Protecting people from illegal harms online

Volume 1: Governance and Risk Management

Statement

Published 16 December 2024

Contents

Section

1. 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.

What are we trying to achieve

1.1 As set out previously in this statement, as we take on our new responsibilities under the Act we are pursuing four overarching strategic objectives:

a) stronger safety governance and risk assessment in online services;
b) online services are designed and operated with safety in mind;
c) greater choice for users so they can have more meaningful control over their online experiences, and
d) greater transparency regarding the safety measures services use, and the action Ofcom is taking to improve them, to build trust.

1.2 In this volume we set out and explain the decisions related to **governance and risk assessment** which we are taking in this statement. 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.

What decisions have we made towards this objective

1.3 A number of the regulatory documents we are finalising today will contribute to our strategic objectives in relation to **governance and risk assessment**:

a) **The Register of Risk** sets out our detailed assessment of the causes and impacts of the online harms addressed in this statement. To aid service providers with their risk assessment, we have extracted the main points from the Register into our **Risk Profiles**, which provide a short overview of the factors and features that give rise to risks of different harms**.** The Risk Profiles and the more detailed analysis in the Register should inform service providers' risk assessment, thereby helping to improve the quality of risk assessment in the sector. Chapter 2 explains the methodology we used when producing the Risk Profiles and the Register. The **Register of Risks** and **Risk Profiles** can be found on our website.
b) Our **Risk Assessment Guidance** explains how we recommend providers should go about complying with their duty to assess the risk of illegal harms occurring on the services they operate. This document will contribute to our strategic objective by setting an expectation that service providers follow good practice in conducting a 'suitable and

sufficient' risk assessment. Chapter 3 explains our approach to the Risk Assessment Guidance which we have decided to take. Our **Risk Assessment Guidance** can be found here.

c) Our **Record Keeping and Review Guidance** explains how service providers can comply with their duties to keep written records of their risk assessments and the measures taken to comply with the safety duties. Once again, by outlining good practice in record keeping, this document will support risk assessment accuracy and enable adherence to governance standards amongst service providers. Chapter 4 explains the approach to the Record Keeping and Review Guidance which we have decided to take. The **Record Keeping and Review Guidance** can be found here**.**

d) Our Codes include a number of **governance measures**. These measures are based on best practice in sectors that have a mature and well-established culture of robust risk assessment and governance as well as on the existing secondary literature on governance. We are not currently satisfied that governance in the tech sector is universally good enough. We consider that applying these measures to service providers will materially improve governance in the sector and help to focus senior decision making on matters relating to online safety. Chapter 5 explains what these governance measures are and why we have chosen to include them in our Codes. The relevant **governance measures** can be found in our Illegal Content Codes of Practice for user-to-user and search services.

2. Register of Risks and Risk Profiles

Summary of findings from our assessment of the causes and impacts of online harm (Register of Risks)

2.1 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 on a service.

2.2 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.

2.3 Over the past two years we have conducted an extensive analysis of the causes and impacts of illegal online harms. As part of our analysis, we reviewed thousands of sources from hundreds of research organisations, academic institutions, online service providers, government, law enforcement and civil society organisations.

2.4 In the November 2023 Consultation, we published an initial draft of this analysis. This drew on research that we have commissioned, a comprehensive review of the existing published evidence and engagement with a wide variety of stakeholders. The over 200 responses to the November 2023 consultation contained a large body of additional evidence on and insights into the illegal harms in scope of the Act. Over the past year, we have analysed these responses in detail and conducted follow up research, analysis and stakeholder engagement to deepen our understanding of the harms.

2.5 Taken together, the work we have done over the past two years shows that illegal online content is widespread and, in many cases, growing in prevalence. For example, Ofcom's own research found that 87% of adult internet users report having encountered a scam or fraud online and 25% of these people have lost money as a result.[1] Almost a fifth of children experienced sexual solicitation from adults they have chatted with online. In a recent report the NSPCC stated more than 7,000 Sexual Communication with a Child offences were recorded by Police in 2023/24, an 89% increase since this offence came into force in 2017/18.[2]

2.6 Given the breadth of the risks online, anyone can experience harm in some capacity, just like anyone can be a victim of crime offline – 63% of UK internet users reported having encountered potentially harmful content online in the past four weeks. But in most

[1] Ofcom, 2023. Online Scams & Fraud Research. [accessed 21 November 2024].

[2] NSPCC, 2024. Online grooming crimes against children increase by 89% in six years. [accessed 14 November 2024].

instances, the risks people face online are not equal, with children and people with certain protected characteristics most likely to be affected and certain kinds of harm being significantly more prevalent among certain groups.

2.7 For example, 16% of minority ethnic internet users have encountered 'hateful, offensive or discriminatory content', compared to 11% of all internet users. Similarly, studies have shown that women are five times more likely to be victims of intimate image abuse. Generally, the more protected characteristics or vulnerabilities someone has, the greater the risk of harm they face from priority illegal harms in the Act. In this sense, online harms mirror and potential amplify wider challenges in society.

2.8 Since the impact of the harms we have looked at can be extremely severe, our work has never been more critical than it is today. Harm is not limited to the online world and profoundly affects people's lives. This was brought sharply into focus by the widespread violent riots that took place in the wake of the tragic murders in Southport, and the role of risky characteristics of some service providers in the spread of illegal content that stirred up hatred, provoked violence and spread false information.[3] Furthermore, our updated Register of Risks demonstrates that harms, such as online grooming, can cause lifelong negative psychological impacts for victims. Additionally, the harm experienced may not be limited to the individuals who directly encounter risks or illegal content online. It also impacts those who are the victims of the subsequent actions of those people who have encountered illegal content or participated in illegal activity online. In some cases – in relation to those radicalised towards terrorism, for example, or the erosion of trust in democratic processes caused by state-sponsored disinformation campaigns –these harms have a wider societal impact.

2.9 We have seen how all types of services can pose a risk of harm from the priority illegal content addressed by the Act, individually and when used together to facilitate criminal activity. While many service providers have made significant investments in tackling online harms in recent years, our evidence shows these have not yet been sufficient. For instance, encrypted cloud-based file-sharing services are still used by offenders to store and share CSAM; social media services with large numbers of targetable users and with highly effective content recommender systems remain susceptible to the proliferation of incitements of violence and hatred viewable by millions of users in a matter of hours. Our analysis and engagement with stakeholders also shows how the risks to the UK public evolve in step with the online landscape and are arguably more significant now than they ever have been.

2.10 Our work in the past year shows that the kinds of illegal harm we have looked at occur on services of all types. Services as diverse as social media services, dating services, marketplaces and listings services, search services, user-to-user pornography services, and file-storage and file-sharing services are all used to disseminate some of the types of illegal content and facilitate the offences we have looked at in the Register of Risks. While certain characteristics of a service might make some kinds of harm more likely to occur on or via a particular service, all services in scope of the Act carry some level risk.

2.11 For instance, the size of a service does not dictate the level of risk but influences how and what kinds of illegal harm are more likely to manifest. Bad actors use both large and small services to spread illegal content, though the way in which they use these services is likely

[3] Ofcom, 2024. Open letter to UK online service providers. [accessed 29 November 2024].

to differ. For example, terrorists often use large services to disseminate propaganda to large audiences, but often use small services for more covert activities such as recruitment, planning and fundraising.

2.12 Although a very wide range of service types pose risks of the priority illegal harms in the Act, there are certain service types that appear to play a particularly prominent role in the spread of priority illegal content. For example, our updated analysis suggests that video-sharing services are more likely to pose a particularly high risk with regards to hate offences. Similarly, our analysis shows that there is a particularly high risk of filesharing services hosting CSAM.

2.13 Similarly, certain 'functionalities' stand out as posing particular risks because of the prominent role they appear to play in the spread of illegal content and the commission and facilitation of priority offences:

- **End-to-end encryption:** Offenders often use end-to-end encrypted services to evade detection. For example, end-to-end encryption can enable perpetrators to circulate CSAM, engage in fraud, and spread terrorist content with a reduced risk of detection.
- **Pseudonymity and anonymity:** In most cases where offenders are using online services to engage in illegal activity, hiding their identity is incredibly important as it supports their ability to evade detection. There is also some evidence that pseudonymity (where a person's identity is hidden from others through the use of aliases) and anonymity can embolden offenders to engage in a number of harmful behaviours with reduced fear of consequences who otherwise might not have. For example, while the evidence is contested, some studies suggest that pseudonymity and anonymity can embolden people to commit hate speech. At the same time, cases of harassment and stalking often involve perpetrators creating multiple fake user profiles to contact individuals against their will and to circumvent blocking and moderation.
- **Livestreaming:** There are many examples of terrorists livestreaming attacks, this can in turn incite further violence. The use of livestreaming remains a persistent feature of far-right lone attackers, many of whom directly reference and copy aspects of previous attacks. Similarly, perpetrators can exploit livestreaming functionality when abusing children online.
- **Content recommender systems:** Content recommender systems are commonly designed to optimise for user engagement and learn about users' preferences. Where a user is engaging with harmful content such as hate speech or content which promotes suicide, there is a risk that this might result in ever more of this content being served up to them.

2.14 Illegal harms and the risk factors which cause them are changing all the time as technology develops and society evolves. The recent emergence of generative AI provides a particularly clear example of this. As well as bringing important benefits, generative AI creates new risks across a variety of kinds of illegal harm including CSAM, terrorism, fraud and foreign interference.

