Your response

Question (Volume 2)	Your response
Question 6.1: Do you have any comments on Ofcom's assessment of the causes and impacts of online harms? Do you think we have missed anything important in our analysis? Please provide evidence to support your answer.	*[Is this answer confidential? No]* We are pleased to see 'Proceeds of Crime' and 'Fraud and financial services' offences being included as separate offences within the risk profiles. We also agree with Ofcom's assessment of the causes and impacts. Specifically, we agree with the assessment of the prominent and harmful role that social media services including online marketplaces and messaging services hosted by firms with inadequate controls for vetting users and monitoring activity play in enabling the spread of priority illegal content/online harms/offences including but not limited to, the recruitment of money mules. **The scale of this type of harm is supported by our own data which showed that in 2023, more than 84% of purchase scams, impersonation scams and investment scams combined and reported to us, originated from Meta platforms in particular.**
Question 6.2: 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.	

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]* Governance and accountability processes are key to identifying and managing online safety risks and we agree that governance and accountability measures should not only be based on services' size with regards online harms but that measures should reflect the risks and impacts of types of harms perpetrated via individual services. We also believe that the introduction of yearly risk assessments as outlined by Ofcom is key – as is the ongoing and continuous proactive identification and management of new and emerging risks outside the yearly cycle.

Question (Volume 3)	Your response
	However, on the basis that performing risk assessments in line with the four steps recommended by Ofcom is not mandatory, the effectiveness of risk assessments performed may be negatively affected. Where risk assessments are however performed in line with Ofcom's recommendations, these may be seen as a 'tick box' exercise on the basis that services that follow Ofcom guidance will be deemed to be compliant, regardless of the comprehensiveness of those risk assessments. **Key to risk assessments and specifically for fraud related illegal harms, social media platform services should not be allowed to finalise their own risk assessments without directly incorporating data from banks (See question 8.3).** In addition, we also recommend that there should be a requirement for services to publish comprehensive and meaningful transparency reports, similar to Payment Systems Regulator fraud data being published by the banking sector. As a minimum, these transparency reports should include data on the following: ➢ Volumes and types of illegal harms (i.e. scams) being enabled via their platforms as reported by users ➢ Volumes of illegal harms being reported by 'others' ie via trusted flaggers ➢ Volumes and types of complaints received, time to resolve and actions taken ➢ Time taken to take down reported scam posts ➢ Financial losses reported by victims ➢ Outputs from regular sample testing performed This is important to ensure that services are held accountable for the enforcement of their own standards and for complying with the new regime. It would also ensure that users and the public also have visibility of online illegal harms and how these are being managed and mitigated and can hold services to account too as well as making informed decisions as to whether to use certain services or not.

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]* We have no comments.
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]* We should not only rely on services marking their own homework. This potential future measure should be seen as **mandatory** as the effective implementation of the new Online Safety regime by services is pivotal to the reduction of online harms and as such, all 'levers' enabling the achievement of this goal should be used. **Based on data held by TSB, our view of how widespread and impactful online harms are does not align with the view held by certain social media platforms services themselves and it is our opinion that to date, actions taken by those services to address the current fraud epidemic have fallen short despite the evidentiary data we hold being available.** We therefore recommend the adoption, as 'mandatory', of third party independent audits at the earliest opportunity. We note the limitations cited (costs and effectiveness) by Ofcom however, audits by third parties are key to independently and holistically assessing the effectiveness of services' ongoing mitigation and management of illegal contents risks, the robustness of risk assessments performed, as well as the effectiveness of services' own internal monitoring and assurance functions especially in cases where services choose to comply with duties in 'other ways' i.e. they design their own measures (Option 1) or they choose to 'pick and mix' (Option 3) as opposed to following all measures proposed by Ofcom (Option 2). Not only would this additional measure provide a level of assurance to services being audited themselves (and as such, ought to be sought by them), but it may also go

Question (Volume 3)	Your response
	some way towards public trust in those services being improved.
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]* This potential future measure should be considered as being **mandatory** because remuneration is an important factor in determining behaviour and better risk management so those named as accountable to the most senior governing body for compliance with illegal content duties should be held to account and measures linked to remuneration should go beyond simply demonstrating compliance with the new regime, i.e. they should evidence positive online safety outcomes. In addition, 'incentives' more generally are important too. If the right incentives to stop fraud being enabled in the first place are introduced, these too can lead to the right changes being identified and implemented. For example, as a result of the right incentives being present (i.e. the reimbursement of victims and the impact of potential regulatory actions), banks have significantly invested in their fraud controls over the past few years with TSB specifically introducing its self imposed Fraud Refund Guarantee which created an ongoing incentive to maintain such strong controls. As the appointed regulator, Ofcom will also need to ensure that prompt and effective enforcement action is taken where appropriate.
Question 9.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* As per our response to question 8.1, we believe that recommendations made to services should not necessarily be based on services' size, but instead, based on the scale of risk potentially enabled/being enabled, especially for those services for which rich and actionable data on the extent of online illegal harms currently being enabled is available.

