

# **Consultation response form**

Please complete this form in full and return to <a href="mailto:protectingchildren@ofcom.org.uk">protectingchildren@ofcom.org.uk</a>.

Consultation title	Consultation: Protecting children from harms online
Organisation name	Welsh Government

# Question

# Your response

Volume 2: Identifying the services children are using Children's Access Assessments (Section 4).

Do you agree with our proposals in relation to children's access assessments, in particular the aspects below. Please provide evidence to support your view.

- 1. Our proposal that service providers should only conclude that children are not normally able to access a service where they are using highly effective age assurance?
- 2. Our proposed approach to the child user condition, including our proposed interpretation of "significant number of users who are children" and the factors that service providers consider in assessing whether the child user condition is met?
- 3. Our proposed approach to the process for children's access assessments?

Confidential? - N

We agree that it is important any age assurance technologies are 'highly effective' as many in current use do not fulfil this criteria. Further, many Part 3 service providers may not know the (true) age of their users and therefore mustn't be able to make assumptions about who is accessing their services. As many children and young people access services for which they are not the intended audience, the child user condition is important to include.

We also support Ofcom's decision not to propose a numerical threshold for *significant number* - context is indeed necessary and 'significant' may not be a comparatively large number.

Due to the speed at which online issues and trends can propagate, we believe it is important that services keep their assessments under review. A service that wasn't previously likely to attract a significant number of children could become one in a relatively short space of time. Therefore, we agree with the requirement (4.50) that providers carry out a new assessment in response to new evidence about a significant increase in the number of children using the service.

#### Volume 3: The causes and impacts of online harm to children

#### **Draft Children's Register of Risk (Section 7)**

#### Proposed approach:

- 4. Do you have any views on Ofcom's assessment of the causes and impacts of online harms? Please provide evidence to support your answer.
- a. Do you think we have missed anything important in our analysis?
- 5. Do you have any views about our interpretation of the links between risk factors and different kinds of content harmful to children? Please provide evidence to support your answer.
- 6. Do you have any views on the age groups we recommended for assessing risk by age? Please provide evidence to support your answer.
- 7. Do you have any views on our interpretation of non-designated content or our approach to identifying non-designated content? Please provide evidence to support your answer.

#### **Evidence gathering for future work:**

- 8. Do you have any evidence relating to kinds of content that increase the risk of harm from Primary Priority, Priority or Non-designated Content, when viewed in combination (to be considered as part of cumulative harm)?
- 9. Have you identified risks to children from GenAl content or applications on U2U or Search services?
- a) Please Provide any information about any risks identified

#### Confidential? - N

From a digital resilience perspective, the assessment appears robust and aligns with evidence we have heard from online safety experts.

We agree with the additional characteristics and support Ofcom's view that certain service types such as social media are more likely to lead children to harmful content. We also agree that recommender systems can be responsible for actively promoting negative content to children and young people.

We believe effective reporting and content moderation are hugely important in tackling online harm and we agree with Ofcom's assessment that many children encounter content they don't want to see and consequently attempt to report.

From a modern slavery perspective, the previous consultation on protecting people from illegal harms online set out the ways in which online content is connected to human trafficking offences, including county lines exploitation, and the wider impacts of this exploitation.

However, with this consultation, we could not find evidence of causes and harms relating to human trafficking offences, including criminal exploitation and county lines in section 7. Criminal gangs use social media platforms such as snapchat as recruitment channels – for example posting material of children in trap houses with weapons, money, and new trainers. These platforms are used to groom children and recruit them into criminal gangs – at which point children may be coerced and blackmailed into carrying drugs and weapons, and their movements tracked.

County lines gangs are using social media to target, groom, and exploit children as young as 11. According to an <u>article by Unseen UK</u>, "between 2017 and 2022, online grooming crimes surged by 82%, with 73% involving platforms like Snapchat, Facebook, Instagram and WhatsApp."

