and carrying out research on our products and how people use them so that we can identify new ways that we can tackle emerging threats. A really important part of that is the expertise that we get from partners such as those who wrote to us last year.

Q274 **Julie Elliott:** One of the recommendations that was set out in that letter from the children's charities was calling on you to publish your internal research, reputational reviews and risk assessments. Have you done that?

Richard Earley: Yes. As we set out in the response, we have published a tremendous amount of our internal research. In fact, we have a research division within Facebook—within Meta, as it is now called, I should say—that has published hundreds of individual papers since 2018 at least. In addition to that, what we believe is essential is not just the work that we can do internally to understand how people are using our platforms, as I said, but also how we can partner with academics and other researchers outside of our organisation to support them in their research. In 2021, we took part in more than 300 peer reviewed academic research papers, which were published in the course of that year.

Q275 **Damian Green:** Good morning to our new witnesses. I want to ask about Ofcom's powers under the proposed new regime and how they might affect you. Richard, your evidence to the Joint Committee suggests on reading it that you want to limit what evidence Ofcom could request. Why?

Richard Earley: I am not certain if that was a recommendation that we made in our submission, either to the PLS Committee or to yourselves. What we did say is that, of course, the powers that Ofcom has are very broad and we will need to make sure when we are collaborating with it that we are also respecting all of the other legal requirements and our obligations to our users, in addition to Ofcom itself.

On the powers that Ofcom is granted in the Bill and the role that Ofcom will play in the Bill, we think that that is at the heart of the reasons why this Bill can be, if it is properly implemented, very effective in bringing additional public confidence in the work of internet companies to keep the people who use their platforms safe, and also to involve greater expert experience, understanding and oversight of the work that we do.

Q276 **Damian Green:** To point you to what you appeared to be saying, you said you were, "currently working with some of the leading academic institutions to figure out what the right rules are to allow access to data in a privacy-protective way". Whose privacy are you trying to protect?

Richard Earley: Apologies, I did not quite catch what you were referring to the first time. Yes, what we did recommend there is actually something that the PLS Committee has taken up, which is that the Government should ask Ofcom to begin immediately a report, which is currently in the draft Bill, that would look at how platforms such as ours that hold data on behalf of our users can better collaborate with external researchers to enable them to fulfil their research objectives using the data we hold.

I just mentioned to your colleague that we have participated in around 300 externally reviewed studies in 2021, but those studies are often very difficult to agree and to move forward with, because there is a very complicated interplay between the privacy obligations that we have to the people who use our services, and our responsibilities set out in law in this draft Bill to ensure that we are taking steps to protect them that is in line with the latest understanding and research being carried out by experts. Our recommendation is that Ofcom and the ICO would be excellently placed to lead some work looking at ways to reduce further any barriers to that sharing of information.

Q277 **Damian Green:** The principle on which you are basing this is that the information should be shared as long as it is kept safe; is that right?

Richard Earley: Not quite. Whenever somebody uses one of our services, they do so in accordance with our privacy policy, which is part of the terms and conditions that anyone signs up to when they first use one of our services. That policy places on us very legitimate and understandable restrictions on how we can share their information or the information that we collect about them with external third parties, for very good reasons. We have seen in the past attempts to use this sort of data to carry out harm. The issue we face is how—

Damian Green: Not by Ofcom, you would not be. I think that we can trust Ofcom enough not to use data in a harmful way.

Richard Earley: That is correct. Here we are talking about partnerships with academics to help them study the ways that people use our services.

Damian Green: Yes. Presumably you would only work with academics who would not use it in a harmful way, I would hope. I am not clear who in this hypothetical situation is going to cause the harm.

Richard Earley: Over the last year or two there have been several very high-profile cases, where academics have called on platforms such as ours to provide greater information and access to the data that we hold about our users, so that they can study the way that different harms, or risks of harms, might manifest on our platforms. It is unblocking those challenges that come from the clash or the tension between our obligations to protect our users' privacy and our desire to make sure that we are advancing understanding across the industry of how to tackle harm that this report would seek to address.

