Consultation response form

Your response

Question	Your response
Consultation question 1:	
Do you have any views on our audit-based assessment, including our proposed principles, objectives, and the scoring system? Please provide evidence to support your response.	**Recommendations:** We recommend balancing these assessment processes and principles by: • Ensuring that Technology Notices cannot be issued to require pre-encryption monitoring of communications transmitted over E2EE services. • Ofcom should make clear that no Technical Notice may be issued that requires a provider of an E2EE service to build any encryption backdoor. • Assessing candidate technologies 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 prototypes developed for the Government's Safety Tech Challenge. • That accuracy thresholds are considered alongside security risks, human rights, and freedom of speech impact as part of the assessment and accreditation of technologies. Accuracy thresholds that may appear high would still result in millions of false positives given the volume of messages sent on a daily basis. Previous rounds of consultations for Ofcom's Guidance have not appropriately acknowledged the multitudinous, and often necessary, benefits of encryption. Instead, they identified end-to-end encryption (E2EE) as a risk factor for multiple offences listed in the Online Safety Act. Similarly, messaging services had been identified as a specific risk in the draft Children's Risk Assessment Guidance, with encrypted and ephemeral messaging identified as of particular concern related to "increas[ing] risk of harm related to violent content and bullying content." In the same vein, the proposed guidance in this consultation on "policy proposals for minimum standards of accuracy for accredited technologies, and guidance to providers" with respect to technology notices also fails to acknowledge and account for the critical role that E2EE plays in modern life, protecting children and adults alike, as well as governments, businesses, critical infrastructure, and other institutions. Encryption is more important than ever as Internet-based crime is one of the fastest growing security threats. E2EE, the most secure form of encryption, ensures that sensitive, confidential information transmitted online remains confidential. This is specifically because E2EE messages can

Question	Your response
	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. Indeed, E2EE play a uniquely positive role that E2EE in providing safety and security to children[1] and adult users, effectively protecting them against a number of real-world harms including stalking, retaliation for reporting abuse, impersonation, and so on. Moreover, E2EE is an essential tool protecting the UK's national security and government operations. It 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. From our reading of this and previous Consultation documents, we identify a significant conflict in Ofcom's proposed approach. Ofcom has previously acknowledged that proactive measures under S.10 do not apply to E2EE services. Ofcom has also acknowledged that proactive measures would not apply where they are not technically feasible without compromising the security of a service, and stated that this meant E2EE services.[2] We interpreted this to mean that Ofcom would not recommend accredited technology for an E2EE service under the S.121 Technology Notices because there is no technology available that meets the requirement. As such, it should have confirmed that Ofcom would not use the Technology Notice mechanism to enforce a requirement for CSEA scanning on an encrypted service. This would be in line with Lord Parkinson's statement in Parliament in Lords Report stage.[3] Without the existence of a proactive measure that is technically feasible, Ofcom could not use the powers outlined in the Online Safety Act. However, the current consultation states explicitly that it "does not take a view on…[t]he extent to which there is technology available that could be used to identify or prevent users encountering terrorism or CSEA content in any particular deployment scenarios, for example end-to-end encrypted environments."[4] While we continue to believe that Technology Notices cannot – and according to Ofcom's previous statements, would not – be issued against end-to-end encrypted services, we must reiterate our concerns in the event of a reversal.

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

[2] See Volume 4, Section 14, notably in Section 14.16 of Ofcom's 2024 Consultation: "Protecting people from illegal harms"

[3] Lord Parkinson, House of Lords Hansard, Column 2363, 19 July 2023.

[4] Ofcom, Technology Notices to deal with terrorism content and/or CSEA content: Consultation on policy proposals for minimum standards of accuracy for accredited technologies, and guidance to providers Consultation, pg. 14 (16 Dec. 2024).

