Ofcom

Protecting people from illegal harms online

Summary of each chapter

Statement

Published 16 December 2024

Contents

Section

Summary of each chapter

We want to make navigating and understanding the decisions set out in this Statement as easy as possible. This document sets out a high-level summary of each chapter of our Illegal Harms Statement to help stakeholders navigate and engage with our document. Our decisions and the underlying rationale are set out in full in the relevant Statement chapters.

This is the first of several decisions we will be publishing under the Online Safety Act. Our full regulatory roadmap and strategy is available on our website.

Index

Context Volume – Background to our Statement on Illegal Harms

Volume 1 – Governance and Risk Management

Volume 2 – Service design and user choice

Volume 3 – Transparency, trust and other guidance

Regulatory documents

Annexes

Context Volume: Background to our Statement on Illegal Harms

Introduction, our duties, and navigating the Statement

This chapter provides a high-level introduction to our Statement. We detail our duties and functions, the scope of the Statement, set out next steps and explain to stakeholders how to navigate our Statement. To help improve the accessibility of our document we have included suggestions for which parts of our Statement different types of stakeholder might find most useful.

Overview of Illegal Harms

This chapter summarises the main duties the Act creates. The Act gives online service providers a range of duties. The main ones relating to illegal content are for providers to assess the risk of harm arising from illegal content or (for a user-to-user (U2U) service) illegal activity on their service, and take proportionate steps to manage and mitigate those risks.

The Act lists over 130 'priority offences'. U2U service providers will need to take steps to prevent users encountering content amounting to one of these offences. Search service providers will need to minimise the risks of users encountering content that amounts to one of these offences. U2U and search service providers only need to take actions that are proportionate. Where appropriate, we have grouped these priority offences into broad groups such as terrorism, hate offences, child sexual exploitation and abuse, and fraud and financial offences.

Providers also have duties in relation to non-priority illegal content of which they are aware.

Overview of regulated services

This chapter explains which types of services are in scope of the Act. The Act places new legal requirements on providers of the following three types of internet service: user to user (U2U) services; search services; and pornographic content.

The duties in the Act apply to providers of services with links to the UK regardless of where in the world they are based. The number of online services subject to regulation could total more than 100,000 and range from some of the largest tech companies in the world to very small services. Providers in scope of the Act come from a diverse range of sectors, including, but not limited to, social media, dating, gaming, search and adult services.

The online space is one of rapid innovation. We know that new types of U2U and search services will emerge, a good example being the on-going developments in generative AI.[1] This has implications for our work. These include the importance of varying expectations depending on the type of service we are dealing with. We will not expect the same for a small low-risk service as we do for the largest

[1] See Ofcom, 2024, Open letter to UK online service providers regarding Generative AI and chatbots. https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/open-letter-to-uk-online-service-providers-regarding-generative-ai-and-chatbots/. [accessed 10 November 2024].

or riskiest services. We will also need to adapt our approach and expectations over time to reflect the emergence of new technologies and types of U2U or search services.

Our approach to developing Codes measures

The Online Safety Act places a new duty of care on online service providers to protect their users from harm and to place safety by design at the heart of their platforms. Ofcom's Codes of Practice (Codes) are one of the main tools we have to get providers to make improvements in these areas.

The measures in our Codes will result in change in three key areas:

- stronger safety governance and risk management;
- specific changes to design to improve safety; and
- increased choice for users so they can control their online experiences.

This chapter gives an overview of how we have approached developing this first edition of the Codes. It also addresses a number of key cross-cutting issues where we have had feedback to our consultations, and signals where further information on these may be found in the documents.

1. We have prioritised introducing the Codes as soon as possible to protect users

As we explain in our Register of Risk, illegal online content is widespread and causes significant harm. Given the scale and urgency of the challenge, it is important for us to make the Codes enforceable as soon as possible so that we can drive compliance and act against breaches at the earliest opportunity. Parliament set Ofcom a deadline of 18 months after the Act achieved Royal Assent for the first two phases of work to be completed, and we launched our first consultation, on illegal harms, less than two weeks from the Act passing.

As part of our Consultation we received suggestions for several new measures to the Codes. Where we were able to incorporate these because we already have sufficient evidence, we have done so. For measures where we need to do more assessment and gather further evidence, we have decided to launch a further consultation in the Spring of 2025. We have removed some measures from smaller low risk services, where the evidence we received suggested they were not proportionate.

Online safety is extremely dynamic, with harms, technology and best practice in user safety evolving quickly. This iterative approach to the Codes – banking what we already have, and building further over time – is central to our approach now and in the future.

2. We have taken a risk-based approach to our Codes

Our key focus is the extent to which measures can reduce risks to users. The Act requires us to ensure measures are proportionate, and we recognise that the size, capacity, and risks of services differ widely. We therefore do not take a one-size-fits-all approach. Instead, we have set out what types of service we think should use specific safety measures to comply with their duties, with the most extensive expectations on the riskiest services.

The size of a service and its user base is one indicator of risk. But there are some services which are inherently risky even when their reach is small. We are clear in our Codes that all providers of high-risk services must take robust steps to protect people, whether they are small or large.

Our Risk Assessment Guidance explains that providers should assess whether they are high, medium or low risk for each harm. The measures the Codes recommend vary depending on the outcome of this assessment. In general, the Codes recommend more onerous measures for high-risk services than for low-risk services. Consistent with this risk-based approach, we have chosen to apply a number of the most robust measures in our Codes to high-risk services even where they are small. This includes the measures on grooming and the measures relating to the detection of CSAM.

At the same time, it is reasonable to expect more of the largest providers than of smaller providers. We have in some cases imposed significant obligations on large low-risk providers because of their greater resources and reach. However, we have sought to minimise the regulatory burden on small low-risk services.

3. We have included a mix of cross-cutting and harms-specific measures in Codes

Many of the design changes will address multiple risks. In other cases, specific changes will address particular harms. Our Codes therefore include:

- A series of measures focused on governance - taken together with our risk assessment guidance these are intended to embed the building blocks of good risk management in providers;
- A series of cross-cutting measures, for instance content moderation measures, which look to ensure that a broad range of service providers take important actions that will address all illegal harms; and
- Some harm-specific measures which aim to address in a targeted way some of the most serious harms online, including child sexual abuse material (CSAM), grooming, terror, and fraud.