Q278 **Damian Green:** You will know that one of the groups on the other side of the argument, or in a different place in this argument, is the ISPs. They think that in some way, if Ofcom cannot get the information from you, it will go to the ISPs, which already have to provide a lot of information. Do you recognise the force of their worry?

Richard Earley: I am not familiar with that argument, I must be honest. When it comes to the information that we will be providing to Ofcom under the terms of the Bill as it is currently drafted, as some of my colleagues on this panel were saying earlier, that is essentially the type of information that will enable them to assess and hold to account the work that we do to meet our obligations to our users, both in the ways that we build and design systems and products that people use, and the systems that we use to try to protect them and find harmful content.

That is a subject of tremendous investment by our company. We have spent more than $15 billion on building these sorts of technologies over the last 10 years. Enabling Ofcom to have a greater oversight of that and to bring its expertise to it will certainly help, as I said before, to increase the public confidence in the system and help us discover new ways of reacting to emerging threats.

Q279 **Clive Efford:** Richard, sorry to focus on you, but I want to follow up the answer you just gave regarding Ofcom's access to data. You seem to be saying that, in order to protect the privacy of your users, you are concerned about people who are undertaking research being able to use that and treat that information in confidence. Then you seem to transpose that over to Ofcom. What is your concern about Ofcom having access to information that it needs to perform its role as a regulator?

Richard Earley: I apologise if I have confused the Committee in my answers. My series of responses about the report regarding researchers' access to information held by platforms was about that recommendation, interactions between platforms and external researchers. When it comes to Ofcom, the powers that are granted to Ofcom are, of course, in keeping with the role that it will be required to do, which is to assess or to verify the risk assessments that we carry out, to seek extra information from us about the ways that we develop policies and to build tools to enforce those policies, all of which I think, provided that the powers granted to Ofcom are in keeping with what is in the Bill, we support.

Q280 **Clive Efford:** That leads to the second point that my colleague just made, which is about the ISPs being concerned that they will have to shoulder the burden of providing the information because you will not. You are saying that if any request comes from Ofcom that is compliant with and in accordance with its powers of enforcement, you would just comply—there would be no second guessing.

Richard Earley: Again, I cannot speak to what the internet service providers are concerned about, but certainly with this legislation, as with any legislation, when we receive a request from a Government or a regulator for any user data or our own data, we assess it against local law and comply to the extent that we can. I am confident that that will be the case with this law as well.

Q281 **Clive Efford:** This question is open to anyone but I will start with you, Richard, if you do not mind. *The Wall Street Journal* raised concerns about a lack of a culture of compliance among big tech companies in areas such as platform design and when addressing illegal content. Since that was raised, what steps have you taken to address that point?

Richard Earley: I am afraid that I do not recognise that characterisation of our company at least, as not having a culture of compliance. We engage very proactively with many regulators and policymakers in this country, as we do around the world. What we have been saying for many years is that we would welcome greater guidance through legislation from Governments such as the UK about how to carry out those different tasks that I described of building services for our users and helping to make sure that they use them in a safe way.

Q282 **Clive Efford:** You know the articles that I am referring to when I talk about *The Wall Street Journal*. There were a number of findings about Facebook, in particular, and the issues around whitelisting. I will not list them all because it would take me a long time, but it was quite a number. I am surprised that you say that you do not recognise that characterisation, because these seem to be factual issues that have been reported by *The Wall Street Journal*. My question is: what steps have you taken in response?

Richard Earley: Let me start by addressing the claims that you mentioned from last year. Many of the claims that have been made about our company, based on a selection of leaked documents, seriously misrepresent the seriousness with which we take the duties I have just been describing to keep our users safe. We have no commercial or moral incentive to do anything other than provide the maximum number of people with the most positive experience on our platforms. The people who use our services do not want to see harmful content. If they see it, they stop using our platforms. The advertisers who provide the money that we use to run our services also, very vocally, do not want to see any sort of harmful content on our platforms. We saw this very visibly in 2020.

The reality is that every day people attempt to use our services to carry out harm towards our users. Every day, the types of techniques they use to do that and the technology that they use to do it changes and evolves. We have a responsibility to continue to study and understand any weaknesses or risks that our services produce. It is that research that underpins the documents that have been published. As I said, we have continued to make several of those documents public since those claims were made, and we will continue to publish research on the steps that we are taking to tackle the issues raised.

