Meta

June 16th, 2022

Chairs, Online Safety Bill Committee
House of Commons
London SW1A 0AA

Dear Sir Roger and Christina,

Thank you again for the recent opportunity to give evidence to your Committee about the Online Safety Bill.

In response to your call for evidence, and your offer at my evidence session to further consider our views in writing, I share below some concerns Meta has about potential unintended consequences of the current text of the Bill. We have also provided our recommendations for how these unintended consequences could be mitigated.

Given the extreme pressures on your Committee's time, we have focused these recommendations solely on what we see as the top priority issues for Parliament to consider before the Bill is passed. We would be happy to discuss any additional questions that the Committee may have, or any particular points in our recommendations if helpful.

Our teams are already in touch, and as discussed separately with the Clerks will follow up with Committee members directly on some of the other issues that were raised during our evidence session.

Kind regards,

Richard Earley

Meta's Submission to the Online Safety Bill Committee

A. Introduction

Regulating the content people generate online, through videos, posts, and comments, requires new and innovative regulatory frameworks. These frameworks must ensure the online world is a safe place but also one where people's data is protected and their fundamental rights to privacy and expression are respected. A place where online service providers act reasonably and proportionately, taking their share of the responsibility to ensure this complex balance is struck.

Meta will continue to be a constructive partner to governments as they weigh the most effective, democratic, and workable approaches to meeting this challenge. Over the last five years, Meta has supported the UK Government's development of the Online Harms framework through evidence sessions, written submissions, ministerial discussions, and multistakeholder in-person technical sessions.

We share the UK Government's stated policy objectives, to make the internet safer while protecting the vast social and economic benefits it brings to billions of people each day. It is important not to lose sight of those benefits, but to have them squarely at the heart of how the UK approaches the Online Safety Bill. Global connectivity has improved economies, grown businesses, reunited families, and raised money for charity.

Millions of UK small businesses use Meta's platforms to reach customers and grow. Almost 1 in 4 of these small businesses say that the growth they have achieved using Facebook, its free tools, and the personalised advertising that it enables, has led them to hire at least one new employee. 35 million Brits now use Facebook Groups to connect every month, and 39% of Brits say the most important group they are a part of now operates primarily online. During the pandemic, over one and a half million people in the UK joined Coronavirus Mutual Aid groups on Facebook.

But at the same time, the internet has made it easier to create, find and spread content that could contribute to harm, like hate speech, misinformation, and terror propaganda. At Meta we have eighteen years' experience in tackling these issues through establishing policies, building tools and technologies, and producing guides and resources, all in partnership with experts both within and outside our company.

If designed well, new frameworks for regulating harmful online content can contribute to the internet's continued success by articulating clear, predictable, and balanced ways for government, companies, and civil society to share responsibilities and work together. Designed poorly, these efforts may stifle free expression, slow innovation and quickly become redundant.

While the title of the Bill has shifted from 'online harms' to 'online safety', it still attempts primarily to establish a systems-focused framework, which aims to require service providers to use proportionate systems and processes to address harmful content, an aim which we support. We believe this approach is the best way to ensure an appropriate balancing of safety, privacy, freedom of expression, and other values, and have long been on record calling publicly for legislation following this mode.

However, over the long process of developing this Bill, it has been drawn away from this outcome-based approach. Parts of it, which we outline below, are now ambiguous and contradictory, or risk undermining user privacy and security. This has the potential to make the Bill less effective and less workable. Other parts of the Bill have become more narrowly focused and risk being prescriptive. Our experience is that the more prescriptive the regulatory requirements and the greater the focus on individual pieces of content, the slower and more complex the decision making for content reviews. Ultimately, under the risk of large penalties, the Bill as it stands risks creating an incentive for the mass removal of *any* content which may fall foul of these rules. This runs counter to the Online Safety Bill's ambition to protect people's rights to freedom of expression.

Narrow requirements focused on specific content could also quickly become outdated and restrictive in the rapidly changing environment of online communication, raising the question of how future-proof the framework may be.

