Your response

Volume 2: The causes and impacts of online harm

Ofcom's Register of Risks

Question 1:	
i)	Do you have any comments on Ofcom's assessment of the causes and impacts of online harms?
Response: No	
ii)	Do you think we have missed anything important in our analysis? Please provide evidence to support your answer.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 2:	
i)	Do you have any views about our interpretation of the links between risk factors and different kinds of illegal harm? Please provide evidence to support your answer.
Response: No	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Volume 3: How should services assess the risk of online harms?

Governance and accountability

Question 3:	
i)	Do you agree with our proposals in relation to governance and accountability measures in the illegal content Codes of Practice?
Response: No	
ii)	Do you think we have missed anything important in our analysis? Please provide evidence to support your answer.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: no)	

Question 4:	
i)	Do you agree with the types of services that we propose the governance and accountability measures should apply to?
Response: yes	
ii)	Please explain your answer.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: yes (all)	

Question 5:	
i)	Are you aware of any additional evidence of the efficacy, costs and risks associated with a potential future measure to requiring services to have measures to mitigate and manage illegal content risks audited by an independent third-party?
Response: In the section on internal assurance and compliance functions, Ofcom rejects options that include elements of external audit or scrutiny in favour of those that rely on internal assurance mechanism. The main basis on which Ofcom rejects options that include external scrutiny is uncertainty over cost.	

We believe that platforms should not mark their own homework, and that there is significant risk in delaying the introduction of any future plan to require services to have measures to mitigate and manage illegal content risks audited by an independent third-party. While we agree that all platforms should ensure that they have established appropriate internal governance frameworks including internal monitoring and assurance functions, it is not credible to describe or consider these as independent. It is important that Ofcom ensures that there are genuinely independent monitoring, reporting and assessment processes.

We agree that the risk of bias is mitigated to some extent by the framework of governance and accountability proposed by Ofcom, which we strongly support. Nevertheless, we believe that it would be unrealistic to assume that incentives for employees could be so effectively demarcated as to render them genuinely independent. We note that Institutional Shareholder Services (ISS) uses a number of criteria to assess the independence of a member of a UK company's Board of Directors*. The first of these criteria is that for a director to be considered independent, they must not have been an employee of the company or group during the previous five years. This is a relevant parallel, we believe, and puts into context what we see as an overly hopeful expectation of independence in Ofcom's proposal, notwithstanding appropriate governance arrangements.

We have also provided proposals below as to how to establish appropriate and cost-effective monitoring, reporting and assessment processes in response to other Volumes.

We believe that our proposal for credible, independent, third-party scrutiny is commensurate with standard auditing requirements for all companies and, for example, with prudential regulation requirements for large financial services institutions. We cannot see any reason to delay implementation of this requirement.

We do not believe that Ofcom's stated reason for not pursuing options which include external scrutiny – that it does not understand the costs of implementing them – are particularly persuasive. This is especially the case since it has not assessed the costs of other options. To the extent that this is the main or only reason Ofcom is not minded to take forward such proposals, we believe it is incumbent on Ofcom to obtain such cost information.

Finally, we would stress that the costs of these audits would be borne by the platforms, and not by Ofcom. Nonetheless, we would not expect these costs to be disproportionate set against the substantial revenues and profits generated by the platform industry. Moreover, platforms have benefited from excess revenues and profits that have been generated through hosting the very illegal and harmful content that the Online Safety Act seeks to eradicate.

* ISS - Institutional Shareholder Services, 2024; Proxy Voting Guidelines, Benchmark Policy Recommendations, United Kingdom and Ireland; https://www.issgovernance.com/file/policy/active/emea/UK-and-Ireland-Voting-Guidelines.pdf

ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: no	

Question 6:	
i)	Are you aware of any additional evidence of the efficacy, costs and risks associated with a potential future measure to tie remuneration for senior managers to positive online safety outcomes?
Response: No	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Service's risk assessment

Question 7:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Specifically, we would also appreciate evidence from regulated services on the following:

Question 8:	
i)	Do you think the four-step risk assessment process and the Risk Profiles are useful models to help services navigate and comply with their wider obligations under the Act?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 9:	
i)	Are the Risk Profiles sufficiently clear?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Do you think the information provided on risk factors will help you understand the risks on your service?
Response: n/a	
iv)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
v)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Record keeping and review guidance

Question 10:	
i)	Do you have any comments on our draft record keeping and review guidance?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 11:	
i)	Do you agree with our proposal not to exercise our power to exempt specified descriptions of services from the record keeping and review duty for the moment?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Volume 4: What should services do to mitigate the risk of online harms

Our approach to the Illegal content Codes of Practice

Question 12:	
i)	Do you have any comments on our overarching approach to developing our illegal content Codes of Practice?
Response: No	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 13:	
i)	Do you agree that in general we should apply the most onerous measures in our Codes only to services which are large and/or medium or high risk?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 14:	
i)	Do you agree with our definition of large services?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 15:	
i)	Do you agree with our definition of multi-risk services?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 16:	
i)	Do you have any comments on the draft Codes of Practice themselves?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 17:	
i)	Do you have any comments on the costs assumptions set out in Annex 14, which we used for calculating the costs of various measures?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Content moderation (User to User)

Question 18:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Content moderation (Search)

Question 19:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Automated content moderation (User to User)

Question 20:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 21:	
i)	Do you have any comments on the draft guidance set out in Annex 9 regarding whether content is communicated 'publicly' or 'privately'?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Do you have any relevant evidence on:

Question 22:	
i)	Accuracy of perceptual hash matching and the costs of applying CSAM hash matching to smaller services;
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	

iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 23:	
i)	Ability of services in scope of the CSAM hash matching measure to access hash databases/services, with respect to access criteria or requirements set by database and/or hash matching service providers;
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 24:	
i)	Costs of applying our CSAM URL detection measure to smaller services, and the effectiveness of fuzzy matching for CSAM URL detection;;
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 25:	
i)	Costs of applying our articles for use in frauds (standard keyword detection) measure, including for smaller services;
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 26:	
i)	An effective application of hash matching and/or URL detection for terrorism content, including how such measures could address concerns around 'context' and freedom of expression, and any information you have on the costs and efficacy of applying hash matching and URL detection for terrorism content to a range of services.
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Automated content moderation (Search)

Question 27:	
i)	Do you agree with our proposals?
Response: n/a n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a n/a	

User reporting and complaints (U2U and search)

Question 28:	
i)	Do you agree with our proposals?
Response: No	
ii)	Please provide the underlying arguments and evidence that support your views.

Response:

In Chapter 16 of the Consultation, Reporting and complaints, Ofcom deals with reporting of complaints to the platforms. We underscore in our response to Volume 6 of the Consultation below the need for comprehensive periodic reporting of complaints and actions from the platforms to Ofcom. We have not seen a proposal from Ofcom on the need for such periodic reporting, nor any comments relating to such periodic reporting.

Expand Trusted Flaggers

On reporting from users to platforms, we agree that it is appropriate for Ofcom to establish dedicated reporting channels (DRCs) for 'trusted flaggers.' In paragraph 16.242 of the Consultation, Ofcom proposes that there be a maximum of 7 trusted flaggers, and that this number relates to cost efficiency when developing and maintaining the reporting function. These trusted flaggers are named: HM Revenue and Customs (HMRC), Department for Work and Pensions (DWP), City of London Police (ColP), National Crime Agency (NCA), National Cyber Security Centre (NCSC), Dedicated Card Payment Crime Unit (DCPCU), and the Financial Conduct Authority (FCA).

While we are not best placed to comment on issues relating to other minority groups, we would note that antisemitism and antisemitic content are often characterised by the use of subtle formulations, nuances, conspiracy theories, tropes and conflations. We believe that none of the designated trusted flaggers identified by Ofcom has demonstrated a sufficiently thorough understanding of antisemitic content, and as such, cannot be replied upon to appropriately identify and report such content to the platforms.

There should therefore be more than 7 trusted flaggers designated by Ofcom.