The Codes represent a package of measures which we recommend that service providers should take to comply with their duties under the Act. The Act stipulates that the Codes are a ‘safe harbour’. This means that Ofcom must treat providers who choose to implement all applicable measures as complying with their relevant duties under the Act. The measures in our Codes are required by the Act to be clear and sufficiently detailed for providers to understand what they entail in practice.

Volume 1: Governance and Risk Management

Introduction to the volume

In this volume, we set out and explain the decisions related to **governance** and **risk assessment**. Ensuring service providers do a good risk assessment and put in place robust governance to manage the risk of online harms is a strategic priority for us. Putting in place strong governance and risk assessment procedures is a pre-requisite for protecting people from illegal content online. Only by monitoring the risk of harms on a given service, how they occur, what features and functionalities enable them, and how they impact users, can providers effectively detect and manage these risks.

Ofcom's Register of Risks and Risk Profiles

Our assessment of the causes and impacts of online harms

The **Illegal Harms Register of Risks ('Register of Risks')** is our assessment of the causes and impacts of illegal online harms based on the evidence that we have gathered over the past four years. The Register of Risks presents our full, sector-wide risk assessment of where and how illegal harms manifest online and the characteristics of services that are relevant to the risks of harm. It forms part of our duty under the Act to assess the factors that can cause a risk of harm to individuals across regulated services.

Service providers are required under the Online Safety Act ('the Act') to carry out risk assessments. Ensuring providers to do so to a high standard is one of our strategic objectives and the Register of Risks is an important resource for achieving this. It is intended to act as a central resource for service providers when they are conducting their risk assessments, providing a clear understanding of how harms manifest online and how specific characteristics of services and users play a role. We have distilled the key findings of the Register of Risks into a series of tables in the Risk Assessment Guidance summarising which factors are associated with a risk of which types of illegal content. We refer to these tables as **'Risk Profiles'.**

Our **Register of Risks** and **Risk Profiles** can be found here.

Risk Assessment Guidance

What is this chapter about?

The Act requires us to produce guidance for service providers to help them meet their illegal content risk assessment duties – our Risk Assessment Guidance for Service Providers. This chapter sets out the decisions we have made regarding that guidance.

What have we decided?

Our Risk Assessment Guidance sets out a four-step risk assessment process for service providers to follow. This involves providers: (i) understanding the kinds of illegal content they need to consider in their risk assessment; (ii) assessing the likelihood and impact of encountering these types of content on their service; (iii) deciding what measures to take to mitigate these risks; and (iv) reporting on, reviewing and updating the risk assessment.

Our guidance also explains what amounts to a 'suitable and sufficient' risk assessment, and how a service provider should record their risk assessment. It is a requirement in the Act that services keep their risk assessments up to date, and we have set clear expectations regarding how services may meet this duty. We have also provided guidance about what amounts to a significant change as a trigger to carry out a new risk assessment.

Why are we making these decisions?

The four-step methodology in our Risk Assessment Guidance will support service providers to meet their legal obligations to carry out a risk assessment which is 'suitable and sufficient'.

It is based on best practice approaches to risk assessments across a range of industries with a mature approach to risk management and incorporates all the elements of the illegal harm risk assessment duty. The evidence we have seen suggests that doing a good risk assessment is critical to achieving good safety outcomes.

Record-keeping and review guidance

What is this chapter about?

Providers of regulated user-to-user (U2U) and search services have duties to make and keep written records of their risk assessments and the measures they take to comply with several duties set out in the Act. Service providers also have a duty to regularly review their compliance with relevant duties specified in the Act. This chapter explains the decisions we have taken about how they can fulfil these duties.

What have we decided?

We have made the following decisions for all providers of U2U and search services:

- We have adopted the guidance set out in the Record-Keeping and Review Guidance;
- It includes guidance that written records should be retained in accordance with the provider's record retention policies, or a minimum of **three years**, whichever is the longer; and
- We are **not** exercising our power to exempt services from the record-keeping or review duties.

Why are we making these decisions?

Our guidance helps service providers to comply with their record-keeping and review duties by explaining the requirements and providing guidance on best practice. Following our guidance should enable service providers to track and evidence their compliance with the relevant duties and help with reviewing risks and monitoring improvements over time.

We have decided a minimum three-year record retention period is appropriate. This aligns with similar requirements in the EU's Digital Services Act (DSA). We consider that this period is sufficient for ensuring the availability of records if retrospective problems are identified and should allow providers to show how they have responded to the evolution of risks over time.

We have confirmed our decision not to exercise our power to exempt certain types of services from any or all the record-keeping and review duties. This is because we consider there is not currently a sufficiently strong evidence base to justify any exemption for any given description of service. We are satisfied that compliance with the record-keeping and review duties is not unduly onerous. Further, it is good practice for all providers to keep records and regularly review their compliance with their safety duties, particularly in the early days of the new regime, when providers' understanding of their obligations is likely to be evolving.

Governance and accountability

What is this chapter about?

In our Codes we recommend that service providers take a number of steps to ensure that they have appropriate governance arrangements in place for tracking and managing online safety risks. In this chapter we explain what these recommendations are and why we have made them.

What have we decided?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICU A1/ ICS A1	The provider's **most senior governance body** in relation to the service should carry out and record an **annual review of risk management activities** having to do with illegal harm, as it relates to individuals in the UK, including as to risk that is remaining after the implementation of appropriate Codes of Practice measures. The review should include how developing risks are being monitored and managed.	• Providers of large U2U services. • Providers of large general search services.
ICU A2/ ICS A2	Providers should **name an individual accountable to the most senior governance body** for compliance with the illegal content safety duties and the reporting and complaints duties.	Providers of all services.
ICU A3/ ICS A3	Providers should have **written statements of responsibilities for senior managers** who make decisions about the management of risks having to do with illegal harm in relation to individuals in the UK.	• Providers of large U2U services. • Providers of large general search services. • Providers of multi-risk services.
ICU A4/ ICS A4	Providers should have an **internal monitoring and assurance function** to provide independent assurance that measures taken to mitigate and manage the risks of harm to individuals identified in the risk assessment are effective on an ongoing basis. This function should report to, and its findings should be considered by, either: a) the body that is responsible for overall governance and strategic direction of a service; or	• Providers of large multi-risk U2U services • Providers of large multi-risk search services.