Q283 **Clive Efford:** Can I come to you, Niamh, with a similar question? What steps have you taken, or did you take any steps, in response to the issues that were highlighted in a number of articles about compliance by big tech companies and the design of their platforms? Have you recently taken any steps to address the issues?

Niamh McDade: Thanks for the question. The specific article you are referencing is not something that I have been across, so I would be happy to come back to the Committee with any information regarding direct action that we have taken in relation to it. It is not something I have been specifically across.

More broadly, on your question around compliance, as a platform we are constantly evolving our rules and policies in terms of the environments that we are working within and the challenges that are faced by our users and communities across the world. The top line here is that, as a platform, we have absolutely not been waiting, for example, for the role of Ofcom to come in to take action and ensure that our rules and policies are effective and enforced as appropriate.

We continue to report transparently on this through our transparency reports, and looking into, I guess, the enforcement of our rules and policies, but also the culture of the platform and how it works and plays out for individuals using it. We provide access to the Twitter API to tens of thousands of researchers—and we have done for over a decade now—providing a wealth of research with it; hence, why you will often hear researchers reporting about Twitter, as well directly.

On the whole, our purpose is to serve the public conversation. To the point that Richard made, as a platform we lose out when there are negative conversations taking place, when it is not a platform that is conducive to the health of the public conversation. That may not always be people in agreement, that may not always be nice conversations, but conversations that are within the terms and conditions, and the rules of Twitter as a platform. Ensuring that we are compliant to our rules and policies but also working closely with regulators in place—indeed, we look forward to working with Ofcom in this space—are key priorities for us.

Q284 **Clive Efford:** Iain, do you have anything to add?

Iain Bundred: Yes. First, I think that transparency is important in this area. YouTube publishes a transparency report that sets out its removals but also its appeals—where we got it wrong—including both successful appeals and not. I do not think it was particularly in relation to that press cycle, but more generally we have introduced a new metric since I last spoke to the Committee, which is our violative view rate metric. That looks at the number of videos that we have removed that have subsequently proved to be violative, so it has been served to a user. That gives you a sense of the number of views that are content that should not have been served.

In that case, that figure fluctuates but it has been coming down over time. On the point around whether our enforcement efforts are working, if we look at our VVR in the last quarter, it was 0.9 to 0.11. That is nine or 10 or 11 views in every 10,000, and that is trending downwards. We work towards that. That number will of course fluctuate, but we have sought through transparency reporting to clarify where we are and how users are potentially being exposed to these sorts of harms.

Elizabeth Kanter: I would add that we, of course, comply with local laws and customs. The culture of our company is built around safety and this safety-by-design approach is what guides our principles in everything that we do. With regards to transparency, similarly to YouTube and other platforms, we publish transparency reports that go into the way in which we enforce our community guidelines. We also publish reports about the requests that we get from Government for user information or user data to provide further transparency to our users.

I am also not overly familiar with the articles that you mentioned, but, broadly speaking, I think that we are in line with what the others have said about approaching both the current regulations and the future regulation. We will comply with what comes in once the Bill is in force. I am happy to come back to you on those articles if you would like to share them with me.

Clive Efford: All right, thank you. I will leave it there.

Q285 **Simon Jupp:** Good morning to the new panellists and welcome back to the ones who were on the previous panel. I want to discuss what you guys do cross-platform when it comes to identifying and addressing instances of child grooming. That could be either systematic or on a case by case basis. Could I start with you, Richard?

Richard Earley: Yes, absolutely. I hope that it goes without saying that these attempts to use our platform for harm are some of the most disturbing and horrendous that we see. As I am sure you would expect, we have a range of approaches that we use on our platforms and work that we have done outside our platforms, too.

On the platforms that we operate, we have, of course, very clear policies about what kinds of behaviour and content is not acceptable on our services. Importantly, we draft those rules in close collaboration with people who work directly with survivors of this sort of activity, and with law enforcement and those who are trying to intercept it.