- 10. Do you have any specific evidence relevant to our assessment of body image content and depressive content as kinds of non-designated content? Specifically, we are interested in:
- a) (i) specific examples of body image or depressive content linked to significant harms to children,
- b. (ii) evidence distinguishing body image or depressive content from existing categories of priority or primary priority content.
- 11. Do you propose any other category of content that could meet the definition of NDC under the Act at this stage? Please provide evidence to support your answer.

Section 7.6 documents the risk of harm from exposure to violent content, with subsection 7.6.27 stating that "encountering violent content can encourage specific behaviours relating to violence, such as carrying weapons" and section 7.6.28 stating there is "evidence to suggest that violent content can directly contribute to, or trigger, acts of violence." From a modern slavery perspective, it should be noted that the desensitisation and normalisation of this violence can also contribute to a young person's vulnerability to exploitation by criminal gangs, where violent activity and gang culture has been glamorised and normalised through exposure to online content.

There also seems to be a gap in the proposals which covers online content that is used to target children with the intent of criminal exploitation and county lines.

From a mental health, suicide and self-harm perspective, we support the proposals and the suggested requirement placed on providers to protect children from online harm. This is consistent with the principles within both the Welsh Government's draft Mental Health and Well-being Strategy and draft Suicide and Self-harm Prevention (SSHP) Strategy.

Noting the specific references to online harm, the draft SSHP Prevention Strategy makes specific reference to the need to enhance online safety. As part of our monitoring framework, we are keen to work with Ofcom and the UK Government to monitor the impact of the legislation on suicide and self-harm in Wales.

#### **Draft Guidance on Content Harmful to Children (Section 8)**

- 12. Do you agree with our proposed approach, including the level of specificity of examples given and the proposal to include contextual information for services to consider?
- 13. Do you have further evidence that can support the guidance provided on different kinds of content harmful to children?

#### Confidential? - N

We believe it is helpful to have specific examples, but the guidance must emphasise these are indicative and that services need to consider more than only what's mentioned here. We think it is useful to include contextual factors as so much of what can be harmful is nuanced.

Following feedback from industry experts, we would encourage Ofcom to make clear that the content referred to within the guidance (the three kinds of con-

- 14. For each of the harms discussed, are there additional categories of content that Ofcom
- a) should consider to be harmful or
- b) consider not to be harmful or
- c) where our current proposals should be reconsidered?

tent in 62(4), 62(6) and 62(7) of the Act) is not restricted to visual content. Helplines have reported an increase in reporting harms caused by audio content (specifically audio pornography). We believe that the guidance would benefit from setting out clearly that 'content' may be visual and or audio.

From a community cohesion perspective, under 'Table 8.6.3: Descriptions and examples of content that incites hatred and is harmful to children' it would be useful to include reference to more extremist hateful views, which go beyond discrimination. The references to dehumanisation/non-humans go some way to capturing this, but perhaps more explicit mention of extremist hateful abuse, for example expressions of support for the murder or genocide of groups of people. There is a difference between pushing the idea that some people are a sub-class to be treated inferior and calling for and celebrating the death of a group of people.

Under the content description 'Defending or legitimising threatening action against a group of people', it would be useful to also see explicit reference to the 'encouragement' of threatening action as part of the description.

Under the content description 'Content which objectifies and demeans a person on the basis of their listed characteristic', could a reference to defamatory language/claims be added. This would capture the harder end of the tropes or narratives around some groups of people, which go beyond demeaning stereotypes. For example, calling a Muslim person a terrorist.

From a LGBTQ+ policy perspective, we want to ensure children are protected children from harmful content online.

In the current social climate, there are those who suggest that the provision of information on LGBTQ+ subjects is harmful to children. For example, with trans experiences, some would argue that children viewing content which explains the transition process or provides information on surgical routes (which are not available in the UK to children) is "encouraging self-

harm" as they reject the notion that such interventions for trans people of any age are an appropriate response to gender dysphoria.

Therefore, it is important that decisions on what content is defined as harmful is based on fact and evidence and removed from outside influence.

It could be argued that by not allowing children to be informed about LGBTQ+ issues, alongside other content promoting inclusion and understanding, this may be harmful in itself.