2.15 The functionalities we describe above are not inherently harmful and can have important benefits for users. End-to-end encryption plays an important role in safeguarding privacy online. Pseudonymity and anonymity can allow people to express themselves and engage freely online. In particular, anonymity can be important for historically marginalised groups such as members of the LGBTQ+ community who wish to talk openly about their sexuality

or explore gender identity without fear of discrimination or harassment. Recommender systems benefit internet users by helping them find content which is interesting and relevant to them.

2.16 The constant emergence of new risks makes it vital that services conduct regular risk assessments. It also makes robust corporate governance particularly important. Where services have good governance arrangements in place with clear accountability for managing risks, they are more likely to detect and appropriately manage emerging risks. In addition to recommending measures to address specific harms, a key focus for us as the Online Safety regime comes into force will, therefore, be ensuring that services conduct robust risk assessments and have appropriate governance arrangements in place.

2.17 Feedback provided in response to the draft Register of Risks in the November 2023 Consultation was extensive, with a wide range of recommendations and supporting evidence provided regarding the links we should draw between certain kinds of illegal harm and the characteristics of online services. As well as providing new and previously uncited evidence, stakeholders highlighted in detail where they felt our analysis and conclusions did not align with their own evidence, experience or views.

2.18 In response, we have made hundreds of additions and edits to the Register of Risks to incorporate new evidence, update our conclusions in response to this evidence, and to clarify numerous points. This includes filling gaps in our analysis in relation to specific kinds of illegal harm and providing clearer explanations on broad themes that are relevant across illegal harms, such as the role of business models and generative AI. We have summarised these updates chapter by chapter in the Register of Risks Annex, where we have also set out the small number of instances where we have not made a change despite requests to do so from stakeholders.

The role and importance of Risk Profiles

2.19 The Act also requires Ofcom to produce 'Risk Profiles' to support service providers with their Risk Assessment.[4] The Risk Profiles are a vital tool to help providers understand the most important findings in our Register of Risks in a practical and meaningful way. Service providers must take account of our Risk Profiles when they conduct their risk assessments, and this will ensure they understand the characteristics of their services that may be particularly relevant to each kind of illegal harm. The Risk Profiles are crucial in translating the breadth and depth of the evidence in our Register of Risks for providers in a way that tangibly improves online safety for at-risk users.

2.20 We have carefully considered how best to design the Risk Profiles to be as effective as possible. A significant challenge for us was to design a tool that both extracted the most important findings from the Register of Risks, and did so in a way that was practical, simple and easy to follow by thousands of potentially regulated service providers encompassing a broad spectrum of service types and sizes.

2.21 In the November 2023 Consultation, we proposed presenting the Risk Profiles as two accessible summary tables for U2U and Search services, which list what functionalities and

[4] Section 98(5) of the Act.

other characteristics of service providers are associated with which harms.[5] The information in the Risk Profiles is a distillation of key findings from the Register of Risks about links between certain harms and risk factors. We further proposed that the Risk Profiles should focus on the best evidenced links between risk factors and harms and should not include links referred to in the Register of Risks where they were not as strongly evidenced. We considered this approach would help providers focus their analysis on the most important risk factors.

2.22 In light of consultation responses, we carefully considered whether this was still the best approach, revisited alternative options for how the Risk Profiles could be framed and considered in detail stakeholders' views for and against our proposals. As explained in Register of Risks Annex 2: Updates to the Register of Risk and Risk Profiles, we remain of the view that our proposed approach of producing two overarching U2U and Search Risk Profiles presented as accessible summary tables is the best approach and preferrable to the alternatives we considered.[6]

2.23 We have confirmed our approach that Risk Profiles should not set out all the risk factors from the Register of Risks, but only those we considered to be particularly important for service providers to consider. We have also maintained our approach of only including risk factors and associated illegal harms where the evidence has been particulary strong. This was based on the quality of evidence that demonstrated their role in increasing the risk of harm. We considered this approach would help providers focus their analysis on the most important risk factors.

2.24 As a result of consultation responses to the Register of Risks and a significant expansion of our evidence base, we have now been able to new kinds of illegal harm to almost every U2U risk factor. Simply put, this means service providers, as required under the Act, will need to take account of the risk of more illegal harms in relation to characteristics of their service than we originally proposed. Previously, we had been unable to do so, owing to gaps in our evidence base. Without the valuable insights and evidence shared with us by so many stakeholders and supplemented with further work by our teams, we would not have been able to fill many of these gaps.

2.25 For kinds of illegal harm with an established evidence base, consultation responses have allowed us to expand the risk factors associated with them. For example, now the risk of harm from terrorism offences must be accounted for in risk assessments by any service provider that has functionalities such as anonymous posting, direct messaging, user-generated content searching or recommender systems. Similary for harms where we had notable gaps in our evidence base, consultation responses have helped with some of them. For example, we have added the risk of harm from firearms, knives and other weapons offences for providers that allow file-sharing or storage, annonymous posting or encrypted messaging.

2.26 We believe our final Risk Profiles will result in more thorough and effective risk assessments by service providers given they are the crucial first step in the risk assessment process

[5] Characteristics include a service's user base, business model, functionalities and any other matters we deem relevant to risk. Risk Profiles focus predominately on user base demographics, functionalities and business models. Step 2 of the risk assessment guidance provides information for services on user base size, governance, and systems and processes.

[6] We considered each option against two main objectives: a) our approach should effectively present our evidence on what makes services risky; and b) our approach should be easy for all services to use.

specified in our Risk Assessment Guidance. This should in turn allow providers to reflect potential risks of harm to even more users and be held accountable for their response to those risks.

3. Risk Assessment Guidance for Service Providers

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.

Introduction

3.1 Ofcom has a duty to produce guidance to help service providers to comply with their duty to complete an illegal content risk assessment. All user-to-user (U2U) and search service providers that fall in scope of Part 3 of the 2023 Online Safety Act (the Act) must complete an illegal content risk assessment which is 'suitable and sufficient', and they must take appropriate steps to keep their risk assessment up to date.

3.2 Illegal content risk assessments are a critical part of the online safety regime. The adoption of good practice in risk assessment is fundamental for the industry culture change needed to put user safety at the heart of service design and decision making. As the nature of harm online continually evolves, robust risk management processes should ensure service providers are able to quickly and effectively identify and respond to emerging risks on their services for people in the UK.

3.3 The illegal content risk assessment duties include different elements. **U2U service providers** must assess the risk of users encountering priority illegal content or other illegal content by means of the service, and the risk that the service may be used for the commission or facilitation of a priority offence. They must also assess the nature and severity of harm which may be suffered as a result.

3.4 The Act requires service providers to consider various characteristics of the service – such as its user base, functionalities, business model, and systems and processes – and to also take account of the relevant Risk Profile(s) produced by Ofcom.[7]

3.5 **Search service providers** must also complete an illegal content risk assessment, which assesses the risk of users encountering priority illegal content or other illegal content by means of the service. Likewise, they must assess the nature and severity of the harm which may be suffered. Like U2U providers, search service providers must take account of Ofcom's Risk Profile(s) and evaluate similar characteristics specified in the legislation (though not the user base).

3.6 Overall, **the purpose of conducting a risk assessment is to ensure service providers have an adequate understanding of the risks that arise from illegal content by means of their service, so that they can take proportionate measures to manage and mitigate those risks (as required by the illegal content safety duties).**

3.7 Following the Risk Assessment Guidance for Service Providers (the Risk Assessment Guidance) is not compulsory. However, **we consider that following our proposed guidance will put service providers in a stronger position to comply with their duties.**

3.8 In the consultation, we asked for feedback on our proposals for U2U and search service providers regarding the Risk Assessment Guidance. We have carefully considered all the responses we received.

3.9 This Statement Chapter explains how we have considered stakeholder feedback and explains the decisions we have taken in our approach to the Risk Assessment Guidance. We have set this out in the following structure:

a) The four-step risk assessment methodology
b) Risk Level Tables
c) Evidence inputs
d) A 'suitable and sufficient' illegal content risk assessment
e) Reviewing and updating an illegal content risk assessment

3.10 We have made some changes to the Risk Assessment Guidance following the consideration of consultation responses and other stakeholder engagement, as well as to ensure it is aligned with other regulatory documents, particularly the draft Children's Risk Assessment Guidance, Children's Risk Profiles and the Children's Register of Risks on which we consulted in May.[8]

---

[7] Ofcom's Risk Profiles are covered in Volume 1, Chapter 2: Register of Risks and Risk Profiles. They provide a summary of the factors we consider to be associated with a heightened risk of illegal harms.

[8] The Statement we are publishing today discusses the steps we expect providers to assess the risk of illegal content occurring on their services and to protect people from these risks. Separately, in May we consulted on the Children's Risk Assessment Guidance. The Act sets out that services that are likely to be accessed by children also need to conduct a 'children's risk assessment'. They must assess the risk of harm they pose to children by the kinds of content harmful to children set out in the Act, and how the design of the service affects the level of risk of harm to children.

3.11 As stated in our May consultation, we are working to keep our approach to risk assessment across the phases of the regime as consistent as possible, whilst ensuring that each piece of guidance reflects the requirements for the relevant duties in the Act. As part of this effort, we have taken account of relevant responses to the Protection of Children consultation for the Children's Risk Assessment Guidance for Service Providers.

3.12 We expect to publish a statement with our final decision on the Children Risk Assessment Guidance in April 2025. Our aim is to align approaches across the two pieces of guidance, where appropriate, so that service providers in scope of both sets of duties can apply risk assessment principles consistently to embed a culture of good governance within the organisation at all relevant levels.

Four-step risk assessment methodology

Background

3.13 Ofcom is required under the Act to produce guidance to assist service providers in complying with their illegal content risk assessment duties.

