Ofcom
making communications work
for everyone

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

Global Encryption Coalition Steering Committee (Internet Society; Center for Democracy and Technology; Global Partners Digital)

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

The Global Encryption Coalition (GEC) was launched in 2020 to promote and defend encryption in key countries and multilateral fora where it is under threat. It also supports efforts by companies to offer encrypted services to their users. With more than 300 members in 95 countries, the Coalition is led by a steering committee made up of three global organisations: the Internet Society (ISOC), Global Partners Digital (GPD) and the Center for Democracy and Technology (CDT).

GEC Members and Friends of the Coalition support the GEC's founding statement:

Encryption is a critical technology that helps keep people, their information, and communications private and secure. However, some governments and organisations are pushing to weaken encryption, which would create a dangerous precedent that compromises the security of billions of people around the world.

Actions in one country that undermine encryption threaten us all. As a global coalition, we call on governments and the private sector to reject efforts to undermine encryption and pursue policies that enhance, strengthen and promote use of strong encryption to protect people everywhere. We also support and encourage the efforts of companies to protect their customers by deploying strong encryption on their services and on their platforms.

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

Existing age assurance and age verification technologies may have negative implications for children's privacy, security, and ability to access information. Platforms are deploying a host of methods to verify ages of its users including but not limited to: asking users to self-attest by sharing their date of birth, bank card, or proof of ID before accessing online services; asking for a friend/parent to vouch for a user's age; or deploying facial recognition or facial analysis technologies on a user to predict their age. These methods are likely to require a third party to

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?

conduct an initial assurance step and collect further data on all users, undermining their privacy and the security of their data and identity. For privacy reasons, it is essential that the user present a trustworthy assertion of age that does not result in any communication of its use with the third party. In practice this means that the third party would likely need to generate a token that asserts the user's age, which the user can then present when needed.

The challenge in using tokens is assuring that they are non-transferrable and not easily de-anonymized. Requiring all platforms to verify the ages of all users will likely undermine users' ability to access information and use online services anonymously. Anonymous browsing is not only a tenet of the modern web, but also critical for the safety of users including journalists and human rights advocates, domestic violence survivors, and users seeking sensitive information about their health or sexuality in a region where that may be weaponized by others. Ensuring that these tokens are rolled out within narrow remits is important.

When tokens are transferrable this creates the risk that an older user might "borrow" (potentially through coercion) a younger user's token to pose as an under-age user. Conversely, under-age users might use the tokens of older users to access age-prohibited services. In this way, when tokens are transferrable, they create new risks for the exploitation of young people.

Age assurance technology may also come with its own shortcomings and errors which may amplify existing bias against people with disabilities, people of colour, and gender nonconforming people. These errors risk threatening users' right to free expression and ability to access information freely. For example, an assurance technique may mischaracterize an adult as a child and age-gate them out of accessing critical information. Conversely a child may be mis-identified and given a token of an adult, which may subject them to harms online. Creating mechanisms to audit assurance technology and deployments of tokens is one way to ensure that users' rights to privacy, security, and access to information are not jeopardized.

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)

No

On 29 June 2021 DCMS published guidance titled: Public and private channels: improve the safety of your online platform. The guidance states, among other things, that end-to-end encryption makes it more difficult to identify illegal and harmful content on private channels and recommends removing end-to-end encryption for children's accounts.

We reiterate the importance of strong encryption, including end-to-end encryption, for children. Services that do not offer strong encryption increase the risk of their users, including children, being exposed to a range of harms including scams, fraud, malicious attack, and blackmail. Services should take measures that increase the confidentiality and integrity of the information

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?

that passes through their systems. This remains true for risk assessments for both adults and/or children.

Amendments tabled by Lord Clement Jones to Clause 6 of the Online Safety Bill seek to ensure that the use of risk assessments does not undermine users' privacy and security. The Global Encryption Coalition Steering Committee agrees that providers should not be required to make fundamental technical changes to the encryption they offer to comply with their risk assessment obligations under the bill. This includes the use of "accredited technologies," such as client-side scanning, that violate a user's expectations of end-to-end encryption. Further comments on client-side scanning can be found in response to Question 21.

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)

No

The 2022 DRCF Roundtable on End to End Encryption report states that technological remedies cannot, by themselves, provide a comprehensive solution to ensuring safety in an end-to-end encrypted (E2EE) environment. Businesses and regulators should consider how user safety is incorporated into the design and development of E2EE services. The report additionally highlights the importance of clarity from the regulator in setting expectations for privacy, safety, and security.

Best practice from services include:

- Transparency reports that detail moderation actions, disclosures, and other practices concerning user generated content and government surveillance using qualitative information and aggregated data;
- Notifications to users about government demands for their data and moderation of their content;
- Providing data access consistent with user privacy rights and expectations through intermediaries to independent researchers, journalists, and civil society organisations; and
- Publicly available analysis, assessments, and audits of the service's practices with regards to privacy and respect of user speech.