We believe that for complex and multi-faceted categories of hate crime, among which we would include antisemitism, civil society bodies such as ourselves will need to play a crucial role. The issues will require bodies with the ability to identify illegal or harmful content, and to educate other trusted flagger organisations and the platforms themselves. Prerequisite for this will be deep expertise in monitoring, litigating and educating which a limited number of organisations have.

It is instructive that the CAA has been designated a trusted flagger by certain platforms. We therefore have direct knowledge of the value that we can bring to the process as trusted flaggers, and of the problems that we solve. An important issue to highlight is that government bodies (such as those mentioned as designated trusted flaggers by Ofcom) have frequently shown that they are unacceptably slow in their communication of illegal content to the platforms, resulting in posts remaining online for far too long. In order to be effective, a trusted flagger needs to be able to flag illegal content quickly. Moreover, the ability to recognise emerging patterns of illegal content, and to understand developing tropes, is far more honed among those organisations that specialise in understanding specific categories of content, such as antisemitism.

Moreover, opening up the field of trusted flaggers would fulfil any objectives to involve civil society organisations and institutions. This would bring to bear more culturally diverse perspectives from religious and other special interest and minority groups which can often better understand and reflect the interests of the communities and groups that they represent.

We do not believe that there would be any material incremental cost to Ofcom, and reinforcing platforms' ability to effectively remove illegal content will likely considerably reduce the cost of enforcement for Ofcom.

Data Sharing Framework to further reduce costs and barriers

In an effort to further reduce costs for the industry, we believe that Ofcom's intimation that each of the 7 trusted flaggers would need to engage with each relevant platform for reporting risks being wasteful and could be anachronistic. If 7 trusted flaggers were to each establish DRCs with, say, 6 platforms, then that would require 42 bilateral arrangements to be designed, agreed, and implemented at some considerable cost, presumably to be incurred by both the platforms and the public purse (the trusted flaggers). We therefore propose what we believe would be a much more cost-effective solution for the industry that would simultaneously enable a greater number of trusted flaggers to be designated by Ofcom, and which would involve only a single standard for all such relationships.

We propose that a single, common standard for a set of APIs (application programming interfaces) for reporting be devised by a working group. The working group could be responsible for setting standards for the sharing of information the platforms and trusted flaggers. In reference to our proposal with respect to Volume 6 of the Consultation, we further propose that this group could set standards for reporting of the incidence and treatment of illegal and harmful content from the platforms to Ofcom. Ofcom may wish to oversee this working group.

The concept would be akin to Open Banking, an UK initiative that has established a world-leading framework for setting standards for data sharing within banking and other industries. Open Banking was proposed in the Fingleton Report (Data Sharing and Open Data for Banks[*]) which was commissioned by HM Treasury and the Cabinet Office in 2014 as it sought to improve competition between banks and to lower barriers to entry to the banking industry. The Competition and Markets Authority imposed Open Banking by way of an order following the Retail Banking Market Investigation of 2017. The standards were ultimately agreed by a working group that was later named the Open Banking Implementation Entity (OBIE).

In a similar way, a single set of open, or non-proprietary standards could be agreed for online platforms in such a way as to contain costs for the entire industry, ensuring that the way data is shared between organisations is designed and agreed only once for all participants. Costs for smaller platforms could be set such that they would be proportionate to their size, or possibly de minimis, and would therefore not add additional or prohibitive cost burdens to smaller platforms. It therefore removes a barrier to entry for small and potentially disruptive new platforms, and so supports Ofcom's competition brief. Future iteration and amendments to the standard could also be discussed, agreed and implemented across the board (rather than 42 times in the example above).

The fact that Open Banking was possible in the banking industry, itself riddled with layer upon layer of legacy systems, and with monolithic decision-making structures, implies that younger and more technology-forward online platform industries should be easily capable of developing a functional framework. We believe that a single common standard for APIs would encapsulate the concept of DRCs, and thus would negate the need to establish so many, and the costs of doing so. * Data Sharing and Open Data for Banks -- A report for HM Treasury and Cabinet Office; September 2014, Fingleton Associates and the Open Data Institute.
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:no

Terms of service and Publicly Available Statements

Question 29:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 30:	
i)	Do you have any evidence, in particular on the use of prompts, to guide further work in this area?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Default settings and user support for child users (U2U)