# Volume 4: How should services assess the risk of online harms? Governance and Accountability (Section 11)

- 15. Do you agree with the proposed governance measures to be included in the Children's Safety Codes?
- a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.
- b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.
- 16. Do you agree with our assumption that the proposed governance measures for Children's Safety Codes could be implemented through the same process as the equivalent draft Illegal Content Codes?

#### Confidential? - N

We believe that all Part 3 services, regardless of size, should be accountable and take whatever steps they can to keep children safer online.

It is important these measures are properly considered and not just seen as a tick box exercise. Internal monitoring and assurance must be robust – as with age assurance, it's not enough to have policies / checks in place if they are inadequate or ineffective. Similarly, codes of conduct and training must be effective to ensure that all service staff understand their role.

We support Measure GA5 (horizon scanning / tracking new harms). This is important due to the speed at which new harms arise and we strongly agree with the point raised at 11.146/147 relating to how the risk of harm can change over time / new trends appear.

We support using the same process for both sets of Codes, as long as providers understand they are distinct.

Children's Risk Assessment Guidance and Children's Risk Profiles' (Section 12)

- 17. What do you think about our proposals in relation to the Children's Risk Assessment Guidance?
- a) Please provide underlying arguments and evidence of efficacy or risks that support your view.
- 18. What do you think about our proposals in relation to the Children's Risk Profiles for Content Harmful to Children?
- a) Please provide underlying arguments and evidence of efficacy or risks that support your view.

Specifically, we welcome evidence from regulated services on the following:

- 19. Do you think the four-step risk assessment process and the Children's Risk Profiles are useful models to help services understand the risks that their services pose to children and comply with their child risk assessment obligations under the Act?
- 20. Are there any specific aspects of the children's risk assessment duties that you consider need additional guidance beyond what we have proposed in our draft?
- 21. Are the Children's Risk Profiles sufficiently clear and do you think the information provided on risk factors will help you understand the risks on your service?
- a) If you have comments or input related to the links between different kinds of content harmful to children and risk factors, please refer to Volume 3: Causes and Impacts of Harms to Children Online which includes the draft Children's Register of Risks.

#### Confidential? - N

There may be a risk if Ofcom is solely reliant on providers to 'do the right thing'. How is it proposed that the assessment be verified (especially for smaller services)?

We support point 12.49 that services with a low number of child users could still potentially be high risk, depending on specific circumstances. If user numbers do not meet the threshold of high or medium impact, there could still be a significant number of children affected.

We also agree with decision not to take a non-prescriptive approach (12.129) and that flexibility and scalability are needed (12.130) due to the *dynamic* and complex nature of risk to children in the online environment.

We support the inclusion of independent expert views and external research in *Enhanced inputs* to provide robust assurance. We also support providers consulting service users, and would be keen to see this include children and young people

#### Volume 5 – What should services do to mitigate the risk of online harms

#### Our proposals for the Children's Safety Codes (Section 13)

#### **Proposed measures**

- 22. Do you agree with our proposed package of measures for the first Children's Safety Codes?
- a) If not, please explain why.

#### **Evidence gathering for future work.**

- 23. Do you currently employ measures or have additional evidence in the areas we have set out for future consideration?
- a) If so, please provide evidence of the impact, effectiveness and cost of such measures, including any results from trialling or testing of measures.
- 24. Are there other areas in which we should consider potential future measures for the Children's Safety Codes?
- a) If so, please explain why and provide supporting evidence.

#### Confidential? - N

From a digital resilience perspective, the measures seem sensible. We are strongly supportive of the inclusion of youth voice in the proposals.

We agree with the named areas for future code measures and would like to see these given further consideration when possible.

We haven't identified any other areas but due to the pace of change, we would encourage Ofcom to continue engaging with experts and horizon scanning so that any additional areas can be quickly incorporated.

#### Developing the Children's Safety Codes: Our framework (Section 14)

25. Do you agree with our approach to developing the proposed measures for the

Children's Safety Codes?

