First phase of online safety regulation

Call for evidence

Call for evidence: First phase of online safety regulation – Welsh translation

CALL FOR EVIDENCE:

Publication date: 6 July 2022

Closing date for responses: 13 September 2022

Contents

Section

Annexes

1. Overview

Background

1.1 Ofcom is the UK's communications regulator, overseeing sectors including telecommunications, post, broadcast TV and radio. We regulate online video services established in the UK, including on-demand programme services (ODPS) and video-sharing platforms (VSPs). In line with our statutory duties, we have a programme of work dedicated to promoting and carrying out research on media literacy.

1.2 In February 2020, Government announced it was minded to appoint Ofcom as the regulator for online safety in the UK. We have been working since then to develop our understanding of the opportunities and challenges of online safety regulation, build our internal capability and begin planning for our regulatory approach. Since December 2020, we have been funded by Government to develop and strengthen our capabilities to prepare for this new role, including creating a new Online Safety Policy team, a new Trust & Safety Technology function and growing our Enforcement, Legal, Research and Insight, and Data teams. The Online Safety Bill, which will give Ofcom new regulatory powers, was introduced to Parliament on 17 March 2022. We're working hard to get ready for our new role, and we look forward to continued close engagement with stakeholders as we develop our policies and plans.

1.3 We have been informed by our experience in regulating video-sharing platforms (VSPs), for which we already have powers under transposed European legislation. We are also continuing to cultivate effective relationships with agencies and regulators in the UK (for example, through the Digital Regulation Cooperation Forum) and in other jurisdictions, and to participate in global conversations and forums that seek solutions to problems of online safety.

1.4 Alongside this call for evidence, we have also published a roadmap, setting out our current thinking about our plan for implementing online safety regulation.

The Online Safety Bill

1.5 The Online Safety Bill, as currently drafted, will require services which host user-generated content and search engines to have systems and processes for protecting individuals from certain types of harm online, and require pornography providers to ensure children are not normally able to encounter pornographic content.

1.6 It provides for a number of new duties for regulated services, which are set out in more detail in our roadmap document. These include duties:

a) to assess risks on their services and to have proportionate systems and processes to protect users and individuals from illegal content;

b) to put in place additional protections for children, defined as anyone under the age of 18, against content that is harmful to them;

c) for 'Category 1' services to have, and consistently apply, their own terms and conditions on certain types of content that are 'legal but harmful' to adults;[1]

d) for 'Category 1' and 'Category 2a' services to have proportionate systems and processes to prevent fraudulent advertising;[2]

e) for 'Category 1' and 'Category 2b' services to publish reports on the incidence of illegal and harmful content on their services;[3] and

f) to have regard to the importance of protecting freedom of expression and users' privacy. 'Category 1' services would have enhanced duties in this area, including to undertake an impact assessment.

1.7 The new regulation will apply to a wide range of user-to-user and search services, with different reach, sizes and risk levels. Regulated services will include (while not being limited to) social media platforms, video-sharing platforms, forums, messaging apps, some online games, cloud storage and sites hosting pornographic content. More detail on the scope of the regime can be found in our roadmap.

1.8 Under the Online Safety Bill, Ofcom will gain information gathering and enforcement powers, including powers to impose substantial fines. Ofcom must produce a range of publications establishing the new framework, including a sector risk assessment (including risk profiles), risk assessment guidance for industry, codes of practice outlining steps that companies may take to comply with their duties, and other regulatory guidance, in areas including transparency reporting and child access assessments.

Our call for evidence

1.9 This call for evidence is focused on the matters that we currently anticipate will be included in our first consultation in 2023, as set out in more detail by our roadmap. The focus is therefore on assessment of the risk of harm from illegal content, mitigations around illegal content, child access assessments and transparency requirements. We anticipate issuing a subsequent call for evidence on legal but harmful content which, as set out in more detail by our roadmap, will be phased in at a later time following secondary legislation.

1.10 Through the questions set out in Annex 1, we are seeking evidence from a wide range of stakeholders, to strengthen our understanding of the range of approaches and techniques that platforms can employ to help them meet their online safety duties. We would

[1] Category 1 services are the highest reach user-to-user services with the highest risk functionalities, with transparency requirements, a duty to assess risks to adults of legal but harmful content, requirements relating to fraudulent advertising and a variety of other duties.

[2] Category 2a services are the highest reach search services, with transparency and fraudulent advertising requirements.

[3] Category 2b services are other services with potentially risky functionalities or other factors, with transparency requirements, but no other additional duties.

welcome evidence on the efficacy of the measures that can be taken to mitigate harm, their costs and the practicability of implementing them – where relevant reflecting on how this may differ depending on service characteristics.

1.11 We would like to hear from providers whose services are likely to fall within scope of the Online Safety framework, from the full range of services in scope, as well as regulators, academics, civil society organisations, consumer representatives and other stakeholders with interest and expertise in this area. We would like to hear from businesses, organisations and groups that are able to provide evidence around the questions in Annex 1. We have signposted where our questions are aimed at all stakeholders, and where they are primarily directed towards service providers given that they request information about current practice, costs and assessment metrics.

1.12 We envisage that the evidence provided in response will be valuable in preparing future reports and initiatives under our media literacy powers and, assuming the Online Safety Bill passes, be of use in informing how we carry out our functions. These include drafting codes of practice and regulatory guidance which will set out steps that companies can take to comply with their duties under the Bill. This call for evidence is one part of our preparations for taking on our new duties, alongside a wider programme of research and extensive stakeholder engagement.

1.13 As we are exploring the ways that services may comply with their duties under the Bill, which is yet to be approved by Parliament, this document and our questions should not be seen as an indication or statement of policy intent, but rather as an opportunity for input.

Next steps

1.14 Our call for evidence will remain open for 10 weeks from publication and we request responses back by 5pm on 13 September 2022.

1.15 Ofcom will not receive any new powers until the Online Safety Bill has received Royal Assent. Further details about our plan for consultation and implementation are provided in our roadmap.

