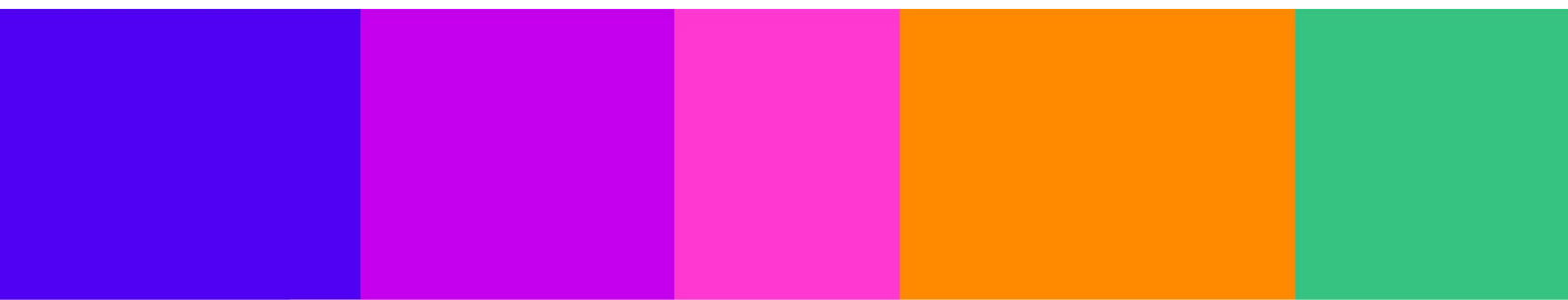# Consultation response form

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

Consultation title	Consultation: Protecting children from harms online
Organisation name	Nexus

Your response

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 1. We agree with this proposal. As detailed in our response to the Ofcom 'Guidance for Service Providers Publishing Pornographic Content' consultation, it is vital that the age assurance technology is accurate at determining the age of a user, is bespoke for the service type, and should be constantly tested for accuracy and reliability. In Annex 10 of this consultation, Ofcom provides a variety of highly effective age assurance technology recommendations in accordance with Measures AA1 And AA6 of Ofcom's Children's Safety Codes- this has been specified so as not to allow child users to slip through the gaps in age assurance programmes as well as to cater to different service setups, capacities, etc. The Digital Trust and Safety Partnership published Guiding Principles and Best Practises for Age Assurance Methods, detailing the importance of identifying, evaluating, and adjusting for risks to youth to inform proportionate age assurance methods, conducting layered enforcement options, and ensuring that the technology is accessibly, risk-appropriate, and effective[1]. 2. We agree with this proposal. As the Section 4 guidance points out, the use of the phrase "significant number of users who are children" should be interpreted not as a flat numerical number, but instead as contextualised by the service itself and the risk of harm to child users. As Ofcom's own research shows, the tope shows for children in 2022 were *Squid Game*, a series about financially burdened people who participate in deadly children's games for the chance to win money, *You*, a thriller that follows a man who stalks his partners and eventually murders them, and *Grey's Anatomy*, a medical drama that includes themes such as abuse, violence, and sexual content[2]. What this research shows is that children are accessing content that can be classified as a risk of harm

---

[1] https://dtspartnership.org/wp-content/uploads/2023/09/DTSP_Age-Assurance-Best-Practices.pdf

[2] https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/childrens-media-lives-2022/childrens-media-lives-2022-summary-report.pdf

Question	Your response

to their wellbeing but is not necessarily advertised for a child audience. This case study supports Ofcom's interpretation of "significant number of children" in order to ensure that potentially risky services are not ruling themselves out on the children's risk assessment.

3. We agree with this proposal. The process is rigorous and intentionally wide-ranging to capture all aspects of service provision, business practises, and child user activity. As Ofcom notes in Section 4, children are also likely to access content, services, and platforms that are not specifically targeted towards their age demographic. By recommending to services to apply the second criterion first in their children's risk assessments, services will be much more likely to be equipped to identify significant child user engagement.

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.

Confidential? – N

4. We would like to see further detail on the kind of harmful content, as defined in the Act under Section 7, Table 7.1 for Pornographic Content, as shown in Section 8, Table 8.1, Section 1.