In addition to that, these policies mean nothing if we do not have tools to enforce them. We have invested a tremendous amount of time and resources in building state-of-the-art technology to find instances of this behaviour taking place, looking at both signals from what people post and behaviours of users on our services, and in removing them from the platform.

Something that goes to the way that we believe that industry should collaborate more closely together is the classifier, as we call it, which is a name for the technology that we use to identify adults who might be engaging in behaviours that are suspicious or potentially in breach of our rules. We made that classifier open source a few years ago so that any company that was looking to implement similar safety controls on its own services could build on that classifier to protect itself.

Taking another step, two years ago the industry, under the umbrella of the Technology Coalition, launched a project called Project Protect, which is about bringing together more than 20 organisations and tech companies. It has brought together research funding to fund research into these sorts of activities and to collaborate or move forward on the technology that I described to find and remove these behaviours, and prevent them from happening.

Q286 **Simon Jupp:** Could I just interrupt? Can you explain what you mean by "open source"? What is that data? How is it accessed by the platforms—without giving too much of it away because we do not want everyone to access this information? How does that also align with GDPR rules?

Richard Earley: I am sorry for using a jargon term. It is too much time working in this company.

By "open source", what I mean is that we have essentially a program, an amount of code, that we use to scan different surfaces on our platforms. It is effectively like a machine that runs on our platforms in order to find these behaviours. We published, in a complicated code format, the underlying instructions that that machine uses to identify, find and surface grooming behaviour.

There is certainly no obstacle to more people finding it. It is fully available on our website. I would be happy to share with the Committee the link to both the page where we describe this work and some of the examples of how it has been used.

Q287 **Simon Jupp:** Forgive me for perhaps not understanding; maybe I need some more coffee. This is not suggesting that you share the data of specific users who have been seen to break these guidelines and engage in child grooming, for example?

Richard Earley: That is correct. Where we find someone, for example, sharing sexual exploitation and abuse imagery of children, we report that directly to the National Centre for Missing and Exploited Children in the US, which then works with international partners to provide that to law enforcement.

Simon Jupp: Thank you. The same initial question to Niamh.

Niamh McDade: It is an extremely important point to raise. Much of the industry collaboration here is aligned with that of the other platforms, starting with the Technology Coalition that Richard mentioned. We work closely with ISPA. We also work with the National Centre for Missing and Exploited Children, as well as the Internet Watch Foundation within the UK.

The top line for us as a platform is that any CSEA material is completely prohibited across the platform, and we proactively work to protect children around the world on this issue with different partners within different areas that we work in across the globe, in a variety of different languages as well. We have had policies in place specifically covering this for a long time and a specific reporting route for this, too.

To give you an idea of the scale of action taken here, tens of thousands of accounts are actioned every year. In fact, in the last reporting period, from July to December 2020, we had almost 500,000 reports actioned in relation to CSEA rules and policies. We are proactively taking enforcement here where anything is reported. Yes, we take action, but there is always, in addition to that, proactive measures in place to ensure that this type of material has no place across the platform.

Simon Jupp: Iain, can I come to you?

Iain Bundred: Obviously, this abhorrent material is exactly where platforms need to work together. Niamh just mentioned some of the ways that we share information through NCMEC—the National Centre for Missing and Exploited Children—the Internet Watch Foundation and so on. At Google and YouTube, we look at this on a four-step basis.

The detection point is important and we work hard around technology to try to bring the best-in-class tools. At YouTube, it was our engineers who created the hash-matching software called CSAI Match. Where previous CSAM content has been uploaded in the past, we use digital fingerprints to identify that and we share that with other platforms through the NCMEC system.

We also have machine classifiers that can identify never-before-seen CSAM imagery. In H1 2020 our partners used that content safety API to classify more than 6 billion images. That is technology that we make free to use for all platforms, including the 24,000 businesses that will be in scope. They can all access this if they want to. It is really important that we work on the detection point.

The second step is around reporting, and I have mentioned already some of the work through NCMEC, which then also reports to the likes of the National Crime Agency. Thirdly, we go to deter, thinking about how we can put in place products and systems on all Google products to make sure that we are deterring this sort of activity, as well as, finally, collaborating with all the platforms to learn of the risks and challenges.