1.16 Our existing duties require us to consult widely on any proposals or decisions which amend or impose new regulatory obligations. We consider that the complexity and novelty of this new regulatory regime will benefit from close engagement with stakeholders. This call for evidence is our first formal opportunity to do this ahead of our 2023 consultation and we look forward to continued close engagement as we develop our policies and plans.

A1. Questions

A1.1 We are seeking evidence to strengthen our understanding of the range of approaches and techniques platforms can employ to help them meet their proposed duties under the Online Safety Bill ('the Bill').

A1.2 In line with the plan set out by our roadmap, a large focus of the first phase of our work will be on the duties relating to illegal content. Where we refer to 'illegal content' in the questions below, we mean both 'priority' and non-priority illegal content as defined by the Bill. 'Priority' offences are listed in Schedules 5 (Terrorism), 6 (Child Sexual Exploitation and Abuse) and 7 (other offences) of the Bill, broadly encompassing offences including:

a) terrorism;

b) child exploitation and abuse;

c) encouraging or assisting suicide;

d) threats to kill;

e) public order offences, harassment, stalking and fear or provocation of violence;

f) drugs and psychoactive substances;

g) firearms and other weapons;

h) assisting unlawful immigration;

i) sexual exploitation;

j) unlawful sexual images, including extreme pornography and disclosing private sexual photographs with intent to cause distress;

k) proceeds of crime including, for example, concealing and facilitating the acquisition of criminal property;

l) fraud; and

m) financial services including, for example, false claims to be authorised under the Financial Services and Market Act.

A1.3 We have set these offences out in full in Annex 2. Non-priority offences are other offences not specified within the Bill, where the victim or intended victim is an individual.[4]

A1.4 Our questions are divided into sections, each covering an element of the new regulatory framework proposed by the Bill. Under each question, we have provided prompts to expand on the areas in which we are particularly interested. You do not need to respond to every question or prompt. Where our questions are targeted at providers of online services, we have indicated accordingly. Please provide evidence to support your

[4] Excluding offences relating to the infringement of intellectual property rights, the safety or quality of goods, the performance of a service by a person not qualified to perform it; or offences under the Consumer Protection from Unfair Trading Regulations 2008 (S.I. 2008/1277).

responses; clearly evidenced and reasoned submissions will be most valuable in improving our understanding of the questions below.

A1.5 Our normal practice is to publish non-confidential versions of responses on our website. As such, you should specify if your response or a part of it is confidential, where necessary. If someone asks us to keep part or all of a response confidential, we will treat this request seriously and try to respect it. Please refer to Annex 3 for further instructions on submitting a response.

A1.6 If you are a business, organisation or group with expertise and relevant evidence around the questions below, and would like to get in touch with us about this call for evidence, please contact os-cfe@ofcom.org.uk.

Preliminary question

A1.7 Our first question asks respondents for information about themselves, so that we can categorise responses to later questions.

A1.8 The Bill applies to a wide range of 'user-to-user services', defined as 'an internet service by means of which content that is generated directly on the service by a user of the service, or uploaded to or shared on the service by a user of the service, may be encountered by another user, or other users, of the service'. It will also apply to search services, defined as 'an internet service that is, or includes, a search engine'. The Bill also imposes duties on providers of pornographic content, defined as services which publish or display material produced primarily for the purpose of sexual arousal. Any service which has significant numbers of UK users or which is targeted at the UK market will have new duties, and must comply with the new law.

For all respondents

Q1. Please provide a description introducing your organisation, service or interest in Online Safety.

For providers of online services, please provide information about:

- The type of service and functionalities you provide;[5]
- Number of users globally, and in the UK;
- Global and UK revenues; and
- Your business models and revenue generation.

Please indicate where this information is confidential.

Risk assessment and management

A1.9 Risk assessment and management will be a cornerstone of the regime and is a core requirement of the Bill. Part 7, Chapter 3 of the Bill sets out the requirements on Ofcom to carry out a risk assessment to identify and assess risks of harm to individuals presented by user-to-user and search services, and to identify and assess the characteristics of different kinds of services that are relevant to those risks. Ofcom must develop 'risk profiles' for user-to-user and search services which relate to the risks of harm. Ofcom must publish its risk assessment (in a register of risks) and risk profiles, keeping both up to date.

A1.10 As well as codes of practice setting out recommended steps for compliance with the safety duties, Ofcom must prepare and publish guidance for providers of regulated services to assist them in complying with their duties to carry out their own risk assessments. In undertaking their risk assessments, services must take account of the risk profiles prepared by Ofcom.

[5] Within the Bill, 'functionalities' of user-to-user services include: creating a user profile, including an anonymous or pseudonymous profile; searching within the service for user-generated content or other users; forwarding content to, or sharing content with, other users of the service; sharing content on other internet services; sending direct messages to or speaking to other users of the service, or interacting with them in another way (for example by playing a game); expressing a view on content, including, for example, by applying a "like" or "dislike" button or other button of that nature, applying an emoji or symbol of any kind, engaging in yes/no voting or rating or scoring content in any way (including giving star or numerical ratings); sharing current or historic location information with other users of the service, recording a user's movements, or identifying which other users of the service are nearby; following or subscribing to particular kinds of content or particular users of the service; creating lists, collections, archives or directories of content or users of the service; tagging or labelling content present on the service; uploading content relating to goods or services; applying or changing settings on the service which affect the presentation of user-generated content on the service; accessing other internet services through content present on the service (for example through hyperlinks). 'Functionalities' of search services include: a feature that enables users to search websites or databases; and a feature that makes suggestions relating to users' search requests (predictive search functionality).

For all respondents

Q2. Can you provide any evidence relating to the presence or quantity of illegal content on user-to-user and search services?

We are interested in briefings, investigations, transparency reports, media investigations and research papers that provide more evidence about how such content might vary across different services or types of service, or across services with particular groups of users, features or functionalities.