	b) an audit committee. This independent assurance may be provided by an existing internal audit function	
ICU A5/ ICS A5	Providers should **track evidence of new kinds of illegal content** on the service, **and unusual increases** in particular kinds of illegal content or illegal content proxy, or (U2U services only) equivalent changes in the use of the service for the commission or facilitation of priority offences	• Providers of large U2U services. • Providers of large general search services. • Providers of multi-risk services.
ICU A6/ ICS A6	Providers should have a **Code of Conduct that sets standards and expectations** for individuals working for the provider around protecting United Kingdom users from risks of illegal harm.	• Providers of large U2U services. • Providers of large general search services. • Providers of multi-risk services.
ICU A7/ ICS A7	Providers should secure that **individuals** working for the provider who are **involved in the design and operational management of the service are trained in the service's approach to compliance** with the illegal content safety duties and the reporting and complaints duties, sufficiently to give effect to them	• Providers of large U2U services. • Providers of large general search services. • Providers of multi-risk services.

Why are we making these decisions?

The evidence we have assessed shows that where providers put in place robust governance arrangements they are likely to be able to manage online safety risks better. The governance measures we are recommending reflect best practice in a range of sectors with a mature approach to governance and risk management. We expect them to embed principles like accountability, oversight, independence, transparency and clarity of purpose into providers' operations, leading to well-functioning governance and organisational design processes. This should lead providers to better understand and anticipate risks, increasing the likelihood that risks to users will be prioritised appropriately, and factored into user safety decisions. We consider that this will result in people being better protected from illegal content online.

Volume 2: Service Design and User Choice

Introduction to the volume

Our Codes work together to create an overall safer experience for users, and cover three broad areas.[2]

In this volume, we set out and explain our decisions that contribute to ensuring **online services are designed and operated with safety in mind** and that there is improved **choice for users to enable more control over their online experiences.** Each measure set out in this volume contributes to one of these two strategic objectives.

Content moderation

What is this chapter about?

Content moderation is when a service provider reviews content to decide whether it is permitted on its service and takes appropriate action to handle it. Content moderation is used by providers to address a wide variety of illegal content as well as legal content that does not comply with a service's terms of service (which we call 'illegal content proxy'). This chapter sets out the content moderation measures we are recommending, why we are recommending them, and to which user-to-user (U2U) services they should apply.

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICU C1	Providers should have systems and processes designed to **review and assess content** the provider has reason to suspect may be illegal content (part of its 'content moderation function').	Providers of U2U services.
ICU C2	Providers should have systems and processes designed to **swiftly take down illegal content and/or illegal content proxy** of which they are aware (part of their 'content moderation function'), unless it is currently not technically feasible for them to achieve this outcome.	• Providers of U2U services

[2] The three broad areas are: (1) stronger safety governance in online services (2) online services are designed and operated with safety in mind; and (3) greater choice for users so they can have more meaningful control over their online experiences. For more information on these, see our 'approach to Codes' chapter.

ICU C3	Providers should **set and record internal content policies.**	- Providers of large U2U services. - Providers of multi-risk U2U services.
ICU C4	Providers should **set and record performance targets** for their content moderation function.	- Providers of large U2U services. - Providers of multi-risk U2U services.
ICU C5	Providers should prepare and apply a policy in respect of the **prioritisation of content for review.**	- Providers of large U2U services. - Providers of multi-risk U2U services.
ICU C6	Providers should **resource their content moderation function**, so as to give effect to measure ICU C3 and measure ICU C4.	- Providers of large U2U services. - Providers of multi-risk U2U services.
ICU C7	Providers should ensure **individuals working in moderation** (non-volunteers) **receive training and materials** that enable them to **fulfil their role** in moderating content, including in relation to measure ICU C1, measure ICU C2 and measure ICU C3.	- Providers of large U2U services. - Providers of multi-risk U2U services
ICU C8	Providers should ensure **volunteers** in their content moderation functions **have access to materials** that enable them to **fulfil their role** in moderating content, including in relation to measure ICU C1, measure ICU C2 and measure ICU C3	- Providers of large U2U services. - Providers of multi-risk U2U services.

We note that we have made changes to some of our measures to be clearer about how they apply when content take down is not currently technically feasible.

Why have we made these decisions?

Effective content moderation systems are able to identify, and prioritise the swift removal of illegal content. Content moderation therefore plays a hugely important role in combatting illegal content. Providers with ineffective content moderation functions may face increased risk of harm on their services. Our analysis suggests that harm to users will be reduced where providers set content policies, resource and train their content moderation teams appropriately and take into account the likely severity of content and the risk the content will be encountered by a high number of UK users when deciding what potentially harmful content to prioritise for review. Given the diverse range of providers in scope of the new regulations, a one-size-fits-all approach to content moderation would not be appropriate. Instead of making very specific and prescriptive recommendations about content moderation, we have therefore decided to make a relatively high-level set of recommendations which would allow services considerable flexibility about how to set up their content moderation teams. We have focussed the most rigorous proposals in this area on services which are large or multi-risk. This will help ensure that the impact of the measures is proportionate.

Similarly, the flexibility built into our proposals will make it easier for providers to carry them out in a way which is cost-effective and proportionate for them.

Search moderation

What is this chapter about?

Search moderation is used by providers to review search content and where relevant, take action to minimise the risk of a wide variety of illegal content from being presented to users in search results, as well as legal content that is covered by their publicly available statement (an illegal content proxy). This chapter sets out the search moderation measures we are recommending, why we are recommending them, and to which search services they should apply.