- a) If not, please explain why.
- 26. Do you agree with our approach and proposed changes to the draft Illegal Content Codes to further protect children and accommodate for potential synergies in how systems and processes manage both content harmful to children and illegal content?
- a) Please explain your views.
- 27. Do you agree that most measures should apply to services that are either large services or smaller services that present a medium or high level of risk to children?
- 28. Do you agree with our definition of 'large' and with how we apply this in our recommendations?
- 29. Do you agree with our definition of 'multi-risk' and with how we apply this in our recommendations?
- 30. Do you agree with the proposed measures that we recommend for all services, even those that are small and low-risk?

#### Confidential? - N

We agree it is sensible to find synergies where possible to avoid duplication of effort and help providers understand what information to provide. We also agree that whilst there should be alignment, it's important that codes relating to Children's safety remain distinct given the specific requirements around children's safety online.

In reference to 'large' services (14.47) – we're assured in the proposed approach that number of users isn't the only measure and are keen that impact is considered too given the severity of harm that could be caused even to a relatively small number of child users. We also support the approach whereby even supposedly low risk providers must apply some measures.

We strongly support Ofcom taking swift enforcement action against non-compliance by any providers (14.70). Providers must be held accountable and, where needed, stringent enforcement processes must be in place so that decisive action can be taken promptly.

#### Age assurance measures (Section 15)

31. Do you agree with our proposal to recommend the use of highly effective age assurance to support Measures AA1-6? Please provide any information or evidence to support your views.

#### Confidential? - N

We fully support HEAA and are keen to see much greater robustness in age assurance online. Service providers must know what they need to assess i.e. have adequate guidance / steer and are also aware of the consequences of non-compliance.

- a) Are there any cases in which HEAA may not be appropriate and proportionate?
- b) In this case, are there alternative approaches to age assurance which would be better suited?
- 32. Do you agree with the scope of the services captured by AA1-6?
- 33. Do you have any information or evidence on different ways that services could use highly effective age assurance to meet the outcome that children are prevented from encountering identified PPC, or protected from encountering identified PC under Measures AA3 and AA4, respectively?
- 34. Do you have any comments on our assessment of the implications of the proposed Measures AA1-6 on children, adults or services?
- a) Please provide any supporting information or evidence in support of your views.
- 35. Do you have any information or evidence on other ways that services could consider different age groups when using age assurance to protect children in age groups judged to be at risk of harm from encountering PC?

We strongly support *Measure US5 - Signposting children to support* so that as well as reporting, there is follow up positive action offered to help children and young people.

However, it would be useful to know how this will be determined. Will support services be offered in user's preferred language / geographical region?

We also support efforts to apply relevant recommender system measures to protect children and young people from PPC / PC.

#### **Content moderation U2U (Section 16)**

- 36. Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.
- 37. Do you agree with the proposed addition of Measure 4G to the Illegal Content Codes?
- a) Please provide any arguments and supporting evidence.

#### Confidential? - N

We support all steps to improve content moderation such as ensuring moderation is well-resourced and has realistic performance targets. We particularly support CM1 - Content moderation systems and processes are designed to swiftly take action against content harmful to children.

We agree volunteer moderators should be provided with the necessary information to be able to perform

the role effectively and that this should be included in the Illegal Content codes.

#### Search moderation (Section 17)

- 38. Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.
- 39. Are there additional steps that services take to protect children from the harms set out in the Act?
- a) If so, how effective are they?
- 40. Regarding Measure SM2, do you agree that it is proportionate to preclude users believed to be a child from turning the safe search settings off?

The use of Generative AI (GenAI), see Introduction to Volume 5, to facilitate search is an emerging development, which may include where search services have integrated GenAI into their functionalities, as well as where standalone GenAl services perform search functions. There is currently limited evidence on how the use of GenAl in search services may affect the implementation of the safety measures as set out in this code. We welcome further evidence from stakeholders on the following questions and please provider arguments and evidence to support your views:

- 41. Do you consider that it is technically feasible to apply the proposed code measures in respect of GenAl functionalities which are likely to perform or be integrated into search functions?
- 42. What additional search moderation measures might be applicable where GenAl performs or is integrated into search functions?

