## Your response

Question (Volume 2)	Your response
Question 6.1:	Is this answer confidential? No
Do you have any comments on Ofcom's assessment of the causes and impacts of online harms? Do you think	The risks connected to functionalities that certain platforms and services have – e.g., storage, anonymity, encryption – are well clear and well-explained in OfCom's documents, and it is heartening to see that these affordances are not viewed as inherently harmful, and that their benefits are highlighted.
we have missed anything important in our analysis? Please provide evidence to support your answer.	The main issue I see in this approach is how or whether these risks are quantified and then linked to governance: for example, a paper by Valverde ( <a href="https://jour-nals.sagepub.com/doi/10.1177/096466399900800202">https://jour-nals.sagepub.com/doi/10.1177/096466399900800202</a> ) shows how legal notions of 'risks of harm' fail to define risks' probabilities and the likelihood of harms happening, allowing several stakeholder groups to see and lobby for – certain content and users to be viewed as 'risky', inadvertently playing into political agendas, for the sake of being seen to be regulating harms.
	If things stay as they stand, this approach means identification of risks presents potential to have devastating effects on platforms and their users. Instead, quantifying probabilities of risk, giving platforms a threshold according to which they have to act, might be beneficial.
	Additionally, regulators should also identify over-compliance as a risk. A lot of governance focuses on harms and risks of harm, compelling platforms to act on said risks and harms but not on the connected surveillance and censorship they carry out in this process. To truly protect ALL users, risk assessments should include the possibility that platforms may want to scrap affordances or the hosting of content connected to these risks, compelling them to respect freedom of expression during harm reduction.
Question 6.2:	[Is this answer confidential? No]
Do you have any views about our interpreta- tion of the links be- tween risk factors and different kinds of illegal harm? Please provide	It is heartening to see references to protected categories when it comes to users who are at risk of harm and beneficiaries of certain affordances. Once again, it'd be beneficial to consider risk factors even in the realm of censorship, to see which users would be affected by potential over-compliance (which should also be seen as a risk and a harm).  Previous research on platform governance (e.g., <a href="https://jour-nals.sagepub.com/doi/10.1177/20563051231155103">https://jour-nals.sagepub.com/doi/10.1177/20563051231155103</a> ) has found that

Question (Volume 2)	Your response
evidence to support your answer.	sex workers, women and LGBTQIA+ folks are often affected by platforms' over-compliance with laws and regulations. These groups – as per the policy workshop I led with several of their members <a href="https://www.themoderationar-cana.com/">https://www.themoderationar-cana.com/</a> files/ugd/dfdcfd 860b7a61f3244965b7d9cbd6881a4045.pdf - would happily flag themselves as a protected category to receive more direct help during the moderation of their content ( <a href="https://jour-nals.sagepub.com/doi/10.1177/14614448241228544">https://jour-nals.sagepub.com/doi/10.1177/14614448241228544</a> ). However, it is often these groups – and therefore these protected categories - who face harassment in the form of malicious flagging, through the weaponization of tools meant to mitigate risks for other users ( <a href="https://jour-nals.sagepub.com/doi/10.1177/14614448241228544">https://jour-nals.sagepub.com/doi/10.1177/14614448241228544</a> ). Protected categories due to gender identity, sexuality, race, disability but also due to work — with sex workers being some of the most marginalised workers in society: <a href="https://dsq-sds.org/index.php/dsq/article/view/9097">https://dsq-sds.org/index.php/dsq/article/view/9097</a> - need to be at the forefront of regulators' minds when it comes to the possibility of platforms' over-compliance and of user harassment.

Question (Volume 3)	Your response
Question 8.1:  Do you agree with our proposals in relation to governance and accountability measures in the illegal content Codes of Practice? Please provide underlying arguments and evidence of efficacy or risks to support your view.	[Is this answer confidential? No]  The proposals outlined by OfCom seem proportionate and focused on accountability and transparency, which is much needed in platform governance. However, said accountability is once again focused on the governance of harms, and not on the unintended consequences of harm governance.  As per the Valverde paper, and also this study: <a href="https://www.taylorfrancis.com/chapters/edit/10.4324/9781003200871-51/violence-feminist-potential-content-moderation-carolina-ysabel-gerrard?context=ubx&amp;refld=5fc80f96-fd42-4793-a7b0-032da7bb0f60">https://www.taylorfrancis.com/chapters/edit/10.4324/9781003200871-51/violence-feminist-potential-content-moderation-carolina-ysabel-gerrard?context=ubx&amp;refld=5fc80f96-fd42-4793-a7b0-032da7bb0f60</a> , defining harms matter. The harms defined in your initial session once again largely focus on user-perpetrated or platform-enabled harms, but not on platform-perpetrated harms ( <a href="https://law.yale.edu/sites/default/files/area/center/justice/reimagining-social-media-governance-harm-accountabil-ity-and-repair.pdf">https://law.yale.edu/sites/default/files/area/center/justice/reimagining-social-media-governance-harm-accountabil-ity-and-repair.pdf</a> ), such as censorship. Further discussion on the unintended consequences of governance is needed.