Secondly the Bill contains numerous tensions and contradictions in what it asks regulated services to do, without sufficient clarity on how these should be balanced - the most obvious example being how companies should police legal speech set against protecting free speech. The Bill's core safety duties require regulated services to take steps to find and remove illegal and harmful content. However, these duties have been placed into tension with a series of countervailing obligations to take care not to remove content where it falls under certain categories. This includes content 'of democratic importance', or content 'created for the purpose of journalism', or where to do so would conflict with rights to freedom of expression or privacy—but crucially, these terms are poorly specified and the balance of risks not made clear, rendering the framework unpredictable when it comes to compliance.

Thirdly the Bill does not contain sufficient protections for people's privacy and security, especially in messaging services and protection for encryption. Indeed, the Bill appears to treat private messaging the same as public social media, despite these being fundamentally different services with different risk profiles, different safety approaches and different consumer expectations.

Lastly, the Bill also creates a significant power of intervention for the Government in the form of the Use of Technology notices regime, without sufficient safeguards for privacy. Without changes to the Bill, this risks services being forced to undermine the privacy and security of UK citizens' communications through application of content monitoring obligations.

The Committee should seek to add much needed clarity about key definitions in the Bill and resolve contradictions, and remove provisions that could lead to negative unintended consequences for people in the UK and for their privacy and security.

B. Areas of Concern

1. Private Messaging 4
2. Use of Technology Notices 6
3. User Empowerment and Verification 8
4. Risk Assessments 11
5. Protections for Journalistic Content 12
6. Age Management 13
7. Fraudulent Ads 15
9. Powers of the Secretary of State 16

1. Private Messaging

Overview

At present the Bill makes no distinction between public social media and private messaging services. This means the latter can be designated Category 1 services, and so made subject to the same obligations that will be placed on public social media sites of that Category. Such an outcome would fundamentally change the nature of some private messaging services and make compliance extremely challenging.

Concerns

Recognising that privacy is a fundamental right, the 2019 Online Harms White Paper made clear that the Government understood public and private communications are different. Therefore, harms in private communication would be subject to a tailored set of requirements to ensure steps to address harms were effective and proportionate while respecting peoples privacy and data security. Making this a very clear distinction seemed to be the minimum response required given that the Government itself noted, in its summary of written responses to the White Paper consultation, that “overall respondents opposed the inclusion of private communication services in scope of regulation”.

However, in the final Bill, the distinction between public social media and private messaging has been lost, with all services in scope now referred to as ‘user-to-user services’. In our view, this risks causing a number of significant unintended consequences.

Under the Bill as it is currently drafted, providers of private messaging services could be forced to meet the following obligations:

- putting measures in place to prevent users coming into contact with legal but harmful content;
- providing dedicated reporting systems for journalistic content and special protections for content of democratic importance;

- introducing systems to verify adult users and to prevent verified users from coming into contact with non-verified users if they chose to.

It is unclear how private messaging services could meet these obligations without in effect scanning all private messaging. Many obligations do not make sense in the context of 1-1 or private group conversations. Private messaging is not the same as public social media—for example, in a service like WhatsApp, there is no discoverability, meaning you cannot search for other users and you need someone's phone number to contact them. WhatsApp also does not promote content to users - content is not 'posted', there are no algorithms or ads, and conversations carry a high expectation of privacy.

The Bill also does not take account of end-to-end encrypted (e2ee) services, including that the above obligations would be in clear conflict with the fundamental premise of e2ee, namely that only the sender and intended recipients of a message can know or infer the contents of that message. Attempting to apply these obligations to e2ee messaging services risks people's private messages being constantly surveilled and censored for legal but harmful content. This would be particularly concerning in a situation where users may de facto be required to register their identity in order to be able to use these services.

It's important to note that the Committee has heard how encryption is vital to protect everyone's online safety, including children. Experts on children's digital safety are clear that there is no inherent tradeoff between privacy and security: in a Unicef report on encryption and child safety, the authors state that "disagreements around platform end-to-end encryption has inadvertently created a perceived conflict between a child's right to privacy and the right to protection from sexual abuse and exploitation. However, the goal of ensuring that children's rights are safeguarded in the digital age involves fulfilment of their rights to both privacy and protection from sexual abuse and exploitation".