Question 31:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 32:	
i)	Are there functionalities outside of the ones listed in our proposals, that should explicitly inform users around changing default settings?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)

Response: n/a

Question 33:	
i)	Are there other points within the user journey where under 18s should be informed of the risk of illegal content?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Recommender system testing (U2U)

Question 34:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 35:	
i)	What evaluation methods might be suitable for smaller services that do not have the capacity to perform on-platform testing?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

We are aware of design features and parameters that can be used in recommender system to minimise the distribution of illegal content, e.g. ensuring content/network balance and low/neutral weightings on content labelled as sensitive.

Question 36:	
i)	Are you aware of any other design parameters and choices that are proven to improve user safety?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Enhanced user control (U2U)

Question 37:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 38:	
i)	Do you think the first two proposed measures should include requirements for how these controls are made known to users?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 39:	
i)	Do you think there are situations where the labelling of accounts through voluntary verification schemes has particular value or risks?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

User access to services (U2U)

Question 40:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Do you have any supporting information and evidence to inform any recommendations we may make on blocking sharers of CSAM content? Specifically:

Question 41:	
i)	What are the options available to block and prevent a user from returning to a service (e.g. blocking by username, email or IP address, or a combination of factors)?
Response: n/a	
ii)	What are the advantages and disadvantages of the different options, including any potential impact on other users?
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 42:	
i)	How long should a user be blocked for sharing known CSAM, and should the period vary depending on the nature of the offence committed?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

There is a risk that lawful content is erroneously classified as CSAM by automated systems, which may impact on the rights of law-abiding users.

Question 43:	
i)	What steps can services take to manage this risk? For example, are there alternative options to immediate blocking (such as a strikes system) that might help mitigate some of the risks and impacts on user rights?
Response: n/a	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Service design and user support (Search)

Question 44:	
i)	Do you agree with our proposals?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Cumulative Assessment

Question 45:	
i)	Do you agree that the overall burden of our measures on low risk small and micro businesses is proportionate?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 46:	
i)	Do you agree that the overall burden is proportionate for those small and micro businesses that find they have significant risks of illegal content and for whom we propose to recommend more measures?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 47:	
i)	We are applying more measures to large services. Do you agree that the overall burden on large services proportionate?
Response: Yes, but insufficient – see response to question 28	

ii)	Please provide the underlying arguments and evidence that support your views.
Response: as per response to question 28	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: no	

Statutory Tests

Question 48:	
i)	Do you agree that Ofcom's proposed recommendations for the Codes are appropriate in the light of the matters to which Ofcom must have regard?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Volume 5: How to judge whether content is illegal or not?

The Illegal Content Judgements Guidance (ICJG)

Question 49:	
i)	Do you agree with our proposals, including the detail of the drafting?
Response: n/a	
ii)	What are the underlying arguments and evidence that inform your view?
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 50:	
i)	Do you consider the guidance to be sufficiently accessible, particularly for services with limited access to legal expertise?
Response: n/a	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Question 51:	
i)	What do you think of our assessment of what information is reasonably available and relevant to illegal content judgements?

Response: In the introduction to Volume 5 (26.1), Ofcom states:

In the introduction to Volume 5 (paragraph 26.1 of the Consultation), Ofcom states:

> *"Section 192 of the Act requires services to take down content where they judge there to be 'reasonable grounds to infer' it is illegal, using 'reasonably available information' to make this judgement."*

Notwithstanding our concerns regarding platforms and moderators using loopholes around what constitutes 'reasonable grounds to infer' and 'reasonably available information,' we consider the tests in Ofcom's Illegal Content Judgments Guidance to be well founded.

While platforms must respect users' right to freedom of expression within the law, we believe it appropriate that Ofcom is vigilant as the new regime becomes established. If there are too many instances of illegal or harmful content not being subject to takedown, especially in certain categories, then Ofcom may wish to reassess its approach to content moderation. This may involve education by accredited institutions that are expert in specific content categories such as islamophobia or antisemitism.