5. We agree with Section 7's analysis of risk factors and content harmful to children. In particular:

For Pornographic Content

- We agree that girls are more likely to experience harmful sexual behaviours, receive unwanted images, and be the subject of pornographic content.
- We agree that boys are more likely to be targeted and influenced by pornographic content
- We agree that young LGBTQIA+ people use pornographic content to learn about LGBTQIA+ sexual relationships, oftentimes due to a lack of inclusive Relationships and Sexuality Education.
- We welcome the recognition that Generative AI and deepfake technology is playing an increased role in exposing children to harmful content.

Question	Your response
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 GenAI content or applications on U2U or Search services? a) Please Provide any information about any risks identified 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.	• We also welcome Ofcom's recognition of the role that direct messaging, tagging, commenting, and commercial ad functionalities play in exposing children to harmful content. For Abuse and Hate Content, and Violent Content • We agree that girls are more likely to experience misogyny and sexism due to their gender online • We agree that children from diverse racial background are more likely to experience racism online • We agree that LGBTQIA+ children are more likely to experience homophobia and transphobia online • Protected groups are more likely to experience abuse and hate content, and violent content through social media, video sharing, livestreaming, and gaming. We agree that the use of anonymous profiles, private vs open profiles, direct messaging and tagging can increase the risk of children being exposed to harmful content. 6. We agree with the age groups that Ofcom have designated. When considering life stages, parental/guardian involvement, online presence, and age-specific risks, the proposed age groups represent the experiences of children online. 7. It is important that, whilst identifying content that presents a risk of harm to children, services do not allow content that does not present as Primary Priority (PPC) or Priority (PC) to be included in a service's assessment of risk of harm to children. Please refer to our answer to Question 49 for further analysis. 8. We would like to highlight the growing use of "Sextortion", which involves "the threat of sharing images or videos – often 'nudes' or sexually explicit content – to extort money or force someone to do something against

Question	Your response
	their will"[3]. The Revenge Porn Helpline reported that sextortion cases increased by 54% in 2023 compared to 2022, with 825 support requests coming from children aged 13-18, and 170 requests from children under the age of 13[4]. The PSNI reported that sextortion cases had jumped from 5-10 reports per month in 2019 to 75-80 cases per month in 2023, with the majority of victims being young men between 18 and 23 years old[5]. The National Crime Agency issued an urgent warning following a "considerable increase in global cases of financially motivated sexual extortion... a large proportion of cases have involved male victims aged between 14-18. Ninety-one per cent of victims in UK sextortion cases dealt with by the Internet Watch Foundation in 2023 were male"[6].

Through our work with children and young people, we have come across a concerning trend using "Live" features on TikTok, Instagram, and Facebook to target children and young people with sexual content.

9. Please see the below research on GenAI risks to children:

From the Internet Watch Foundation Report on AI Child Sexual Abuse Imagery[7]-

- "In total, 20,254 AI-generated images were found to have been posted to one dark web CSAM forum in a one-month period".
- "Perpetrators can legally download everything they need to generate these images, then can produce as many images as they want – offline, with no opportunity for detection. Various tools exist for improving and editing generated images

---

[3] https://www.victimsupport.org.uk/crime-info/types-crime/sextortion/#:~:text=Sextortion%2C%20also%20known%20as%20webcam,threats%20or%20other%20unfair%20means.)

[4] https://revengepornhelpline.org.uk/assets/documents/revenge-porn-helpline-report-2023.pdf?_=1714738699

[5] https://www.newsletter.co.uk/news/crime/around-75-80-sextortion-cases-being-reported-to-psni-each-month-this-year-public-asked-to-be-on-your-guard-4408853

[6] https://www.nationalcrimeagency.gov.uk/news/nca-issues-urgent-warning-about-sextortion#:~:text=The%20unprecedented%20warning%20comes%20after,to%2010%2C731%20the%20year%20before.