Our proposals

3.14 At consultation we proposed a four-step methodology to help service providers produce a 'suitable and sufficient' risk assessment to help them meet the requirements of the illegal content risk assessment duties, and associated duties (such as those on record-keeping and review). The four-steps in the proposed risk assessment methodology were:

a) Understanding the harms;
b) Assessing the risk of harms;
c) Deciding on measures to mitigate risks;
d) Report, review and update risk assessment.

3.15 Our proposals were designed to help service providers establish a risk assessment cycle which fits within their existing business practices. They were supported by the Governance measures proposed in our Illegal Content Codes of Practice. This is to recognise that good organisational governance structures and accountability are key to reinforcing a risk-based approach in business operation and decision-making.

3.16 The risk assessment methodology presented in the guidance was designed to be:

- based on best practice from risk-based industries;
- flexible and support the range of service providers in scope of these duties under the Act, and;
- adapted for service providers to assess risks as they exist at the time of their risk assessment.

Stakeholder responses

3.17 The stakeholder responses to our proposed four-step methodology broadly raised the following themes:

- **Support for the flexibility of proposed methodology:** Many stakeholders were generally supportive of the proposed four-step methodology.[9] Specifically, Microsoft and LinkedIn noted the clear and practical guidance it provides as well as the flexibility that it allows providers.[10] Snap added that the four-step methodology is the same approach they have taken for their Digital Services Act (DSA) risk assessment and report.[11] Further, Skyscanner welcomed the ability to assign a kind of illegal content a 'negligible' risk level where the service's functionalities would not allow for the harm to be presented by means of the service.[12] They said this recognises that many services lack the functionalities recognised by Ofcom as risk factors, making the risk assessment process less burdensome for service providers. Service providers who responded to our consultation on our proposed Children's Risk Assessment Guidance for Service Providers which took the same approach, also welcomed the flexibility it allowed.[13]
- **Insufficient focus on residual risk:** We received some stakeholder responses which suggested that the risk assessment methodology did not adequately support service providers to consider any pre-existing controls to manage risk as part of their assessment, resulting in an inaccurate picture of risk on their service.[14] These stakeholders expressed concern that the proposed methodology focused too much on "theoretical inherent risk" rather than "real residual risk". These respondents stated that Ofcom's approach recommends assessing the likelihood and impact of illegal harm without considering existing measures and mitigations, with UK Interactive Entertainment (Ukie) stating that the goal of risk assessments "should be to assess the actual risk of illegal content appearing on the service in question, not the risk of illegal content appearing on the service absent any mitigation measures."[15] Similarly, Airbnb highlighted that this approach could lead services to overestimate harms by not considering existing mitigations, contradicting Ofcom's commitment to proportionality.[16] [✂][17]
- **Criticisms of the methodology:** The proposed methodology was described as overly prescriptive by some stakeholders, in that it would limit the ability of service providers to align risk assessment processes under the Act with existing organisational risk management processes.[18] Reddit highlighted that being prescriptive has the potential to

[9] LinkedIn response to November 2023 Illegal Harms Consultation, p.4; Microsoft response to November 2023 Illegal Harms Consultation, p.4; Mid Size Platform Group response to November 2023 Illegal Harms Consultation, p.4; OnlyFans response to November 2023 Illegal Harms Consultation, p.3; Skyscanner response to November 2023 Illegal Harms Consultation, p.8; Snap response to November 2023 Illegal Harms Consultation, p.8.
[10] LinkedIn response to November 2023 Consultation, p.4; Microsoft response to November 2023 Consultation, p.4.
[11] Snap response to November 2023 Consultation, p.7.
[12] Skyscanner response to November 2023 Illegal Harms Consultation, p.6.
[13] Scottish Government response to May 2024 Protection of Children Consultation, p.10; Nexus Northern Ireland response to May 2024 Protection of Children Consultation, p.10; ParentingFocus response to May 2024 Protection of Children Consultation, p.17.
[14] Booking.com response to November 2023 Illegal Harms Consultation, p.8; eBay response to November 2023 Ofcom Consultation, p.3; Mid Size Platform Group response, p.4; Roblox response to November 2023 Ofcom Consultation, p.6; techUK response to November 2023 Illegal Harms Consultation, pp.11-12; UK Interactive Entertainment response to November 2023 Illegal Harms Consultation, p.7.
[15] UK Interactive Entertainment response, p.7.
[16] Airbnb response to November 2023 Consultation, p.4.
[17] [✂].
[18] Google response to November 2023 Illegal Harms Consultation, p.14; techUK response, p.2.

discourage innovation and progress in the development of new safety tools and methods.[19] They suggested increased flexibility of the risk assessment process on those grounds. On the other hand, 5Rights Foundation said the risk assessment guidance proposal placed too much onus on the proportionality and cost to services rather than focusing on supporting services to carry out holistic and robust assessments.[20]

- **Alignment with other regimes:** Stakeholders welcomed Ofcom's alignment with existing online safety regimes so far and encouraged extending this approach where possible. Specifically, Microsoft explained that mutual recognition among regulators could help achieve a "global and standardised approach, supporting effective and consistent safety information-gathering across jurisdictions."[21] In their response to the Protection of Children consultation, TikTok also welcomed the efforts to align our approach to the Risk Assessment Guidance with existing regimes.[22] X proposed mutual recognition between regulators on the grounds that it would allow for greater efficiency and focus on combating user-facing harms as well as reducing the burden on service providers.[23]

Our consideration of stakeholder responses

3.18 In considering the responses from stakeholders to our proposed methodology, we have noted the following:

- **Support:** We note the broad support from stakeholders for the four-step methodology, with many positive responses regarding the balance between supporting service providers to meet their risk assessment duties while building in flexibility for the range of service providers in scope.
- **Criticism of the methodology:** We considered feedback which expressed concern that the methodology could damage the pace of innovation or that service providers could struggle to incorporate the steps into their existing business functions. Regarding 5Rights Foundation's concern that the risk assessment placed too much emphasis on proportionality, we have written the guidance in line with the requirements set out in the Act. It is important to ensure that our guidance is proportionate overall, considering the range of services in scope of the risk assessment duties, and the resources available to them. We have used the guidance to set Ofcom's expectations for how service providers should meet these requirements, and where proportionate set higher expectations for larger services or those facing greater risks.
- **Ensuring flexibility and alignment**: We considered stakeholder requests to align the risk assessments for the Act with comparable global regimes. While the duties under the Act are distinct and service providers must meet them, we consider that the four-step framework is flexible and can align with existing processes and approaches.
- **Clarity on the role of existing measures, and on residual risk:** We considered stakeholder concerns about the relevance of existing controls as part of risk assessments:

[19] Reddit response to November 2023 Consultation, p.23.
[20] 5Rights Foundation response to November 2023 Illegal Harms Consultation, p.13.
[21] Microsoft response to November 2023 Consultation, p.4.
[22] TikTok response to May 2024 Protection of Children Consultation, p.2.
[23] X (formerly known as Twitter) response to November 2023 Consultation, p.3.

> The purpose of the Risk Assessment Guidance is to help service providers determine the level of risk of illegal content to users as it actually exists (taking into account any mitigations they have in place). For this, service providers should use the Risk Profiles to understand the risks that may be associated with the functionalities and characteristics of their service. They should then consider, using evidence from the operation of their service, how any existing controls and other aspects of the design and operation of their service affect the actual level of risk to users.
> Some responses raised concerns that our guidance would require service providers to consider the 'theoretical' risk posed by its service's functionalities rather than actual risk to users. This fed into arguments regarding the emphasis of the Risk Assessment Guidance on what several stakeholders termed 'hypothetical risk'. However, the Act requires all service providers to consult the Risk Profiles and assess 17 kinds of priority illegal content and other illegal content. Even if a provider believes they are low risk for a kind of illegal content, they need to assess this content to demonstrate with evidence that they are low or negligible risk. Service providers need to do this to demonstrate they are complying with the requirements in the Act. Step 1 and Step 2 of the guidance is intended to support service providers to meet this requirement.

3.19 When assessing their risk levels, service providers should consider existing controls that are in place at the time of the risk assessment. For example, it could consider itself to be the lower risk level (medium or low, as appropriate) after evaluating functionalities or other characteristics (using Risk Profiles in Step 1 and after considering other characteristics in Step 2), and supported with evidence that the controls in place at the time of the risk assessment are effective and demonstrably reduce the level of such risks.

Our decisions

3.20 We have decided to retain the four-step methodology as part of our Risk Assessment Guidance for service providers. This decision confirms the proposals made in our November 2023 consultation, which we explain in the section entitled 'Benefits and effectiveness'.

3.21 The four steps are:

- Step 1: Understand the kinds of illegal content that need to be assessed.
- Step 2: Assess the risk of illegal harm.
- Step 3: Decide measures, implement and record.
- Step 4: Report, review and update.

3.22 We have adjusted Step 1 to clarify that service providers should assess the risk of harm arising from illegal content. We have added this explanation at the top of the Risk Assessment Guidance to help service providers understand their duties.

3.23 In response to stakeholder feedback and to clarify parts of our proposal, we have made some minor changes to our guidance.

3.24 We have improved the title of Step 1 to make it clearer that services should assess the risk of harm arising from illegal content.

3.25 We have also adjusted Step 2 to explain more clearly how service providers should consider existing controls as part of determining a risk level for each kind of priority harm. In this step, service providers should use evidence to assess each kind of priority illegal content and assign a risk level. Our intent is for services to assess risk as it exists on the service,

factoring in whatever mitigations it is deploying at the time of the risk assessment. This includes how the design or operation of the service increases or reduces the risk of kinds of harm.