Recommendations

In light of this, it is clear that the Bill should not treat private messaging as though it were analogous to public social media. As private messaging is both technically different and subject to different user expectations, service providers must by necessity take different steps to protect the users of their private messaging services.

The Committee should amend the Bill to return to the Government's original distinction between public social media and private messaging. The most complete way to do this would be to make clear that private messaging services, including encrypted services, cannot be designated as Category 1 service. However, Ofcom should also be required to prepare a dedicated Code of Practice for private messaging services, in order to clarify the differing ways that they can meet their obligations under the Bill. This aligns with the recommendation from the PLS Committee that there is an urgent need to clarify how encrypted private messaging services can be compliant with the Bill's requirements

2. Use of Technology Notices

Overview

S.103 of the Bill gives Ofcom the ability to issue 'technology notices', requiring private messaging services to put in place 'accredited' technology for the purpose of detecting and removing child sexual exploitation and abuse content (CSEA). It is unclear how this would be possible in an encrypted messaging service, and would have significant privacy, security and safety implications for users. These powers also appear to lack sufficient safeguards in terms of how and when they might be used.

Concerns

In our view, s.103 should not grant powers to force scanning of encrypted messages. Ofcom having the power to force companies to scan private messages risks compromising encryption which is a critically important technology across the digital economy, it also undermines the fundamental promise of e2ee that only the sender and intended recipients can know or infer the contents of their messages. This will have chilling effects on privacy and freedom of speech.

If the Bill passes as is and Ofcom does have powers to impose scanning technologies, at a minimum there must be requirements for Ofcom to consider the technical feasibility, security implications, impact on privacy and cost of compliance with the accredited technology to be mandated by the technology notice. Further, there is no independent judicial oversight, or other substantive and procedural safeguards on the use of this power, which is something we'd expect to see in legislation that could have such far reaching implications for privacy and security of citizens and could also mandate a company to change their product(s)

Ian Stevenson (OSTIA) and Adam Hildreth (Crisp) both claimed in their oral evidence on 24 May that there are approaches to scanning private message contents that don't undermine e2ee. Whilst they noted the privacy concerns, this view is contrary to the prevailing view of cybersecurity experts. Current technologies to scan messages introduce security risks into cryptographic systems, increasing the likelihood that hackers and criminals can gain access to people's messages, and are fundamentally incompatible with an e2ee messaging service.

Moreover, this type of message scanning is not effective at stopping the proliferation of child sexual exploitation and abuse (CSEA) or catching abusers, and may actively harm children by denying them security and privacy. In 'Bugs in our Pockets', a coalition of experts in cryptography, computer science and policy state that client-side scanning technology "by its nature creates serious security and privacy risks for all society while the assistance it can provide for law enforcement is at best problematic. There are multiple ways in which client-side scanning can fail, can be evaded, and can be abused."

There is no way to make scanning technologies work for 'good' purposes only. As demonstrated by the use of the software Pegasus, there are already many efforts underway around the world by Governments to invade the privacy of their own citizens as well as those of other countries. With states like China and Russia seeking to monitor conversations and clamp down on political dissidents, the Bill shouldn't do anything that makes that even easier.

Concerns surrounding client-side scanning, and the importance of end-to-end encryption, are set out length in the independent human rights impact assessment of Meta's end-to-end encryption plans conducted by Business for Social Responsibility, which concluded that "the deployment of client-side scanning technologies as they exist today should not be pursued, as doing so would undermine the cryptographic integrity of end-to-end encryption and constitute a disproportionate restriction on privacy and a range of other human rights."

In summary, the types of technologies mooted as 'accredited technologies' would require widespread and systematic scanning or monitoring of citizens' private communications. At the same time they would be fundamentally incompatible with end-to-end encrypted private messaging, and would create vulnerabilities in services on which they were implemented that could be exploited.