IMPORTANT: Under this section we are NOT seeking links to or copies/screenshots of content that is illegal to hold, such as child sexual abuse. Deliberately viewing such images may be a criminal offence and will be reported to the police.

For providers of online services

Q3. How do you currently assess the risk of harm to individuals in the UK from illegal content presented by your service?

Please provide information about:

- how risks from illegal content are identified (including any relevant internal processes, policies and documents);
- the extent to which you factor in evidence on users' behaviour and age to consider potential risks; and
- the exacerbating and mitigating risk factors you consider (for example, relating to your **users**, **business model** or **features and functionalities** which may have an impact on the risk of harm).

Q4. What are your governance, accountability and decision-making structures for user and platform safety?

As part of your answer, please outline how different teams may consider user safety risks across different business functions such as product development, management, engineering, public policy, safety, legal, business development and marketing.

Please consider:

- How are risks to user safety escalated and acted upon in your organisation? What senior management oversight is in place?
- How are staff trained to understand how their own roles and responsibilities can potentially create risks to user safety?
- How do you ensure consistency in consideration of user safety across teams?

Terms of service and policy statements

A1.11 The Bill sets a number of expectations around regulated user-to-user services' terms of service, including that they cover specified types of content, such as priority illegal content, and measures, such as the use of proactive technologies. Terms of service are also expected to be clear and accessible to users and consistently applied. Search services will be required to outline similar provisions in public policy statements.

For all respondents

Q5. What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements?

Please submit evidence about what features make terms or policies clear and accessible.

For providers of online services

Q6. How do your terms of service or public policy statements treat illegal content? How are these terms of service maintained and how much resource is dedicated to this?

Please outline as part of your answer:

- What you currently cover in your terms of service and public policy documents in relation to illegal content and why, including any reference to what is considered illegal content, any proactive measures to identify it, and sanctions applied to users who are in breach;
- Whether you consider, when drafting these documents, the accessibility needs of different user groups and affected persons, including children, parents or groups with certain characteristics which may put them at a higher risk of harm from content on the service;
- Evidence of the process, time and any costs involved in developing these terms; and
- Whether you have any evidence about how users engage with your terms of service, or whether they understand what they mean in practice.

Reporting and complaints

A1.12 Under clauses 17 and 27 of the Bill, regulated services will be required to operate systems allowing users and affected persons to easily **report** content they consider to be illegal.

A1.13 Under clauses 18 and 28, regulated services will also be required to operate **complaints procedures**, allowing users and affected persons to complain about content they consider to be illegal; or if they consider that the provider is not complying with its duties, or that their content has been removed or ability to make use of the service has been restricted unduly.

For all respondents

Q7. What can providers of online services do to enhance the transparency, accessibility, ease of use and users' awareness of their reporting and complaints mechanisms?

Please submit evidence about what features make user reporting and complaints systems effective, considering:

- reporting or complaints routes for registered users;
- reporting or complaints routes for non-registered users; and
- reporting routes for children and adults.

For providers of online services

Q8. If your service has *reporting or flagging* mechanisms in place for illegal content, or users who post illegal content, how are these processes designed and maintained?

We are interested in obtaining evidence on:

- how users report content on your service (including the mechanisms' location and prominence for users, and any screengrabs you could provide) and whether this is separate from your complaints procedure;
- whether users need to create accounts to access reporting and flagging;
- what type of content or conduct users and non-users may flag or report, including any specific lists or categories;
- the key choices and factors involved in designing these mechanisms;
- how you ensure that reporting and flagging is user-friendly and accessible;
- whether any particular consideration is given to the needs of children, including children of different ages;
- whether any particular consideration is given to the needs of different user groups, including users with certain characteristics which may put them at a higher risk of harm from content on the service (if so, which user groups you have considered);
- the cost involved in designing and maintaining these mechanisms;
- whether your reporting and flagging mechanisms are effective, in terms of identifying illegal content, and how you determine this; and
- whether you use trusted flaggers (and if their reports are handled differently).

Q9. If your service has a *complaints* mechanism in place, how are these processes designed and maintained?

We are interested in obtaining evidence on:

- how users can complain about content on your service (including any screengrabs you could provide) and whether there is any separate procedure from your reporting mechanisms above;
- the type of complaint that can be made, including any categories or lists (including if the same categories are used for complaints);
- whether complaints can be made in cases when content is de-prioritised or taken down proactively by technology, or about freedom of expression and privacy issues;
- who can make a complaint;
- whether there are different procedures for different categories of users, including children or content creators vs. affected persons;
- the key choices and factors involved in designing your complaints procedure;
- the cost involved in designing and maintaining this process;
- whether any record of complaints is kept, and if so what; and
- any appeals processes if complainants are unsatisfied with the outcome of the complaint, when these can be accessed, how they work, and whether they are independent from the complaints processes.

Q10. What action does your service take in response to *reports* or *complaints*?

We are interested in obtaining evidence on:

- any variation in the number and actionability (i.e. the proportion that result in a takedown or other action in response) of reports or complaints across type of harm;
- what proportion of reports are considered, and what proportion are acted upon;
- what proportion of complaints are considered, and what proportion are acted upon;
- whether any reports or complaints are expedited or directed to specialist teams, the criteria for this, and the cost involved in facilitating this;
- if users identify illegal content, whether law enforcement or other third parties are informed;
- how the validity of reports is triaged and assessed, including how malicious reporting is identified, handled and/or mitigated;
- how users and content creators are informed as to whether any and what action has been taken, and why, as a result of material they or others have reported or flagged;
- what happens to the content while it is being assessed/processed; and
- any internal or external timeframes or key performance indicators (KPIs) for acting on or determining reports or complaints.

Moderation

A1.14 Under their duties set out in Part 3 of the Bill, regulated services will be required to: use proportionate (in reference to both the findings of the most recent risk assessment and size and capacity of the service provider) measures to effectively mitigate and manage the risks of harm to individuals; operate the service using proportionate systems and processes to prevent or minimise the risk of individuals encountering certain content; and, for user-to-user services, remove illegal content when they become aware of it. This includes through content moderation.