3.26 With this adjustment, we are clarifying that **the risk assessment should be an assessment of risk as it exists at the point in time of the service provider's assessment cycle.** Service providers must then identify relevant safety measures recommended from the Codes of Practice to implement.

3.27 **The guidance makes clear that service providers should not place excessive weight on the impact of existing controls in their assessment and should take a precautionary approach where there is uncertainty or inconclusive evidence of the effectiveness of existing controls**. We recognise the importance of considering existing systems and processes to accurately assess risk, as long as this does not lead to service providers underestimating risk and misapplying safety measures recommended in Codes. Our changes show that existing controls and processes are important factors in determining actual risk level, as part of the full range of relevant factors and evidence described in the guidance.

3.28 We have also amended the Record Keeping and Review Guidance for service providers, requiring them to record in their records of risk assessments how existing controls have affected the risk levels assigned to particular types of illegal harm.[24]

Our reasoning

Benefits and effectiveness

3.29 We consider effective risk management to be a critical factor for organisations to achieve good outcomes, both in terms of user safety and wider business objectives (commercial, reputational or related to sustainability and corporate responsibility). Where risk management systems are absent, inadequate, or inconsistently applied, there can be serious risks that an organisation will be unable to anticipate or respond to adverse events, or to protect users from harm.

3.30 There is a broad range of evidence to support this conclusion. Research on risk management from Milliman, commissioned by Ofcom, found that it is best practice for firms of all sizes to have a proportionate risk management system and governance framework in place.[25] Through analysis of other sectors such as Health and Safety, Financial Services, Cyber Security and Human Rights it is apparent that a strong culture of risk management is integral to operating successfully.[26] The Government HM Orange Book supports this, saying that risk management "enhances strategic planning and prioritisation,

[24] See Volume 1, Chapter 4 and the Record Keeping and Review Guidance.
[25] Milliman, 2023. Report on principles-based best practices for online safety Governance and Risk Management, p.4.
[26] Financial Conduct Authority, Handbook Chapter 7: Risk Control. [accessed 9 October 2024]; Health and Safety Executive, Managing risks and risk assessment at work. [accessed 9 October 2024]; National Institute of Standards and Technology's (NIST), Framework for Improving Critical Infrastructure Cybersecurity. [accessed 9 October 2024]; BSR, 2021. Human Rights Assessment: Identifying Risks, Informing Strategy. [accessed 9 October 2024]; Danish Institute of Human Rights, Human rights impact assessment guidance and toolbox [accessed 9 October 2024].

assists in achieving objectives and strengthens the ability to be agile to respond to the challenges faced" in successful organisations.[27]

3.31 Ofcom has also reviewed a wide range of literature on best practice in risk assessments and risk management and what this looks like in sectors with a mature culture of risk management as the ones mentioned above.[28] The literature we reviewed and the case studies we undertook showed that there is broad consensus that a structured approach to risk management is critical for successful implementation and that the key elements of such an approach were as follows:

Table 3.1: The key elements of good practice

Element	Activities	Illustrative outputs
Identifying risks	Exercises to identify risks that may affect an organisation, even if the risk is unlikely to occur or to materially impact business operations. This may involve interviews or surveys with relevant stakeholders, evidence-based methods such as literature reviews and analysis of historical data, scenario analysis and structured examination techniques such as Hazard & Operability Analysis (HAZOP) or Structured What If Technique (SWIFT).[29]	A risk register which provides an exhaustive list of potential risks, classified by category or type.
Assessing risks	Conducting risk assessments and evaluating risks. This includes determining the likelihood and potential impact of events taking place that could affect or disrupt business operations. This typically feeds into an exercise to determine the severity or significance of events.	A risk assessment which scores, maps or evaluates risks according to pre-determined criteria; documentation which highlights priority risks an organisation faces based on the outcome of a risk assessment exercise (i.e. risks which have the most severe or significant consequences).
Managing risks	Putting in place risk mitigations, and internal and external controls that seek to reduce the likelihood of the events occurring, or to manage and mitigate their impact on business objectives.	Risk management plans detailing the controls in place to manage and mitigate risk. Plans should include consideration of any unintended consequences that controls may trigger.

[27] UK Government, 2023. The Orange Book, Management of Risk - Principles and Concepts, p.3. [accessed 9 October 2024].
[28] UK Government, 2023. The Orange Book, Management of Risk - Principles and Concepts. [accessed 9 October 2024]; International Organization for Standardization (ISO). ISO 31000 Risk Management. [accessed 9 October 2024].
[29] International Electrotechnical Commission (IEC) and International Organisation for Standardization 31010 on Risk Assessment Techniques, 2019.

Element	Activities	Illustrative outputs
Reporting risks	Ensuring that risk assessments and decisions on implementation of controls are recorded. Embedding risk management processes into governance structures by ensuring that risk management activities are regularly reported to risk governance bodies, and that there is effective oversight of risk across an organisation.	Records of risk assessments and measures taken to manage risks. Policies and documentation laying out governance processes for risk management, including decision-making and oversight functions.

3.32 The available evidence therefore suggests that providers that follow a four-step process will develop a good understanding of the risks their service pose and that this will improve their ability to mitigate these risks. We expect this to deliver significant benefits.

Costs

3.33 As we noted in our November 2023 Consultation, service providers will incur costs in implementing our four-step risk assessment methodology. These costs will vary substantially across service providers. As we explain below, evidence and analysis will vary depending on the size and complexity of a service(s). We would expect costs for small services with few risks to be relatively limited. Conversely, providers of large, complex services with many risks could incur substantial costs. Costs may be higher for service providers that do not have existing risk management processes in place.[30] Our proposed methodology is intended to be scalable and flexible depending on a service's risk levels, size and resources in order to minimise the cost burden.

Conclusion

3.34 Our assessment suggests that the four-step process is an effective means of assessing risk and that adopting it will confer significant benefits. As set out above, it is not possible to quantify the precise costs, as they are likely to vary substantially, with small and low risk providers incurring relatively limited costs and large providers with more risks potentially incurring greater costs. Despite this, we have concluded that a good risk assessment is fundamental to achieving good safety outcomes and so it is proportionate to expect providers to follow the four-step process. The evidence we received in consultation responses broadly supports this conclusion.

3.35 It is also important to consider that the illegal content risk assessment duties are imposed by the Act. To meet these legal obligations, service providers will need to incur appropriate costs to carry out suitable and sufficient risk assessments. Given the need to set up effective risk assessment and management processes for compliance, we consider that the four-step process is likely to be the best way for providers to ensure their risk assessments are suitable and sufficient.

3.36 For the reasons set out above, we have clarified in the Risk Assessment Guidance that **service providers should assess risk as it exists on the service at the point of the risk assessment.** We believe these changes clarify the intention of the guidance and better support service providers to meet the risk assessment duties in the Act. We believe

[30] Volume 3, November 2023 Illegal Harms Consultation, p.58.

considering the impact of any existing controls alongside evidence of harm on their service will lead to more robust conclusions about the risk of harm and improve the overall quality of risk assessments. These clarifications had already been factored into our proposals for the Children's Risk Assessment Guidance, which we consulted on in May 2024.

Risk Level Tables

Background

3.37 As part of their risk assessment, service providers need to attribute a level of risk for each of the 17 kinds of priority illegal content and other illegal content. The risk level they assign to each kind of priority illegal content and other illegal content will impact the safety measures they are recommended to implement in our Codes of Practice. This is because many Codes measures are contingent on specific levels of risk. To help providers assess the risk level of their services appropriately, we included Risk Level Tables to help inform this judgement.

Our proposals

3.38 In our consultation we included a 'General Risk Level Table'. This provides service providers with guidance as to the circumstances in which it is likely to be appropriate for them to assess themselves as posing a high, medium, low or negligible risk of harm for a particular kind of illegal content respectively. For U2U service providers, we also provided three additional tables for assessing the risks of specific child sexual exploitation and abuse (CSEA) offences. These covered CSAM imagery, CSAM URLs and Grooming. The Risk Level Tables we consulted on can be found at Tables 6, 7, and 8 of the Draft Risk Assessment Guidance.[31]

3.39 Our policy intention for these tables was to help service providers inform their judgement on levels of risk, rather than to be interpreted as definitive criteria. We considered that the Risk Level Tables would also help to ensure consistency in the calibration of risk on any service. The risk level could be high, medium, low or negligible for each kind of priority illegal content and other illegal content. Providers could inform their assessment with relevant evidence from core and enhanced inputs and proportion of risk factors for a kind of illegal harm.

Stakeholder responses

3.40 There were various responses on the draft Risk Level Tables. The main themes were:

- **Insufficient focus on actual risk**: Some responses said the draft Risk Level Tables placed weight on hypothetical risk factors, rather than whether a service might have actual risks.[32] For example, Roblox argued the presence of many or several risk factors should not automatically lead to medium or high risk but should instead take account of how a service's functionalities may affect the risk of harm.[33]

[31] Annex 5: Draft Risk Assessment Guidance for Service Providers, November 2023 Illegal Harms Consultation, pp.23-29.
[32] Airbnb response to November 2023 Consultation, p.5; Booking.com response to November 2023
[33] Roblox response to November 2023 Consultation, p.7.