#### Confidential? - N

As before, we agree 'search services have a duty to allow users to make complaints about content that is harmful to children and to take 'appropriate action' in response to such complaints'. However, we would ask what would be considered 'appropriate' and how will this be decided?

We recognise the difficulty for search services to know which users are children given no sign up is currently required to use their services. If the age of a user is unknown, will the default be to assume that the user could be a child and therefore measures apply?

Whilst many users may opt to remain on default settings, there is a concern it would still be relatively easy for a child user to change this setting (17.45) and there would be no measures in place, including nudges, to prevent this.

In relation to the argument that content warnings could potentially make content more appealing to children and young people, we think that this would be worth exploring further. This would enable Ofcom to either implement stronger measures on search services or to rule out these specific measures and consider other options.

We support measures such as downranked content not being on first page of results and agree that child users should not be able to turn off safe search functions (SM1 & 2).

#### **User reporting and complaints (Section 18)**

- 43. Do you agree with the proposed user reporting measures to be included in the draft Children's Safety Codes?
- a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.
- b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.
- 44. Do you agree with our proposals to apply each of Measures UR2 (e) and UR3 (b) to all services likely to be accessed by children for all types of complaints?
- a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.
- b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.
- 45. Do you agree with the inclusion of the proposed changes to Measures UR2 and UR3 in the Illegal Content Codes (Measures 5B and 5C)?
- a) Please provide any arguments and supporting evidence.

#### Confidential? - N

From a digital resilience perspective, we welcome steps to improve users' ability to report and for providers to take decisive action on reports – greater transparency and accountability is a key part of helping children and young people to stay safe online.

We support the proposed measures and agree with reasons given as to why child users do not engage with complaints processes (18.5 – 18.9).

We support greater transparency from all services and giving complainants more information about the complaints process and what they can expect. UR3(b) only suggests what actions the provider may take and whether the complainant should expect to hear the outcome of their complaint – could this be tightened up to give more surety?

It's important that children who choose to communicate in Welsh can report in their preferred language too (18.132). Just as complainants using English will be able to provide context to their complaint, complainants using Welsh should be provided with the same opportunity. Providers will need to ensure they can respond to such complaints swiftly so that reporting in Welsh does not cause any delays either in responding or actioning take down of the content.

We understand the reasons why Ofcom decided not to ask service providers to collaborate with specialist children's organisations when designing their complaints processes. However, we believe there is value in children's organisations working up examples of good practice that could be shared. This might offer providers useful steer of what good looks like without being too prescriptive.

We agree that the additional recommendations should be added to both the draft Children's Safety Codes and the draft Illegal Content Codes.

# Terms of service and publicly available statements (Section 19)

- 46. Do you agree with the proposed Terms of Service / Publicly Available Statements measures to be included in the Children's Safety Codes?
- a) Please confirm which proposed measures your views relate to and provide any arguments and supporting evidence.
- b) If you responded to our illegal harms consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.
- 47. Can you identify any further characteristics that may improve the clarity and accessibility of terms and statements for children?
- 48. Do you agree with the proposed addition of Measure 6AA to the Illegal Content Codes?
- a) Please provide any arguments and supporting evidence.

#### Confidential? - N

We agree that providers' terms of service should be accessible and understood by their child users. We also agree that age-appropriate user support materials are needed (US6) so that children can clearly understand the tools and functionality available to them. Being able to see and understand this, will help develop their digital literacy skills too.

We agree with the proposed addition of Measure 6AA to the Illegal Content Codes.

#### Recommender systems (Section 20)

- 49. Do you agree with the proposed recommender systems measures to be included in the Children's Safety Codes?
- a) Please confirm which proposed measure your views relate to and provide any arguments and supporting evidence.
- b) If you responded to our illegal harms consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.

