

Call for evidence response form

Please complete this form in full and return to os-cfe@ofcom.org.uk

Title

Second phase of online safety regulation: Protection of children

Full name

✂
✂

Contact phone number

✂

Representing (select as appropriate)

Organisation

Organisation name

PAPYRUS Prevention of Young Suicide

Email address

✂
✂

Confidentiality

We ask for your contact details along with your response so that we can engage with you on this consultation. For further information about how Ofcom handles your personal information and your corresponding rights, see [Ofcom's General Privacy Statement](#).

Your details: We will keep your contact number and email address confidential. Is there anything else you want to keep confidential? (select as appropriate)

Nothing

Your response: Please indicate how much of your response you want to keep confidential (select as appropriate)

None

For confidential responses, can Ofcom publish a reference to the contents of your response? (select as appropriate)

Yes

Your response

Question 1: To assist us in categorising responses, please provide a description of your organisation, service or interest in protection of children online.

Is this a confidential response? (select as appropriate)

No

PAPYRUS Prevention of Young Suicide is the national charity dedicated to the prevention of young suicide in the UK. We exist to reduce the number of young people who take their own lives, by shattering the stigma surrounding suicide and equipping young people and their communities with the skills to recognise and respond to emotional distress.

We have campaigned for nearly 20 years about online safety and specifically about sites, portals and forums which encourage suicide, glamorise it, give access to information about means and even sometimes sell such means online.

We are in support of The Bill in principle as it recognises the role social media channels and other online services play in keeping children and young people safe, but we feel it could go further.

Question 2: Can you identify factors which might indicate that a service is likely to attract child users?

Is this a confidential response? (select as appropriate)

No

Social media channels in general are attractive for children as a way to connect with peers. TikTok, Instagram and Snapchat are particularly attractive to children as a way to find information, create their own content, follow celebrities and influencers and seek support and validation.

Content linked to popular TV shows, music or sports are widely distributed and digested by children and are often shared with sound bites or music clips. Children can find a particular piece of music or sound clip that they like and find thousands of pieces of content that has been created with this. This can result in content being discovered which isn't something the user would search for but has included a popular or 'trending' soundbite.

Question 3: What information do services have about the age of users on different platforms (including children)?

Is this a confidential response? (select as appropriate)

No

Currently age verification is sporadic. Tik Tok asks users were given 5 months to prove that they are over the age of 13 by sharing a selfie of themselves with their passport, or a photograph with a trusted adult. Whilst this is a positive step forward, it can be circumnavigated by simply setting up a new account each time the account is disabled.

Facebook and Instagram don't appear to have anything in place that is robustly verifying age and therefore their proposed moves to limit certain content, so it only appears to over 18s is floored.

Meta is proposing that certain filters can be added to accounts of users under the age of 18 to protect them from legal but harmful content. Whilst this again is a positive move, it automatically removes this filter once the user turns 18 and so they will start to receive the previously blocked content without any changes being made by the user. We believe that this change should only be made by the user rather than an automatic feature of the platform. Where this facility to filter content is available, it should be used to protect all users.

Question 4: How can services ensure that children cannot access a service, or a part of it?

Is this a confidential response? (select as appropriate)

[Please select]

Question 5: What age assurance and age verification or related technologies are currently available to platforms to protect children from harmful content, and what is the impact and cost of using them?

Is this a confidential response? (select as appropriate)

[Please select]

Question 5: What age assurance and age verification or related technologies are currently available to platforms to protect children from harmful content, and what is the impact and cost of using them?

Is this a confidential response? (select as appropriate)

No

Harmful content is still available online via search sites such as recipe sites for those looking for how to take their own life. Access to these must be restricted to all users. Sites such as Wikipedia provide users with lots of information around suicide method, including step by step instructions based on height, weight and even on pain rating. Whilst users searching for such content will be often met with offers of help via signposting, this is due to organisations like PAPYRUS paying for Google AdWords to have their services intercept the search and not due to the search sites putting anything in place to support the user.

On social media sites such as Facebook, users can search for suicide related content and will again be met with offers of support from the platform. However, this can be bypassed, and harmful content can be accessed. Particularly worrying are the number of groups that users can join where members share their emotional state and talk about ending their lives. Whilst some members are encouraging users to reach out for support, others are encouraging them to end their lives. If these groups must exist, then human moderators should be monitoring what is being said and direct support offered to those in need. Children should not be able to access them under any circumstances.

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

Is this a confidential response? (select as appropriate)

No

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

Question 8: How do services currently assess the risk of harm to children in the UK from content that is harmful to them?

