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	Internet Society, Internet Society UK England Chapter

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? – Y / N
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?	4.a. The Draft Children's Register of Risk makes several references to encryption "exacerbating potential harm." We recommend balancing these statements by acknowledging the positive role that End-to-end encryption (E2EE) plays in providing safety and security to children[1] and

[1] "Parents' Guide to Encryption." Global Encryption Coalition, 12 Mar. 2024, www.globalencryption.org/parents-guide-to-encryption/.

Question	Your response
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 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,	adult users, effectively protecting them against a number of real-world harms including stalking, retaliation for reporting abuse, impersonation, and so on. E2EE messages can only be read by the sender and recipient. They cannot even be read by the service provider. This ensures a guarantee of privacy, security, confidentiality, as well as authenticity that it has come from the sender who it says it has come from (and not a spoof) and that the message cannot have been changed or altered by anyone. With billions of people reliant on digital communications to speak not only to friends but also to government bodies, their health provider, and their bank, this level of security is important. Confidential? – N

Question	Your response
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.	
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? – Y / N
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? Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.	15. Ofcom's Guidance document in the previous round of consultation identified E2EE as a risk factor for multiple offences listed in the Online Safety Act. Similarly, messaging services are identified as a specific risk in the draft Children's Risk Assessment Guidance, with encrypted

Question	Your response
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?	and ephemeral messaging identified as particularly "increas[ing] risk of harm related to violent content and bullying content." Volume 4 identifies that services that are "multi-risk for content harmful to children" take measures GA3 (Written statement of responsibility for senior members who make decisions relating to management of child safety risks); GA4 (Have an internal monitoring and assurance function to provide independent assurance); GA6 (Have a code of conduct that sets standards for employees around protecting children); and GA7 (Ensure staff involved in the design and operational management of service are sufficiently trained in approach to compliance with children's safety duties). We re-emphasize that treating encryption as a risk-factor undervalues the fundamental premise of E2EE in providing safety and security for all users on the platform. That is over 40 million users in the UK and some two billion users globally. Encryption is more important than ever as Internet-based crime is one of the fastest growing security threats. End-to-end encryption, the most secure form of encryption, ensures that sensitive, confidential information transmitted by billions of people online every day remains confidential and out of the hands of criminals. End-to-end encryption also helps prevent spies, terrorists, and hostile governments from accessing and exploiting confidential communications of government officials, and penetrating computer systems and databases that could cause wide-scale, systemic disruptions to economies, infrastructure, and security. Designating encryption as a compounding risk factor (multi-risk) that would oblige services to four measures (GA3, GA4, GA6, GA7) sends a mixed message to service providers. Indirect pressure on providers would effectively circumvent exceptions for E2EE laid out in the Online Safety Act, implicitly pushing service providers not to roll-out encryption on their services. We are concerned that this pressure could ultimately introduce new systemic weaknesses and vulnerabilities, putting users at risk and causing economic harm. Confidential? – N

Question	Your response
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	Confidential? – Y / N

Question	Your response
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.	
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.	23. We are glad to see that thought has been given to the difficulties of moderating content in various environments, and that the Center for Democracy and Technology research on content moderation in E2EE systems[2] has been cited as a source. Ofcom's preferred approach is to specify "factors services should have regard to in designing [their] systems and processes" which is preferable to setting blanket requirements across differing systems, especially in encrypted systems. We would like to provide further evidence as to the technical feasibility of client-side scanning, a technology that has been discussed extensively as a potential solution to content detection obligations under the Online Safety Act. In our recently published research "Preemptive monitoring in End-to-End Encrypted Services"[3] we identify several factors on which the technical feasibility of client-side scanning will depend, but which are not considered in the legislation. The risks we identify include: • Attacks on personal sensitive data when third-party servers collect data about an individual's

[2] "Outside Looking in: Approaches to Content Moderation in End-to-End Encrypted Systems." *Center for Democracy and Technology*, 21 June 2023, cdt.org/insights/outside-looking-in-approaches-to-content-moderation-in-end-to-end-encrypted-systems/.

[3] "Preemptive Monitoring in End-to-End Encrypted Services." *Internet Society*, July 2024, www.internetsociety.org/resources/doc/2024/preemptive-monitoring-e2ee-services/.