- **Size should not determine risk level:** Many responses also said that size and user numbers alone should not determine a service's risk level.[34] Roblox suggested this was particularly the case where the service does not offer functionality that promotes content virality.[35]

- **Risk Level Tables are too mechanical and set too low a bar for medium risk**: Some responses said the draft Risk Level Tables were too mechanical and would catch too many services as being medium risk when this was not appropriate. Examples of these responses included:

  > Trustpilot said a single piece of evidence would count as evidence of harm occurring, and as a result a medium level of risk could be assigned. They said that this was not a meaningful threshold.[36]

  > Microsoft said the Risk Level Tables implied almost any risk could potentially be assigned at least a medium risk level.[37]

  > Ukie said that an isolated example of an identified harm materialising on a service should not automatically lead to a medium or high-risk level being assigned. It suggested that there should be more consideration of factors such as the frequency of such examples arising, and the measures implemented to reduce the harm.[38] [39]

- **Inappropriate threshold for image-based CSAM Risk Level Tables:** Two respondents said that the mere fact that a service (a) allows for image or videos to be uploaded, and (b) has two or more relevant risk factors associated with image-based CSAM (including direct messaging) does not necessarily indicate that a service is likely to pose a medium risk of image-based CSAM.[40] Booking.com argued that instead, likelihood should be assessed by looking at frequency or potential occurrence of impact, taking into account the nature of the platform.[41] Similarly, one response considered it simplistic and inappropriate for the image-based CSAM Risk Level Tables to suggest that any file-storage or file-sharing service which allows images to be uploaded, posted or sent should be classified as high risk for image-based CSAM.[42]

- **Not enough emphasis on severity of harm**: Some responses said there was not enough emphasis on the severity of harms. For example, the Samaritans said that the seriousness of certain harms (for example, suicide and self-harm) is such that there can be a devastating impact even if they reach only a small number of people, and hence severity of potential harm should be added as an indicator of high risk.[43]

Our consideration of stakeholder responses

[34] Google response to November 2023 Consultation, p.19; LinkedIn response to November 2023 Consultation, p.19; OSAN response to November 2023 Illegal Harms Consultation, p.88; Roblox response to November 2023 Consultation, p.7; Spotify response to November 2023 Illegal Harms Consultation, p.11.
[35] Roblox response to November 2023 Consultation, p.7.
[36] Trustpilot response to November 2023 Illegal Harms Consultation, pp.18-19.
[37] Microsoft response to November 2023 Consultation, p.9.
[38] UK Interactive Entertainment response to November 2023 Consultation, p.7.
[39] UK Interactive Entertainment response to May 2024 Protection of Children Consultation, pp.18-19.
[40] Booking.com also provided the example of direct messaging as a risk factor. Airbnb response to November 2023 Consultation, p.10; Booking.com response to November 2023 Consultation, pp.10-11.
[41] Booking.com response to November 2023 Consultation, pp.10-11.
[42] Apple response to November 2023 Illegal Harms Consultation, pp.7-8.
[43] Samaritans response to November 2023 Illegal Harms Consultation, p.4.

3.41 Considering responses from the consultation, we have made various changes to the four Risk Level Tables. These changes are intended to make them clearer and more accurately reflect our policy intent. We consider the changes address the main concerns raised by stakeholders, as set out above. This should ensure consistency among service providers on their assessment of risk and ensure measures are recommended when appropriate. The main changes to the General Risk Level Tables are:

- **More emphasis on evidence of illegal content**: We have made clearer that the tables are intended to capture the actual risk on services at the time the risk assessment is undertaken. They are not intended to capture only theoretical or inherent risk (as discussed in paragraph 3.18 of this chapter). We have also given more emphasis in the Risk Level Tables to evidence of actual harm. For example, we have been explicit that evidence of a material amount of illegal content of a particular kind of harm being present on a service is a strong indicator that the service is at least medium risk, and could be high risk. We would normally expect all service providers with such evidence to be medium or high risk for that kind of harm. However, as set out in the General Risk Level Tables, we consider that in some cases services may pose a high risk of certain kinds of harm even if there is no concrete evidence of the harm having occurred in the past. Some types of harmful content and conduct (for example grooming) are hard to detect. Whilst reports of harm occurring provide strong evidence that the service poses a risk, absence of evidence of a particular harm occurring cannot always be taken as proof that the risk of that harm occurring is insignificant.
- **Less emphasis on risk factors and size:** We recognise that assessments of how much risk a service poses are necessarily context-specific. Our risk tables identify some factors which we consider generally make a service likely to pose a medium or high risk of harm. However, we recognise that there may be factors which reduce the risks a service poses. Where such factors exist, it is open to providers to conclude their services are low risk even where they exhibit other factors which might otherwise make it high risk. In such a scenario we would, however, expect them to be able to justify their decision. We are also clear that a large number of users is not necessarily – by itself – a strong indicator of higher risk, especially compared to evidence of substantial illegal content being present or of a severe impact on users or affected individuals.
- **Occasional incidence of harm does not necessarily imply medium risk**: We have made clear that occasional occurrences of a particular kind of illegal content on a service do not necessarily mean the service is medium risk, provided there is limited scope for harm to impact users or other affected individuals. For example, this could be the case where the harm impacts very few users or other individuals, and the severity of that harm is likely to be low.
- **Likely severity of harm matters**: We have made it more explicit in the Risk Level Tables that the severity of harm is relevant to the impact. If a kind of harm, and the way that it is encountered by users on the service, would have a severe impact on users or other affected individuals, then even if the number of individuals affected is low, a service could be high risk.

3.42 We have also made changes to the additional CSEA Risk Level Tables that U2U service providers should consider (relating to CSAM imagery, CSAM URLs and Grooming). Taking account of responses, we have changed these tables to give more emphasis to evidence of illegal harm (as with the General Risk Level Table), and to make clear that, even if a service

has the relevant risk factors, they can still be assessed as low risk, if they can demonstrate that they have systems and processes in place that sufficiently reduce the risk.

Our decisions

3.43 We will keep the Risk Level Tables in the guidance to help service providers assign a risk level for each kind of priority illegal harm as it exists on the service at the time of the risk assessment. These tables are: a General Risk Level Table to help service providers, as well as additional guidance and specific tables for U2U service providers for CSAM imagery, CSAM URLs and Grooming offences.

3.44 Taking account of stakeholder responses, we have amended the Risk Level Tables to place more emphasis on evidence of illegal content, gathered through core or enhanced inputs, and evidence regarding existing controls on services. We have also drawn out how the amount of illegal content identified in evidence drawn from the service's operation affects the risk level. These changes are intended to clarify the policy intent of the risk assessment and address feedback from stakeholders from the consultation that the tables could lead service providers to erroneous conclusions.

3.45 We have also provided hypothetical examples to illustrate how to use the Risk Level Tables (these are included in Appendix B of the Risk Assessment Guidance for Service Providers). We have done this to help the calibration of risk to be appropriate and consistent with the measures recommended in Codes, and for different U2U services. The final version of the Risk Level Tables can be found at Part 3, Section 3 of the Risk Assessment Guidance for Service Providers.

Our reasoning

Benefits and effectiveness

3.46 We consider that following this methodology to assign a risk level for each kind of priority illegal content and other illegal content will deliver important benefits.

3.47 Providing service providers with criteria to determine risk levels will help them to reach better conclusions about their risk levels. This will enable them to more accurately assess the risks their services pose. The more accurately providers can identify the risks their services pose, the better able they will be to mitigate these risks. Better risk mitigation will result in people in the UK experiencing less harm on online services.

3.48 We also provide specific tables setting out how to assign risk levels for CSAM and grooming offences which we expect providers of U2U services to use. We single out these kinds of illegal content because different types of offences are associated with specific risk factors and because we have measures in Codes where the application of the measure depends on whether the service has a specific kind of CSEA risk. Additional guidance will assist service providers in making an accurate judgement of their level of risk for these kinds of illegal content and inform their approach to complying with their duties.

Costs

3.49 The Risk Level Tables are designed to help service providers assign risk levels for each kind of illegal harm, and assessing the level of risk for each kind of harm is a requirement of the risk assessment duty. We do not think that using the criteria set out in the tables will incur a significant additional burden for service providers seeking to comply with their risk assessment duties.

Conclusion

3.50 We believe that the Risk Level Tables will help service providers use criteria to assign risk levels accurately for each kind of harm. We think they will help service providers to meet their legal duties, will fulfil the policy objective we have set out, and represent a proportionate set of recommendations to service providers.

Evidence inputs

Background

3.51 Service providers need accurate and robust evidence to carry out their risk assessment, and to be confident in their conclusions about the level of risk of each kind of priority illegal content and other illegal content on their service. In general, the better the quality of the evidence services use, the more accurate their assessment of risks is likely to be, and the better placed they will be to protect their users from illegal content. The level and kind of evidence inputs will vary by service, by harm, and other factors.

3.52 The evidence considered by a service is a key to assigning a level of risk to each kind of priority illegal harm.

Our proposals

3.53 Our proposed Risk Assessment Guidance included guidance to help service providers identify the kinds of evidence inputs that they should use to make judgements about the risk of harm to users on their service.

3.54 Given the range of service providers in scope of this duty, we split the guidance on the kinds of evidence different service providers should consider into 'core' and 'enhanced' inputs. We also included instructions for service providers to navigate the kinds of evidence they should consider when assessing risk on their service.

3.55 We proposed that service providers should consider any available information that is relevant to assess risk. At a minimum, we proposed a list of 'core' evidence inputs which all service providers should consider. They are the kinds of evidence which all service providers have access to, including Ofcom resources such as the Register of Risks.

3.56 We also proposed a list of 'enhanced' inputs, consisting of additional evidence. We explained that we would expect providers of larger services, and service providers which identified several specific risk factors, to use at least some enhanced inputs. Our proposed guidance noted that some service providers may already hold the additional evidence we included in our proposed list of enhanced evidence inputs, but some other service providers should consider gathering it, to improve their understanding of risk on the service.[44]

Table 3.2: Our proposed summary of evidence inputs

Type	Overview of inputs
Core inputs	• Risk factors identified through relevant Risk Profile (Step 1) • User complaints and reports

[44] Annex 5: Draft Risk Assessment Guidance for Service Providers, November 2023 Illegal Harms Consultation, pp.18-19.