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICS C1	Providers should have systems and processes designed to review, assess and where relevant take **appropriate action** in relation to **illegal content or illegal content proxy** of which they are aware (a 'search moderation function').	Providers of search services.
ICS C2	Providers should **set and record internal content policies**.	- Providers of large general search services. - Providers of multi-risk search services.
ICS C3	Providers should **set and record performance targets** for their search moderation function.	- Providers of large general search services. - Providers of multi-risk search services.
ICS C4	Providers should prepare and apply a policy in respect of the **prioritisation of search content for review.**	- Providers of large general search services. - Providers of multi-risk search services.
ICS C5	Providers should **resource their search moderation function**, so as to give effect to measure ICS C2 and measure ICS C3.	- Providers of large general search services. - Providers of multi-risk search services.
ICS C6	Providers should ensure **people working in search moderation** (non-volunteers) **receive training and materials** that enable them to **fulfil their role** in moderating search content, including in relation to measure ICS C1 and measure ICS C2	- Providers of large general search services - Providers of multi-risk search services.

Why have we made these decisions?

Effective search moderation systems are able to identify and enable providers to make timely and accurate decisions about search content that is suspected to be illegal, and take appropriate action to protect users. Providers with ineffective search moderation functions may face increased risk of harm on their services. Our analysis suggests that harm to users will be reduced where providers set content policies, resource and train their search moderation teams appropriately and take into account the likely severity of content and the risk that it will be encountered by a high number of UK users when deciding what content to prioritise for review. We do not think a one-size-fits-all approach to search moderation would be appropriate. Instead of making very specific and prescriptive recommendations about search moderation, we have therefore decided to make a relatively high-level set of recommendations which would allow providers considerable flexibility about how to set up their search moderation teams. We have focussed the most rigorous proposals in this area on providers which are large or multi-risk. This will help ensure that the impact of the measures is proportionate. Similarly, the flexibility built into our proposals will make it easier for providers to carry them out in a way which is cost-effective and proportionate for them.

Automated content moderation

What is this chapter about?

Services use automated tools, often in tandem with human oversight, to make content moderation processes more effective at identifying and removing illegal content or content in breach of their terms of service. As these tools allow services to identify large volumes of harmful content more quickly, they are critical to many services' attempts to reduce harm. This chapter sets out the automated content moderation measures we are recommending, why we are recommending them, and to which user-user (U2U) services they should apply.

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICU C9	Providers should ensure that **hash-matching technology is used** to detect and remove child sexual abuse material (CSAM). This involves analysing images and videos communicated publicly on the service and comparing a digital fingerprint of that content to digital fingerprints of previously identified CSAM.	• Providers of large U2U services which are at medium or high risk of image-based CSAM. • Providers of U2U services which are at high risk of image-based CSAM and have more than 700,000 monthly active UK users. • Providers of U2U services which are at high risk of image-based CSAM and are file-storage and file-sharing services.
ICU C10	Providers should **detect and remove content** communicated publicly on the service which matches a URL on a list of	• Providers of large U2U services which are at medium or high risk of CSAM URLs.

	URLs previously identified as hosting CSAM.	• Providers of U2U services which have more than 700,000 monthly active UK users and are at high risk of CSAM URLs.

We have also decided **not to recommend** a measure for providers to use standard keyword detection to identify content that is likely to amount to a priority offence concerning articles for use in frauds.

Why have we made these decisions?

The circulation of CSAM online is increasing rapidly. Child sexual abuse and the circulation of CSAM online causes significant harm, and the ongoing circulation of this imagery can re-traumatise victims and survivors of abuse. Hash matching and URL detection can be useful and effective tools for combatting the circulation of CSAM. While the decisions we are taking today will impose significant costs on some services, we consider these costs are justified given the very serious nature of the harm they address. To ensure that the costs are proportionate, we propose targeting these measures at services where there is a medium or high risk of image-based CSAM or CSAM URLs.

In principle, we consider that, even where they are very small, it would be justified to recommend that services which are high-risk deploy these technologies. However, we have decided to set user-number thresholds below which services will not be in scope of the measure. This is because to implement hash matching and URL detection services will need access to third party databases with records of known CSAM images and lists of URLs associated with CSAM. There are only a limited number of providers of these databases, and they only have capacity to serve a finite number of clients. Setting the user-number thresholds we have should ensure that the database providers have capacity to serve all services in scope of the measure. Should the capacity of database providers expand over time, we will look to review whether the proposed threshold remains appropriate. The evidence we have assessed shows that file sharing services play a particularly significant role in the sharing of CSAM. Therefore, we have decided that all high risk file sharing services should be in scope of our hash matching measure regardless of size.

In the November 2023 Consultation, we proposed recommending the use of standard keyword detection to identify content likely to amount to a priority offence concerning articles for use in frauds to providers of large user-to-user services which are at medium or high risk of fraud. We acknowledged that there is a range of more sophisticated automated tools that service providers may use to detect harmful content (e.g. natural language processing or machine learning). However, we did not have sufficient evidence on the costs and efficacy of these alternative tools to justify recommending their use. Having assessed relevant stakeholder feedback to the consultation, we have decided not to proceed with the measure at this stage. We are instead focusing our efforts on exploring a broader and more flexible measure regarding the use of automated content moderation technologies (including AI), on which we intend to consult in Spring 2025.

Automated search moderation

What is this chapter about?

Search services use automated moderation tools to identify large volumes of harmful content more quickly, and these are therefore critical to many services' attempts to reduce harm. This chapter sets out our recommendation of a measure for services to take steps to remove URLs identified as hosting child sexual abuse material ('CSAM') from search results, why we are recommending it, and to which search services it should apply.

What decisions have we made?

We are recommending the following measure:

Number in our Codes	Recommended measure	Who should implement this
ICS C7	Providers should take action to ensure that users do not encounter, in or via search results, search content present at or sourced from URLs on a list of URLs previously identified as hosting CSAM.	Providers of general search services.

Why have we made this decision?

The circulation of CSAM online is increasing rapidly. The evidence presented in the Register of Risk shows that perpetrators often use search services to access CSAM. As we explained above, child sexual abuse and the circulation of CSAM online causes significant and lifelong harm and the ongoing circulation of this imagery can re-traumatise victims and survivors of sexual abuse. URL detection is an effective and well-established tool for combatting the circulation of CSAM on search services. The largest search services are already using it to address CSAM. Whilst the use of URL detection imposes some costs we consider these are justified given the severity of the harm they address and the significant benefits of limiting exposure to known CSAM.

Reporting and complaints

What is this chapter about?