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

Strong encryption, including end-to-end encryption, should be activated for, and communicated to children in a clear and understandable way:

- Providers should offer encryption to users "by default". Several private messaging services offer encryption but require that users first go to their settings to activate it. In other cases, encryption is on "by default" for certain sub-services (such as one-to-one chats) but not others (such as group chats).

  When encryption is not on "by default" it can create a false sense of security for users that may wrongfully believe that the confidentiality and integrity of the information they are sending and receiving is protected.

- Providers that offer strong encryption should clearly label their services in a manner that is accessible to a wide range of age groups and does not require technical knowledge. This includes using messaging tools, colour labels, and symbols. Clear labelling is especially important for providers that may offer encryption for certain sub-services but not others.

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)

No

As described in the DRCF Roundtable on End to End Encryption report (January 2022), there are a variety of functionalities that can help prevent harm. These include a 'safety by design' approach that focuses on ensuring users can control the data they receive and share. For example the report suggests:

*"a safety by design approach focusing on preventing online services being used for illegal activity; User controls for blocking or verifiable reporting within E2EE environments; Flagging and removing accounts that violate platform standards (in a transparent manner); The use of non-content signals such as metadata to identify and address suspicious behaviour (*where the interception, retention and searching of related communications data should be analysed by reference to the same safeguards as those applicable to content- * our addition); and accessing the end-user device".*

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)

No

Client-side scanning – which is a method that scans message contents on the user's phone, tablet, or mobile device, either on the user device or on a remote server – should not be employed to prevent harm. This is because, as noted in the paper "Breaking Encryption Myths," client side scanning increases the "attack surface" for encrypted communications by creating additional ways to interfere with communications - including by manipulating the database of prohibited content. The method is disproportionate and is too easily misused to scan for content beyond the original purpose it was created for. This threatens mission-creep from authorities, could allow hostile state actors to surveil the communications of persons of interest, and creates opportunities for criminals to hack communication channels, including those of children.

These systems are also prone to false positives and negatives, and create the conditions for censorship and undue interference with user rights to freedom of expression and privacy. By breaking the expectation of privacy between sender and receiver, client-side scanning breaks the end-to-end encryption trust model, directly putting users' confidentiality at risk, and indirectly undermining trust in online services.

Tech Against Terrorism's latest report on encryption includes an overview of the security risks, privacy violations, jurisdictional challenges and longer-term normative risks of breaking end-to-end encryption through client-side scanning (see pages 82 - 84).

In 2021, Apple abandoned plans to apply client-side scanning on its devices, with the aim of addressing the use of its services for CSEA, as it was deemed disproportionate, insecure and unworkable.

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

Is this a confidential response? (select as appropriate)

No

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

Services currently have access to the following mitigation tools and may consider using a combination of these in increasing their protections for children:

- Deploying counter speech against harmful speech, whether through funding or supporting counter speech projects and initiatives, or through developing automated tools which can generate effective counter speech;
- Redirecting users who are searching for or consuming illegal or damaging content, such as terrorist content or CSEA, towards alternative content such as helplines or resources;
- Ensuring that private or encrypted services have clear and accessible user complaint mechanisms that allow users to report content shared on the private or encrypted channel that they think violates the terms of service. This ensures that online service providers can continue to provide end-to-end encryption, which provides security to online activities and communications and protects data from potential malicious actors – which is particularly important for the protection of vulnerable groups, including LGBTQ+ persons, survivors of domestic violence and human rights defenders – while also ensuring that illegal or harmful content is not left unchecked on those channels;
- Allowing users to customise their own moderation rules beyond what is prohibited in the terms of service, such as Twitter's Bodyguard tool, which allows users to set their own moderation rules;
- Allowing users to block content from particular people or groups, or on particular topics, or content from unverified or anonymous accounts, such as Twitter's Block Party tool;
- Allowing users to limit their own discoverability, or to have invisible or anonymous accounts;
- Developing software that helps users to review, document and export repeated instances of illegal or harmful content online, such as Google Jigsaw's Harassment Manager tool;
- Allowing users to flag what they believe are underage accounts;
- Implementing additional privacy-by-default settings for children's accounts, such as only allowing their content or profile to be visible to or engaged with by their friends or contacts;
- For younger children, developing parental controls to allow adults to have control over what types of content is encountered, particularly for vulnerable children.

As a note on methodology: a critical factor in mitigating risk and harm from illegal content is clarity about the problem to be solved. The Online Safety framework runs the risk of failing to define the problem with sufficient clarity, with the result that proposed solutions don't work, and/or have unintended and harmful consequences.

First, it is overwhelmingly unlikely that any single technical "fix" will successfully neutralise all societal ills that materialise on the internet. For many of the Government's intended aims in this policy area, not all of which include criminality, the appropriate intervention is not technical at all, but a matter of user education, awareness-raising, and digital and ethical literacy.

A disproportionate focus on technical "fixes" is not the most effective solution to what are essentially societal problems.