	• User data • Retrospective analysis of incidents of harm • Register of Risks (optional for user-to-user services) • Other relevant information (including any other characteristics that apply to your service that may increase or decrease risks of harm)
Enhanced inputs	• Results of product testing • Results of content moderation systems • Consultation with internal experts on risks and technical mitigations • Results of previous interventions to reduce online safety risks • Views of independent experts • Internal and external commissioned research • Outcomes of external audit or other risk assurance processes • Consultation with users • Results of engagement with relevant representative groups

Stakeholder responses

3.57 We did not receive any responses which objected in principle to our approach to segment evidence inputs into 'core' and 'enhanced'. However, we received feedback from stakeholders on our proposals relating to evidence inputs on the following themes:

- **Limitations of core inputs:** In their response, 5rights Foundation commented that Ofcom's 'core' evidence inputs do not adopt a thorough and ex ante approach, especially considering that in some cases this may be the only evidence type used by service providers.[45]
- **Introduction of new evidence types and re-categorisation of inputs:** A number of respondents also recommended other types of evidence for either the 'core' or 'enhanced' evidence categories as well as re-categorising some 'enhanced' inputs as 'core'. This included:
  - The National Society for the Prevention of Cruelty to Children (NSPCC) expressed that all internally available information should be classed as a 'core' input for service providers to be best equipped to judge the efficacy of their current approach and identify any areas where measures are not sufficient and require improvement.[46] They stated that it is both proportionate and reasonable to expect any service providers that hold this data to use it in order to produce a fully accurate and comprehensive risk assessment. This was echoed by End Violence Against Women (EVAW), who suggested the onus should be on service providers to provide evidence that ensures a reliable and robust risk assessment.[47]
  - Support for the inclusion of evidence from external stakeholders (and their representatives) and other relevant experts within 'core' evidence rather than 'enhanced' was quite broad.[48] Some stakeholders suggested the need to include the perspectives of external groups to cover any information discrepancies left after

[45] 5Rights Foundation response to November 2023 Illegal Harms Consultation, p.16.
[46] NSPCC response to November 2023 Illegal Harms Consultation, pp.9-10.
[47] EVAW response to November 2023 Illegal Harms Consultation, p.14.
[48] Barnardo's response to November 2023 Illegal Harms Consultation, p.7; [✂]; Mencap response to November 2023 Illegal Harms Consultation, p.4; TSB Bank response to November 2023 Illegal Harms Consultation, p.5; UK Finance response to November 2023 Illegal Harms Consultation, p.5.

consideration of service providers' own evidence.[49] This argument was furthered by Mencap, who said that this would provide service providers with additional evidence and insight that the service otherwise might not be able to gather internally, as well as allowing marginalised groups to make service providers aware of risks.[50]

- The Centre for Competition Policy argued that consultation with other service providers should be considered another 'enhanced' evidence input, highlighting that significant value could be found in communication and sharing of insight among service providers.[51] Cifas also argued that consultation with 'relevant industry groups' should come under 'enhanced' inputs.[52]
- Lloyds Banking Group suggested that evidence should include reports from affected parties. They said that the guidance should clarify that trends in fraud reports from customers and affected third parties should fall under 'other evidence'.[53]
- 5Rights Foundation and the Online Safety Act Network (OSAN), in their responses, also discussed product testing being part of 'enhanced' evidence as opposed to 'core'.[54] 5rights Foundation specifically highlighted the importance of considering the risk associated with a product, feature, or functionality before it is introduced as part of best practice.[55] For similar reasons, they also suggested that consultation with internal experts on risks and technical mitigations should also be considered a core input for providers of larger services.

- **Privacy concerns:** In contrast, some concerns were raised around the use of detailed 'enhanced' information, with stakeholders commenting this required more personal data than necessary, and highlighting legal privilege issues.[56] [57] We provide the following responses as detailed examples:

  - In their response, Google expressed concern towards Ofcom's guidance regarding evidence.[58] They stated that some enhanced, quantitative evidence types such as user complaints and content moderation reports are not necessarily collated by providers currently in a way that would allow for them to serve as evidence for risk assessments. They also highlighted that service providers may not hold or be able to infer some forms of user demographics and characteristics and request clarity on the collection of personal data for the purpose of the risk assessment.
  - The Information Commissioner's Office (ICO) also commented on the use of user data in their response, stating that the guidance on user demographics as evidence could lead to an assumption amongst some service providers that they need to collect this

[49] NSPCC response to November 2023 Consultation, p.10; UCL Gender and Tech Research Group response to November 2023 Illegal Harms Consultation, p.5; Caroline Are response to November 2023 Illegal Harms Consultation, p.4.
[50] Mencap response to November 2023 Illegal Harms Consultation, p.4.
[51] Centre for Competition Policy response to November 2023 Illegal Harms Consultation, p.12.
[52] Cifas response to November 2023 Illegal Harms Consultation, p.3.
[53] Lloyds Banking Group response to November 2023 Illegal Harms Consultation, p.7.
[54] 5Rights Foundation response to November 2023 Consultation, p.16; OSAN response to November 2023 Illegal Harms Consultation, p.19.
[55] 5Rights Foundation response to November 2023 Consultation, p.16.
[56] The ICO and Google discussed the issue around personal data. Google response to November 2023 Consultation, pp.18-19; ICO response to November 2023 Consultation, p.7.
[57] Snap discussed the issue around legal privilege. Snap response to November 2023 Consultation, p.7.
[58] Google response to November 2023 Consultation, pp.19-20.

personal data if they do not already do so, thus carrying privacy implications for users.[59]

- **Risk profiles as evidence:** Roblox argued that, when services review Risk Profiles as one part of several 'core' evidence inputs during their risk assessments, services should be free to determine the weight they give to the contents of individual Risk Profiles and be able to determine that certain conclusions set out in the Risk Profiles do not apply to them without having to consult enhanced evidence inputs in every instance. This will allow organisations to focus on addressing more likely/more impactful risks, rather than stretching resources to prove certain risks are low risk.[60]

Our decisions

3.58 We have decided to confirm the proposals we made in our November 2023 consultation regarding 'core' and 'enhanced' evidence inputs. We believe that this approach will help service providers to fulfil their legal duties under the Act, as we explain further in the section entitled 'Benefits and effectiveness'.

3.59 Taking into account stakeholder feedback, we have considered how help service providers understand expectations regarding the kinds of evidence inputs they should consult to assess the risk of illegal harm on their service. We explain these changes below.

3.60 We have added to the descriptions of some inputs to explain how they can improve an understanding of a harm. Specifically:

- using the findings of other internal risk assessment or audit processes to consider a harm,
- pointing service providers to specific resources that may help them consider the risk of illegal harm to children and vulnerable users on their service,
- including examples of the kinds of metrics they might gather to assess the impact of algorithms on their service relating to kinds of harmful content.

3.61 We have also added language in the Risk Level Table to help service providers understand the role of the core and enhanced evidence inputs in assigning a risk level, including the role of any risk factors identified through relevant Risk Profiles. Service providers should consult all evidence necessary for them to understand the risk of harm to users on their service and be able to point to this evidence in their record to evidence their conclusions for risk level for each kind of priority illegal content. Evidence of harm on the service remains the key determining factor in assigning a level of risk to each kind of priority illegal harm.

3.62 To help service providers navigate gathering data for a risk assessment with privacy concerns, we have collaborated with the ICO to identify resources to signpost. We have also made it clear in the guidance that service providers are not expected to gather additional personal data on users to produce a 'suitable and sufficient' assessment.

3.63 Some of the changes and updates to guidance relating to evidence inputs were also reflected in minor drafting and clarificatory changes in the Children's Risk Assessment

[59] ICO response to November 2023 Consultation, p.7.
[60] Roblox response to November 2023 Consultation, pp.8-9.

Guidance as we are working to ensure the two pieces of guidance are aligned in policy intent and methodology where appropriate.

Table 3.3: Summary of evidence inputs published in our Risk Assessment Guidance

Type	Overview of inputs
Core inputs ***All service providers should consider***	• Risk factors identified through relevant Risk Profile (Step 1) • User complaints and reports • User data (for example age, language, groups at risk) • Retrospective analysis of incidents of harm • Relevant sections of Ofcom's Register of Risks, to understand the context of the risk factors in Risk Profiles [61] • Evidence drawn from existing controls • Other relevant information (including other characteristics of your service that may increase or decrease risk of harm, such as existing controls)
Enhanced inputs ***Should be considered by large service providers and those who have identified multiple specific risk factors for a kind of illegal content.***	• Results of product testing • Results of content moderation systems • Consultation with internal experts on risks and technical mitigations • Views of independent experts • Internal and external commissioned research • Outcomes of external audit or other risk assurance processes • Consultation with users • Results of engagement with relevant representative groups
NB. *These are* ***not exhaustive*** *examples -* ***may vary by service or business model****.* ***See also Part 3, Section 2 of the Risk Assessment Guidance: 'Evidence inputs'.***	

Our reasoning

Benefits and effectiveness

3.64 As we noted in our November 2023 Consultation, we believe that a key part of producing a 'suitable and sufficient' risk assessment is for it to be grounded in relevant evidence. If providers use appropriate evidence for their risk assessments, this will ensure that the assessment accurately reflects risks as they exist on a service. This will in turn make them better placed to identify and put in place mitigations which protect their users from illegal content.