This chapter sets out the measures we are recommending in relation to reporting and complaints, why we are recommending them, and to which user-to-user (U2U) and search services they should apply.[3]

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICU D1 / ICS D1	All providers of U2U and search services should have **complaints systems and processes** that enable prospective complainants (UK users, affected persons and (for search services) interested persons) to make relevant complaints in a way that will secure appropriate action in relation to them.	Providers of all services.
ICU D2 / ICS D2	Providers should design and operate complaints procedures so that they are **easy to find, access, and use**.	Providers of all services.

[3] We note, that in this Statement, we are also including decision on two further measures proposed in our May 2024 Consultation on Protecting Children from Harms Online ('May 2024 Consultation').

ICU D3	Providers should ensure their reporting tool for specific content enables to access information that **informs complainants** if they will share information about a complaint **with another user**, and what information will be shared. This includes information relating to the original complaint and complainant if the other user subsequently appeals.	• Providers of large U2U services which are likely to be accessed by children. • Providers of U2U services at medium or high risk of any kind of illegal harm, which are likely to be accessed by children.
ICU D4 / ICS D3	Providers should **acknowledge receipt of a complaint** and provide complainants with an **indicative timeframe** for when a complaint might be resolved.	• Providers of large U2U services. • Providers of large general search services. • Providers of services at medium or high risk of any kind of illegal harm.
ICU D5 / ICS D4	Providers should inform a complainant about the **possible outcomes of a complaint**, including whether the service will update the complainant on the outcome.	• Providers of large U2U services likely to be accessed by children. • Providers of large general search services likely to be accessed by children. • Providers of services at medium or high risk of any kind of illegal harm, which are likely to be accessed by children.
ICU D6 / ICS D5	Providers should give complainants the **option to opt out of receiving non-ephemeral communications** about a complaint from the service provider.	• Providers of large U2U services. • Providers of large general search services. • Providers of services at medium or high risk of any kind of illegal harm.
ICU D7 / ICS D6	Providers should handle complaints about suspected illegal content in accordance with their **content prioritisation processes** and content moderation functions, **or promptly** if the recommendations about prioritisation and targets do not apply to that provider.	Providers of all services.

ICU D8 / ICS D7	Providers should determine **complaints which are appeals** and monitor against performance targets for time taken to determine and for accuracy. The provider should have a prioritisation policy for appeals.	• Providers of large U2U services. • Providers of large general search services. • Providers of multi-risk services.
ICU D9 / ICS D8	Providers should determine **complaints which are appeals** promptly.	• Providers of U2U services that are neither large nor multi-risk. • Providers of search services that are neither large general search services nor multi-risk.
ICU D10 / ICS D9	For **complaints which are appeals**, if a provider reverses a decision that content was illegal content, it should: **reverse the action** taken (so far as appropriate and possible); **adjust any relevant moderation guidance** if appropriate; and take appropriate steps to secure that the use of automated moderation technology does not result in the **same content being taken down** / search content no longer appearing in search results or being given a lower priority in the ranking of search results, again.	Providers of all services.
ICU D11	For complaints about the use of proactive technology in breach of terms of service, providers should **inform complainants of the action the provider might take and the complainant's right to bring proceedings.**	Providers of U2U services.
ICS D10	For complaints about the use of proactive technology in breach of policies, providers should **inform complainants of the action the provider might take**.	Providers of search services.
ICU D12 / ICS D11	Providers should **nominate an individual or team** to ensure that **all other relevant complaints** are directed to the appropriate individual or team for processing.	Providers of all services.

ICU D13 / ICS D12	A provider may disregard a relevant complaint (that is not an appeal) if it has a policy that sets out the information and attributes that would indicate that a relevant complaint is manifestly unfounded. It must make decisions in accordance with this policy and review the application of the policy annually.	Providers of all services.
ICU D14 / ICS D13	Providers should establish and maintain a **dedicated reporting channel for trusted flaggers** (including at least the specified list of public bodies) to use to report, at a minimum, fraud.	• Providers of large U2U services that are at medium or high risk of fraud. • Providers of large general search services that are at medium or high risk of fraud.

Why have we made these decisions?

Complaints are important mechanisms for providers to become aware of harmful content. The decisions we have taken today will help ensure that reporting and complaints functions operate effectively. We consider this will make providers better able to identify and remove illegal content, thereby reducing harms to users. We have included a provision in our Codes allowing providers to disregard complaints which are manifestly unfounded ('spam' complaints). This exception means providers can focus their resources on taking appropriate action against legitimate complaints, and do not need to review complaints that are part of a co-ordinated attack or have been submitted by malicious actors.

Dedicated reporting channels provide an easy way for expert 'trusted flaggers' to report problems to providers. These can play a valuable role in improving detection of illegal content, therefore reducing harm to users. In principle dedicated reporting channels could be used to address a wide range of harms. In this first version of our Codes we have focused our recommendations regarding dedicated reporting channels for trusted flaggers on fraud. That is because we have received specific evidence indicating that organisations with expertise in fraud often find it difficult to report known scams to providers and that the creation of a dedicated reporting channel would play an important role in addressing this problem.

Recommender systems

What is this chapter about?

Content recommender systems are algorithmic systems used to curate personalised feeds of user-generated content and to aid the organic discovery of such content. The evidence we have suggests that these systems can play a role in increasing the risk of users encountering certain types of illegal content. Through the responsible monitoring of these systems, service providers can manage some of this risk.

This chapter discusses steps service providers can take to help them better understand the risks their content recommender systems pose. The chapter presents a measure designed to give service providers a methodical way of monitoring the risk of ongoing design adjustments to their content recommender systems by collecting safety metrics when they carry out on-platform testing.

What decisions have we made?

We are recommending the following measure:

Number in our Codes	Recommended measure	Who should implement this
ICU E1	Providers should, when carrying out on-platform testing of content recommender systems, **collect additional safety metrics when making design adjustments**.	Providers of U2U services that: • carry out on-platform testing of their content recommender systems; and • are at medium or high risk of two or more specified kinds of illegal harm.

Why have we made these decisions?

Many service providers carry out on-platform tests when they are making adjustments to the design of their recommender systems. Typically, these focus on the impact that design adjustments have on user engagement with the service. We are recommending that service providers incorporate safety metrics into their on-platform tests. This will help providers better understand whether a design adjustment to their recommender systems might contribute to illegal content risk and, if so, how and why. This will enable them to make more informed choices about the design of their content recommender systems, and be better placed to manage risks associated with these algorithms. This should help reduce the amount of illegal content disseminated by content recommender systems.