The technology notice provision also risks setting a dangerous precedent globally. Passing such a law opens the door for other countries to follow suit - potentially without robust safeguards and with a wider scope of application. This is not just a theoretical risk for citizens in other countries, but also for UK residents who rely on private messaging services, including ours, to communicate with others around the world. Parliament should consider what its position would be if a foreign government took a power to force a company to scan the private messages of UK citizens. By taking such a power themselves, the UK Government risks empowering others to do just that.

Recommendations

Allowing the Technology Notices to apply to encrypted content through the forced application of scanning technology would break end-to-end encryption in private messaging. To make these provisions workable, the Committee should therefore include in the Bill explicit protections for e2ee, or make clear that Technology Notices cannot be used for the purpose of accessing the contents of private messages.

If the above recommendations are not adopted, the Committee should at the very least, due to the significant impact the issue of a Technology Notice would have on users, introduce greater procedural and substantive safeguards to the process including: a general duty of privacy to govern all decisions taken including consideration of security impact; judicial double-lock; and a full merits-based appeal process, as was preferred by the consultation responses to the Online Harms White Paper.

3. User Empowerment and Verification

Overview

Section 57 ("user identity verification") requires only Category 1 services to offer only adult users the option to verify their identity. This is linked to Section 14 ("user empowerment duties"), which requires only Category 1 services to offer those adult users who have verified their identity under. S. 57 the ability to "filter out" non-verified users.

These provisions did not undergo any prior industry consultation and will have a number of negative effects. In particular the interdependence between S. 14 (User Empowerment Duties) and the identity verification requirement of S. 57 risks dividing the UK online space into three strata: verified adults, unverified adults, and users under 18s.

This risks cutting off those who cannot or do not wish to share their identity online (including young people, victims of domestic abuse, and underrepresented groups) from taking full part in online conversations, without solving the problem of online abuse which unfortunately many people engage in openly.

Concerns

The interdependence between Section 14 and Section 57 renders the Bill unworkable in its current form, and does not represent the right balance between prescriptiveness and flexibility.

Section 57 ("user identity verification")

This section explicitly refers to identity verification, and not a more flexible standard (e.g., "identity assurance/authentication"), which suggests that Category 1 services will likely be required to collect, process, and store significantly more sensitive and confidential data in order to comply with such a strict requirement.

Though Section 57 stipulates that such verification need not require additional documentation, the reality is that verification standards to be established by Ofcom will necessarily require a high threshold to meet the "verification" obligation on the face of the Bill. "Verification" has a clear meaning in ordinary usage, and we understand from statements by the Government and the Minister that the intention is in fact for Section 57 to link identity to formal ID.

We know that many individuals who use online platforms do not have access to formal ID documentation or have very good reasons not to want to share this (e.g. survivors of domestic abuse, the transgender community). Moreover, 'badging' of individuals (i.e. indicating they have provided us with verification) could lead to an assumption that these users are more trustworthy, when in fact identity verification does not solve for bad actors or bad behaviour online.

There is a range of current and in-development technologies, that could satisfy a balanced requirement for identity assurance. As currently drafted, this requirement is too rigid and would likely restrict accessibility for many segments of the population.

All online platforms should be afforded the flexibility to take the most appropriate measures to ensure that user accounts are authentic, or that a user's identity can be reasonably assured, based on the circumstances of particular situations.

Section 14 ("user empowerment duties")

As a direct consequence of how Section 14 is linked to Section 57 (i.e. users who choose to have their identity verified can then choose to "filter out" non-verified users), the Bill will result in a number of significant, albeit unintended, consequences for the online environment in the UK, and how UK adults and young people experience that environment:

(a) S. 14 is not applicable to all online platforms. Any services which are not categorised by the Secretary of State and placed in the Ofcom register of Category 1 services will be exempt. Therefore, harmful content can, and will, still proliferate and individuals who use smaller services will not have the same controls. As was noted at 2nd Reading, the Bill should regulate based on risk, not just on size.