Moreover, Ofcom's understanding of the extent of content moderation failure can only be optimised through the level transparency that we propose in our response to Volume 6 of the Consultation.

ii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: no

Volume 6: Information gathering and enforcement powers, and approach to supervision.

Information powers

Question 52:	
i)	Do you have any comments on our proposed approach to information gathering powers under the Online Safety Act?
Response: Yes	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: In paragraph 28.49 of the Consultation, Ofcom states in respect of 'information powers', *"We will use our information gathering powers in a way that is proportionate to the use to which the information will be put and will only issue an information notice where we require information to exercise an online safety function or to decide whether to do so. We expect to use our power to issue statutory information notices regularly from the outset of the regime. Any information notices we issue will clearly set out the purpose of the request and why we require the information. We do not anticipate using our other information gathering powers such as skilled persons reports and powers of entry, inspection and audit as often, and these will typically be reserved for more serious cases."* We believe that Ofcom should avail itself of a major opportunity to impose transparency, enable effective monitoring, and thus ensure the effectiveness of the Online Safety Act. Ofcom should require periodic reporting by the platforms of the incidence reports of illegal or harmful content, the number of takedowns, and the number of complaints (e.g., relating to requested takedowns that were not fulfilled). If this is reported by each platform, not only can their relative performances be put into context, but it will also provide a picture of a substantial part of the whole industry in the UK. Ofcom appears to have favoured asserting its powers to request information from platforms in an *ad hoc* manner, for example by use of information notices, audit notices and warrants. We believe that given the scale of harm at issue, the balance between transparency and a "least intrusive approach" sought by Ofcom is far from optimal. By "least intrusive" we might expect a situation similar to that which we see today whereby there is a shroud of opacity surrounding each platform and the broader industry with respect to the prevalence of illegal and harmful content, and the platforms' effectiveness in dealing with it. Nobody knows how much illegal or harmful content is posted, because it is not routinely monitored, measured or reported, let alone identified. Any credible information that could reveal the company-level or aggregate scale of	

illegal and harmful content on platforms simply does not exist. We believe that Ofcom should assert its right to intrude further, and not merely search where the torch happens to shine.

Ofcom must at the very least strive to know what is going on in the industry that it regulates, and it must further strive to understand the impact of the legislation that it is implementing. Without relevant statistics, Ofcom's understanding of the industry it must regulate will be acutely impaired. It risks leaving the industry disproportionately opaque, and apparently the industry will benefit from the "least intrusive" approach. We believe that effectiveness of the Online Safety Act itself can only be assessed if we understand the outcomes.

Further, we note that Ofcom could have proposed an industry body or working group to promote best practice between the platforms to combat illegal activity by their users. This working group, to which we refer in our response to Volume 4, which would be responsible for, setting standards for the sharing of information between the platforms and trusted flaggers, could also define protocols for sharing of information between the platforms and Ofcom. Alternatively, these responsibilities could be discharged by two separate working groups.

Given our concerns with Ofcom's proposals as articulated above, we ask that Ofcom amends its proposals as follows.:

Ofcom should require regulated firms to report periodically:

- the number of reports of illegal content, broken down by category (antisemitic, Islamophobic, transphobic etc.), and as reported by users and trusted flaggers respectively
- the number of incidents, broken down as above, where content has been removed
- the number of incidents where illegal content is identified and removed by the platform without intervention, e.g., by AI bots (again broken down as above)
- complaints (various relevant measures)

We acknowledge that dialogue may result in appropriate refining and extending the scope of periodic reporting. Nevertheless, Ofcom expressed objectives (paragraph 28.49 of the Consultation) to "ensure that interventions are evidence-based, proportionate, consistent, accountable and transparent in both deliberation and outcome." We believe that failing to demand information along the lines that we suggest will likely result in a failure to meet Ofcom's objectives, especially in terms of consistency, accountability and transparency.

Breakdown by Category

It is crucial that reporting is broken down by category e.g., CSAM, harassment, and the different hate categories, for example, antisemitic hate, Islamophobic hate, transphobic hate etc. Identifying inconsistencies between platforms and categories may point to lacunae in certain platforms' understanding or implementation with respect to particular issues. There may notwithstanding be valid reasons for differences, but without shining a light, critical comparison

will not be possible. Ofcom could require that platforms remedy any such unwarranted inconsistencies through more rigorous application of moderation techniques, including through education of moderators on specific topics.