U2U settings, functionalities and user support

What is this chapter about?

This chapter describes a series of measures we are recommending to tackle online grooming, why we are recommending them, and to which providers of user-user (U2U) services they should apply.

What decisions have we made?

We are recommending the following measures (The measures detailed below apply in relation to users aged under 18):

Number in our Codes	Recommended measure	Who should implement this
ICU F1	Providers should implement **safety defaults** for child user accounts, which target particular **functionalities**[4] and that restrict the visibility and engagement between **child users** and other **users**.	Providers of U2U services which have an **existing means to determine the age or age range of a particular user** of the service and have relevant functionalities, if they are: • at high risk of grooming, or; • are large services and at medium risk of grooming.

[4] Measure ICU F1 will target functionalities which have been identified as risk factors for grooming harm on U2U services. These include network expansion prompts, direct messaging, connection lists and automated information displays.

ICU F2	Providers should give **child users** relevant **supportive information** at critical points[5] in a child user's journey to allow child users to make more **informed choices**. Providers should ensure the messaging is prominently displayed, and is clear and easy for children to understand.	Providers of U2U services which have an **existing means to determine the age or age range of a particular user** of the service and have relevant functionalities, if they are: • at high risk of grooming, or; • are large services and at medium risk of grooming.

Why have we made these decisions?

Child sexual exploitation and abuse (CSEA) is a serious crime which can have a severe and lifelong impact on children and communities. Grooming for the purpose of sexual abuse involves a perpetrator establishing communications with a child to enable their abuse and exploitation both online and offline.[6] In online spaces, it can often lead to the exchange of self-generated indecent images and financially motivated sexual extortion, which would also mean that a range of other specific child sexual abuse material (CSAM) offences have been committed.

Strategies that perpetrators deploy to groom children frequently include: sending scattergun 'friend' requests to large volumes of children; infiltrating the online friendship groups of children they have succeeded in connecting with; and sending unsolicited direct messages to children they are not connected with. Taken together, the measures described in this chapter will make it more difficult for perpetrators to adopt these strategies and empower child users to make informed choices about their online interactions. This would therefore make grooming more difficult, thereby combating CSEA.

Search settings, functionalities, and user support

What is this chapter about?

Search services can act as a gateway to illegal content that exists online. In particular, search functionalities and features that have been designed to optimise search results can inadvertently make it easier for users to encounter illegal content in those results.

There are steps that service providers can take to reduce these risks and make it less likely that users encounter illegal content through their service. In addition, providers can offer supportive information to users to allow them greater choice and control over their experiences on search services. This will help to reduce the likelihood that users seek out illegal content and mitigate the risk of harm they may face as a result.

This chapter set out and explains the rationale for a series of measures we are recommending providers of search services take to protect people from illegal content, and to which search services they should apply.

[5] As part of measure ICU F2, we specify four critical points in a child user's online journey, where the child user may take a decision which impacts their engagement with a user or users. See 'Measure on support for child users' in volume 2: chapter 8: U2U settings, functionalities, and user support, for more details.

[6] 'What is Online Child Sexual Abuse and Exploitation?', Child Exploitation and Online Protection (CEOP). [accessed 31 October 2024].

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICS F1	Providers should offer users a means to **easily report predictive search suggestions** which they believe can **direct users towards priority illegal content**. If a **clear and material risk** is identified, the provider should take appropriate steps to ensure that the reported predictive search suggestion **is not recommended to any users.**	Providers of large general search services that use a predictive search functionality
ICS F2	Providers should detect and **provide warnings and support resources** in response to **search requests** where the wording clearly indicates that the user may be seeking to encounter **child sexual abuse material** (CSAM).	Providers of large general search services
ICS F3	Providers should provide **crisis prevention information** in response to **search requests** that contain **general queries regarding suicide** and queries seeking **specific, practical, or instructive information regarding suicide methods.**	Providers of large general search services

Why are we making these decisions?

Our first measure (ICS F1) will reduce barriers to reporting predictive search suggestions that can direct users to encounter priority illegal content. This will raise providers' awareness of problematic search suggestions and enable them to ensure that they are no longer recommended to users. In doing so, this measure will reduce the likelihood of users being prompted to run those searches, and encountering illegal content as a result.

Our second measure (ICS F2), on CSAM warning messages, is designed to deter potential perpetrators from accessing CSAM via the search results by providing them with resources that may help them refrain from committing CSEA offences. By reducing searches for CSAM, the measure may also reduce the harm inflicted on child victims by the subsequent re-viewing and re-sharing of this content.

Our third measure (ICS F3), on crisis prevention information, will effectively disrupt user search journeys to minimise the risk of those users encountering illegal suicide content, and minimise the risk of harm should users encounter such content.

Terms of service and publicly available statements

What is this chapter about?

Terms of service ('terms') and publicly available statements ('statements') typically lay out the rights and responsibilities that a service provider and the users of their service have towards one another. Terms and statements are important to ensure transparency around the steps service providers are taking to protect users from illegal content. They are a tool for users to better understand the risk of using a service. This chapter sets out the measures relating to terms of service and publicly available statements we are recommending, why we are recommending them, and to which user-to-user (U2U) and search services they should apply.

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICU G1/ ICS G1	Providers should include provisions in their terms and statements regarding **the protection of individuals** from illegal content, any **proactive technology used**, and information on **how complaints are handled and resolved**.	Providers of all services.
ICU G2/ ICS G2	Providers should **summarise the findings** of their most recent **illegal content risk assessment** in their terms and statements.	Providers of Category 1 and Category 2A services.
ICU G3/ ICS G3	Providers should ensure that provisions included in **terms and statements** regarding the protection of individuals from illegal content **are clear and accessible**.	Providers of all services.

Why are we making these decisions?

These decisions are intended to ensure that users understand the risks they face on relevant services and the measures service providers are taking to protect them from these risks. This will enable them to make more informed choices about what services to use. Not only will this allow users to better protect themselves from harm, but it may also generate reputational incentives for service providers to improve their safety measures.