As you have heard in evidence in this Committee, I do not think that this is a problem that is going away. In the UK alone we saw the child safety policies on YouTube being the No. 1 issue for removals in our last reporting quarter. Around 33% of the 60,000 videos removed in the UK were for child safety breaches. This is a real challenge that we have to keep working on, and it will be through a mixture of collaboration across the platforms, working with law enforcement agencies and, of course, continuing to innovate technology in this area.

Simon Jupp: Elizabeth, finally to you.

Elizabeth Kanter: We take a three-step approach to this area. First, we start, of course, with prevention. We have a zero-tolerance approach to CSAM, CSEA and grooming, per our community guidelines. We have expert teams within our trust and safety team. Our head of trust and safety sits in Dublin, but we have a team globally looking at this. We have trained experts who can look for predatory behaviour in our app. They will look at those signals, remove that content and, importantly, report anyone who violates our community guidelines by posting any content that may be grooming, CSEA or CSAM.

The other thing that we do is we also have an innovative approach to our direct messaging. We recognise that people who want to engage in predatory behaviour and grooming might use direct messaging to engage with younger users and try to take them off our platform on to another platform. No direct message can be sent between two people who are not connected. To send or receive a direct message you must be connected to the other person. We do not allow unsolicited messages to be sent. We also do not allow under-13 to 15-year-olds to send or receive a direct message. We think that is important. Lastly, we do not allow any sort of attachment other than comments or a TikTok video to be sent by direct messaging. We think that is a strong approach in preventing grooming through direct messaging.

Lastly, on the point that others have made about collaboration, we are also participants in the Technology Coalition. We are involved in the Internet Watch Foundation. I am proud that I sit on the board of the Internet Watch Foundation. We have a holistic approach to this—tools on the platform, things that we are providing to empower our users and protect them, and then, importantly, this aspect of collaboration.

Something that I think has not been mentioned yet is the importance of taking a cross-industry approach to this issue, whether that is with law enforcement or the tech platforms, remembering also the whole value chain of players in this space. We need to involve those smaller players that might be more prone to bad actors taking advantage of their platforms, and ensuring that they are doing the right thing as well. We need to bring those smaller players into the conversation.

Q288 **Simon Jupp:** Thank you all for outlining what you do and the processes you have in place. According to the Children's Commissioner, it does not seem like it is enough. The Children's Commissioner is asking the Government to add duties into law to address these concerns. Listening to the answers from all of you, there is perhaps more that could be done cross-platform, if it was possible to share more data—if it was applicable and was allowable by law, for example. What do you think could be enhanced in law to protect users and to carry out what you are trying to do to protect them better in the future? I will ask for brief answers given time constraints, and I will go to Iain first.

Iain Bundred: The point I would make is that the age appropriate design code has only recently come into force. I saw that the prelegislative scrutiny Committee recommended that Ofcom review that interaction. I think that is important, because a number of the changes that platforms have made to be compliant with AADC has driven up the experiences and the work that is being done.

Of course, we are not waiting for regulation on this. I have already mentioned the great work being done specifically on child sexual abuse. You can always do more, but I think that we are all working very hard on this. The point, I guess, as we look ahead to the Online Safety Bill, will be how we ensure that all 24,000 businesses in scope are putting in place high standards in order to keep children safe online. We are all supportive of that.

Niamh McDade: Thank you for the question. I think that the important point here is on that collaborative piece. It is about how we can work together as an industry. What we do not want to see is bad actors coming to one platform, being pushed off by the rules there and moving on to another. That is exactly what we want to protect against. It is about working in collaboration with the platforms here, but more broadly, as Iain mentioned, the wider ecosystem of businesses impacted. That will be affected by the Online Safety Bill, the age appropriate design code, and the Information Commissioner's Office. It is looking at this as a wider piece—how we join up the dots, and how we join up our approaches in a way that ensures that bad actors are not getting through the cracks and making their way on to platforms or, indeed, the internet.