Question	Your response
	Ofcom's proposed audit-based assessment, principles, and scoring system do not sufficiently account for the material risks involved. We believe client-side scanning technology is the technology that Ofcom acknowledges does not exist. **<u>Systemic Risks to E2EE</u>** If Ofcom were to change course and issue a Technical Notice requiring a provider to integrate an accredited client-side scanning technology into their E2EE services, this would violate the core privacy and trust guarantees that E2EE provides. It would undermine privacy protections like authenticating communicants and the integrity of the underlying E2EE communications and ensuring that no one other than the intended communicants can access the contents of those communications. Additionally, it is technically infeasible to implement client-side scanning technology due to issues of inherent systemic risk and the violation of user trust. A systemic weakness or vulnerability is one that extends beyond the targeted device or service that an individual user is using and is implemented such that any other user could be affected.[5] Any measure to screen the content of messages on an E2EE platform would introduce systemic risk, compromising devices and systems and leading to unauthorised access to data. It would increase risk for service providers and users. The outcome would be an unsafe and untrustworthy online environment and a new canvas for criminals to exploit. <u>E2EE Backdoors, and Adversarial Attacks Associated with Perceptual Hashing</u> Ofcom would require providers to use proactive measures such as perceptual hashing techniques.[6] Perceptual hashing creates a digital fingerprint of images uploaded and checks them against a database of images classified as illegal. Perceptual hashing could be implemented on the provider's server by creating a backdoor into the system to decrypt the message for scanning.[7] A backdoor is a form of exceptional access to allow for interception of messages.[8] Importantly, a backdoor represents a vulnerability point that not only would be used by law enforcement but also criminals and hostile foreign governments that seek unauthorised access. The alternative proposal is to implement the perceptual hashing system on the users' devices. The sales pitch is that it does not require a backdoor to decrypt the messages because it will intercept the user's communications as they are being uploaded, before the encryption process begins. This is commonly known as "client-side scanning". Importantly, depending on the implementation of

[5] Parliament of the Commonwealth of Australia, Review of the amendments made by the Telecommunications and Other Legislation Amendment (Assistance and Access) Act 2018, 4.56

[6] As outlined in Volume 4 of the 2024 Ofcom Consultation: "Protecting people from illegal harms online."

[7] For example, encrypting each message with a key known to the provider, rather than the keys on the devices of the communicating parties

[8] Jeff Wilbur, Ryan Polk. "A Backdoor Is a Backdoor Is a Backdoor." Internet Society, 24 Mar. 2020, www.internetsociety.org/blog/2020/03/a-backdoor-is-a-backdoor-is-a-backdoor/.

Question	Your response
	client-side scanning technology, it could also require building an encryption backdoor. If the technology were to go beyond preventing flagged content from being sent at the outset, and require that flagged communications be reported to a third party, such as law enforcement or safety agencies (e.g., NCA), enabling that report-out function would require building an encryption backdoor. In our recently published research "Pre-emptive monitoring in End-to-End Encrypted Services"[9] we identify several other 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 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. There are workarounds to perceptual hashing as well, such as the risk that perpetrators of CSEA and other illegal images evade the scanning software by flooding the system with false positives, overloading law enforcement authorities with material that is not illegal. Alternatively, they would generate false negatives—images that "match" as CSEA material that is not—that can slip through the scanner. **<u>Additional Risks Associated with Client-Side Scanning</u>** Client-side scanning is not a viable solution for content moderation in encrypted environments due to issues of inherent systemic risk and the violation of user trust. For example, putting the hash algorithm onto the client device would open it up to reverse engineering. The average user's expectations of privacy would be violated while criminals and hostile state actors would encounter little more than a speed bump that they would quickly develop techniques to circumvent. Client-side scanning risks other unintended consequences. It creates new opportunities for attackers to target the database, for example, by inserting unauthorised material for scanning. Academic researchers in the UK suggest that facial recognition could be surreptitiously inserted.[10] The update mechanism could be subverted to install malicious software like what happened in the Solar Winds cyber-attack. Putting the database on the device increases the "attack surface" that bad

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

[10] Jain, S., Cretu, A., Cully, A., and de Montjoye, Y., 2023. Deep perceptual hashing algorithms with hidden dual purpose: when client-side scanning does facial recognition.