(b) The Bill will have the practical consequence of separating young Brits (U18) from those who are over 18 because S. 14 and S. 57 only apply to adults. As such, young people will not have the option to be verified and therefore will be automatically non-verified. Therefore, verified adults who choose to "filter out" non-verified users, will also "filter out" all users who are under 18. For example, if a hypothetical candidate for UK public office is using an online platform to reach their constituents and their families, but has exercised their option under S. 14 to "filter out" non-verified users, young people will not be able to use those online platforms to engage with that candidate. This concern was raised by several voices at 2nd Reading, noting the real risk of excluding young Brits from being able to participate in the digital world freely and safely.

(c) The combination of S. 14 and S. 57 will fragment the UK internet into several strata - individuals who are under 18 (for the reasons noted above), verified individuals over 18, and unverified individuals over 18. The fact that these strata are as a result of user's choice does not unfortunately change the reality that this will compound the very challenges the Government is seeking to address by creating walled gardens of unverified and verified users which will not prevent the type of abuse that these Sections are attempting to solve for. For example, if a candidate for political office is verified and has "filtered out" unverified users, they may still be subject to harmful behaviour by bad actors who have chosen to verify themselves.

(d) S. 14 presupposes that identity verification will mitigate such behaviour because people are less likely to engage in this manner when such conduct is no longer anonymous. However, we know this not to be the case, and bad actors will continue to engage in such conduct even after they have been verified. In contrast we have worked extensively to understand what tools and user controls do have an impact on preventing harmful behaviour, and have set out in previous submissions what these are—for example comment controls; comment warnings; Direct Message controls; the ability to block and restrict accounts; and the ability to limit interactions from accounts that don't follow or only recently followed you. These Sections run the real risk that verified users will be subject to doxxing (for example, individuals might publicly call out people who are engaging in a manner that is abusive, or which they do not like, in an effort to compel them to cease such behaviour).

(e) The coupling of S. 14 and S. 57 will have the practical effect of separating UK adult users from non-UK adult users in the rest of the world. This is because the Bill only applies to individuals in the UK and therefore individuals outside of the UK will not

have the ability to be verified pursuant to S. 57 (i.e. all rest of world users will be non-verified). For example, individuals in France who wish to engage with a verified UK political candidate because they are running on a platform around increased exports of their local produce to France, will be blocked from doing so.

(f) In addition, the cumulative effect of S. 14 and S. 57 in its practical implementation and the uncertainty which it brings to the UK regulatory environment, is that it will have a significant impact on the digital economy. This will be entirely contrary to the Government's ambition to make the UK "the tech capital of the world", as was noted at 2nd Reading.

In the Bill's provisions on risk assessments and obligations on mitigation measures, online platforms will already be required to assess and address the risks of anonymous accounts, as per the recommendation by Clean Up The Internet, which in written testimony suggested that, "the Bill could require platforms to demonstrate to the regulator that through their design, systems and processes they have taken reasonable steps to mitigate the risks associated with anonymous accounts."

Our view is therefore that S. 14 and S. 57 do not create any additional benefits or assurances for individuals above and beyond those obligations, and, in fact, will result in substantial negative outcomes.

Recommendation

However, we fully appreciate the Government's ambition in wanting to give users of online services greater powers over their experience, and the option to provide greater information about who they are to help others.

So we recommend that the Committee:

(1) amend S. 57 to reflect existing Government positions and standards around "identity assurance/authentication" (i.e. GPG45, and ICO guidance on ID assurance).

(2) separate S. 14 from S. 57 by deleting the link whereby verified (or "authenticated") users can 'filter out' content from non-verified (or "non-authenticated") users, while retaining all the provisions around user control over illegal and harmful content (i.e. removing S. 14 (6), (7) and (9) but retaining S. 14 (1) - (5)).

4. Risk Assessments

Overview

The Bill requires providers to conduct up to three separate types of risk assessments: (i) illegal content risk assessments (Section 8), (ii) childrens' risk assessments (Section 10), and (iii) adults' risk assessments (Section 12).