Question (Volume 3)	Your response
Question 8.2:  Do you agree with the types of services that we propose the governance and accountability measures should apply to?	[Is this answer confidential? No] Yes.
Question 8.3:  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?	[Is this answer confidential? No]  An interesting case involving the British Journal of Medicine shows the issues that using third-party providers can cause. The BMJ published an investigation into vaccines which, despite having a clickbait title, did reveal important issues with storage in certain pharmaceuticals companies. Yet, a third party company Facebook used to fact-check stories deemed this to be misinformation, leaving the BMJ unable to appeal because the moderation issue highlighted was not picked up by Facebook themselves, but by their fact-checking third-party, affecting the Journal's image and reads following the misinformation label applied to their content (read more here: <a href="https://www.bmj.com/content/376/bmj.o95">https://www.bmj.com/content/376/bmj.o95</a> ). Third parties need to be held accountable and vetted, and need to be reachable through appeals just like platforms do.
Question: 8.4:  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?	[Is this answer confidential? No]  Commercial content moderators are already often paid by the moderation they perform on each content, which does not encourage them to spend lengthy amounts of time and/or to 'over-think' each decision (see: Ghost Work, <a href="https://ghostwork.info/">https://ghostwork.info/</a> ). Associating a financial reward to something which should be at the heart of platform governance— the need to create a healthy, safe environment for expression and work—may easily encourage workers at platforms to go overboard. This is something we're already seeing: platforms over-complied with US legislation making them liable for promoting trafficking and sex work, and they made sure their algorithms picked up anything remotely sexual as a result—not just to be seen to comply with laws, but also to avoid alienating their advertisers: <a href="https://www.tandfonline.com/doi/full/10.1080/23268743.2021.1974311">https://www.tandfonline.com/doi/full/10.1080/23268743.2021.1974311</a> .  Online safety already has a financial dimension, and tying it to bonuses

Question (Volume 3)	Your response
	would only encourage more censorship. See here: <a href="https://northumbri-ajournals.co.uk/index.php/IJGSL/article/view/1258">https://northumbri-ajournals.co.uk/index.php/IJGSL/article/view/1258</a>
Question 9.1:	[Is this answer confidential? No]
Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	Yes, the proposals seem to ensure a fair and accountable system to review risks. However, as shown by the evidence above, risk assessments tend to bring stakeholders to over-comply with risk management, often affecting marginalised users. Representatives of groups most affected by censorship – the aforementioned sex workers, LGBTQIA+ users, activists etc. – should be included in expert interviews about outcomes and upon reviews.
	It is also worth wondering whether platforms will actually 'tell on them- selves' by showing regulators the risks they are spotting. Is it always in their interest to do so, or will they prefer letting something fly under the radar until they find a solution?
Question 9.2:	[Is this answer confidential? No]
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?	Yes, but once again censorship and over-compliance are not included.
Question 9.3:	[Is this answer confidential? No]
Are the Risk Profiles sufficiently clear and do you think the information provided on risk factors will help you understand	No. Further information, complete with examples and anonymised case studies would provide additional clarity.

Question (Volume 3)	Your response
the risks on your service? <sup>1</sup>	
Question 10.1:	[Is this answer confidential? No]
Do you have any comments on our draft record keeping and review guidance?	It is essential that platforms are transparent about their records and share them with regulators towards education and keeping tabs of different processes even when employees leave. Once a year may be too long a time in the tech world – some practices and affordance changes result in overnight decision-making. I am fully aware of the need to avoid burdening workers, but once every quarter may be a more realistic and thorough approach.  Separately, in my roundtables with de-platformed and harassed users, they expressed interest in accessing their case records and a record of decisions made on their content. This should be made available to all users, off-platforms, to be able to contest decisions (see here: <a href="https://www.the-moderationar-cana.com/files/ugd/dfdcfd-860b7a61f3244965b7d9cbd6881a4045.pdf">https://www.the-moderationar-cana.com/files/ugd/dfdcfd-860b7a61f3244965b7d9cbd6881a4045.pdf</a> ).
Question 10.2:	[Is this answer confidential? No]
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?	Yes.