[7] https://www.iwf.org.uk/media/q4zll2ya/iwf-ai-csam-report_public-oct23v1.pdf

Question	Your response
	until they look exactly like the perpetrator wants".
	• "Most AI CSAM found is now realistic enough to be treated as 'real' CSAM".
	• "There is now reasonable evidence that AI CSAM has increased the potential for the re-victimisation of known child sexual abuse victims, as well as for the victimisation of famous children and children known to perpetrators".
	• "AI CSAM offers another route for perpetrators to profit from child sexual abuse".
	The National Crime Agency said in their 2023 National Strategic Assessment that "We have also begun to see hyper realistic images and videos entirely created through Artificial Intelligence. The use of AI for child sexual abuse will make it harder for us to identify real children who need protecting, and further normalise abuse. And that matters, because we assess that the viewing of these images – whether real or AI generated - materially increases the risk of offenders moving on to sexually abusing children themselves. There is also no doubt that our work is being made harder, as major technology companies develop their services, including rolling out end to end encryption, in a way that they know will make it more difficult for law enforcement to detect and investigate crime and protect children"[8].
	The Stanford Internet Observatory conducted an investigation that identified "hundreds of known images of child sexual abuse material in an open dataset used to train popular AI text-to-image generation models"[9]. The report found that "Models trained on this dataset, known as LAION-5B, are being used to create photorealistic AI-generated nude images, including CSAM"[10].
	The UK Safer Internet Centre reported that "schoolchildren in the UK are now using AI to generate indecent im-

[8] https://www.nationalcrimeagency.gov.uk/news/director-general-graeme-biggar-launches-national-strategic-assessment
[9] https://cyber.fsi.stanford.edu/news/investigation-finds-ai-image-generation-models-trained-child-abuse
[10] Ibid.

Question	Your response
	ages of other children, with experts warning urgent action is needed to help children understand the risks of making this sort of imagery…. Children may be making this imagery out of curiosity, sexual exploration, or for a range of other reasons, but images can quickly get out of hand and children risk "losing control" of the material, which can then circulate on the open web". The Centre also says that "this imagery can have many harmful effects on children – and warns it could also be used to abuse or blackmail children"[11].
	10. Not applicable to our area of expertise.
	11. We are satisfied with the proposed NDC categories.
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? 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?	Confidential? – N 12. We agree with the proposed approach. Providing a non-exhaustive, illustrative list of example content that may or may not constitute PPC, PC, or NDC balances clarity for services to identify potentially harmful content with applicability, thereby not limiting services to only the examples provided. We welcome Ofcom's guidance on how content can be highly subjective and context specific, meaning that different kinds of harms can vary in nature due to the presentation of the content and the specific nature of the user and poster. We also welcome Ofcom's efforts to differentiate between content harmful to children and recovery content, which can be beneficial for children and other users who are on or beginning a recovery journey. Ofcom provides guidance for services to determine the nature of the content, as discussed above. For specific content sections: • We have raised a point of clarification in our answer to Question 14 pertaining to Pornographic Content.

---

[11] https://saferinternet.org.uk/blog/children-must-understand-risk-as-uk-schools-say-pupils-abusing-ai-to-make-sexual-imagery-of-other-children

Question	Your response
	- We agree that, for Abuse and Hate Content, that hate and abuse can overlap with Violent Content, particularly in the case of abuse, hate, and violence towards women and girls. - We also agree that hate and abuse can be present regardless of whether the individual or group targeted holds the listed characteristic, or they are perceived to hold said characteristic. This is important to recognise, as stereotypes, typecasting, and false and misinformation can generate hate and abuse towards individuals and groups with listed characteristics. - We welcome Ofcom's recognition of gender-based violence and sexual violence in the Violent Content section, particularly including content that - Justifies and/or defends the use of sexual violence - Commends domestic abuse as a means to 'control' women - Argues that victims and survivors of sexual assault must bear some responsibility 13. As discussed in our answer to Question 12, we agree with the proposal to include codewords, hashtags, substitute terms/phrases, sounds, pornographic GIFs, sexualised emojis, and comments as elements for services to consider as content that poses a risk of harm to children. The Internet Watch Foundation has a curated Keyword List that compiles words, phrase, and codes for concealing child sexual abuse material online[12]. 14. We have a point of clarification for 8.2 Guidance on Pornographic Content- in 8.2.7, Ofcom outlines the elements that will make it more or less likely for the content to be deemed as having the principal or sole purpose of sexual arousal, which includes "Very strong references to sexual behaviour: Use of language associated with sexual activity and pornography e.g. 'milf', 'horny'". Then, in

[12] https://www.iwf.org.uk/our-technology/our-services/keywords-list/

Question	Your response
	point 8.2.8, Ofcom states that "However, if content depicts an individual using pornographic language but they are fully clothed and not carrying out or simulating sexual activity this would likely not be judged to be pornographic". We would appreciate some clarification around when the use of language strongly associated with sexual activity and pornography is considered to be pornographic content. We have found from our engagement with young people that they don't often use explicit, sexualised language and are more likely to use acronyms, emojis, and other code words to convey sexualised content. We would therefore also like clarification on the definition and context of non-sexualised GIFs and emojis, and the guidance for services to interpret these functions in the context of the post.