Elizabeth Kanter: It is a similar answer in terms of sharing best practices between platforms about how they detect predatory behaviour and look for signals from grooming gangs or other actors. The other thing that we know the Technology Coalition does, which we participate in, is that it conducts threat assessments about either foreseeable risks or known risks. That is potentially something that we can be doing on a more proactive and regular basis with others in the industry. There is a lot of great work being done at the moment through different organisations, and perhaps it is just trying to do that on a more regular basis and being more proactive about looking at the risks that may come up in the future.

Richard Earley: On your question about what legislative changes could help in this area, it is worth pausing to remember that there are a significant number of legislative changes either taking place right now in the UK, such as the Bill we are discussing, or that have recently entered into force, such as the children's code, or that are foreseen for the next year or so, such as the Government's advertising programme.

A valuable step forward in helping companies and regulators to manage this landscape was the creation a year and a half ago, I think, of the digital regulation co-operation forum—the DRCF—which brings together the heads of the relevant regulatory organisations that have oversight of these different frameworks. I know that it is already meeting regularly, but greater collaboration with us as the tech companies to help us to understand how its thinking is evolving and how it might see opportunities that we have not seen would be a beneficial step for all of us.

Q289 **John Nicolson:** Good afternoon. Can I address a few questions to you, Ms McDade from Twitter? We all agree, don't we, that the racist abuse that is sent to black footballers is repugnant? What has Twitter done to restrict it?

Niamh McDade: Thank you for the question here. You are bringing up a very pertinent issue and something that is a key priority for us as a platform.

There is no place for racism on Twitter. Unfortunately, we have seen that. We have rules and policies in place to protect users against seeing racism, and users using racist terms and slurs in the first place. We are constantly evolving our policies in this space to keep up to date with bad actors and to work around this. Rest assured that it is a priority for us as a platform, and we take this issue as of utmost importance and utmost seriousness.

Q290 **John Nicolson:** Of course, one of the problems is that Twitter has adopted a definition that allows Mickey Mouse, anonymous accounts to be designated not anonymous. Can you tell us what the verification process is that you do?

Niamh McDade: Yes, absolutely. When you sign up to Twitter you are asked for three pieces of information. The first is your date of birth. The second would be an email address and/or a mobile phone number. Then you are asked to verify each of these pieces of information when you sign up to an account.

Q291 **John Nicolson:** Okay. Let's examine that. Clean up the Internet created an account. I will tell you the name of the account and let's check that verification process that you do. The name of the account was "Mickey Mouse", the date of birth given was 2001, and the email address that was offered was the following: mickeymouseisnotreallymyname@gmail.com. The phone was a pay-as-you-go SIM card, which cost 99 pence. It was bought solely for the purpose of creating this account. Just in case your ever-vigilant guys in Twitter were not completely across the case, "Mickey Mouse" decided to send a tweet to Marcus Rashford. The tweet, which I am holding up here, said, "Squeak". Your system is not really working, is it?

Niamh McDade: Thank you for sharing that example. It is not one that I have been across myself. Please, I would urge you to send that over for review and we can certainly follow up with the Committee on what action we have taken there.

Q292 **John Nicolson:** Come on, come on, come on—it is so ludicrous. I can see some of you trying not to smile at the sheer absurdity of it: a Mickey

Mouse account entitled "mickeymouseisnotmyrealname". There is no point in thanking me for offering it up. There is something deeply flawed at the heart of your system that somebody can take the Michael out of you to such a ludicrous degree and you have not even heard about it. You are turning up here at a Select Committee not knowing about this, despite the fact that it was sent to Marcus Rashford, the country's best-known black footballer.

Niamh McDade: As I have said, thank you for raising it. I would be very happy to look into this with the relevant teams. They may already be across it as well. I myself do not work on the content moderation team. We have a dedicated team that looks into this and works with specific tweets similar to that in question.

The broader concern here that you have raised, and it is one that I know we have discussed in detail with this Committee in the past, is around anonymity, pseudonymity and the ability to use a pseudonym on the platform. There are a number of reasons that we want to protect the ability to use a pseudonym across the platform. Of course, there are a minority of bad actors out there who do use anonymous and pseudonymous names in order to—

Q293 **John Nicolson:** They were not bad actors, that was a good actor pointing out the absurdity of your system. Of course, the point remains that what is wrong with your system—apart from the fact that it is just hopeless, chaotic and does not work—is that you have tried to manipulate the description of an anonymous account. Somebody like that, Mickey Mouse there, is designated as a non-anonymous account under your current system. Your system does not work.