<sup>&</sup>lt;sup>1</sup> If you have comments or input related the links between different kinds of illegal harm and risk factors, please refer to Volume 2: Chapter 5 Summary of the causes and impacts of online harm).

Question (Volume 4)	Your response
Question 11.1:  Do you have any comments on our overarching approach to developing our illegal content Codes of Practice?	[Is this answer confidential? No]  It is understandable that OfCom will stress platforms' regulation of obvious illegal harms. Yet, as a regulator, OfCom should also focus on how platforms' steps to mitigate or fight those illegal harms influence their treatment of user expression – the limitation of which is also a harm. More information on how OfCom expect platforms to comply with freedom of expression during risk management should be included in the Codes of Practice.
Question 11.2:  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?	[Is this answer confidential? No] Yes.
Question 11.3:  Do you agree with our definition of large services?	[Is this answer confidential? No] Yes.
Question 11.4:  Do you agree with our definition of multi-risk services?	[Is this answer confidential? No] Yes.

Question (Volume 4)	Your response
Question 11.6:  Do you have any comments on the draft Codes of Practice themselves? <sup>2</sup>	[Is this answer confidential? No] As above.
Question 11.7:  Do you have any comments on the costs assumptions set out in Annex 14, which we used for calculating the costs of various measures?	[Is this answer confidential? No]  Platforms should be compelled to also enforce freedom of speech – there is a risk that, while being asked to undertake costly moderation of contentious content, they may scrap borderline content altogether.
Question 12.1:  Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No]  Yes. It is heartening to see the need to strike a balance between taking down illegal harm and making sure accurate content moderation decisions are made. It is also great to see there's an emphasis on moderator education — the same education should be extended to users, who should receive more specific information about the violations they committed (see here: <a href="https://www.themoderationar-cana.com/_files/ugd/dfdcfd_860b7a61f3244965b7d9cbd6881a4045.pdf">https://www.themoderationar-cana.com/_files/ugd/dfdcfd_860b7a61f3244965b7d9cbd6881a4045.pdf</a> ).  It would be interesting to receive more information about how trusted flaggers' work is evaluated, and accountability measures for said flaggers.
Question 13.1:  Do you agree with our proposals? Please provide the underlying argu-	[Is this answer confidential? No] Yes.

<sup>&</sup>lt;sup>2</sup> See Annexes 7 and 8.

Question (Volume 4)	Your response
ments and evidence that support your views.	
Question 14.1:	[Is this answer confidential? No]
Do you agree with our proposals? Do you have any views on our three proposals, i.e. CSAM hash matching, CSAM URL detection and fraud keyword detection? Please provide the underlying arguments and evidence that support your views.	Yes, but any automated process should be complemented by swift, effective and direct appeals with a human platform worker at the other end to prevent over-enforcement and mistakes. See here: <a href="https://www.themoderationar-cana.com/">https://www.themoderationar-cana.com/</a> files/ugd/dfdcfd 860b7a61f3244965b7d9cbd6881a4045.pdf.
Question 14.2:	[Is this answer confidential? No]
Do you have any comments on the draft guidance set out in Annex 9 regarding whether content is communicated 'publicly' or 'privately'?	No.

Question	Your response
(Volume 4)	
Question 14.3:	[Is this answer confidential? No]
Do you have any	No.
relevant evi-	
dence on:	
The accu-	
racy of	
percep-	
tual hash	
matching	
and the	
costs of	
applying	
CSAM	
hash	
matching	
to	
smaller	
services;	
The abil-	
ity of ser-	
vices in	
scope of	
the	
CSAM	
hash	
matching	
measure	
to access	
hash da-	
ta-	
bases/ser	
vices,	
with re-	
spect to	
access	
criteria	
or re-	
quire-	
ments	
set by	
database	
and/or	

Question	Your response
(Volume 4)	·
hash	
matching	
service	
provid-	
ers;	
The costs	
of apply-	
ing our	
CSAM	
URL de-	
tection	
measure	
to	
smaller	
services,	
and the	
effective-	
ness of	
fuzzy match-	
ing <sup>3</sup> for CSAM	
URL de-	
tection;	
• The costs	
of apply-	
ing our	
articles	
for use in	
frauds	
(standard	
keyword	
detec-	
tion)	
measure,	
including	
for	
smaller	
services;	
and	
3114	

\_

 $<sup>^3</sup>$  Fuzzy matching can allow a match between U2U content and a URL list, despite the text not being exactly the same.