Proportionate, and Commensurate with FCA Approach

This approach is commensurate with a reporting approach taken by the FCA, for example. The Investment Firms Prudential Regime (IFPR) applies to all MiFID investment firms (chiefly investment managers), including small firms. Relevant firms must report data on a periodic basis (some quarterly, some semi-annually, and some annually) including profit and loss measures, balance sheet measures, assets under management. Details of errors and complaints must also be reported periodically. Daily trading activity on most exchanges must be reported within 1 day.

We use this example to show the extent of obligation and cost on some financial firms, including small firms, which is nonetheless considered proportionate. Therefore requiring often much better resourced online platforms to report a more restricted set of data on a periodic basis (not merely at the specific *ad hoc* request by Ofcom) should not be considered prohibitively onerous or disproportionate. In any event, any robust compliance framework implemented by a platform should easily be able to produce these data points with minimal human intervention or resource requirement. Likewise, any cost of data-gathering by Ofcom may even be outweighed by savings from having to make fewer expensive *ad hoc* interventions and from less enforcement activity.

Platforms should not mark their own homework

This data will provide vital information as to the extent of compliance with firms' obligations to remove illegal content, and will allow comparison between firms. Without these reports, it will be nigh on impossible to assess whether each firm is taking its responsibilities seriously. We do not believe that the keeping of internal records will be sufficient in this respect.

Without casting light on the relative performance of each firm, it will be impossible for firms, or for Ofcom, to judge the ongoing performance of each firm or the industry. Without this reporting, firms would be marking their own homework. Lack of comparison would mean that Ofcom will have no means of judging how effectively firms have implemented their compliance frameworks with respect to the Online Safety Act. It would also mean that firms could more easily deliberately avoid compliance and evade detection of that avoidance.

We believe that periodic reporting is vital for Ofcom to be able to monitor compliance with, and the impact of, the Online Safety Act.

Information Asymmetry

We believe that a failure to require periodic reporting would result in a perpetuation of the information asymmetry between Ofcom and regulated services that Ofcom is seeking to address (paragraph 28.2 of the Consultation), with each firm potentially knowing far more about its performance than Ofcom or anyone else. Reporting of these data cannot be considered

commercially or competitively sensitive. Indeed, the lack of requirement to report will likely bring about **moral hazard**, with a risk that failure to enforce regulation could encourage users to post illegal content with impunity.

Moreover, there is a real danger of a sort of **adverse selection**, in that a platform which is less diligent in identifying and taking down illegal content will disproportionately attract illegal content and so concentrate and magnify the problem. In the absence of due enforcement, what are arguably akin to proceeds of crime could help the firms to become more commercially successful. Thus a lack of such reporting risks explicitly incentivising platforms to evade compliance with the Online Safety Act, and at the expense of platforms that do comply.

Ofcom might also consider making the data publicly available, either by firm or in aggregate. We believe that public disclosure of the data would ensure that platforms' records are scrutinised by interested parties such as ourselves. We would strongly support making the data public since the issue ought not to be commercially sensitive for platforms. In fact, we would be sceptical of the motives of any platform that would seek to keep such data confidential.

iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: No	

Enforcement powers

Question 53:	
i)	Do you have any comments on our draft Online Safety Enforcement Guidance?
Response: No	
ii)	Please provide the underlying arguments and evidence that support your views.
Response: n/a	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: n/a	

Annex 13: Impact Assessments

Question 54:	
i)	Do you agree that our proposals as set out in Chapter 16 (reporting and complaints), and Chapter 10 and Annex 6 (record keeping) are likely to have positive, or more positive impacts on opportunities to use Welsh and treating Welsh no less favourably than English?
Response:	
ii)	If you disagree, please explain why, including how you consider these proposals could be revised to have positive effects or more positive effects, or no adverse effects or fewer adverse effects on opportunities to use Welsh and treating Welsh no less favourably than English.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	