Volume 4: How should services assess the risk of online harms?

Governance and Accountability (Section 11)

Question	Your response
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 15. We agree with the proposed governance measures. There is a balance of accountability between governance, senior leadership accountability, internal monitoring, and staff policy implementation to create a holistic approach to governance and accountability. As we stated in our submission to Question 3 of the Ofcom Consultation on 'Protecting People from Illegal Harms Online', designating and training senior members of staff to make decisions on online safety as well as track evidence of risk in their services is vital to adapting to new challenges and maintaining accountability. We would like to reiterate, however, that assessing risk is often complex and nuanced, and needs supported by other forms of specific training on CSE, CSA, Safeguarding, Child Protection, etc. 16. We agree that implementing the same process for the Children's Safety Codes and the Illegal Content Codes would ensure consistency across service provision, accountability, and innovation. We also appreciate the emphasis on the detrimental effects of poor governance and accountability on children and vulnerable people, especially if there is a disjointed approach across different codes. The proposed codes are in line with a user-

Question	Your response
	centred approach and has been informed by children and vulnerable people's safety.

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

Question	Your response
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	Confidential? – N 17. Overall, the proposals are detailed, extensive, and accessible for service providers. It is clear that risk of harm for children needs to account for and recognise each kind of harmful content, the likelihood of a child interacting with the harmful content on the service, the effectiveness of existing safety measures, the impact the content on children- both directly and indirectly- and the severity and reach of the content. The Guidance on record keeping of Risk Assessments and Ofcom's enforcement powers are clear and precise. 18. We believe that the proposals in relation to the Children's Risk Profiles are comprehensive and informative. The Profiles provide a succinct guide to assist services with identifying high risk factors in different aspects of their content. As we discussed in our answer to Question 2, it is vital that Ofcom creates guidance and assessments that remove any guesswork on behalf of services and directly highlight the harms that children are at risk of coming into contact with and therefore ensures that services completing their Risk Assessments are not self-eliminating their services. Ofcom's Risk Profiles will describe how the risk arises and what kind of content is relevant to perpetuating said risks by directly recognising the links between service capability and child user experiences. 19. The 4-step model is clear, concise, and informative. Each step provides the background rationale contained in the Online Safety Act, the explanation of each supplementary guidance document, and the expectations for

Question	Your response
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.	actioning each step, which continues to build consecutively as you move through each step. There is great emphasis placed on not just assessing harms but also assessing how services are used, the risk of repetitive and cumulative harms, and the duty to recognise and respond to changes, triggers, and new developments. 20. We believe that the guidance on the children's risk assessment duties is clear and precise, providing direct information on legal duties for services and how to carry out the obligations step by step, as well as the purpose behind each step. 21. Not relevant to our service.

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 22. We agree with the measures proposed in the Children's Safety Codes, with some recommendations included in our answer to Question 24. 23. Not applicable to our service. 24. We would recommend that, under measure PCU E3, in *Signposting child users to support* of the *Protection of Children Code of Practice for user-to-user services*, Ofcom include the following in red: "Services likely to be accessed by children where there is a medium or high risk of bullying content, eating disorder content, self-harm content or suicide content, sexual abuse content, domestic abuse content, content containing or pertaining to violence against women and girls, and content containing or pertaining to racism, ableism, homophobia, transphobia". We would also recommend that, under measure PCS E3 in *Features, functionalities and user support* of the *Protection of Children Code of Practice for search services,* Ofcom include the following in red:

Question	Your response
	"Provision of crisis prevention information relating to suicide, self-harm, eating disorders, sexual abuse content, domestic abuse content, content containing or pertaining to violence against women and girls, and content containing or pertaining to racism, ableism, homophobia, transphobia".