Niamh McDade: With that specific account, there may well be indicators of who that individual is. As I mentioned, a date of birth, a phone number or an email address will be provided. It is also worth flagging that using a pseudonym on the platform does not always equal abuse, just as using a real name does not always equal someone not being abusive across the platform.

Q294 **John Nicolson:** I did not say it did, but clearly it often does. It frequently does. Black footballers, in particular, are being targeted by the most terrible abuse. If you cannot protect somebody like Marcus against this kind of abuse—a Mickey Mouse account calling itself "Mickey Mouse" and sending a message, "Squeak" at you—what hope is there for all the anonymous folk who just get the daily misery of abuse if they use your platform?

Niamh McDade: Abuse and racism is a clear violation of the Twitter rules and policies.

John Nicolson: We know that, in theory but not in practice.

Niamh McDade: As I have mentioned, I do not have the details of this specific case. I would be very happy to follow up in more detail with the

Committee and yourself on this. It is just not something I have further information on at the moment. I can say that abuse and racism are a clear violation of the Twitter rules and policies. Where we are made aware of this and where we come across it proactively, we do enforce it. We have engaged with a range of other mechanisms to help support users as well.

Q295 **John Nicolson:** Twitter has been saying this for years. Let me give you another example. Simon Fell MP is the Conservative Party Member of Parliament for Barrow and Furness. He questioned Twitter's head of UK policy when she appeared in front of the Home Affairs Committee recently. He asked the same question: "What qualifies as anonymous?" The response given by Twitter was that 99% of the accounts are verifiable. That means that they have provided at least one and in most cases two pieces of personal information. She defined these as full name, date of birth, email address and phone number. She said that you have to verify to get on the service.

In order to test that, an account was set up under the name of Simon Fell MP—again, another ludicrous example of how badly you do at this. The account was created, "Simon F Fell". The date of birth was given as 9/11/01, making the MP 20, which I think you will find is unusual. I think that Mhairi Black is the only one who has reached that threshold since the 17th century. The address "fellsimon@gmail.com" was offered and no phone number at all was taken up, which allowed "Simon Fell" to tweet Simon Fell.

Niamh McDade: Impersonation, which sounds like what you are referencing there, would again be a breach of the Twitter rules. I cannot speak to that specific case directly, but it does sound like a case of impersonation. More broadly, if that impersonating account was again to breach Twitter rules—

John Nicolson: It certainly is a case of "No ! Sherlock". It is obviously a case of impersonation. It was set up to be a case of impersonation to show how completely inadequate your systems are. The person impersonating, who again was a good faith actor doing it in order to demonstrate the hopelessness of Twitter's creaking apparatus as you move towards legislation that will finally put some controls on Twitter, did not even have to add a phone number—not even a phone number that you can buy for 99 pence down at the local newsagent. They did not offer any phone number. They were allowed to tweet impersonating an MP. I have had the experience of that. I have had people tweeting using my name. You do not do anything about it.

I read the guidelines, which are very clear, but you do not actually implement them. You just write back saying, "This does not violate our community standards". I do not know how long it is going to take for the penny to drop. That is why MPs cross-party, from the left to the right, have finally decided that you guys need control, because we have lost faith in your ability to self-control. Okay, no answer to that. Back to you, Chair.

Chair: Thank you, John. Finally, Steve Brine.

Q296 **Steve Brine:** It is always a pleasure to follow my colleague Mr Nicolson, who has highlighted some holes in the colander. Let me see if I can find another one.

We have heard from the NSPCC and the Internet Watch Foundation about children unable to report images of themselves because the reasons for their distress in those images and their parents' distress do not meet the legal thresholds. Should people be able to withdraw consent when posting their images then causes them distress? I will start with you, Richard.

Richard Earley: I am very sorry; I am afraid I did not quite understand the question. Could you repeat it?