3.65 The literature on best practice in risk management which we have reviewed makes clear that analysing relevant information about risk and harm is critical to assessing risk accurately and implementing appropriate mitigations.[62] Our approach guides service providers to use evidence to meet their legal obligations under the Act, and therefore will be beneficial for them to identify risks and mitigate harm on their service.

[61] A service provider can consult the Register of Risks to better understand different kinds of priority harm, for instance those which they identify risk factors for when consulting Risk Profiles.

[62] Milliman, 2023. Report on principles-based best practices for online safety Governance and Risk Management; UK Government, 2023. The Orange Book, Management of Risk - Principles and Concepts. [accessed 9 October 2024]; International Organization for Standardization (ISO). ISO 31000 Risk Management [accessed 9 October 2024].

3.66 We have considered the benefits from using the enhanced inputs rather than relying exclusively on the core inputs. We conclude enhanced inputs to be particularly relevant for large services and multi-risk services. For the reasons set out above, use of the enhanced inputs will improve the quality of service providers' risk assessments thereby resulting in improvements to their ability to manage risk.

3.67 Enhanced inputs will also allow multi-risk services, who face more complex risks, to benefit from robust evidence-based judgements.

3.68 Conversely, while it is important that these service providers assess and manage risk effectively, because they have fewer users and/or pose fewer risks, the marginal benefit of improving their risk assessment by using enhanced input is smaller than is the case with large and/or multi-risk services.

Costs

3.69 We expect the incremental cost of using core inputs to be low. This is because they are necessary to undertake activities required under the Act or involves the use or processing information or data that they might already have for the functioning of the service.

3.70 We recognise enhanced inputs may incur more significant costs to services providers. We have not been able to quantify these costs, but consider they could be material. However, we note that service providers have the discretion to choose the lowest cost and most beneficial enhanced input(s) to fill any evidence gaps they have to produce a suitable and sufficient risk assessment.

Costs

3.71 [this paragraph is intentionally left blank]

3.72 [this paragraph is intentionally left blank]

Conclusion

3.73 Given the importance of ensuring that risk assessments are evidence-based and the relatively modest cost of obtaining such evidence, we consider it is proportionate for us to expect all service providers at a minimum to use the core inputs we have listed. Similarly, we consider that it is proportionate to expect providers of services which are large or multi-risk to use the relevant enhanced inputs given the substantial benefits of them doing so. We have decided not to make a firm recommendation about how many of the enhanced inputs they should use and which ones they should use. We expect providers to exercise a degree of judgment about this and select whichever of the enhanced inputs are most relevant to their assessment.

3.74 The case for setting a blanket expectation is less clear for smaller services which are not multi-risk to use the enhanced inputs. As explained above, we consider the benefits of this to be smaller. Moreover, the costs may prove more challenging for some of them to absorb. However, because of the fundamental requirement to do a good risk assessment, such services should consider enhanced inputs when they do not have sufficient clarity about

their risk levels. We therefore conclude for providers of services which are neither large nor multi-risk to consider enhanced inputs if analysis with the core inputs leaves material ambiguity about their risk levels.

3.75 The changes we have made to the language in our guidance will clarify expectations for service providers about how to consider evidence in relation to their duty to produce a suitable and sufficient risk assessment. These decisions demonstrate the importance of gathering evidence in assessing the risk of harm on the service.

'Suitable and sufficient'

Background

3.76 The Act requires all in scope service providers to produce an illegal content risk assessment which is 'suitable and sufficient'. This is a specific requirement in the Act, section 9(2) for user-to-user providers and section 11(2) for search providers. We have sought to produce Risk Assessment Guidance which supports service providers to fulfil this duty.

Our proposals

3.77 In our consultation, we explained that a 'suitable and sufficient' risk assessment is an illegal content risk assessment which meets all of the requirements set out in the Act. The guidance has been developed to support them to meet this overarching requirement and meet the specific requirements set in the Act.

3.78 We set out the definition in the introduction of the document and made reference to the various elements that constitute a 'suitable and sufficient' assessment through the proposed guidance.[63] Particularly in Step 1, where we made clear service providers must separately assess 17 kinds of priority illegal content, and other illegal content (including non-priority offences) for the assessment to be 'suitable and sufficient', and in Step 2 where we outlined the evidence service providers will need to use to make a 'suitable and sufficient' assessment.

3.79 We also explained that due to the range of service providers in scope, what is 'suitable and sufficient' could vary substantially by service. We said we expected that larger service providers, for instance, would need to use and gather a wider range of evidence inputs to understand the risk of harm on their service.

Stakeholder responses

3.80 We received feedback from a range of stakeholders welcoming our guidance to help them meet the risk assessment duties set out in the Act. However, we received some feedback seeking clarification on the expectation that the risk assessment be 'suitable and sufficient', specifically following these themes:

- **Alignment with other regimes:** We received some consultation responses that highlighted that some stakeholders were unclear what would meet the requirement of a 'suitable and sufficient' risk assessment. For example, LinkedIn and Microsoft queried

[63] Annex 5, Draft Risk Assessment Guidance for Service Providers, November 2023 Illegal Harms Consultation p.5.

whether previous risk assessments completed to comply with other regulatory regimes, such as under the Digital Services Act or Digital Trust and Safety Partnership Safe Framework, would be accepted by Ofcom as 'suitable and sufficient'.[64] In their response to the Protection of Children consultation for our proposed Children's Risk Assessment Guidance, Microsoft said that the method and substance of conducting risk assessments should be aligned across regimes.[65]

- **Flexibility of the duty:** Skyscanner agreed with Ofcom that what constitutes a 'suitable and sufficient' illegal content risk assessment is a context-specific requirement. They stated that this should allow service providers flexibility to meet the requirement of a 'suitable and sufficient' risk assessment based on their characteristics.[66]
- **Requests for proportionality:** Some stakeholders highlighted the importance of a flexible and proportionate approach to the methodology.[67] [✂] encapsulated this by saying "We believe that a more effective approach would be for Ofcom to consider each service on a case-by-case basis, taking into account the service's specific functionalities and user base to establish an overall level of risk".[68]
- **How this will be assessed and enforced upon:** Several comments also highlighted the potential need for Ofcom to produce guidance on how it intends to assess the suitability and accuracy of completed risk assessments, including how enforcement activities will be initiated in the cases of services failing to meet pre-identified levels (in the respondent-proposed guidance) of compliance.[69] [70] Barnardo's, for example, expressed concerns over service providers not taking action to mitigate risks they are aware of and assigning a lower risk level than appropriate.[71] They stated that external input and auditing would ensure proper oversight of the process. Similarly, EticasAI commented that establishing specific thresholds and metrics is necessary for implementation and enforcement.[72] This point was echoed by OSAN, who highlighted the limitations of services' harm mitigating measures as well as the need for the development of appropriate metrics by service providers.[73] The Global Network Initiative built on this to request more information on whether or not - and how - completed risk assessments will be published.[74]

Our consideration of stakeholder responses

[64] LinkedIn gave the example of the Digital Services Act and Microsoft provided the example of the Digital Trust and Safety Partnership Safe Framework. LinkedIn response to November 2023 Consultation, p.4; Microsoft response to November 2023 Consultation, p.4.

[65] Microsoft response to May 2024 Protection of Children Consultation, p.6.

[66] Skyscanner response to November 2023 Consultation, p.7

[67] Name withheld 5 response to November 2023 Illegal Harms Consultation, p.4; techUK response to November 2023 Consultation, p.2; Trustpilot response to November 2023 Consultation, p.4.

[68] Name withheld 5 response to November 2023 Consultation, p.8

[69] Barnardo's response to November 2023 Consultation, p. 7; Eticas.ai response to November 2023 Illegal Harms Consultation, p. 4; NSPCC response to November 2023 Consultation, p.11.

[70] Association of Police and Crime Commissioners response to May 2024 Protection of Children Consultation, p.9; Centre of Excellence for Childrens' Care and Protection response to May 2024 Protection of Children Consultation, p.7.

[71] Barnardo's response to November 2023 Consultation, p.7.

[72] Eticas.ai response to November 2023 Consultation, p.4.

[73] OSAN response to November 2023 Consultation, pp.30-39.

[74] The Global Network Initiative response to November 2023 Illegal Harms Consultation, p.5.

3.81 Having considered these responses, we make the following points:

- **Our view on aligning with other risk assessment processes:** We considered responses on whether the 'suitable and sufficient' requirements could be likened to the standards required by other online safety regimes such as the DSA. The duty to produce a 'suitable and sufficient' risk assessment is a distinct and crucial part of the Act, and the standard comes directly from the legislation. Ultimately, providers' risk assessments must meet the requirements of the Act. While there are similarities between the DSA and the Act, there remain important differences between including the kinds of illegal content they cover, the role of Risk Profiles, and the use of data and information relevant to service users in the UK. This means that it is unlikely that a risk assessment produced solely for the purposes of the DSA would meet the suitable and sufficient requirements in the Act without modification. However, elements of a DSA risk assessment may be relevant to service providers' analysis. Service providers should use our guidance to determine what they need to do to meet their risk assessment duties under the Act.
- **Flexibility of the duty and proportionality:** We welcome service providers' feedback that our approach to what constitutes a suitable and sufficient risk assessment should be proportionate to the service undertaking the assessment. A suitable and sufficient illegal content risk assessment must meet all the requirements of the Act, but we agree that risk assessments will look different between across service providers.
- **How this will be assessed and enforced**: Our Online Safety Enforcement Guidance sets out how Ofcom will approach enforcement of the duties and requirements imposed under the Act and explains how we will exercise our enforcement powers. Attaching specific metrics to how service providers should demonstrate compliance with the illegal content safety duties would risk overcomplicating the definition, and make it harder for service providers to comply. Decisions about whether to take enforcement action are made on a case-by-case basis having regard to our statutory duties and all matters that appear to be relevant, including the priority factors set out at paragraph A3.9 in the Enforcement Guidance.