As evidenced by the following research:

- Ofcom's *Children's Media Lives 2022* Report, "around a third of the participants were engaging with content about racism"[13] and a research study conducted in January 2024 found that Black children and teenagers who experienced racism online were at risk of developing post-traumatic stress disorder[14].
- Stonewall's 2017 School Report found that 40% of LGBTQ+ young people have been the target of homophobic, biphobic and transphobic abuse online[15].
- Cornell University research found that social media and online platforms were being used to amplify microaggressions, spread misinformation, and discriminate against people with disabilities[16].
- Internet Matters conducted a report on Online Misogyny, finding that "half (50%) of boys aged 15-16 and over half (55%) of girls aged 15-16 believe that the online world has made misogyny worse"[17].

[13] https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/childrens-media-lives-2022/childrens-media-lives-2022-summary-report.pdf
[14] https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2813624
[15] https://www.stonewall.org.uk/resources/school-report-2017
[16] https://news.cornell.edu/stories/2022/10/online-microaggressions-strongly-impact-disabled-users
[17] https://www.internetmatters.org/wp-content/uploads/2023/09/Internet-Matters-Online-misogyny-and-image-based-abuse-report-Sep-2023-2.pdf

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.	Confidential? – N 25. We agree with the approach, which covers proportional expectations for a variety of services, the comprehensive rights assessment, the research base, and the similarities to the Illegal Harms Codes.
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.	26. We agree with this approach. It is important that the implementation of the Online Safety Act captures both legal and illegal content as a measure to protecting children online. As discussed throughout our consultation response, children access harmful content on a variety of social media apps, online websites, chatrooms, video games, etc. which are not always illegal by design. In order to effectively protect child users, we agree that there needs to be two separate Codes that work together to protect children.
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?	27. We recognise the need for a levelled approach to measures for different services according to risk, size of service, functionalities, and type of service. This will ensure that services that are assessed to be medium to high level of risk, are of a larger service user base, and/or present multi-risks are held to a higher standard of measures and duties.
28. Do you agree with our definition of 'large' and with how we apply this in our recommendations?	28. We agree with this definition.
29. Do you agree with our definition of 'multi-risk' and with how we apply this in our recommendations?	29. We agree with this definition.
30. Do you agree with the proposed measures that we recommend for all services, even those that are small and low-risk?	30. We agree with this proposal, as it is important that all services have terms of service, user reporting processes and content moderation processes that comply with child safety.
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. 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?	Confidential? – N 31. Nexus agrees with the proposed criteria for the following reasons: • Technical accuracy- It is vital that the age assurance technology is accurate at determining the age of a user. This includes constant and varied testing of the technology. • Robustness- This criterion speaks to the multitude of different technologies and situations where age assurances need to be flexible and bespoke, responding to different technological environments and circumvention tactics. • Reliability- Service providers should be exercising due diligence in selecting and outsourcing age assurance technologies. These systems should be reliably sourced and reproduced. • Fairness- The guidance deals fairly with the concerns of bias, wrongful exclusion, and discrimination. Service users should ensure that they are not preventing adults from accessing legal content, and as such this criterion emphasises the need for age assurance technologies to be tested with diverse backgrounds, needs, and datasets. • We also emphasise the need for employing the expertise to continually scope for software that might be able to bypass the age assurance technology as this landscape would move quickly. 32. We agree with the scope of the services captured by AA1-6. 33. Not applicable to expertise. 34. Nexus agrees with the Impact Assessment. The assessment clearly states that there may be costing involved in complying with regulatory standards, however, these are the minimum expectations and as such are proportional for small and micro businesses as well as larger businesses. There is also a degree of flexibility codified in the guidance to ensure that service providers are instituting age assurances in the most cost-effective and appropriate way. The age assurances will also have a

	positive impact on adult service users, who will be able to access content through age verification. 35. Not applicable to our expertise.
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 36. We agree with the proposals on content moderation for user-to-user services. Section 16 is clear on the serious risks of ineffective content moderation, which underpins the explanations of each measure. There is a balance between what services should be doing through the service they offer and what internal work they should do to train, resource, and support their service moderators. Ofcom's recommendations include cost-effective, proportionate measures for all services, with extra requirements for large and multi-risk services whereby there is a greater need for protecting service users. We also welcome Ofcom's commitment to an additional consultation later this year on automated content moderation and detection tools, as we see a growing trend in online technology turning to automated features, AI content and recommendation functions, and automated customer support. 37. We agree with this proposal, as the Illegal Content Codes will guide services to perform similar recommended actions as the Children's Safety Codes, therefore formally including the measure on the provision of materials to volunteer moderators will benefit moderators identifying illegal content online.
Search moderation (Section 17)	
38. Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.	Confidential? – N 38. We agree with the proposals in this Section. Downranking, blurring, and filtering PCC and PC content when the user is believed to be a child will protect the user from harmful content and any recommender system prompts. As in our answer to Question 31, we would like to emphasise the need for services to employ rigorous