For all respondents

Q11. Could improvements be made to content moderation to deliver greater protection for users, without unduly restricting user activity? If so, what?

Please provide relevant evidence explaining your response to this question. Please consider improvements in terms of user safety and user rights, as well as any relevant considerations around potential costs or cost drivers.

For providers of online services

Q12. What automated moderation systems do you have in place around illegal content?

In particular, please consider:

- Do you use hash-matching to identify any forms of known illegal content?
- Do you use any other automated techniques to identify any forms of illegal content?
- If you use in-house automated techniques, how are these developed and trained? What training data was used and how it was sourced?
- Do you use any third-party datasets or providers and, if so, which?
- If you use a third-party provider, did you consider there to be choice in the market? Why did you opt for your chosen provider?
- What happens to content once it is identified by automated means?
- Can you provide any evidence as to how effective these techniques are, in terms of reducing harm to users, and how this may vary by harm?
- How do you assess the performance (including wrongful takedown rates, precision and recall) of automated moderation techniques?
- What safeguards are employed to protect user privacy and/or to avoid over takedown?
- Are there certain types of automated techniques which have strengths over others, or which are better suited to certain harms or types of content over others?
- What are the barriers and costs involved in deploying these techniques?

Q13. How do you use human moderators to identify and assess illegal content?

In particular, please consider:

- How do you determine the level of human moderation required by your platform, including by type of content?
- Are moderators employed by the service, outsourced, or made up of volunteers?
- Are moderators vetted, and how?
- What coverage do moderators provide (e.g. on weekends or overnight, UK time)?
- What training and support is provided to moderators? For instance, are certain moderators specialised in certain harms or speech issues? Is training updated, and when? Are moderators trained to identify illegal content?
- If you make use of automated moderation, how do human moderators and automated systems work together, and what is their relative scale? How do you guard against automation bias?
- What are the costs involved in these practices? In the absence of specific costs, please provide indication of cost drivers (e.g. moderator location) and other relevant figures (e.g. number of moderators employed, how many items the platform moderates per day).
- How do you assess the accuracy and consistency of human moderation teams?

Actioning content and sanctioning users

A1.15 As well as content moderation, the duties set out in paragraph A1.14 encompass taking down content and policies on user access to the service, or particular content present on the service.

For all respondents

Q14. How are sanctions or restrictions around access (including to both the service and to particular content) applied by providers of online services?

Please provide evidence around the application and accuracy of sanctions/restrictions, and safeguards you consider should be in place to protect users' privacy and prevent unwarranted sanction.

For providers of online services

Q15. In what instances is illegal content removed from your service?

Please outline the circumstances in which illegal content is removed and how this may differ by type of illegal content.

Q16. Do you use other tools to reduce the visibility and impact of illegal content?

Please consider providing the following information:

- the circumstances in which content is actioned and how;
- whether these tools are specified in public user terms or policies;
- evidence of whether your users understand these tools' existence and application;
- what safeguards you use to ensure they are applied consistently and fairly;
- evidence on the costs of developing and maintaining these tools; and
- evidence on the effectiveness of these tools, in terms of reducing harm.

Q17. What other sanctions or disincentives do you employ against users who post illegal content?

Please outline:

- What type of sanctions do you apply and in what instances?
- Do you take steps to prevent users who have been banned from a service re-accessing it and, if so, what? What are the costs involved in this?
- Are users informed about sanctions and, if so, how?
- What evidence can you provide around their efficacy, in terms of reducing harm?
- What safeguards do you use to ensure users are sanctioned consistently and fairly?

Design and operation of the service, including functionalities and algorithms

A1.16 The duties set out in paragraph A1.14 also encompass measures relating to the design of functionalities, algorithms and other features of the service, and provision of user support measures.

A1.17 We are interested in understanding services' approach to design as a way to mitigate harm, including how safety is considered in the design of products and functionalities. The Online Safety Bill provides a number of examples of 'functionalities'; for user-to-user services, this includes, for example, the ability of a user to have an anonymous profile, to like or dislike content, to share location information with other users, and to forward content to other users. An example of a functionality on a search service is auto completion, where a search engine predicts the rest of a query that a user has begun to type.

A1.18 We are interested in hearing about aspects of design that could be considered to be preventative of harm; this could include both providing new features to users (e.g. allowing control over what they encounter), or restricting functionality (e.g. limiting discoverability).

For all respondents

Q18. Are there any functionalities or design features which evidence suggests can effectively prevent harm, and could or should be deployed more widely by industry?

Please provide relevant evidence explaining your response to this question.

For providers of online services

Q19. To what extent does your service encompass functionalities or features designed to mitigate the risk or impact of harm from illegal content?

In particular:

- Can you identify any functionalities or features on your platform which you consider can effectively prevent harm?
- What control do users have over what they are shown or is delivered to them?
- To what extent are the features or functionalities you identified reliant on other technology – for instance, age verification or ID verification mechanisms?
- What costs or cost drivers are involved in developing these features or functionalities?
- Is safety incorporated into the product design and development process? To what extent do you consider evidence about user behaviour when developing features or functionalities intended to enhance user safety?
- How do you measure and what evidence can you provide around the impact and effectiveness of these techniques, in terms of reducing harm to users?
- How do you assess the impact of these techniques on users' privacy and minimise the risk of over restriction?

Q20. How do you support the safety and wellbeing of your users as regards illegal content?

In particular:

- Do you provide support through your platform (e.g. signposting to resources)?
- Do you provide support off-platform (e.g. funding or facilitating programmes)?
- How effective are these types of interventions, in terms of minimising harm from and impact of illegal content on users? What evidence do you have to show this?
- What are the costs involved in implementing the support measures you have described?