Question (Volume 4)  • An effective application of hash matching and/or URL detection for terrorism content, including how such	
<ul> <li>An effective application of hash matching and/or URL detection for terrorism content, including</li> </ul>	
tive application of hash matching and/or URL detection for terrorism content, including	
plication of hash matching and/or URL de- tection for ter- rorism content, including	
of hash matching and/or URL de- tection for ter- rorism content, including	
matching and/or URL de- tection for ter- rorism content, including	
and/or URL de- tection for ter- rorism content, including	
URL de- tection for ter- rorism content, including	
tection for ter- rorism content, including	
for ter- rorism content, including	
rorism content, including	
content, including	
including	
how such	
measures	
could ad-	
dress	
concerns	
around	
'context'	
and free-	
dom of	
expres-	
sion, and	
any infor- mation	
you have	
on the costs and	
efficacy	
of apply-	
ing hash	
matching	
and URL	
detection	
for ter-	
rorism	
content	
to a	
range of	
services.	

Question (Volume 4)	Your response
Question 15.1:  Do you agree with our pro- posals? Please provide the un- derlying argu- ments and evi- dence that sup- port your views.	[Is this answer confidential? No]  Yes, but once again consider the challenges of over-enforcement and allow users to appeal and reach human moderators / workers to address this issue.
report content, the weaponization of flagging from malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor be considered – this is particularly effective towards silencing malicious actor by considered – this is particularly effective towards silencing malicious actor by considered – this is particularly effective towards silencing malicious actor by considered – this is particularly effective towards silencing malicious actor by considered – this is particularl	While it is important to establish trusted flaggers and to allow audiences to report content, the weaponization of flagging from malicious actors needs to be considered – this is particularly effective towards silencing marginalised communities, particularly if their content is already frowned upon by platforms, even generating flagging scams. See:  1. https://journals.sagepub.com/doi/10.1177/14614448241228544 2. https://journals.sagepub.com/doi/10.1177/13548565231218629 3. https://www.propublica.org/article/instagram-fraudster-ban-influencer-accounts 4. https://oversightboard.com/decision/BUN-IH313ZHJ/ 5. https://www.vice.com/en/article/k78kmv/instagram-ban-restore-service-scam.  On the opposite side of the spectrum is the fact that often, platforms do not share information as to whether user flags are being taken up, and definitely do not treat content flagged by marginalised communities as violating (e.g.
	see <a href="https://www.oversightboard.com/news/1376420189678927-oversight-board-overturns-meta-s-original-decision-in-post-in-polish-targeting-trans-people-case/">https://jour-nals.sagepub.com/doi/10.1177/01634437221140531</a> ), showing concerning double standards between content moderation of posts by dominant and marginalised groups: <a href="https://www.brennancenter.org/our-work/research-re-ports/double-standards-social-media-content-moderation?ref=welcometo-hellworld.com">https://www.brennancenter.org/our-work/research-re-ports/double-standards-social-media-content-moderation?ref=welcometo-hellworld.com</a> .  The issue of trusted flaggers does not necessarily remove this issue, further empowering state actor to surveil activists and users whose work is disproportionately policed by law enforcement (namely, sex workers). As such, transparency in the choice of trusted flaggers, the ability to include marginal-ised communities in the process — either through direct collaboration or