Risk assessments must be conducted within 3 months of Ofcom's guidance, kept continuously up-to-date, and a further risk assessment must be conducted before making any "significant change" to "any aspect of a service's design or operation". As well as being unclear, depending on the interpretation of these terms this could make for essentially continuous updating of risk assessments, adding unnecessary regulatory 'red tape' and putting the UK out of step with emerging best practice including in the EU's Digital Service Act (DSA).

The issues to be considered and techniques to be used in carrying out risk assessments are also set out at length on the face of the Bill. As technology evolves, having such extensive details fixed in the Bill risks the framework quickly falling out of date. As such these details would be better left to Ofcom to set out in Codes of Practice.

Concerns

The requirement to conduct a further risk assessment before making any "significant change to any aspect of a service's design or operation" is vague. Meta works tirelessly to improve its services every day, and is constantly introducing new innovations, including minor tweaks, intended to better protect our users from emerging harms. Making each of these changes subject to an additional assessment is likely to have a chilling effect on UK-based innovation and/or incentivise providers to prioritise the development and introduction of new technologies in product features in jurisdictions other than the UK. It could also have the unintended consequence of making services less able to quickly respond to emerging threats for UK-based users.

The onerousness of these requirements is ultimately unnecessary to achieve the objective of having providers identify and mitigate risks arising from the services, as this could be accomplished through much less frequent reviews that still considered all of the same risk factors.

The Bill's requirements are also out of step with similar risk assessment obligations under, for example, the EU's DSA. The DSA only requires providers to undertake risk assessments on an annual basis, and consider "systemic risks" as opposed to the disparate range of factors that must be considered under the Bill as drafted.

This divergence means that providers will be unable to implement compliance solutions at scale, and could lead to the UK becoming a less competitive jurisdiction for online services, in contrast to the Government's ambition to make the UK "the tech capital of the world", as was noted at Second Reading.

Recommendation

To address the risk of these provisions of the Bill quickly going out of date as technology and best practice evolves, the Committee should amend the Bill so that the details of what steps services must take in carrying out risk assessments are set out by Ofcom in codes of practice.

We understand that it is the Government's intention that risk assessments should only be updated following major innovations that result in fundamental changes to a service. The Bill

currently uses ambiguous terminology such as 'significant change' which could lead Ofcom to diverge from this policy intention. The Committee should amend the Bill to clarify this intention, or more simply should make risk assessments an annual rather than continuous requirement, and avoid the need for an ambiguous 'trigger' for additional ad hoc risk assessments altogether.

5. Protections for Journalistic Content

Overview

Section 16 of the Bill requires Category 1 services to take special steps to protect journalistic content, including a dedicated process to report such content that is wrongly removed. While we recognise that news is a public good, and Meta has made significant investments to help support journalism in the UK, the duty in the Bill to protect journalistic content is vague and risks introducing a significant element of subjectivity and therefore complexity, while simultaneously excluding a broad swath of this content (i.e. news publisher content) from the Bill's safety duties

Concerns

Firstly, we are concerned that the Government is putting obligations on private companies to make extremely complex and real time assessments about what constitutes journalistic content which could be impossible to implement consistently. We would also question whether it is appropriate for platforms to be defining what counts as journalistic content.

Secondly, we are also concerned that these requirements may inadvertently give a level of protection to bad actors, which may expose users to an increased risk of harm. Therefore, it will also be placed in direct tension with the other duties such as addressing legal content that is harmful to children or adults. In-scope services must disclose in the Terms of Service the methods they use to identify journalistic content, which will make the system easier to game. This is especially the case when it comes to content from users who assert that they are operating as citizen journalists. We have already seen instances where some users claim they are citizen journalists in an attempt to reduce the likelihood of Facebook (and other platforms) taking action against them or their content, ensure their content is given enhanced protection, and try to circumvent our policies - our community standards and third party fact checking programs. Due to this very reason, the broadness of this exemption has worried other organisations such as Full Fact, IMPRESS and Hacked Off.