Q21. How do you mitigate any risks posed by the design of algorithms that support the function of your service (e.g. search engines, or social and content recommender systems), with reference to illegal content specifically?

In particular:

- If different from your risk assessment process outlined in response to Q3, how do you assess the risk from algorithms central to the function of your service?
- What safeguards do you have in place to mitigate the risks posed by algorithms (e.g. testing them before they are put into use, and monitoring their performance in real world settings)?
- What are the costs involved in implementing these safeguards? In the absence of specific costs, please provide indication of the key cost drivers.
- How do you measure the effectiveness of these safeguards, in terms of reducing harm to users?
- What information can you provide to demonstrate the effectiveness of such safeguards?
- How do you assess the impact of these safeguards on users' privacy and minimise the risk of over restriction?

Child protection

A1.19 Use of age assurance or age verification may be one way that providers could seek to fulfil their illegal content duties, outlined in paragraph A1.14 above, particularly in relation to child exploitation and abuse, such as grooming content.

A1.20 All regulated services will be required to conduct a children's access assessment to determine if services are likely to be accessed by children (Part 3, Chapter 4 of the Bill), defined as anyone under the age of 18. The Bill states that providers may only conclude that it is not possible for children to access a service, or part of it, if there are systems or processes in place – for example, age verification, or another means of age assurance – that achieve the result that children are not normally able to access the service or part of it.

A1.21 If a service, or a part of the service, is likely to be accessed by children, providers will need to conduct a children's risk assessment and comply with the duties to protect children's safety online (clauses 10, 11, 25 and 26 of the Bill). Part 5 of the Bill sets out the separate duty on in-scope online services related to non-user-generated or 'provider' pornographic content. Services hosting such content must ensure that children are not normally able to encounter pornographic content (for example, by using age verification). As explained in our roadmap, we do not currently anticipate consulting in relation to these requirements until later, after the relevant secondary legislation has been made, but would be interested in views as to whether the use of age verification/age assurance would differ in relation to these duties compared to illegal content.

For all respondents

Q22. What age assurance and age verification technologies are available to platforms, and what is the impact and cost of using them?

In particular, please provide evidence explaining:

- how these technologies can be assessed for effectiveness or impact on users' safety;
- how accurate these tools are in verifying the age of users, and effective in preventing children from accessing harmful content;
- steps that can be taken to mitigate any risk of bias or exclusion that may result from age assurance and age verification tools;
- the costs involved in implementing such technologies; and
- the safeguards necessary to ensure users' privacy and access to information is protected, and over restriction is avoided.

Q23. Can you identify factors which might indicate that a service is likely to attract child users?

For providers of online services

Q24. Does your service use any age assurance or age verification tools or related technologies to verify or estimate the age of users?

In particular, please provide information explaining:

- the nature of the age assurance or age verification tools you have in place, including whether you engage third party providers;
- the costs involved in implementing such technologies;
- how accurate these tools are in verifying the age of users, and effective in preventing children from accessing harmful content;
- how your age assurance policies have been developed and what age group(s) they are intended to protect;
- what kind of content/activity your age assurance policies and mechanisms are intended to protect children from;
- if the service is tailored to meet age-appropriate needs (for example, by restricting specific content to specific users), how this works;
- the steps you take to mitigate any risk of bias or exclusion that may result from the age assurance and age verification tools you have in place; and
- any safeguards in place to ensure users' privacy and access to information is protected, and over restriction is avoided.

Q25. If it is not possible for children to access your service, or a part of it, how do you ensure this?

Q26. What information do you have about the age of your users?

Do you gather, or share with any third parties (and if so which), any information that can assist in estimating a user's age, either at the point a user first accesses the service or subsequently? If so, how do you do this and why?

Transparency

A1.22 Under Part 4, Chapter 3 of the Bill, certain platforms will be required to publish annual transparency reports including information set out by Ofcom for each provider. Ofcom will also be required to publish transparency reports based on platforms' transparency reports, its own research, and third-party research. Ofcom is interested to hear from the full range of stakeholders about what information would be most useful to include in platforms' transparency reports, Ofcom's own transparency reports, and Ofcom's report about researcher access.[6]

For all respondents

Q27. For purposes of transparency, what type of information is useful/not useful? Why?

In particular, please consider:

- Any evidence of public information positively or negatively affecting online user safety or behaviours, how this information is used, and by whom;
- What information platforms should make available, considering frequency, format and intended audiences;
- What information Ofcom should make available through its transparency report, considering frequency, format, intended audiences and potential use cases by external stakeholders;
- The benefits and/or drawbacks of standardised information and metrics; and
- Any negative impacts or potential unintended consequences of publishing certain types of information, and how these may be mitigated.

Other

For all respondents

Q28. Other than those in this document, are you aware of other measures available for mitigating risk and harm from illegal content? We would be interested in any evidence you can provide on their efficacy, in terms of reducing harm to users, cost and impact on user rights and user experience.

[6] Schedule 8 of the Bill sets out the matters which may be covered in platforms' transparency reports.

A2. Priority offences

A2.1 For ease of reference, we have copied below the priority offences outlined in Schedules 5, 6 and 7 of the Online Safety Bill, dated 28 June 2022.

Schedule 5, terrorism offences

A2.2 An offence under any of the following provisions of the Terrorism Act 2000—

a) section 11 (membership of a proscribed organisation);

b) section 12(1) (inviting support for a proscribed organisation);

c) section 12(1A) (expressing an opinion or belief supportive of a proscribed organisation);

d) section 12(2) (arranging a meeting supportive of a proscribed organisation);

e) section 13(1A) (publishing image of uniform of proscribed organisation);

f) section 15 (terrorist fund-raising);

g) section 16(1) (use of money or property for terrorist purposes);

h) section 16(2) (possession of money or property for terrorist purposes);

i) section 17 (involvement in terrorist funding arrangements);

j) section 18 (laundering of terrorist property);

k) section 54(1) (providing weapons training);

l) section 54(3) (inviting another to receive weapons training);

m) section 56 (directing a terrorist organisation);

n) section 58 (collection of information likely to be of use to a terrorist);

o) section 58A (publishing information about members of the armed forces etc);

p) sections 59 to 61 (inciting terrorism outside the United Kingdom).