Question (Volume 4)	Your response
	through training from relevant community members – can therefore be helpful.
Question 17.1:	[Is this answer confidential? No]
Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	Yes, users need clear, understandable and accessible ToS. These would benefit from anonymised examples of violating posts to better educate users.  Platforms should inform users of all the enforcement decisions they make on content, providing clear definitions of violating content in the ToS, clear notifications and explanations of violating content to users upon violation, and transparency about internal policies. The deletion of swathes of kink and sex positive accounts from Instagram in the summer of 2023 is a case in point to show the damage unclear ToS and internal policies can make: it is likely these accounts were deleted as per IG's "implicit solicitation" policy, something users struggle to understand and that platforms apply subjectively:  https://www.dazeddigital.com/life-culture/article/60228/1/instagram-keeps-banning-sex-positive-and-kink-accounts-censorship-creators.  Further, platforms need to be clearer in their ToS and notifications to users about their enforcement of demotion / shadowbanning techniques, which they tend to often deny or share little information about. See below:  1. https://www.tandfonline.com/doi/full/10.1080/14680777.2021.1928 259  2. https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1974 311  3. https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1994 624.  Platforms need to be compelled to be transparent for users and regulators alike to hold them accountable for their governance.
Question 17.2:  Do you have any evidence, in particular on the use of prompts, to guide further work in this area?	[Is this answer confidential? No] No.

Question (Volume 4)	Your response
Question 18.1:  Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No]  I agree that children should have additional protection and that certain affordances should be mitigated for them. However, given platforms' previously opaque and puritan moderation, content that may have been beneficial to teenagers such as sex and pleasure education, sexual health information, mental health support and connections with community of fellow Queer and Transgender users (e.g. see here: <a href="https://jour-nals.sagepub.com/doi/10.1177/13548565231218629">https://jour-nals.sagepub.com/doi/10.1177/13548565231218629</a> ) has been mistakenly censored and hidden from teen users, who rely on it for support they do not receive at home or at school. Any invalidation of affordances needs to protect activism, education and support. For some kids, accessing this content can be a key to safety, consent, or a way to escape families that do not agree with their identities. Creating barriers to their access to it can be harmful.
Question 18.2: Are there functionalities outside of the ones listed in our proposals, that should explicitly inform users around changing default settings?	[Is this answer confidential? No] In my roundtable workshops with users (see here: <a href="https://www.themoderationar-cana.com/">https://www.themoderationar-cana.com/</a> files/ugd/dfdcfd 860b7a61f3244965b7d9cbd6881a4045.pdf), codesgined functionalities to further empower account owners to decide what to see included periodic checks about which content they wished and didn't wish to see, and also allowing creators to age-gate single posts, to avoid their whole accounts being restricted. For example, a sex educator may want teens to access a post about consent, but not a more sexually charged artistic nude picture. Empowering users to do that may allow for less blanket moderation.
Question 18.3:  Are there other points within the user journey where under 18s should be informed of the risk of illegal content?	[Is this answer confidential? No]  Not that I can think of.

Question (Volume 4)	Your response
Question 19.1:	[Is this answer confidential? No]
Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	I agree that the use of recommender systems can be crucial to limit the spread of harmful information. However, once again provision towards the tackling of harms do not account for the harms that demotion via recommended systems can cause to users who directly work through social media, such as content creators: diminished reach causes brand partnerships to fail, creator reward programs to return very little reward, as well as frustration and isolation – an emotional and financial impact. Side effects of demotion / lack of recommendation of harmful content needs to provide mechanisms for redress. See below:  1. https://www.tandfonline.com/doi/full/10.1080/14680777.2021.1928 259
	<ol> <li>https://www.tandfonline.com/doi/full/10.1080/23268743.2021.1974 311</li> <li>https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1994 624.</li> <li>As such, workers' rights need to be built into content moderation, and platforms' 'algorithmic boss' role needs to be acknowledged as potentially harmful to workers. See below:</li> <li>https://journals.sagepub.com/doi/full/10.1177/2057047320959855</li> <li>https://ijoc.org/index.php/ijoc/article/view/15761</li> </ol>
Question 19.2:	[Is this answer confidential? No]
What evaluation methods might be suitable for smaller services that do not have the capacity to perform on-platform testing?	Not sure.
Question 19.3:	[Is this answer confidential? No]
We are aware of design features and parameters that can be used	Not sure.

Question (Volume 4)	Your response
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. Are you aware of any other design parameters and choices that are proven to improve user safety?	
Question 20.1:  Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No]  Yes, but the blocking and restricting does put most of the onus of self-defence on users, and often results in slip ups on the platform side — see here: https://journals.sagepub.com/doi/10.1177/01634437221140531.  As such, protected categories should have access to specific teams within platforms who can help them address specific harassment issues.
Question 20.2:  Do you think the first two proposed measures should include requirements for how these controls are made known to users?	[Is this answer confidential? No] Yes.