Question (Volume 3)	Your response
	For example, data on the fraud harms enabled via social media platforms is currently available and includes volumes of victims and financial losses enabled by them as well as levels of reimbursement borne by the financial services sector (see our response to question 6.1).
Question 9.2: 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?	*[Is this answer confidential? No]* The four step risk assessment process and the risk profiles, based on best practice industry frameworks common elements, look to be comprehensive and should prove to be useful models for in scope services. On a practical level, and on the basis that 'services will need to determine for themselves what approach they need to take', key to the risk assessment process will be the inclusion of additional and readily available evidence, such as that held by 'relevant representative groups'. Currently, this is only recommended in cases where it is 'deemed by services that other evidence does not provide them with a 'sufficiently good understanding of their risk levels'' whilst it is also noted that 'risk assessment should, as far as possible, be based on relevant evidence'. We firmly believe that the inclusion of such evidence, especially for some social media platforms services, should be **mandatory** especially since Ofcom has indicated that the presence of harmful content may not be a sign of non compliance with the new regime. Failure to do so would affect the effectiveness of yearly risk assessments being performed and would therefore materially limit the prevention of online harms.
Question 9.3: Are the Risk Profiles sufficiently clear and do you think the information provided on risk factors will help you understand the risks on your service?[1]	*[Is this answer confidential? No]* Key will be for services to ensure that a complete and comprehensive list of potential online harm risks, classified by category and/or type, is identified and considered as part of risk assessments to ensure a complete and accurate evaluation of those risks.

[1] 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 3)	Your response
	Only then will the correct mitigants be identified.
Question 10.1: Do you have any comments on our draft record keeping and review guidance?	*[Is this answer confidential? No]* We have no comments.
Question 10.2: 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?	*[Is this answer confidential? No]* We have no comments.

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]* We have no comments.
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]* Please see our responses to questions 8.1 and 9.1. We do not believe that the word 'onerous' is being used appropriately on the basis that it implies that some measures are deemed to be excessively complex, extreme, troublesome and/or unjustified when in fact they are anything but.

Question (Volume 4)	Your response
	We believe that those measures referenced should be applied to services based on the actual and demonstrable harm enabled by them. Whilst the identification of services which meet the criteria of large and/or medium will be straightforward, the identification of 'high risk' services is therefore key. And for those smaller services where there is no data supporting the evidence of online harms to date, measures should be proportionate and less 'onerous'.
Question 11.3: Do you agree with our definition of large services?	*[Is this answer confidential? No]* We have no comments.
Question 11.4: Do you agree with our definition of multi-risk services?	*[Is this answer confidential? No]* We have no comments.
Question 11.6: Do you have any comments on the draft Codes of Practice themselves?[2]	*[Is this answer confidential? No]* We have no comments.
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]* We have no comments.

[2] See Annexes 7 and 8.

Question (Volume 4)	Your response
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]* We have no comments.
Question 13.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* We have no comments.
Question 14.1: 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.	*[Is this answer confidential? No]* We have no comments.
Question 14.2: Do you have any comments on the draft guidance set out in Annex 9 regarding whether content is communicated 'publicly' or 'privately'?	*[Is this answer confidential? No]* We have no comments.
Question 14.3: Do you have any relevant evidence on: • The accuracy of perceptual hash matching and the costs of applying CSAM hash matching to smaller services; • The ability of services in scope of the CSAM hash	*[Is this answer confidential? No]* We have no comments.

Question (Volume 4)	Your response
matching measure to access hash databases/services, with respect to access criteria or requirements set by database and/or hash matching service providers; • The costs of applying our CSAM URL detection measure to smaller services, and the effectiveness of fuzzy matching[3] for CSAM URL detection; • The costs of applying our articles for use in frauds (standard keyword detection) measure, including for smaller services; and • 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.	
Question 15.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* We have no comments.

[3] Fuzzy matching can allow a match between U2U content and a URL list, despite the text not being exactly the same.