A2.3 An offence under section 113 of the Anti-terrorism, Crime and Security Act 2001 (use of noxious substances or things).

A2.4 An offence under any of the following provisions of the Terrorism Act 2006—

a) section 1 (encouragement of terrorism);

b) section 2 (dissemination of terrorist publications);

c) section 5 (preparation of terrorist acts);

d) section 6 (training for terrorism);

e) section 11 (terrorist threats relating to radioactive devices etc).

Inchoate offences

A2.5 An offence of attempting or conspiring to commit an offence specified above.

A2.6 An offence under Part 2 of the Serious Crime Act 2007 (encouraging or assisting) in relation to an offence specified above, or (in Scotland) inciting a person to commit such an offence.

A2.7 An offence of aiding, abetting, counselling or procuring the commission of an offence specified above, or (in Scotland) being involved art and part in the commission of such an offence.

Schedule 6, child sexual exploitation and abuse offences

England and Wales, and Northern Ireland

A2.8 An offence under section 2 of the Obscene Publications Act 1959 relating to an obscene article tending to deprave and corrupt others by encouraging them to commit an offence specified in the paragraphs A2.9, A2.11, A2.12, A2.14 and A2.15.

A2.9 An offence under section 1 of the Protection of Children Act 1978 (indecent photographs of children).

A2.10 An offence under Article 3 of the Protection of Children (Northern Ireland) Order 1978 (S.I. 1978/1047 (N.I. 17)) (indecent photographs of children).

A2.11 An offence under section 160 of the Criminal Justice Act 1988 (possession of indecent photograph of a child).

A2.12 An offence under any of the following provisions of the Sexual Offences Act 2003—

a) section 8 (causing or inciting a child under 13 to engage in sexual activity);

b) section 10 (causing or inciting a child to engage in sexual activity);

c) section 11 (engaging in sexual activity in the presence of a child);

d) section 12 (causing a child to watch a sexual act);

e) section 13 (child sex offences committed by children or young persons);

f) section 14 (arranging or facilitating commission of a child sex offence);

g) section 15 (meeting a child following sexual grooming etc);

h) section 15A (sexual communication with a child);

i) section 47 (paying for sexual services of a child);

j) section 48 (causing or inciting sexual exploitation of a child);

k) section 49 (controlling a child in relation to sexual exploitation);

l) section 50 (arranging or facilitating sexual exploitation of a child).

A2.13 An offence under any of the following provisions of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2))—

a) Article 15 (causing or inciting a child under 13 to engage in sexual activity);

b) Article 17 (causing or inciting a child to engage in sexual activity);

c) Article 18 (engaging in sexual activity in the presence of a child);

d) Article 19 (causing a child to watch a sexual act);

e) Article 20 (sexual offences against children committed by children or young persons);

f) Article 21 (arranging or facilitating commission of a sex offence against a child);

g) Article 22 (meeting a child following sexual grooming etc);

h) Article 22A (sexual communication with a child);

i) Article 37 (paying for sexual services of a child);

j) Article 38 (causing or inciting child prostitution or pornography);

k) Article 39 (controlling a child prostitute or a child involved in pornography);

l) Article 40 (arranging or facilitating child prostitution or pornography).

A2.14 An offence under section 62 of the Coroners and Justice Act 2009 (possession of prohibited image of a child).

A2.15 An offence under section 69 of the Serious Crime Act 2015 (possession of paedophile manual).

Inchoate offences

A2.16 An offence of attempting or conspiring to commit an offence specified in paragraphs A2.8 to A2.15.

A2.17 An offence under Part 2 of the Serious Crime Act 2007 (encouraging or assisting) in relation to an offence specified in paragraphs A2.8 to A2.15.

A2.18 An offence of aiding, abetting, counselling or procuring the commission of an offence specified in paragraphs A2.8 to A2.15.

Scotland

A2.19 An offence under either of the following provisions of the Civic Government (Scotland) Act 1982—

a) section 52 (indecent photographs etc of children);

b) section 52A (possession of indecent photographs of children).

A2.20 An offence under any of the following provisions of the Protection of Children and Prevention of Sexual Offences (Scotland) Act 2005—

a) section 1 (meeting a child following certain preliminary contact);

b) section 9 (paying for sexual services of a child);

c) section 10 (causing or inciting provision by child of sexual services or child pornography);

d) section 11 (controlling a child providing sexual services or involved in pornography);

e) section 12 (arranging or facilitating provision by child of sexual services or child pornography).

A2.21 An offence under any of the following provisions of the Sexual Offences (Scotland) Act 2009—

a) section 21 (causing a young child to participate in a sexual activity);

b) section 23 (causing a young child to look at a sexual image);

c) section 24 (communicating indecently with a young child etc);

d) section 31 (causing an older child to participate in a sexual activity);

e) section 33 (causing an older child to look at a sexual image);

f) section 34 (communicating indecently with an older child etc);

g) section 54 (incitement to commit certain sexual acts outside Scotland).

Inchoate offences

A2.22 An offence of attempting or conspiring to commit an offence specified in paragraphs A2.19 to A2.21.

A2.23 An offence of inciting a person to commit an offence specified in paragraphs A2.19 to A2.21.

A2.24 An offence of aiding, abetting, counselling or procuring the commission of an offence specified in paragraphs A2.19 to A2.21, or being involved art and part in the commission of such an offence.

Schedule 7, priority offences

Assisting suicide

A2.25 An offence under section 2 of the Suicide Act 1961 (assisting suicide etc).

A2.26 An offence under section 13 of the Criminal Justice Act (Northern Ireland) 1966 (c. 20 (N.I.)) (assisting suicide etc).

