

Overview

This document is the first of four major consultations that Ofcom, as appointed regulator of the new Online Safety Act ('the Act'), will publish as part of our work to establish the new regulations over the next 18 months.

It focuses on our proposals for how internet services which enable the sharing of user generated content ('user-to-user' or 'U2U' services) and search services should approach their new duties relating to illegal content. It covers the following areas: the causes and impacts of illegal harms; how services should assess and mitigate the risks of illegal harms; how services can identify illegal content; and our approach to enforcement.

The proposals in this document reflect research we have conducted over the past three years as well as information and evidence gathered through extensive engagement with industry and other experts.

Causes and impacts of illegal harms

We are consulting on our assessment of how the priority illegal harms covered by the Act manifest online, what factors give rise to a risk of these harms and what the impact of the harms is. This analysis underlines the need for services to take action to combat online harms. It shows that a large proportion of the UK population has experienced harm online and that the impact of online harms can, in cases, be extremely severe. Our assessment demonstrates that women, children and groups with protected characteristics are especially likely to be exposed to harm online.

Assessing risk

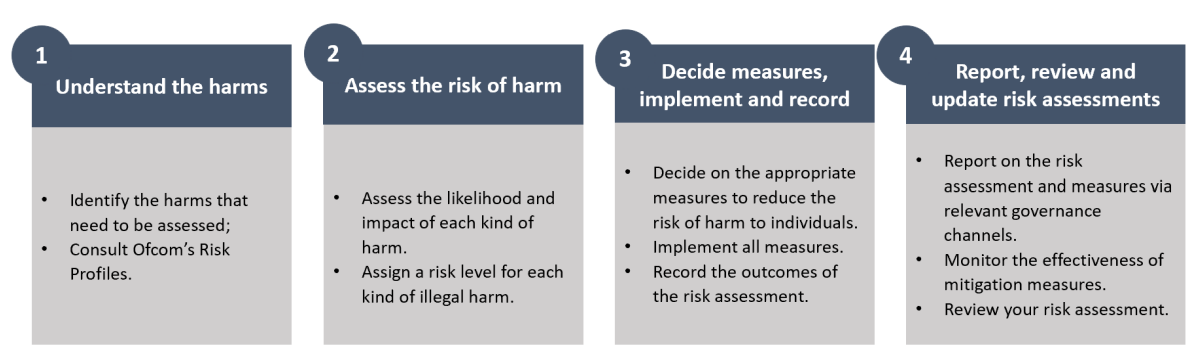
We are consulting on the guidance we propose to give about how regulated stakeholders should assess the risk of illegal harm taking place on their services (the 'Risk Assessment Guidance'), as well as on our proposals about the governance services should put in place to manage risks. This document also consults on guidance about how services should keep adequate records of their risk assessments.

Services should undertake a robust and comprehensive risk assessment. More specifically, we propose that providers should follow the four-step process set out in Figure 1 below when assessing the risk of illegal content on their services.

We are also proposing that services take several steps to ensure that they have strong governance procedures in place to mitigate the risks associated with illegal content. For example, we are proposing that senior governance bodies at large services review the service's risk management activities related to online safety at least annually and that all services identify a named senior executive who is accountable for compliance with the online safety duties.



Figure 1: Proposed four-step process for illegal content risk assessment



Mitigating risk

We are consulting on our illegal harms Codes of Practice, laying out recommended measures that regulated services can take to mitigate the risk of illegal harm. While the Codes are not binding, and so services can choose to take a different approach to meeting their duties, they act as a 'safe harbour'. This means any service that implements the recommendations in the Codes would be deemed to be compliant with its related safety duties.

In the codes we propose to recommend that services put in place a series of measures which, taken together, will help combat all of the priority illegal harms in scope of the Act. These measures include:

- ensuring content moderation teams are appropriately resourced and trained;
- having easy-to-use systems for users to report potentially illegal content and make complaints;
- allowing users to block other users or disable comments;
- conducting tests when they update their algorithms that recommend content to users ('recommender systems') to assess the risk that the changes would increase the dissemination of illegal content; and
- a series of recommended steps to make their terms and conditions clear and accessible.