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?	testing and scoping for any technologies that can enable bypassing or circumventing search moderation and age assurance technologies.
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?	39. Not applicable to our expertise.
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 GenAI services perform search functions. There is currently limited evidence on how the use of GenAI 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:	40. We agree with this proposal. It is imperative that search services are identifying potential child users and institute the relevant safeguards accordingly.
41. Do you consider that it is technically feasible to apply the proposed code measures in respect of GenAI functionalities which are likely to perform or be integrated into search functions?	41. Not applicable to our expertise.
42. What additional search moderation measures might be applicable where GenAI performs or is integrated into search functions?	42. Not applicable to our expertise.

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 43. We agree with the proposed measures for user reporting and complaints. Particularly, we agree that services should create accessible, easy to understand, and transparent complaint and reporting systems that will appeal to children and vulnerable services users and thereby increase the likelihood of harmful content and user profiles being reported. From multiple discussions with young people, they rarely report or use reporting systems. When in discussions with classes of 11–12-year-olds who disclosed regularly receiving unwanted sexual images through DMs on a social media platform, they stated that they block the person and don't report as "what's the point" and because it happens nearly daily, they felt blocking was easier and more instant. b) From our submission to the Ofcom Illegal Harms Consultation: "Nexus would like to highlight the importance of support for children using a service when they identify content that is illegal and harmful. In particular, we would recommend that Ofcom include strict guidance for services to ensure that their complaints procedure is robust, simplified, and accessible; for example, once a complaint has been made by a child user, will the content and/or profile that has been reported be suspended pending investigation? Alongside this, will service platforms provide support information after a complaint has been made? And are there measures for parents, guardians, carers, or a nominated caretaker to make a complaint on behalf of the child? These are only some examples of measures to protect children and young people online that Ofcom can recommend to services as part of their safeguarding measures". It must be clear and easy to understand but more importantly, as simplified a process as possible as young people tell is there's no point in reporting as it's often "over ruled" or it takes too long. 44. We agree with this proposal. As Ofcom states, there is evidence that children are unfamiliar or unaware of the procedures with reporting and whether any identifying information will be included in the report. Children are already facing mounting peer pressure online-Ofcom's Children's Media Literacy 2024 Report found that "87% of users of these apps of this age agree that

	there is pressure to be popular on social media and messaging sites/apps, at least some of the time"[18]- so it is important that any reporting features are clearly laid out and informative for users as to inform their decision to report a post or user. 45. We agree with the proposed changes. Please refer to our answer to Question 44 for further detail.

[18] https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/children-media-use-and-attitudes-2024/childrens-media-literacy-report-2024.pdf

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 46. We agree with the proposed Terms of Service and Publicly Available Statements measures. In order to be accessible and informative, Terms of Service/Publicly Available Statements need to comply with the Online Safety Act, contain clear language, and keep users informed of harmful content and how to get support from the service. b) As we stated in our submission to the Ofcom Illegal Harms Consultation: ""It is important that users be informed about how services treat illegal content" and that these provisions are "designed for the purposes of ensuring usability for those dependent on assistive technologies... [and are] clearly signposted for the general public, regardless of whether they have signed up to or are using the service"[19]. 47. We believe that language, presentation, and length of the document are key considerations for engaging children, young people, and vulnerable users in the Terms of Service. By creating a Terms of Service and/or Publicly Available Statement that is engaging and relates to the interests of young people- such as highlighting the importance of mental health, how to spot harmful content, and where to go for help and how the service will respond to your report- will increase the likelihood of children and young people engaging with the Terms of Service/ Publicly Available Statement. One suggestion that we have found accessible to engage young people is through audio and visual mediums as opposed to text. By engaging our different senses, we can capture the attention and retention of children and young people, as well as educate parents, carers, and teachers through an accessible format. 48. For Category 1 and 2A Services, we also agree that transparency is important, therefore these services

[19] https://www.ofcom.org.uk/__data/assets/pdf_file/0022/271147/volume-4-illegal-harms-consultation.pdf pg.228

	should summarise and publish their most recent children's risk assessment in their Terms of Service and/or Publicly Available Statements in order for potential users to be fully informed.