Threats to kill

A2.27 An offence under section 16 of the Offences against the Person Act 1861 (threats to kill).

Public order offences, harassment, stalking and fear or provocation of violence

A2.28 An offence under any of the following provisions of the Public Order Act 1986—

a) section 4 (fear or provocation of violence);

b) section 4A (intentional harassment, alarm or distress);

c) section 5 (harassment, alarm or distress);

A2.29 An offence under any of the following provisions of the Public Order Act 1986—

a) section 18 (use of words or behaviour or display of written material);

b) section 19 (publishing or distributing written material);

c) section 21 (distributing, showing or playing a recording);

d) section 29B (use of words or behaviour or display of written material);

e) section 29C (publishing or distributing written material);

f) section 29E (distributing, showing or playing a recording).

A2.30 An offence under section 50A of the Criminal Law (Consolidation) (Scotland) Act 1995 (racially-aggravated harassment).

A2.31 An offence under any of the following provisions of the Protection from Harassment Act 1997—

a) section 2 (harassment);

b) section 2A (stalking);

c) section 4 (putting people in fear of violence);

d) section 4A (stalking involving fear of violence or serious alarm or distress).

A2.32 An offence under any of the following provisions of the Protection from Harassment (Northern Ireland) Order 1997 (S.I. 1997/1180 (N.I. 9))—

a) Article 4 (harassment);

b) Article 6 (putting people in fear of violence).

A2.33 An offence under any of the following provisions of the Crime and Disorder Act 1998—

a) section 31 (racially or religiously aggravated public order offences);

b) section 32 (racially or religiously aggravated harassment etc).

A2.34 An offence under any of the following provisions of the Criminal Justice and Licensing (Scotland) Act 2010 (asp 13)—

a) section 38 (threatening or abusive behaviour);

b) section 39 (stalking).

Drugs and psychoactive substances

A2.35 An offence under any of the following provisions of the Misuse of Drugs Act 1971—

a) section 4(3) (unlawful supply, or offer to supply, of controlled drugs);

b) section 9A (prohibition of supply etc of articles for administering or preparing controlled drugs);

c) section 19 (inciting any other offence under that Act).

A2.36 An offence under section 5 of the Psychoactive Substances Act 2016 (supplying, or offering to supply, a psychoactive substance).

Firearms and other weapons

A2.37 An offence under section 1(1) or (2) of the Restriction of Offensive Weapons Act 1959 (sale etc of flick knife etc).

A2.38 An offence under any of the following provisions of the Firearms Act 1968—

a) section 1(1) (purchase etc of firearms or ammunition without certificate);

b) section 2(1) (purchase etc of shot gun without certificate);

c) section 3(1) (dealing etc in firearms or ammunition by way of trade or business without being registered);

d) section 3(2) (sale etc of firearms or ammunition to person other than registered dealer);

e) section 5(1), (1A) or (2A) (purchase, sale etc of prohibited weapons);

f) section 21(5) (sale etc of firearms or ammunition to persons previously convicted of crime);

g) section 22(1) (purchase etc of firearms or ammunition by person under 18);

h) section 24 (supplying firearms to minors);

i) section 24A (supplying imitation firearms to minors).

A2.39 An offence under any of the following provisions of the Crossbows Act 1987—

a) section 1 (sale and letting on hire of crossbow);

b) section 2 (purchase and hiring of crossbow).

A2.40 An offence under any of the following provisions of the Criminal Justice Act 1988—

a) section 141(1) or (4) (sale etc of offensive weapons);

b) section 141A (sale of knives etc to persons under 18).

A2.41 An offence under any of the following provisions of the Criminal Justice (Northern Ireland) Order 1996 (S.I. 1996/3160 (N.I. 24))—

a) Article 53 (sale etc of knives);

b) Article 54 (sale etc of knives etc to minors).

A2.42 An offence under any of the following provisions of the Knives Act 1997—

a) section 1 (unlawful marketing of knives);

b) section 2 (publication of material in connection with marketing of knives).

A2.43 An offence under any of the following provisions of the Firearms (Northern Ireland) Order 2004 (S.I. 2004/702 (N.I. 3))—

a) Article 24 (sale etc of firearms or ammunition without certificate);

b) Article 37(1) (sale etc of firearms or ammunition to person without certificate etc);

c) Article 45(1) and (2) (purchase, sale etc of prohibited weapons);

d) Article 63(8) (sale etc of firearms or ammunition to people who have been in prison etc);

e) Article 66A (supplying imitation firearms to minors).

A2.44 An offence under section 36(1)(c) or (d) of the Violent Crime Reduction Act 2006 (sale etc of realistic imitation firearms).

A2.45 An offence under any of the following provisions of the Air Weapons and Licensing (Scotland) Act 2015 (asp 10)—

a) section 2 (requirement for air weapon certificate);

b) section 24 (restrictions on sale etc of air weapons).

Assisting illegal immigration

A2.46 An offence under section 25 of the Immigration Act 1971 (assisting unlawful immigration etc).

Sexual exploitation

A2.47 An offence under any of the following provisions of the Sexual Offences Act 2003—

a) section 52 (causing or inciting prostitution for gain);

b) section 53 (controlling prostitution for gain).

A2.48 An offence under any of the following provisions of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2))—

a) Article 62 (causing or inciting prostitution for gain);

b) Article 63 (controlling prostitution for gain).

Sexual images

A2.49 An offence under section 63 of the Criminal Justice and Immigration Act 2008 (possession of extreme pornographic images).

A2.50 An offence under section 33 of the Criminal Justice and Courts Act 2015 (disclosing, or threatening to disclose, private sexual photographs and films with intent to cause distress).

A2.51 An offence under section 2 of the Abusive Behaviour and Sexual Harm (Scotland) Act 2016 (asp 22) (disclosing, or threatening to disclose, an intimate photograph or film).