User access

What is this chapter about?

Removing users and accounts that post the most harmful types of illegal content is an effective way of combatting the spread of such content. However, such restrictions need to be considered very carefully given the impact they have on freedom of expression. This chapter sets out the measure we are recommending in relation to the removal of user accounts operated on behalf of proscribed terrorist organisations, why we are recommending it, and to which user-to-user (U2U) services it should apply.

What decisions have we made?

We are recommending the following measure:

Number in our Codes	Recommended measure	Who should implement this
ICU H1	Providers should **remove a user account** from a service if there are reasonable grounds to infer it is **operated by or on behalf** of a **terrorist organisation proscribed by the UK Government.**	Providers of U2U services

Why have we made these decisions?

It very likely that the content generated, shared, or uploaded via accounts operated on behalf of proscribed organisation will amount to an offence. Removing such accounts should make it more difficult for these organisations to communicate and cause harm online. This will help protect users and society at large from the harm caused by online terrorist content.

User controls

What is this chapter about?

This chapter sets out the user controls measures we are recommending, why we are recommending them, and to which user-to-user (U2U) services they should apply.

What decisions have we made?

We are recommending the following measures:

Number in our Codes	Recommended measure	Who should implement this
ICU J1	Providers should offer every registered user options to **block and mute** other user accounts on the service.	Providers of large U2U services that: - *have user profiles and certain user interaction functionalities; and* - *are at medium or high risk of one or more of the following kinds of illegal harm: grooming;* encouraging or assisting suicide (or attempted suicide); hate; harassment, stalking, threats, and abuse; and coercive and controlling behaviour.
ICU J2	Providers should offer every registered user the option of **disabling comments** on their content.	Providers of large U2U services that: - *enable users to comment on content; and* - are at medium or high risk of one or more of the following kinds of illegal harm: grooming; encouraging or assisting suicide (or attempted suicide); hate; harassment, stalking, threats, and abuse.
ICU J3	Providers should provide information to help users understand why and how user profiles are labelled under notable user schemes and monetised schemes. They should also have and consistently apply clear internal policies on these schemes.	*Providers of large U2U services that:* - *are at medium or high risk of fraud or the foreign interference offence; and* - *operate a notable user scheme or monetised scheme.*

Why are we making these decisions?

The measures in this chapter, taken together, aim to ensure that service providers give users access to tools that allow them to determine the content they see on services, who can contact them and who can interact with them, and information that helps them to decide whether to engage with and trust content.

Enabling users to block other users can help them reduce the risk of encountering illegal content. In particular it can play an important role in helping users avoid harms such as harassment, stalking, threats and abuse, and coercive and controlling behaviour. Similarly, allowing users to disable comments can be an effective means of helping them avoid a range of illegal harms including harassment (such as instances of epilepsy trolling and cyberflashing) and hate. These offences are widespread and cause significant harm.

In light of the prevalence and impacts of the harms and the important role we consider the measures could play in tackling them, we consider that the benefits of these measures are sufficient to justify the costs we have identified. There is a degree of uncertainty about some of the costs. In order to ensure that we are acting proportionately, we have decided to target the measures at medium or high-risk large services.

Our measure relating to notable user and monetised profile labelling schemes (ICU J3) should increase user understanding of why profiles are labelled, enabling them to take this context into account when deciding whether to engage with content posted via the account in question. It will provide users with information to reduce the risk of them falling victim to foreign interference or fraud. We have made three minor amendments to the measure, which are set out in the relevant 'Our decision' section.

Combined Impact Assessment

What is this chapter about?

In the preceding chapters in this volume we have assessed the impact of each of the measures we are recommending in this Statement. In this chapter, we assess the combined impact of the recommended measures as a package. Having considered the combined impact on different groups of services, we consider the package of measures to be proportionate.

Statutory Tests

What is this chapter about?

In designing our Codes, the Online Safety Act requires us to have regard to a number of principles and objectives, set out in Schedule 4 to the Act. The Communications Act 2003 also places a number of duties on us in carrying out our functions.

In this chapter we set out the matters to which we must have regard under the Online Safety Act and the Communications Act, and explain the reasons why we think the recommendations in our illegal content Codes of Practice meet them. We provide further information regarding Ofcom's duties relating to the preparation of our Codes in our introduction to the Statement, our Legal Framework (Annex 2), and Annex 4, in which we set out our Equality Impact Assessment and Welsh language assessments.

Volume 3: Transparency and Trust

Introduction to the volume

In this volume, we provide further **transparency** and guidance to providers to help them better protect users. We consider these decisions will help serve the objective of **building trust** in service providers and in the regulatory regime.

It sets out the decisions we have taken in producing the following three guidance documents:

- Our Illegal Content Judgements Guidance (ICJG) will help providers understand what illegal content is and what information they should have regard to when making judgements about content.
- Our enforcement guidance, which has been informed by our experience and track record in other sectors, sets out in clear terms how we will normally approach enforcement under the Online Safety Regime.
- In our guidance on content communicated 'publicly' and 'privately' under the Online Safety Act, we have set out the factors providers should consider when determining whether content is communicated 'publicly' and included some case studies case studies which we intend to assist stakeholders and provide greater clarity.

Ofcom's Illegal Content Judgements Guidance (ICJG)

What is this chapter about?

The Act requires us to provide guidance to service providers about how they can judge whether a piece of content is likely to be illegal and we do this in the Illegal Content Judgements Guidance (ICJG). In making such judgements, the approach to be followed is for service providers to consider whether there are 'reasonable grounds to infer' it is illegal content (that is, that it 'amounts to a relevant offence'), using all relevant information reasonably available to the provider ('reasonably available information') to make this judgement. Definitions of 'reasonable grounds to infer' and 'reasonably available information' are set out in the final guidance.

In this chapter, we set out a high-level summary of what the ICJG proposals for the November 2023 Consultation on Illegal Harms were, and the stakeholder responses we had to those proposals. We then set out the decisions we have made with regard to specific stakeholder responses. We have split this section into two; one focusing on cross-cutting or broader issues, and the other on more offence-specific issues. We have also included separate annexes setting out less substantive responses and changes, and further detail on our original proposals.

What decisions have we made?