Steve Brine: What I am asking about is the activity known as breadcrumbing. Do you know what breadcrumbing is?

Richard Earley: Yes.

Steve Brine: I am asking about the issue known as that, whereby people cannot report images that are distressing to them because they do not meet the legal thresholds of your terms. Should people be able to withdraw their consent when posting of their image causes them great distress?

Richard Earley: I am familiar with the term breadcrumbing, and it is something we have looked at.

Steve Brine: Basically, it is content, isn't it? It is content and activity that is specifically designed to game your content moderation rules.

Richard Earley: Yes, as I understand it and as the safety teams I have spoken to about this understand it. It is an example of what we were talking about previously, which is that as platforms like ours get better at finding and removing those who are engaged in certain types of abuse on our platforms, those bad actors who still want to carry out that abuse try to find new ways to do that.

An example we have seen and a place we have acted in this space is people making inappropriate comments on images of children. The comments themselves might not actually break any of the policies we had in place at the time around sexualisation or abuse. What we did—and again we worked very closely with a number of child safety NGOs on updating this policy—is we expanded our definition of the sorts of posts and comments we remove to include comments that are not themselves inherently sexualising but which do indicate from their context that they are intending to make those remarks about individuals. That policy applies not just to images that are posted or comments underneath images but also to individual accounts, or groups or pages that are set up deliberately to carry out that kind of activity.

Q297 **Steve Brine:** Do you think that that addresses the issue of breadcrumbing, whereby images and comments that appear innocuous to the amateur eye are used to signpost what is illegal activity to others?

Richard Earley: That is the intention of this recent update to our policy, yes, which is to prevent that from happening.

Steve Brine: Niamh, do you have a view on this?

Niamh McDade: Yes, absolutely. To my previous point as well—on child sexual exploitation and abuse, I made the point about the importance of collaboration here to ensure that images and content shared on one platform does not reach another. Breadcrumbing in that way and leading children to other areas of concern—other areas where they could experience abuse—is something that we strive to protect against within our policies, but we are also working closely with organisations across the board on this.

In fact, a lot of our work and the majority of our policies in this space are led by the collaborative work we do with our Trust and Safety Council. Twelve members of this are child-focused organisations. It is not necessarily us as Twitter sitting in a room discussing what is best to protect here. We rely on the expertise of those organisations working in this space and that collaborative piece there. Along with our rules and policies, along with the work and the role we can have as platforms in silo, collaboration is key here. That is something we look forward to moving forward on with each organisation here and with platforms across industry as a whole.

Q298 **Steve Brine:** Finally, do you feel hopeful about the future in this whole online safety space? On the examples that Mr Nicolson pointed out to you, all the other examples that you will be aware of and all the things that we have discussed this morning, is it the case that we can act as Parliament or is it the case that whatever we do, whatever online safety Bill we bring forward, there will be the next version, the next take on breadcrumbing in that particular area, that will get around what we do? Are we dealing with the nature of the beast, Richard, that is social media, where we have to recognise the limits of our ability to effect change as a Parliament?

Richard Earley: Thank you. I want to come in there because I think you raise a very important point, which is one that we also made in our submission to your Committee and the PLS committee.

Steve Brine: I read it.

Richard Earley: I am glad. It is very important. It is a very difficult balance for the Bill to strike between being very clear about what requirements there are upon companies and making sure the Bill stays agile as technology and people's use of the internet changes. There are one or two places where we feel that the Bill does not quite get that balance right. One is in the amount of detail that the Bill currently goes into about the process for carrying out risk assessments, which may be very state of the art right now but could potentially go out of date as both technology and best practice change.

The second is the fact that, at present, we do not feel that the balance of where certain objectives or priority categories are placed in the enforcement regime from the Bill to secondary legislation is quite right. I think that it is worth the Committee looking carefully at that to see if the risks that you just described can be mitigated further through that.

Steve Brine: I think we will leave it there, Chair. We have been going long enough.

Chair: Thank you. That concludes our evidence for today. I wish to thank Iain Bundred, Richard Earley, Elizabeth Kanter and Niamh McDade for their evidence. Thank you very much. That concludes this session.