Proceeds of crime

A2.52 An offence under any of the following provisions of the Proceeds of Crime Act 2002—

a) section 327 (concealing etc criminal property);

b) section 328 (arrangements facilitating acquisition etc of criminal property);

c) section 329 (acquisition, use and possession of criminal property).

Fraud

A2.53 An offence under any of the following provisions of the Fraud Act 2006—

a) section 2 (fraud by false representation);

b) section 4 (fraud by abuse of position);

c) section 7 (making or supplying articles for use in frauds);

d) section 9 (participating in fraudulent business carried on by sole trader etc).

A2.54 An offence under section 49(3) of the Criminal Justice and Licensing (Scotland) Act 2010 (articles for use in fraud).

Financial services

A2.55 An offence under any of the following provisions of the Financial Services and Markets Act 2000—

a) section 23 (contravention of prohibition on carrying on regulated activity unless authorised or exempt);

b) section 24 (false claims to be authorised or exempt);

c) section 25 (contravention of restrictions on financial promotion).

A2.56 An offence under any of the following provisions of the Financial Services Act 2012—

a) section 89 (misleading statements);

b) section 90 (misleading impressions).

Inchoate offences

A2.57 An offence of attempting or conspiring to commit an offence specified in paragraphs A2.25 to A2.56.

A2.58 An offence under Part 2 of the Serious Crime Act 2007 (encouraging or assisting) in relation to an offence specified in paragraphs A2.25 to A2.56, or (in Scotland) inciting a person to commit such an offence.

A2.59 An offence of aiding, abetting, counselling or procuring the commission of an offence specified in paragraphs A2.25 to A2.56, or (in Scotland) being involved art and part in the commission of such an offence.

A3. Responding to this call for evidence

How to respond

A3.1 Ofcom would like to receive responses by 5pm on 13 September 2022.

A3.2 You can download a response form from https://www.ofcom.org.uk/consultations-and-statements/category-1/online-safety-call-for-evidence. You can return this by email or post to the address provided in the response form.

A3.3 If your response is a large file, or has supporting charts, tables or other data, please email it to os-cfe@ofcom.org.uk, as an attachment in Microsoft Word format, together with the cover sheet. This email address is for this call for evidence only, and will not be valid after 13 October 2022.

A3.4 Responses may alternatively be posted to the address below, marked with the title of the consultation:

Online Safety Call for Evidence
Ofcom
Riverside House
2A Southwark Bridge Road
London SE1 9HA

A3.5 We welcome responses in formats other than print, for example an audio recording or a British Sign Language video. To respond in BSL:

- send us a recording of you signing your response. Suitable file formats are DVDs, wmv or QuickTime files; or
- upload a video of you signing your response directly to YouTube (or another hosting site) and send us the link.

A3.6 We will publish a transcript of any audio or video responses we receive (unless your response is confidential).

A3.7 We do not need a paper copy of your response as well as an electronic version. We will acknowledge receipt of a response submitted to us by email.

A3.8 You do not have to answer all the questions in the call for evidence if you do not have a view; a short response on just one point is fine. We also welcome joint responses.

A3.9 It would be helpful if your response could include direct answers to the questions asked in Annex 1 of this document. It would also help if you could explain why you hold your views and provide supporting evidence.

A3.10 If you want to discuss the issues and questions raised in this document, please contact the Online Safety team by email to os-cfe@ofcom.org.uk.

Confidentiality

A3.11 In the interests of transparency and good regulatory practice, and because we believe it is important that everyone who is interested in an issue can see other respondents' views, we usually publish responses on the Ofcom website.

A3.12 If you think your response should be kept confidential, please specify which part(s) this applies to, and explain why. Please send any confidential sections as a separate annex. If you want your name, address, other contact details or job title to remain confidential, please provide them only in the cover sheet, so that we don't have to edit your response.

A3.13 If someone asks us to keep part or all of a response confidential, we will treat this request seriously and try to respect it, but sometimes we may need to disclose responses to fulfil certain legal obligations (e.g. where it is proportionate and fair to do so to enable appropriate consultation, or if we are ordered to disclose them).

A3.14 To fulfil our pre-disclosure duty, we may share a copy of your non-confidential response with the relevant government department before we publish it on our website. This is the Department for Business, Energy and Industrial Strategy (BEIS) for postal matters, and the Department for Culture, Media and Sport (DCMS) for all other matters.

A3.15 Please also note that copyright and all other intellectual property in responses will be assumed to be licensed to Ofcom to use. Ofcom's intellectual property rights are explained further in our Terms of Use. Please also see our Privacy Statement.

Updates on Ofcom publications

A3.16 If you wish, you can register to receive mail updates alerting you to new Ofcom publications.

A4. Response coversheet

BASIC DETAILS

Call for evidence title: **Online Safety**

To (Ofcom contact): **Online Safety team, OS-CFE@ofcom.org.uk**

Name of respondent:

Representing (self or organisation/s):

Address (if not received by email):

CONFIDENTIALITY

Please tick below what part of your response you consider is confidential, giving your reasons why

Nothing ☐

Name/contact details/job title ☐

Whole response ☐

Organisation ☐

Part of the response ☐

If there is no separate annex, which parts? _____

_____

If you want part of your response, your name or your organisation not to be published, can Ofcom still publish a reference to the contents of your response (including, for any confidential parts, a general summary that does not disclose the specific information or enable you to be identified)?

_____

DECLARATION

I confirm that the correspondence supplied with this cover sheet is a formal response to Ofcom's Call for Evidence that Ofcom can publish subject to the confidentiality section above. However, in supplying this response, I understand that Ofcom may need to disclose some information marked as confidential where it is proportionate and fair to do so to enable appropriate consultation, or if Ofcom is ordered to disclose them. If I have sent my response by email, Ofcom can disregard any standard e-mail text about not disclosing email contents and attachments.

Ofcom aims to publish responses at regular intervals; if your response is non-confidential (in whole or in part), and you would prefer us to publish your response only once the response period has ended, please tick here. ☐

Name Signed (if hard copy)