#### Confidential? – N

We agree with the evidence cited and support a much more considered approach to recommender systems, particularly giving children the opportunity to negatively feedback if they see something they don't want to see or hear again. Children should be able to give feedback and be assured that feedback will be acted upon. Therefore, a clear mechanism to set out how feedback would be acted upon should be provided. We believe it's important for children to have greater control over what they see (and don't see).

- 50. Are there any intervention points in the design of recommender systems that we have not considered here that could effectively prevent children from being recommended primary priority content and protect children from encountering priority and non-designated content?
- 51. Is there any evidence that suggests recommender systems are a risk factor associated with bullying? If so, please provide this in response to Measures RS2 and RS3 proposed in this chapter.
- 52. We plan to include in our RS2 and RS3, that services limit the prominence of content that we are proposing to be classified as non-designated content (NDC), namely depressive content and body image content. This is subject to our consultation on the classification of these content categories as NDC. Do you agree with this proposal? Please provide the underlying arguments and evidence of the relevance of this content to Measures RS2 and RS3.
- Please provide the underlying arguments and evidence of the relevance of this content to Measures RS2 and RS3.

We agree providers need to take a precautionary approach on measures calling for them to filter out or reduce the prominence of content 'likely to be' PPC / PC from children's recommender feeds.

We agree that depressive content and body image content should be included in RS2.

# **User support (Section 21)**

- 53. Do you agree with the proposed user support measures to be included in the Children's Safety Codes?
- a) Please confirm which proposed measure your views relate to and provide any arguments and supporting evidence.
- b) If you responded to our Illegal harms consultation and this is relevant to your response here, please signpost

Confidential? - N

We support measures to give children greater control over who they interact with and to what extent.

We strongly support efforts to signpost children and young people to support services whenever and wherever reasonably possible.

to the relevant parts of your prior response.

### Search features, functionalities and user support (Section 22)

- 54. Do you agree with our proposals? Please provide underlying arguments and evidence to support your views.
- 55. Do you have additional evidence relating to children's use of search services and the impact of search functionalities on children's behaviour?
- 56. Are there additional steps that you take to protect children from harms as set out in the Act?
- a) If so, how effective are they?

As referenced in the Overview of Codes, Section 13 and Section 17, the use of GenAl to facilitate search is an emerging development and there is currently limited evidence on how the use of GenAl in search services may affect the implementation of the safety measures as set out in this section. We welcome further evidence from stakeholders on the following questions and please provide arguments and evidence to support your views:

57. Do you consider that it is technically feasible to apply the proposed codes measures in respect of GenAl functionalities which are likely to perform or be integrated into search functions? Please provide arguments and evidence to support your views.

Confidential? - N

We support proposals to provide children and young people with greater ability to report and find support when online. We also support the requirement for appropriate steps to be taken by providers to ensure that a reported suggestion is no longer recommended to any user.

#### **Combined Impact Assessment (Section 23)**

58. Do you agree that our package of proposed measures is proportionate, taking into account the impact on children's safety online as well as the implications on different kinds of services?

#### Statutory tests (Section 24)

- 59. Do you agree that our proposals, in particular our proposed recommendations for the draft Children's Safety Codes, are appropriate in the light of the matters to which we must have regard?
- a) If not, please explain why.

#### **Annexes**

#### Impact Assessments (Annex A14)

- 60. In relation to our equality impact assessment, do you agree that some of our proposals would have a positive impact on certain groups?
- 61. In relation to our Welsh language assessment, do you agree that our proposals are likely to have positive, or more positive impacts on opportunities to use Welsh and treating Welsh no less favourably than English?
- a) 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.

#### Confidential? – N

We agree that measures / proposals should have positive impacts on groups. If any unintended consequences did arise, then we would expect them to be given due consideration.

From a digital resilience perspective, Welsh language users could potentially have more positive experiences and should be encouraged to use their preferred language to engage with services. However, consideration should be given to any current or future Welsh language services which would have a small number of users which would mean they would be in scope for only some of the measures. This could in turn, unintentionally, make Welsh speaking children more vulnerable.

Please complete this form in full and return to <a href="mailto:protectingchildren@ofcom.org.uk">protectingchildren@ofcom.org.uk</a>.