Thirdly, we already have dedicated services for media organisations and journalists in the UK, including a media partnerships team offering training, best practice and guidance around our content policies. These systems have been developed in consultation with our publisher partners, including small and local publishers, and are a better means of supporting them than those proposed in the Bill. Our systems will also adapt and change over time through further feedback and as technology and usages changes, whereas fixing certain processes into primary law risks the Online Safety Bill quickly going out of date.

In conclusion, these complex and potentially overlapping definitions will make it extremely challenging for platforms to design systems and processes that rigorously and consistently identify and appropriately handle the different forms of content.

Recommendations

Given the disproportionate risks that this provision introduces, weighed against the current existence of protections for journalism on many platforms, the Committee should remove this section of the Bill entirely. If any provision is kept in the Bill relating to journalism, this provision should clearly define journalistic content without subjective interpretation, for example by creating a definitive list of 'recognised news publishers' that are entitled to enhanced protection.

6. Age Management

Overview

Sections 10 and 11 require providers of services that are likely to be accessed by users in the UK who are under 18 to use proportionate systems and processes to prevent children of any age from encountering primary priority content (e.g. by using age verification, or another means of age assurance),

However, in order to determine if an individual is over or under 18, we will be required to verify (or 'assure') the age of all individuals using our services. This could necessitate the collection, processing, and retention of more personal data of individuals in the UK, including that of individuals who are under 18.

The Bill also contains a number of provisions that either contradict or do not fully align with existing law in the UK around protecting young people online, most importantly the Age Appropriate Design Code.

Concerns

As currently drafted, the Bill does not contain sufficient clarity as to what the intention is for the inclusion of both age verification and age assurance in the Bill. Section 189 provides a definition of age assurance, but not of age verification—which is the obligation written onto the face of the Bill.

Section 11 (1) specifies "regulated user-to-user services that are likely to be accessed by children" but Section 11 (13) specifies that the duties set out in the section extend only to such parts of a service as it is possible for children to access. The ICO has provided guidance around services which are likely to be accessed pursuant to the Age Appropriate Design Code (AADC), approved by Parliament. Divergence from this standard within the Online Safety Bill will create significant uncertainty for businesses and individuals as to the standard that should be applied in determining whether services are, or are not, in scope for S. 10 and S. 11.

It is also unclear whether Ofcom will establish a definition of or guidance to what is a "proportionate" set of systems and processes that would satisfy the obligations under S. 11. This was a recommendation made by the Pre-Legislative Scrutiny Committee on the Bill: "we recognise the concerns that, without proper guidance, service providers might seek to place disproportionate age assurance measures in place, impacting the rights of both children and adults." The Public Bill Committee should consider whether to take up this recommendation.

The Bill also appears to contradict the existing regulatory frameworks around minors' online experiences. For example, the AADC applies to "information society services likely to be accessed by children" which includes social media platforms. AADC also suggests that compliance with the Code means switching off geolocation services, but the Bill's obligations only apply with respect to U18 in the UK. To ascertain this, we will need geolocation information, putting us in conflict with AADC. The Report of the Committee following Pre-Legislative Scrutiny notes the challenges with the tests put forth in the draft Bill, quoting Common Sense Media that the scope of the Bill should not be more restricted than the AADC and that the "likely" test should be the same in both.

Recommendations

The issues with S. 10 and S. 11 could be best addressed by the Committee via amendments to bring the Bill more into line with the existing AADC guidance on age appropriate design, age assurance, and services which are likely to be accessed by children (U18). This would fulfil the recommendations that were previously made to this effect in the report of the Joint Committee on the Draft Online Safety Bill (the Pre-Legislative Scrutiny Committee)

7. Fraudulent Ads

Overview

Sections 34-36 of the Bill require Category 1 services to take measures to deal with fraudulent advertisements. This encapsulates paid-for advertisements that amount to a fraud offence (as listed in s.36) and that are not user-generated content.

However, the regulation of online advertising including illegal advertising such as fraudulent advertising is subject to a pre-existing consultation in the UK as part of the Online Advertising Programme (OAP), launched in March 2022.