Is this a confidential response? (select as appropriate)

[Please select]

We hear on our helpline, HOPELINEUK, that children are being impacted by the content they are seeing on social media and the wider online environment. We have also heard from many bereaved parents who believe that the content their child viewed online contributed to their suicide. Whilst the online space can be positive for children to share their experiences, there are too many instances where they are searching for content that is helping to validate their emotions but then are sent down a rabbit hole of content which further impacts their mental health and can lead to thoughts of suicide. This content should not be available to anyone, and certainly must not be included within algorithms which identify and push related content to the user.

Question 9: What are the exacerbating risk factors services do or should consider which may have an impact on the risk of harm to children in the UK?

Is this a confidential response? (select as appropriate)

[Please select]

Question 10: What are the governance, accountability and decision-making structures for child user and platform safety?

Is this a confidential response? (select as appropriate)

[Please select]

Question 11: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements for children (including children of different ages)?

Is this a confidential response? (select as appropriate)

No

These need to be accessible for children of different abilities and should form part of the sign up process where understanding is confirmed in a robust way.

Question 12: How do terms of service or public policy statements treat 'primary priority' and 'priority' harmful content?¹

Is this a confidential response? (select as appropriate)

[Please select]

¹ See A1.2 to A1.3 of the call for evidence for more information on the indicative list of harms to children.

Question 13: What can providers of online services do to enhance children's accessibility and awareness of reporting and complaints mechanisms?

Is this a confidential response? (select as appropriate)

[Please select]

Again, these processes need to be made clear in accessible ways to meet the differing needs of children. They should be part of the sign up process where understanding is checked. Regular reminders could appear as notifications on accounts that belong to child users. Also, if algorithms are able to recognise content that is legal but harmful, and they are still appearing in news feeds, then a prompt to report the content should be automatically provided with acknowledgement that if they find the content upsetting they can report it and maybe link to an adults account who will be notified of the action taken.

Question 14: Can you provide any evidence or information about the best practices for accurate reporting and/or complaints mechanisms in place for legal content that is harmful to children, or users who post this content, and how these processes are designed and maintained?

Is this a confidential response? (select as appropriate)

[Please select]

Question 15: What actions do or should services take in response to reports or complaints about online content harmful to children (including complaints from children)?

Is this a confidential response? (select as appropriate)

[Please select]

Question 15: What actions do or should services take in response to reports or complaints about online content harmful to children (including complaints from children)?

Online services should be treating complaints of this nature as a priority and where AI is used to moderate some content, once the complaint is made, human moderators should take over and complete further checks to safeguard other children.

Question 16: What functionalities or features currently exist that are designed to prevent or mitigate the risk or impact of content that is harmful to children? A1.21 in the call for evidence provides some examples of functionalities.

Is this a confidential response? (select as appropriate)

[Please select]

Question 17: To what extent does or can a service adopt functionalities or features, designed to mitigate the risk or impact of content that is harmful to children on that service?

Is this a confidential response? (select as appropriate)

[Please select]

Question 18: How can services support the safety and wellbeing of UK child users as regards to content that is harmful to them?

Is this a confidential response? (select as appropriate)

[Please select]

By removing content from the platform and where any may slip through, offers of support and signposting to be offered with a robust process launched to determine how the content bypassed filters.

Question 19: With reference to content that is harmful to children, how can a service mitigate any risks to children posed by the design of algorithms that support the function of the service (e.g. search engines, or social and content recommender systems)?

Is this a confidential response? (select as appropriate)

[Please select]

Question 20: Could improvements be made to content moderation to deliver greater protection for children, without unduly restricting user activity? If so, what?

Is this a confidential response? (select as appropriate)

[Please select]

Question 21: What automated, or partially automated, moderation systems are currently available (or in development) for content that is harmful to children?

Is this a confidential response? (select as appropriate)

[Please select]

Question 22: How are human moderators used to identify and assess content that is harmful to children?

Is this a confidential response? (select as appropriate)

[Please select]

Question 23: What training and support is or should be provided to moderators?

Is this a confidential response? (select as appropriate)

[Please select]

Question 24: How do human moderators and automated systems work together, and what is their relative scale? How should services guard against automation bias?

Is this a confidential response? (select as appropriate)

[Please select]

Question 25: In what instances is content that is harmful to children, that is in contravention of terms and conditions, removed from a service or the part of a service that children can access?

Is this a confidential response? (select as appropriate)

[Please select]

Question 26: What other mitigations do services currently have to protect children from harmful content?

Is this a confidential response? (select as appropriate)

[Please select]

Question 27: Where children attempt to circumvent mitigations in place on a service, what further systems and processes can a service put in place to protect children?

Is this a confidential response? (select as appropriate)

[Please select]

Question 28: Other than those covered above in this document (the call for evidence), are you aware of other measures available for mitigating the risk, and impact of, harm from content that is harmful to children?

Is this a confidential response? (select as appropriate)

[Please select]