Question	Your response
	actors can exploit,[11] and exposes millions of people's phones to bugging by unauthorised entities. These could include foreign states. Interference with privacy would be collateral damage. Multiple databases to address different content categories would increase the complexity of enforcement, generate extra network traffic, and require extra processing on the device, leading to complex issues around scalability, testing, consistency of data, and governance. **<u>Proportionality and the Human Rights Risks of Ofcom's Proposal</u>** Ofcom acknowledges its obligation to ensure that any technical notice it issues is necessary and proportionate in its impact on privacy and free expression. We are concerned the audit-based assessment proposed in this consultation does not incorporate any test to ensure that no technology which would violate human rights is not accredited in the first instance. The Online Safety Act states that measures implementing the duty in S.10 of the Act (Illegal content safety duties) must be designed in light of the principle of freedom of expression and the importance of protecting the privacy of users.[12] Freedom of expression is defined in the Act as having the meaning intended by Article 10 of the European Convention on Human Rights),[13] which is enshrined in the Human Rights Act 1998. When considering the proportionality of measures, the precise objective should be clear. Any restriction must be lawful, which generally means it must be clearly described in legislation. The lawfulness of the interference must be balanced against the rights of other users whose rights may be arbitrarily interfered with. The measures must be necessary and specific to achieve a legitimate purpose, ideally with a fact-based assessment of their effectiveness.[14] The least intrusive restriction should be used to meet the policy aim and it should be possible to show that other less intrusive measures have been evaluated. The quality of the law and existence of adequate safeguards will be factors. The law must be sufficiently clear so that users will be able to foresee when their communications could be interfered with. The scope of discretion for private actors implementing the measures must also be clear. The legal framework must include adequate safeguards against abuse by State or non-State actors, which should be assessed at each stage of the process. E2EE is one of the most privacy preserving designs for encrypted services, used billions of times daily to protect information flows online (most webpages are encrypted end-to-end from the server to the browser).[15] Providers of E2EE services cannot read messages shared on their platform—they merely transmit them. Only the sender and recipient can read them. The encryption keys are used to encode the message into a random series of numbers and letters at the "endpoint" (a device or piece of software) of the sender and then, on the other side, they are used to

[11] Joint Statement by Europol and the European Union Agency for Cybersecurity (ENISA) of 20 May 2016 on lawful criminal investigation that respects 21st Century data protection in the case of Podchasov v Russia in the European Court of Human Rights, (Application no. 33696/19) Judgment 13 February 2024.
[12] Schedule 4 (10) (1).
[13] S.236.
[14] EDPB-EDPS Joint Opinion 04/2022 on the Proposal for a Regulation of the European Parliament and of the Council laying down rules to prevent and combat child sexual abuse, Paragraph 42.
[15] Husovec, Martin. "Podchasov v. Russia App. No. 33696/19 - Martin Husovec." European Information Society Institute, 28 Sept. 2021, husovec.eu/wp-content/uploads/2021/10/Podchasov-v-Russia-Brief.pdf.

Question	Your response
	decode the message at the endpoint of the receiver. Neither the private key nor the original message is available to the operator. In this way, end-to-end encryption preserves and protects the integrity, authenticity, and confidentiality of people's messages. If one were looking to describe a service that enables and protects the right to privacy, one would probably describe something very like an encrypted service, such as that it could only be read by the sender and recipient. **<u>Privacy and Freedom of Expression</u>** A requirement for E2EE services to scan content would be likely to fail a proportionality test, in light the judgement of 13 February in the European Court of Human Rights, *Podchasov v. Russia*.[16] This case sends a clear signal that the measures Ofcom is proposing would in all probability be unlawful on encrypted services. Although there are some obvious differences in the specific measures that were challenged in the court, the principles set out in the judgement would apply to any instance where a service was asked to break, weaken, or compromise end-to-end encryption, or introduce backdoors, weaknesses, or vulnerabilities. Central to the case was the proportionality test and the court's reasoning as to why breaking end-to-end encryption would be disproportionate. In a nutshell, it is not possible to monitor specific content on an end-to-end encrypted service without creating indiscriminate interference with the privacy of other users who are not the target of the measures. The ECtHR judgement confirmed that regardless of the technology choice, the screening of uploaded content from every user on the system engages privacy and free expression rights. This is because it is not possible to monitor specific users' content, without arbitrarily affecting others on the network. To read the content of one user, providers have to install software—either through a backdoor on the server or on the end-user devices—that will indiscriminately impact all users. The reasoning for the ruling went along these lines: end-to-end encryption is a privacy-protecting tool that protects the integrity and confidentiality of communications and in doing so keeps individuals safe from attacks on their messages by hackers and other malicious actors. The content of the message is protected from everyone, even the platform provider, and only the sender and recipient can read it. This puts a barrier in the way of identifying the targeted material. It can only be identified by intercepting the communication, and reading the message in clear text, which involves either reading it before it is encrypted, or decrypting the message in transit. Potentially, if on-device scanning (client-side scanning) is deployed, the scanning software and database would be held on the users' smartphones. All of this is without the users' consent. Mapping this onto the Online Safety Act, long-standing protections for British citizens against State intrusion into their private lives could be undermined if such measures were required. A UK government-sponsored study of proof-of-concept tools for scanning encrypted services stated:

[16] European Court of Human Rights, Podchasov v. Russia. Judgement 13 February 2024 https://hudoc.echr.coe.int/eng/#{%22itemid%22:[%22001-230854%22]}. See also Third-Party Intervention by European Information Society Institute (EISi) https://husovec.eu/wp-content/uploads/2021/10/Podchasov-v-Russia-Brief.pdf

Question	Your response
	"…from a Human Rights perspective, the confidentiality of the E2EE service users' communications cannot be guaranteed when all content intended to be sent privately within the E2EE service is monitored pre-encryption." [17] Additionally, the current proposal would interfere with freedom of expression and privacy because it does not set a high accuracy threshold as a barrier to accreditation. Billions of people of people around the world rely on encrypted communications services every day. Since there is no technical way to ensure that only messages sent within the UK are subject to scanning, and imposing any such limitation would be impracticable for providers, if not impossible given the globally interconnected aspect of the Internet, the result would likely be the scanning of all global E2EE communications. If an accredited technology had a 99% accuracy rating, then even assuming only 1,000,000,000 messages are sent per day, that would result in 10,000,000 false positives. The real number of messages sent per day is likely hundreds of times higher. Therefore, in the best-case scenario, millions, if not billions, of false positives would be flagged and either blocked from being sent, or subjected to invasive human review. Even more concerningly, a Technology Notice could forward along all these false positives to third parties like law enforcement or safety agencies. These measures would engage privacy rights because they require providers to intercept and scan communications. They are very broad powers, without a warrant or suspicion that the individual has committed a crime. They would engage the right to freedom of expression because they may deter people from speaking, creating a 'chilling effect'. A significant percentage of the population could be affected: two-thirds of the UK adults say that WhatsApp is their main communications service.[18] A legal opinion from a leading barrister concludes that these measures are unlikely to be in accordance with the law and would be open to challenge on the basis of that they would constitute a disproportionate interference with privacy rights.[19] 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. Confidential? – N
Consultation question 2:	

[17] REPHRAIN: Towards a Framework for Evaluating CSEA Prevention and Detection Tools in the Context of End-to-end encryption Environments: a Case Study. February 2023

[18] "Whatsappening in the World of Online Communications?" Ofcom, 25 Oct. 2023, www.ofcom.org.uk/news-centre/2023/whatsappening-in-the-world-of-online-communications.

[19] Index on Censorship/ Matthew Ryder KC: Surveilled and Exposed: how the Online Safety Bill creates insecurity, November 2022.

Question	Your response
Do you have any views on our proposals for independent performance testing, including the two mechanisms for setting thresholds; the approach to testing technologies in categories against particular metrics; and data considerations? Please provide evidence to support your response.	Confidential? – Y/N
Consultation question 3:	
Do you have any comments on what Ofcom might consider in terms of how long technologies should be accredited for and how often technologies should be given the opportunity to apply for accreditation? Is there any further evidence we should consider?	Confidential? – Y / N
Consultation question 4:	

Question	Your response
Do you have any views on how to turn these proposals into an operational accreditation scheme, including the practicalities ofsubmitting technology for accreditation? Is there any additional evidence that you think we should consider? Please provide any information that may be relevant.	
Consultation question 5:	
Do you have any comments on our draft Technology Notice Guidance?	Confidential? – Y / N