Question (Volume 4)	Your response
Question 16.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* 'Reporting and complaints' and 'Dedicated reporting channels for services with risks of fraud' sections. With regards the fraud online harm enabled via social media platforms, 'users who want to complain to social media firms should be able to do so whilst being treated fairly and should get a response in good time'. We support Ofcom's plans that in the first instance, focus should be placed on creating a DRC related to fraud, since there are currently various challenges, as noted by Ofcom, with reporting fraud into online services. This would ensure that trusted information supplied – aggregated and case specific - is acted upon with 'priority and without delay'. However, our preference would be for option 2 articulated in section 'Measure 6' to be implemented, allowing regulated entities to have 'trusted flagger status'. In choosing Option 1, Ofcom proposals state that this is a first step and thinking is underway as to how a process could be implemented where banks have a dedicated reporting channel. **This sounds good in theory however in practice, Fraud isn't like other harms in that victims report fraud to their bank, as shown by available data, and instances will further increase when the new Payment Systems Regulator mandatory reimbursement rules are implemented in October 2024.** This means that: ➢ **For consumers/users**, there is no incentive to report scams to social media platform services on the basis that once scams are reported to banks, these will be investigated and reimbursements made where relevant; ➢ **For banks,** under current proposals, nor would they qualify for 'trusted flagger' status nor would they have access to their own 'dedicated reporting channel' despite becoming the 'affected' party from the point at which scam losses have been reimbursed by them with no options to explore to recoup some or all of the funds from those services who enabled the harm (including financial losses) in the first instance.

Question (Volume 4)	Your response
	Therefore, as stated in earlier responses, it is key that data that can be made available by the financial services industry be shared with services and included without fail by them in their risk assessments, that they have trusted flagger status and access to their own dedicated reporting channel. The concern noted by Ofcom that services could be 'overwhelmed' is overstated on the basis that in our experience, one firm specifically would run this risk although this would be unlikely given its revenue and size. It also seems unreasonable to expect public bodies to use time and resources acting as an intermediary between banks and platforms when this could be done directly between firms. Payment services firms have existing regulatory responsibilities to operate processes to receive fraud reports from their customers and the best approach would be to allow banks to handle this on behalf of the victim, linking in with the relevant enabling platform as required.
Question 17.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* We have no comments.
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]* We have no comments.
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]* We have no comments.

Question (Volume 4)	Your response
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]* We have no comments.
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]* We have no comments.
Question 19.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* We have no comments.
Question 19.2: What evaluation methods might be suitable for smaller services that do not have the capacity to perform on-platform testing?	*[Is this answer confidential? No]* We have no comments.
Question 19.3: 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. Are you aware of any other design parameters and choices that are proven to improve user safety?	*[Is this answer confidential? No]* We have no comments.

Question (Volume 4)	Your response
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]* We have no comments.
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]* We have no comments.
Question 20.3: Do you think there are situations where the labelling of accounts through voluntary verification schemes has particular value or risks?	*[Is this answer confidential? No]* We have no comments.
Question 21.1: Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	*[Is this answer confidential? No]* We have no comments.

Question (Volume 4)	Your response
Question 21.2: Do you have any supporting information and evidence to inform any recommendations we may make on blocking sharers of CSAM content? Specifically: • 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)? What are the advantages and disadvantages of the different options, including any potential impact on other users? • How long should a user be blocked for sharing known CSAM, and should the period vary depending on the nature of the offence committed? • 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. 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?	*[Is this answer confidential? No]* We have no comments. *[Is this answer confidential? Yes / No (delete as appropriate)]*

Question (Volume 4)	Your response
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]* We have no comments.
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]* We have no comments.
Question 23.2: 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?	*[Is this answer confidential? No]* We have no comments.
Question 23.3: We are applying more measures to large services. Do you agree that the overall burden on large services proportionate?	*[Is this answer confidential? No]* We have no comments.
Question 24.1: 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? If not, why not?	*[Is this answer confidential? No]* We have no comments.

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]* We have no comments.
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]* We have no comments.
Question 26.3: What do you think of our assessment of what information is reasonably available and relevant to illegal content judgements?	*[Is this answer confidential? No]* We have no comments.

Question (Volume 6)	Your response
Question 28.1: Do you have any comments on our proposed approach to information gathering powers under the Act?	*[Is this answer confidential? No]* We have no comments.
Question 29.1: Do you have any comments on our draft Online Safety Enforcement Guidance?	*[Is this answer confidential? No]* We have no comments.

Question (Volume 6)	Your response

Question (Annex 13)	Your response
Question A13.1: 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?	*[Is this answer confidential? No]* We have no comments.
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 effects or fewer adverse effects on opportunities to use Welsh and treating Welsh no less favourably than English.	*[Is this answer confidential? No]* We have no comments.

Please complete this form in full and return to IHconsultation@ofcom.org.uk.