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. 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	Confidential? – N 49. We agree with the proposed recommender system measures. As discussed in the draft guidance, recommender systems can create a constant stream of harmful content for child users, making it harder for child users to avoid triggering content. Content that might, on the surface, not appear to be harmful can quickly be used to inform recommender algorithms and push harmful content onto children's feeds, profiles, and 'for you' pages. Machine learning will identify content that is being engaged with and will identify users through hashtags, location, music, etc. and will recommend further content. Children are at risk of coming across content that they may not wish to see, but because they have engaged with it, they will continue to be recommended said content. The following research illustrates the risks of recommender systems: • Investigation by the Institute for Strategic Dialogue found that YouTube's recommender system routinely pushes extremely misogynistic hateful material into boys' feeds[20]. • Amnesty International found that "Between 3 and 20 minutes into our manual research, more than half of the videos in the 'For You' feed were related to mental health struggles with multiple recommended videos in a single hour romanticizing, normalizing or encouraging suicide"[21]. • A report by The Centre for Countering Digital Hate titled 'The Incelosphere' found that "YouTube channels are hosting incel channels with over 136,000 subscribers and Google search surfaces incelosphere sites on searches related to key incel concerns, like improving

---

[20] https://www.isdglobal.org/wp-content/uploads/2022/04/Algorithms-as-a-weapon-against-women-ISD-RE-SET.pdf

[21] https://www.amnesty.org/en/latest/news/2023/11/tiktok-risks-pushing-children-towards-harmful-content/

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.	physical appearance"[22]. The term 'incel' is short-hand for 'involuntary celibate' which has come to represent, as Founder of the Everyday Sexism Project Laura Bates said, "male-only spaces that blame their members' problems on women, promoting a hateful and violent ideology linked to the murder or injury of 100 people in last ten years, mostly women"[23]. • Alongside incels, there is a growing body of so-called 'femcels', who believe that "the toxic and unrealistic body standards that are set upon women by society, they are unable to find a sexual partner or form a romantic relationship"[24]. According to research in the Archives of Sexual Behaviour, femcels were "more interested in their own frustrations than men's frustrations"[25] such as improving their attractiveness and desirability to achieve features including, as Psychology Today reports, a "symmetrical face," "short philtrum," "full lips," "small noses," "positive eye tilt," "smaller foreheads," "neotenous features," and a "smaller jaw/chin"[26]. Glamour Magazine's piece on Femcel Culture also raised concerns of femcel communities engaging in transphobic language, body shaming, and fatphobia[27]. c) As we stated in our submission to the Ofcom Illegal Harms Consultation: We agree with Ofcom's reasoning: "Gathering information about the impact changes to recommender systems have on the dissemination of illegal content will put services in a position to make materially better design choices than they otherwise would. Whilst this measure may impose some costs on services, it may also deliver some countervailing savings as identifying and addressing potential causes of harm upfront may reduce the costs services incur mitigating harm after the fact"[28].

---

[22] https://counterhate.com/research/incelosphere/
[23] Ibid.
[24] https://www.glamourmagazine.co.uk/article/what-are-femcels
[25] https://link.springer.com/article/10.1007/s10508-023-02796-z
[26] https://www.psychologytoday.com/gb/blog/the-modern-heart/202211/how-do-femcels-and-incels-differ
[27] https://www.glamourmagazine.co.uk/article/what-are-femcels
[28] https://www.ofcom.org.uk/__data/assets/pdf_file/0022/271147/volume-4-illegal-harms-consultation.pdf pg.265

	50. Not applicable to our expertise.
	51. Not applicable to our expertise.
	52. We agree with this proposal. Research conducted by The Centre for Countering Digital Hate (discussed in Question 49) found that "Forums in the "incelosphere" network we identify in this report feature on the first page of Google search results for terms associated with body image and unemployment"[29]. This research demonstrates the way that users can quickly fall into harmful content without even searching identifiably harmful terms, words, phrases etc.