In addition to these cross-cutting measures, we propose that relevant services should take a series of targeted steps to combat Child Sexual Exploitation and Abuse (CSEA), fraud and terrorism. These targeted steps include:

- Using a technology called 'hash matching' to detect and remove known CSAM (Child Sexual Abuse Material). Consistent with the restrictions in the Act, this proposal does not apply to private communications or end-to-end encrypted communications. We are **not** making any proposals that would involve breaking encryption. However, end-to-end encrypted services are still subject to all the safety duties set out in the Act and will still need to take steps to mitigate risks of CSAM on their services;
- Taking steps to make it harder for perpetrators to groom children online. For example, configuring default settings so that children do not appear in lists suggesting other users connect with them;
- Deploying keyword detection systems to help find and remove posts linked to the sale of stolen credentials. This should help prevent attempts to commit fraud online. We are also recommending large and high-risk services have dedicated fraud reporting channels;

- Where services operate account verification schemes, they should be transparent about the steps they are taking to verify accounts. This is aimed at reducing the risk of users being deceived by posts on fake accounts and should help address fraud and foreign interference in UK processes such as elections; and
- Blocking accounts run by banned terrorist organisations.

Our analysis suggests that, taken together, these measures will be an effective and proportionate means of tackling the priority illegal harms in scope of the Act. Some of the proposals set out in this consultation would apply to all services. However, we are targeting many of the more onerous proposals only at services which are large and/or high risk. We provide a full list of all the measures we are proposing and who they would apply to in our “consultation at a glance” which is published alongside this consultation.

Our proposals will make a contribution to combatting violence against women and girls online, including by setting out how services should assess their risk of coercive and controlling behaviour, stalking, harassment and threats, and intimate image abuse. Our proposals for cross-cutting measures and some harm specific measures will begin to mitigate these risks. However, we recognise that more work is needed in this area, and we will be publishing draft guidance on how services can combat violence against women and girls in early 2025.

Identifying illegal content

A new legal requirement of the Act is for all services to swiftly take down specific illegal content when they become aware of it. Today we are consulting on our Illegal Content Judgements Guidance (‘ICJG’). This will provide guidance to services on how they can identify whether a piece of content is likely to be illegal.

Enforcement

The Act gives us significant enforcement powers in the event of non-compliance, including the ability to issue fines of up to £18m or up to 10% of the service’s qualifying worldwide revenue (whichever is greater) and to apply for a court order requiring an internet service provider to withdraw access to the service to prevent a significant risk of harm to UK users as a result of its failure.

Next steps

We welcome comments from stakeholders in response to the proposals set out in this document, including any further evidence and supporting information to inform our final decisions. The consultation closes on 23 February 2024, and we invite stakeholders to provide responses by this date.

This is also a statutory consultation under the Act, and we will engage with specific statutory consultees through the consultation process. For further details of the process for making Codes and guidance, please refer to our Legal Framework (Annex 12).

To improve the accessibility of our consultation, we are also publishing a “chapter summaries” document. It explains what each chapter is about, what proposals we are making, why we are making them and what input we would appreciate from stakeholders.

Chapter 1 of this document (Introduction) provides further information on how services can use and navigate the document. Annexes 1-4 explain our approach to consultations and how stakeholders can respond to this consultation.

Once the consultation closes, we will consider stakeholder responses, review our proposals, and publish a statement setting out our final decisions in relation to our consultation proposals, including final versions of our guidance and Codes.

We currently plan to publish our Statement in Winter 2024. Following our Statement, services will have three months to conduct their risk assessment. The Codes will also be subject to a parliamentary approval process. We expect this process to conclude by the end of 2024 at which point the Codes will come into force.

This is one of a number of major consultations we will publish as over the next 12-18 months as we put the new online safety regulations in place. We provide more detail on our roadmap for implementing the Online Safety Act at the following link: [Creating a safer life online for people in the UK](#).