Concerns

Meta's overriding priority is the safety of our users and we therefore take a zero tolerance approach to fraud on our platforms. By its very nature fraud is adversarial and hard to spot, and the perpetrators of fraud are continually searching for ways to subvert the rules, processes and safeguards we put in place to protect our users. This includes working across multiple platforms and throughout the advertising stack.

For these reasons, we believe that the OAP's pre-existing consultative process is the best place to consider the regulation of online advertising. This is because it takes a holistic,

evidence based view across the entire advertising ecosystem. The inclusion of fraudulent advertising in the Bill will set a precedent for the regulation of online advertising, undermining the OAP's consultation process. Furthermore, the Bill's lack of industry consultation in relation to fraudulent advertising risks leading to unintended consequences.

For example, the Bill as currently drafted calls for Category 1 services to "include clear and accessible provisions in their Terms of Service, giving information about any proactive technology used to comply with the above duty, including the kind of technology, when it is used, and how it works."

Meta agrees with the important principle of transparency on which this text is based. But we question the wisdom of requiring platforms to set this information out publicly. Doing so will also notify the very fraudsters we are attempting to stop and help guide them around the measures Meta takes to curb such activity. This is the type of issue that will be exposed in the course of the OAP's review.

The Bill also proposes a stand alone duty on Category 1 services, with no consideration of smaller, but nonetheless high risk services. Meta and other Category 1 services are already taking many of the measures the government believes is necessary as part of this Bill while smaller services are not. These services are susceptible to an influx of fraud as Category 1 services continue to improve their defences. As the Bill is introduced it may create a waterbed effect pushing fraud to these services, where no obligations apply. It is therefore unclear that these measures are targeted in the right place.

Recommendations

Online fraud is a serious issue in the UK and globally following the increase in the use of technology during the COVID pandemic. Meta stands ready to continue engaging in the Home Office's Online Fraud Steering Group (OFSG), or any complementary processes that the Committee may call on the Government to set up. However, the Committee should remove this clause entirely to enable the pre-existing OAP to address fraudulent advertising holistically and with the benefit of industry consultation.

If a version of this clause is retained, it must be amended to remove the obligation on Category 1 services to set out the compliance measures they take with respect to fraudulent advertising, which would in essence, act as a guidebook for fraudsters. The Committee should also broaden the provisions beyond just Category 1 services to smaller services.

9. Powers of the Secretary of State

Overview

Section 39, Section 40, and paragraph 2 of Schedule 10 give the Secretary of State broad powers to direct the enforcement of the Bill's provisions and to amend its coverage. These powers are constrained by only limited oversight by Parliament.

Concerns

While some changes have been made since Pre-Legislative Scrutiny, the Bill still grants the Secretary of State considerable powers to define the parameters of the regulatory regime and intervene in the enforcement of its requirements. These powers go beyond what is necessary to ensure democratic accountability for the regulator and are not constrained in line with other examples of how to mediate the role of elected representatives within the executive. They create the risk that the rules for online speech could be amended to align with 'public policy' with little Parliamentary oversight or evidence base.

In practice, these powers would grant the Secretary of State the ability to set enforcement priorities, designate categories of harmful content, and stipulate compliance recommendations based on little more than political preference, which would move key aspects of the Bill out of the independent regulatory sphere and squarely into the realm of partisan politics.

Companies in scope of the Bill will need consistency and certainty to be able to carry out their duties effectively and without undue regulatory or cost burdens. And Ofcom must have appropriate independence to build and enforce a regime that reflects the best available evidence based on objective, apolitical standards. Undermining Ofcom's independence runs counter to longstanding principles of regulation in the UK, increases the risk of unintended consequences, and sets a worrying precedent for the future of other regulatory bodies in the UK.

Recommendations

The current powers are overbroad and provide the opportunity for both this and any future administration to intervene in the regime by amending codes of practice. These powers in Section 39 should be removed, as should the entirety of Section 40 and the relevant provisions in Paragraph 2 of Schedule 10.