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 to the relevant parts of your prior response.	Confidential? – N 53. We agree with these proposals, and we see the importance of these proposals being rolled out to all users. In the consultation document, it is noted that Measure US2 (Option to block and mute other accounts) is included as Measure 9A in the Illegal Harms Codes, and Measure US3 (Option to disable comments) is included as Measure 9B in the Illegal Harms Codes. However, we would recommend that Measures US1, US4, US5, and US6 are incorporated for all service users as well to protect people from harmful content.

Search features, functionalities and user support (Section 22)

---

[29] https://counterhate.com/research/incelosphere/

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 GenAI to facilitate search is an emerging development and there is currently limited evidence on how the use of GenAI 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 GenAI functionalities which are likely to perform or be integrated into search functions? Please provide arguments and evidence to support your views.

Confidential? – N

54. We agree with the proposals in this Section. We would like to recommend, however, that Measure SD2 is amended to include "crisis prevention information in response to known PCC and PC-related search requests regarding suicide, self-harm, eating disorders, sexual abuse content, domestic abuse content, content containing or pertaining to violence against women and girls, and content containing or pertaining to racism, ableism, homophobia, transphobia".

Our research base for this can be found in our answer to Question 24.

55. From our work with children and young people through our Early Intervention and Prevention workshops, we have found the following:

When looking at our Body Positivity module, the young men and boys we have worked with have spoken about the pressures of "looksmaxxing", a rapidly growing trend that targets mostly boys and young men to critique their physical appearance, score themselves on a scale of 'manliness', and aim to increase their 'scores' to achieve the 'peak male appearance'. Looksmaxxing preys on the insecurities of young men and boys to physically alter themselves using different exercises, filters, and influencer routines and products in order to appear more attractive. Young men and boys are searching TikTok and Google for tips, products, and inspiration for achieving this look, which is then fuelled by algorithmic recommender systems to continue to feed this content, with the risk of young men and boys venturing into incel spaces without recognising it; Mike Nicholson, a former teacher who runs a workshop programme called 'Progressive Masculinity' has said that "the world that these young men and boys are inhabiting is one that is trying to increase their anxieties and potentially lead them down this path that, if you're not careful, can lead to 'incel' ideologies"[30].

In the same module, young women and girls are reporting the pressure to use filters to alter their appearances, using social media to search for content that promotes

[30] https://www.theguardian.com/lifeandstyle/2024/feb/15/from-bone-smashing-to-chin-extensions-how-looksmaxxing-is-reshaping-young-mens-faces

appearance-alerting products and procedures. According to a 2021 survey of 200 teens ages 13 to 21 from ParentsTogether, young people who use beauty filters weekly are more likely to want to have cosmetic surgery and to alter their skin colour[31]. With a recommender search system, young women and girls are at risk of being recommended the kinds of products that claim to alter appearances to achieve a desired attractiveness, such as cosmetic procedures, dangerous dieting information, and extreme exercise. The Campaign Director for ParentsTogether, Amanda Kloer, told Teen Vogue that "there's also a big problem with the kind of content that these algorithms amplify. For example, if you're a young girl and you start an Instagram account, you get so much diet and exercise content, even if that's not something you indicate that you're interested in"[32].

56. Not applicable to our service.

57. Not applicable to our expertise.

---

[31] https://parentstogetheraction.org/wp-content/uploads/2021/09/ParentsTogether_Social-Media-Beauty-Filters-Survey-Results_09-29-21.pdf
[32] https://www.teenvogue.com/story/social-media-filters-how-young-people-see-themselves

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?	Confidential? – N 58. We agree that the proposals are proportionate, risk-aware, and equitable. The proposals account for the higher risk of harm to child users as well as provides flexibility for the range of services to comply with their duties.
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.	Confidential? – N 59. We agree that the proposals are appropriate and meet the requirements of the codes of practise as set out in Schedule 4 to the Online Safety Act.
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 60. We agree with the Impact Assessment. Ofcom's proposals have taken into consideration the listed characteristics under the Equality At 2010, ranging from accessibility of language, assistive technologies, age-appropriate user support materials, safety features for children and those vulnerable and at risk of harm from content and materials including inciting hate, self-harm and eating disorder content, threats of and instances of violence, targeted harassment of protected groups, as well as a focus on protecting people from racist, homophobic, sexist, and transphobic content. 61. We agree that the proposals in this consultation are likely to have a positive impact on the opportunities to use Welsh and to treat Welsh no less favourably than English.

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