We have considered stakeholder responses and made a number of offence-specific decisions which are set out in the following chapter. These include the following:

- **Fraud by false representation**: At consultation, we provided a list of indicators which, when present in specific combinations, would provide reasonable grounds to infer that content amounts to an offence of fraud by false representation. On review, we believe such an approach was too rigid as the indicator list could become quickly out-of-date. As such, we have decided to steer service providers to consider these indicators as illustrative and non-exhaustive. We have also altered the organisation of our list of indicators, basing the four groups of indicators on the four necessary requirements of the offence.

- **Intimate image abuse (IIA)**: At consultation, we proposed that, when content is shared, reposted or forwarded, the state of mind that matters is that of the user sharing, reposting or forwarding. We have decided to strengthen this position in relation to IIA. Absent any evidence that the user reposting, forwarding or resharing content has taken appropriate steps to ascertain consent, it is reasonable to infer that the user does not have a reasonable belief in consent. It follows that if the content concerned is an intimate image which has been shared without consent, it will be illegal content when it is forwarded, shared or reposted. This strengthening of our guidance will provide extra protection to victims and survivors of IIA.
- **Sexual exploitation of adults**: In light of evidence provided to us during the consultation process, we have given additional guidance on how service providers can recognise content related to the sexual exploitation of adults. The ICJG lists a series of risk factors which service providers should consider when assessing whether posts are likely to amount to offences related to the sexual exploitation of adults. Where enough of these risk factors are present, we set out that, absent evidence to the contrary, these posts are illegal content .
- **Encouraging/assisting suicide and self-harm**: We have made numerous changes to these chapters, drawing out the nuance in relation to the language, the vulnerability of users posting potentially illegal content and instances where intent may be able to be inferred.
- **Cyberflashing**: At consultation, we proposed that it would not be possible to infer intent to cause distress, alarm or humiliation in most cases. We have now decided to amend our position and state that service providers can infer intent in most instances, except for in certain, very specific circumstances. This change of position will strengthen protections for victims of cyberflashing.

Alongside this, we have streamlined and refined the ICJG to make it more accessible and better reflective of the law. We have also set out the information we consider reasonably available for service providers in a box for each offence, and added new sections to each chapter to draw out more clearly what types of content may need to be considered for the purposes of risk assessment.

Why are we making these decisions?

We have made these decisions in order to better reflect the law, to make the ICJG is as accessible as possible, and to ensure that our approach is informed by evidence provided by stakeholders. We have at all times sought to strike an appropriate balance between user protection and user rights.

Enforcement powers

What is this chapter about?

This Chapter explains our general approach to regulatory enforcement, how we will approach enforcement under the Online Safety Act (the Act) and introduces our Online Safety Enforcement Guidance (the Guidance).

What decisions have we made?

We have retained the overall approach set out in our November 2023 Illegal Harms Consultation (the November 2023 Consultation) about how we intend to exercise our enforcement powers and have issued the Guidance alongside this Statement. Having considered respondents' comments on our draft Guidance we have made some minor changes, including to clarify how we intend to engage with affected stakeholders, and other entities with relevant expertise, before making an application for business disruption measures.

Why are we making these decisions?

Our approach to enforcement under the Online Safety Regime has been informed by our experience and track record of enforcement in other sectors we regulate. We believe it will enable us to take effective and timely enforcement action in the interests of citizens and consumers, including by driving compliance; protecting users, especially children, from harm; deterring future wrongdoing; and holding wrongdoers to account. The changes we have made provide further clarity for stakeholders on the process we will follow in doing so.

Guidance on content communicated 'publicly' and 'privately' under the Online Safety Act.

What is this chapter about?

Ofcom can recommend that service providers use proactive technology in a Code of Practice to help them to fulfil some of their duties under the Act. We can only recommend such technology to analyse user-generated content (or metadata relating to such content) that is communicated 'publicly'. Service providers looking to apply such a measure in accordance with the Codes will first need to determine which content on their service is communicated 'publicly'.

In November 2023, we consulted on high-level draft guidance to assist providers in determining whether content on their service is communicated 'publicly'. We based this draft guidance on the three statutory factors that Ofcom must consider when deciding whether content is communicated 'publicly' or 'privately' under the Act.

In this chapter, we outline the feedback that we received on our proposed guidance in the November 2023 Illegal Harms Consultation ('November 2023 Consultation') and explain the decisions that we have taken in response.

What decisions have we made?

We have considered stakeholder responses and have made the following decisions:

- We have broadly confirmed our proposed guidance with some additions to improve clarity. In particular, we have included new illustrative case studies.
- We have amended the guidance to make clear that we expect service providers to adopt a consistent approach regarding what content is communicated 'publicly' on their service, and that we consider that maintaining records could help providers to achieve this.
- We have amended the guidance to acknowledge that the fact that content is accessible to less than a substantial section of the public does not mean that it should be automatically considered as communicated 'privately'.

Why are we making these decisions?

The aim of the guidance is to assist providers in determining whether content on their service is communicated ‘publicly’ or ‘privately’, so that they can apply proactive technology measures in accordance with the Codes where appropriate.

Many stakeholders welcomed the draft guidance. While we did receive challenges in some areas, we did not consider there to be sufficient evidence to change our approach. This has led us to largely confirm our proposed guidance.

We have made changes in response to stakeholder feedback suggesting that the draft guidance did not provide sufficient clarity. The addition of case studies, for example, is intended to help service providers better understand how Ofcom would likely approach a holistic assessment of the three statutory factors.

We also received feedback indicating that our position towards content that is accessible to less than a “substantial section of the public” was unclear. Given there can be scenarios where the other statutory factors strongly suggest content is communicated ‘publicly’ in this instance, we have amended the guidance to acknowledge that the fact that content is accessible to less than a “substantial section of the public” does not mean that it should be automatically considered as communicated ‘privately’.

One stakeholder suggested that we should recommend in our record-keeping and review guidance that providers keep a record of how they have assessed whether content is communicated ‘publicly’ or ‘privately’ on their service. We do not consider the record-keeping and review guidance to be the appropriate channel for this. Instead, we have decided to amend our guidance on whether content is communicated ‘publicly’ or ‘privately’ to set out our expectations around taking a consistent approach to the assessment, and how keeping records can help providers to achieve this.