Question (Volume 4)	Your response
Question 20.3:  Do you think there are situations where the	[Is this answer confidential? No]  Meta and Twitter's new process to allow users to pay for verification has great potential for fraud and disinformation. This needs to be reassessed following these changes.
labelling of accounts through voluntary verification schemes has particular value or risks?	
Question 21.1:	[Is this answer confidential? No ]
Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	Yes.
Question 21.2:	[Is this answer confidential? No ]
Do you have any supporting information and evidence to inform any recommendations we may make on blocking sharers of CSAM	No.

Question	Your response
(Volume 4)	
content? Specifi- cally:	
What are	
the op-	
tions	
available	
to block	
and pre-	
vent a	
user from re-	
turning	
to a ser-	
vice (e.g.	
blocking	
by	
usernam	
e, email	
or IP ad-	
dress, or	
a combi-	
nation of	
factors)?	
What are	
the ad-	
vantages	
and dis-	
ad-	
vantages	
of the	
different	
options,	
including	
any po- tential	
impact	
on other	
users?	
How long	
should a	
user be	
blocked	

Your response
real response

Question (Volume 4)	Your response
alternative options to immediate blocking (such as a strikes system) that might help mitigate some of the risks and impacts on user rights?	
Question 22.1:  Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	[Is this answer confidential? No] Yes.
Question 23.1:  Do you agree that the overall burden of our measures on low risk small and micro businesses is proportionate?	[Is this answer confidential? No]  Yes, pending the additional work defining, quantifying risks and creating redress and reversal for the unintended consequences of harm reduction, while also compelling platforms to respect and uphold freedom of expression.

Question (Volume 4)	Your response
Question 23.2:	[Is this answer confidential? No]
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?	As above.
Question 23.3:	[Is this answer confidential? No]
We are applying more measures to large services. Do you agree that the overall burden on large services proportionate?	Yes.
Question 24.1:	[Is this answer confidential? No]
Do you agree that Ofcom's pro- posed recom- mendations for the Codes are ap- propriate in the	Yes, as explained above however more is needed to account for over-compliance and over-moderation, and to provide at-risk and overly-targeted users with direct support.

Question (Volume 4)	Your response
light of the mat- ters to which Ofcom must have regard? If not, why not?	

Question (Volume 5)	Your response
Question 26.1:  Do you agree with our proposals, including the detail of the drafting? What are the underlying arguments and evidence that inform your view.	[Is this answer confidential? / No]  While it is essential that platforms only remove what is illegal, and that they are judged by OfCom according to the illegal content they remove, the fact that they frequently remove legal content that is not harmful needs to be acknowledged. Over-enforcement is barely addressed in these provisions.
Question 26.2:  Do you consider the guidance to be sufficiently accessible, particularly for services with limited access to legal expertise?	[Is this answer confidential? / No]  No – and not just for smaller services without access to legal expertise.  Large-scale platforms have so far struggled with legal definitions, centring  US legislation by applying it to local content – see their reactions to  FOSTA/SESTA here: <a href="https://www.tandfonline.com/doi/full/10.1080/23268743.2021.1993972">https://www.tandfonline.com/doi/full/10.1080/23268743.2021.1993972</a> .  Are platforms going to be compelled by UK law alone? And if so, how?
Question 26.3: What do you think of our assessment of what information is reasonably available	[Is this answer confidential? / No]  I'd say the main issue with most of these approaches is that they assume platforms will act in good faith, something they have not done in cases such as Cambridge Analytica in the past. Information was reasonably available to them in the past, but due to their monopolies and power,

Question (Volume 5)	Your response
and relevant to illegal content judgements?	tackling it is not always at the forefront of their minds to protect their image.

Question (Volume 6)	Your response
Question 28.1:	[Is this answer confidential? / No]
Do you have any comments on our proposed approach to information gathering powers under the Act?	More information on the storing of such information, and the protection of vulnerable users – to the state, to platforms – would be welcome.
Question 29.1:	[Is this answer confidential? / No]
Do you have any comments on our draft Online Safety Enforcement Guidance?	No.

Question (Annex 13)	Your response
Question A13.1:	[Is this answer confidential? / No]
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	Not sure.

Question (Annex 13)	Your response
Welsh no less favourably than English?	
Question A13.2: 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 ef- fects or fewer adverse effects on opportunities to use Welsh and treating Welsh no less favourably than English.	[Is this answer confidential? / No] N/A

Please complete this form in full and return to  $\underline{\text{IHconsultation@ofcom.org.uk}}.$