Question	Your response
	device usage, to match hashes with an individual. • Distributed denial-of-service attacks when alerts are subverted to increase network traffic with false positives and overwhelm a third-party server. • Manipulation of child sexual exploitation and abuse (CSEA) databases when unauthorized material is added, repurposing the database to scan for other forms of content. • Reverse engineering when data processing happens on device and detection, content-matching, and reporting mechanisms can be altered, allowing circumvention by criminals. • Attacks to supress or modify alerts sent for data processing when attackers seek to avoid detection or to create fake alerts. Technical measures to screen the content of messages in E2EE systems introduce systemic risk for both service providers and users, frustrate law-abiding users' intent to communicate privately, and interfere with their ability to do so in practice. Systemic functions for consentless scanning lay the foundations for numerous attacks with serious and widespread impacts. In technical terms, such measures compromise the integrity of devices and systems, increasing the risk of system-wide attacks and of unauthorised access to personal data, whether accidental or malicious. This undermines the trustworthiness of the online environment, with serious economic and cybersecurity implications, and creates new opportunities for criminals to exploit. This will make it harder, not easier, for law enforcement to achieve the stated goals of the Online Safety Act. The identified risks, systemic vulnerabilities, and other factors represent a serious obstacle to meeting the technical feasibility requirement placed on Ofcom by the Online Safety Act. We recommend that candidate technologies in this area be assessed using the framework developed by the National Research Centre on Privacy, Harm Reduction and Adversarial Influence Online (REPHRAIN), which has been comprehensively tested on the

Question	Your response
	prototypes developed for the Government's Safety Tech Challenge. We additionally recommend that client-side scanning be rejected as an approach, on the basis that it is incompatible with a safe over-all finding of technical feasibility. Confidential? – N

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?	29. On Ofcom's use of the term "multi-risk", please see our response to Question 15 in this consultation. Messaging services are identified as a specific risk in the draft Children's Risk Assessment Guidance, with encrypted and ephemeral messaging identified as particularly "increase[ing] risk of harm related to violent content and bullying content." This direct reference to encryption could create indirect pressure on providers, implicitly pushing service providers not to roll-out encryption on their services. Confidential? – N
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? – Y / N

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?	
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? – Y / N
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 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: 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? 42. What additional search moderation measures might be applicable where GenAI performs or is integrated into search functions?	Confidential? – Y / N

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.	43. Improved reporting mechanisms are an important tool for content detection in E2EE environments, as if one participant in a conversation voluntarily elects to share what they have received, then this is not a violation of strong encryption. We suggest that the REPHRAIN evaluation criteria[4] also be used in this context. For user reporting to be effective, users must trust and use it. This can best be achieved if it follows the REPHRAIN criteria and is secure, transparent, data protecting, maintainable, etc. Confidential? – N

[4] "Safety Tech Challenge Fund." *REPHRAIN*, www.rephrain.ac.uk/safety-tech-challenge-fund/. Accessed 11 July 2024.

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? – Y / N
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	Confidential? – Y / N

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 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.	
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.	53.a. We are supportive of the proposed support measures to be included in the Children's Safety Codes. User agency is key to a trustworthy Internet. Ensuring that all users, including children, have the agency to (US1) accept or decline an invite to a group

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.	chat; (US2) block and mute other users' accounts; (US3) disable comments on their own posts; (US4) have access to information when they restrict interactions with other accounts or content; (US5) receive clear signposts to support; (US6) and have access to age-appropriate user support materials increase the range of tools available to users to control their experience on user-to-user services. This includes limiting their exposure to unsolicited and unwanted content. We recommend the addition of a further support measure: "provide children with clear information about third party access to their communications. Services that offer privacy enhancing technologies like E2EE should have clear instructions on its use." Encryption is a tool that empowers users, including children, to control who has access to private and sensitive data. As shared earlier in this submission, we strongly advise that Ofcom dismiss client-side scanning technologies for content detection purposes in end-to-end encrypted environments. Any user-to-user service that does voluntarily employ client-side scanning technologies – as we strongly oppose mandated scanning - should disclose this to users in an age-appropriate manner, clearly informing them that their data will be scanned. Confidential? – N
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 GenAI to facilitate search is an	Confidential? – Y / N

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.	

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?	The summary document states that "The measures proposed in this consultation will necessarily have an impact on the experiences of children and adults and can impact their rights to freedom of expression and other fundamental rights." Any requirement for E2EE services to scan content would be likely to fail a proportionality test, in light of the judgement of 13 February in the European Court of Human Rights, *Podchasov v. Russia*. A key factor in the proportionality assessment for an encrypted service, is the possibility of arbitrary surveillance of users who are not the target of the measures, sometimes referred to as "collateral damage". It's important to consider the big picture, rather than individual measures, and look at the regime that is being created and that Ofcom will oversee. The question is whether it creates "collateral damage" by interfering in an arbitrary way with the rights of innocent users. On an encrypted service, the creation of backdoors and systemic vulnerabilities and weaknesses is known to result in that kind of interference, as the ECtHR stated. Our research on "Preemptive monitoring in End-to-End Encrypted Services"[5] identifies a list of systemic vulnerabilities that client-side scanning would create, which could be exploited. This could result in the arbitrary surveillance of users who are not the target of the measures. In this regard, the technical feasibility and reasonable practicability of any content detection technology must be factored into a proportionality study. Confidential? – N
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?	Confidential? – Y / N

[5] "Preemptive Monitoring in End-to-End Encrypted Services." *Internet Society*, July 2024, www.internetsociety.org/resources/doc/2024/preemptive-monitoring-e2ee-services/.

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? – Y / N

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