Our decisions

3.82 We consider that our guidance about what constitutes a 'suitable and sufficient' risk assessment provides enough clarity for service providers to meet this requirement.

3.83 We have improved the consistency of the language regarding what a 'suitable and sufficient' risk assessment is. We have also signposted through the Risk Assessment Guidance to more clearly highlight this. We have done this in the introduction part of the guidance, where we set out services' legal duties. We have also ensured that the language about a 'suitable and sufficient' assessment is consistent throughout the document.

3.84 We think that our changes make the expectations for how service providers will meet this overarching requirement clearer, helping them to meet this duty.

Our reasoning

Benefits and effectiveness

3.85 The duty to produce a 'suitable and sufficient' risk assessment is a distinct legal requirement of the Act. Our guidance for illegal content risk assessments will help service providers be in the best position to meet this obligation, whilst also giving them flexibility in carrying this process.

3.86 Our proposed methodology was met with broad approval after consultation, which we believe was because our guidance enables service providers to take a flexible and proportionate approach whilst also ensuring a baseline for all service providers to meet the requirements in the Act.

Costs

3.87 As this duty is an overarching requirement of risk assessments, any costs associated will be the same as those needed to follow our methodology and produce a risk assessment that meets the legal duties under the Act.

Conclusion

3.88 Given our assessment and reasoning as established above, we believe that our guidance and methodology both follow best practice and aid service providers in meeting the requirement of a 'suitable and sufficient' risk assessment.

3.89 Furthermore, the changes we have made both to our wording and signposting around 'suitable and sufficient' as well as to other elements of the methodology will provide further clarity for service providers on how to meet this requirement.

Reviewing and updating risk assessment

Background

3.90 There are specific duties in the Act for in scope service providers regarding taking appropriate steps to keep their illegal content risk assessment up to date, including the circumstances in which they must update or carry out a new illegal content risk assessment.

3.91 These duties are:

- A duty to take appropriate steps to keep an illegal content risk assessment up to date, including when Ofcom makes any significant change to a Risk Profile that relates to services of the kind in question.
- Before making any significant change to any aspect of a service's design or operation, a duty to carry out a further 'suitable and sufficient' illegal content risk assessment relating to the impacts of that proposed change.[75]

Our proposals

3.92 In our consultation, we included detail on how we expect service providers to meet these requirements. This was included in the final section of the proposed guidance entitled 'When to review or carry out a new risk assessment'.[76]

3.93 This section included a recommendation that service providers hold a written record which included who was responsible for overseeing the risk assessment, and how frequently they planned to renew it (we also recommended that this period to be no more than 12 months).

[75] Section 9(3)-(4) of the Act for U2U service providers and Section 26(3)-(4) of the Act for Search service providers.
[76] Annex 5: Draft Service Risk Assessment Guidance, November 2023 Consultation, pp.45-51.

3.94 In this part of the proposed guidance, we also explained that service providers also have a duty to review their risk assessment if Ofcom makes a significant change to a Risk Profile relevant to them.

3.95 Service providers also have a duty to carry out a new risk assessment before implementing a significant change to any aspect of the service design or operation. To help give service providers clarity about when this duty is triggered, we proposed some draft guidance as to what constitutes a significant change. Given the diversity of service providers in scope of the Act and the fast-moving nature of the technology, it is not possible for us to anticipate every single eventuality that could constitute a significant change. Therefore, we proposed principles-based guidance rather than specifying bright line rules. This can be found at Table 13 of the Draft Risk Assessment Guidance.[77]

3.96 Our proposed guidance to help service providers determine what a significant change is sets out principles and questions to identify if a proposed change is 'significant', and includes some indicative examples.

Stakeholder responses

3.97 Stakeholder responses to our proposals raised the following key themes:

- Regarding the 12-month review timeline we proposed, we received the following feedback:
  - **Support for the guidance:** Stakeholders including Snap and [✂] were generally supportive of the review measures proposed by Ofcom.[78] We Protect Global Alliance was also supportive of the recommendation for service providers to review their assessment at least every 12 months.[79] We proposed the same 12-month review period in our proposed Children's Risk Assessment guidance and several respondents were also supportive of this proposal.[80]
  - **Suggestions of a more frequent review:** However, Barnardo's said that the proposed 12-month update period is not often enough and runs the risk of not capturing emerging risks in a quickly developing environment, and for the same reason Lloyds Banking Group suggested a 6-month timeline for updating the risk assessment.[81] [82]
  - **Disproportionate for some service providers:** Skyscanner criticised the 12-month review period as being disproportionate in the case of low-risk vertical search services. They argued that the nature of this kind of service means that the risk posed to them is unlikely to change quickly and therefore requiring them to review their compliance as frequently as other services would not be proportionate.[83]
- With regard to our guidance on 'significant change' and what meets this criteria, we received the following responses:

[77] Annex 5: Draft Service Risk Assessment Guidance, November 2023 Consultation, pp.48-51.
[78] Snap response to November 2023 Consultation, p.8; Name withheld 5 response to November 2023 Consultation, p.3. [79] We Protect Global Alliance response to November 2023 Illegal Harms Consultation, p.7.
[80] Centre of Excellence for Childrens' Care and Protection response to May 2024 Protection of Children Consultation, p.7; TikTok response to May 2024 Protection of Children Consultation, p.2; Global Network Initiative response to May 2024 Protection of Children Consultation, p.9.
[81] Barnardo's response to November 2023 Consultation, pp.7-8.
[82] Lloyds Banking Group response to November 2023 Consultation, p.7.
[83] Skyscanner response to November 2023 Consultation, p.12.

- **Disagreement with the threshold:** Some stakeholders pointed out that for any of the examples provided by Ofcom, there could be a large spectrum of changes that could occur, most of which should not meet this threshold.[84] Snap argued that a majority of changes to a service will apply to a substantial proportion of a services' user base and so they do not believe this is a reasonable factor in determining whether a change is significant.[85] They suggested that the key factor in determining whether such changes are significant is whether they impact the risk to users from illegal content on the service.
- **Criteria too burdensome:** Some large and smaller stakeholders were also concerned that the approach to significant change is overly burdensome and disproportionate and called for greater flexibility in the implementation of this requirement.[86] For example, Trustpilot argued that this guidance fails to take account of the realities of how online businesses operate and evolve and there is a significant risk that innovation will be stifled, and that the costs of carrying out a risk assessment prior to enacting product changes will outweigh the benefits.[87]
- **Alignment with other regimes:** Meta and Snap both called for Ofcom to pursue harmonisation with the DSA to clarify what might constitute a 'significant change' under the Act.[88] Meta argued that Ofcom's guidance on what constitutes a 'significant change' was novel and overly broad.[89] In contrast, the DSA allows providers to assess this on their own.
- **Requests for clarity on how it is defined:** Various stakeholders called for improved guidance on the definition of 'significant change', challenging the criteria for triggering a new risk assessment.[90] Some stakeholders also pointed out that significant changes set out in the guidance are based on the change having an 'impact' or 'effect', without specifying how material that impact or effect needs to be for the change to qualify as significant.[91] Meta asked for clarity on the threshold for significance on this basis.[92] Mid Size Platform Group also added that any requirement to carry out a new risk assessment should not be a blanket requirement but instead based on how likely the change is to alter risk.[93]
- **How it applies to recommender systems:** We also received comments from some stakeholders regarding significant change relating to recommender systems. Stakeholders expressed the opinion that where a service concludes, after a thorough evaluation of the effects of its recommender system, that the system itself does not

[84] Booking.com response to November 2023 Consultation, p.10; Google response to November 2023 Consultation, pp.15-16; Skyscanner response to November 2023 Consultation, pp.8-9; Trustpilot response to November 2023 Consultation, p.5.
[85] Snap response to November 2023 Consultation, p.7.
[86] Google response to November 2023 Consultation, p.15; Meta response to November 2023 Consultation, pp.11-12; Skyscanner response to November 2023 Consultation, p.8; Trustpilot response to November 2023 Consultation, p.5.
[87] Trustpilot response to November 2023 Consultation, p.5.
[88] Meta response to November 2023 Consultation, pp.11-12; Snap response to November 2023 Consultation, p.8.
[89] Meta response to November 2023 Consultation, p.11.
[90] Meta response to November 2023 Consultation, p.12; Mid Size Platform Group response to November 2023 Consultation, p.12.
[91] Meta response to November 2023 Consultation, p.12; Mid Size Platform Group response to November 2023 Consultation, p.12.
[92] Meta response to November 2023 Consultation, p.12.
[93] Mid Size Platform Group response to November 2023 Consultation, p.12.

materially affect the risks posed on a service, then the requirement to collect safety metrics to assess whether changes to the system increase the risk of illegal content should not apply.[94] Google also echoed the sentiment that proposed platform changes to recommender systems should be assessed under the same parameters as other changes, stating that "The bar should be 'significant change' in both instances".[95]

Our consideration of stakeholder responses

3.98 Our assessment of these points is as follows:

- **The review period is appropriate:** We have considered the feedback relating to our proposed approach to the timing of how frequently reviews should be conducted – either that 12 months is too frequent or too infrequent. We consider that our guidance for service providers to review the assessment at least every 12 months is the appropriate period to meet their legal obligations, as we discuss further in the section entitled 'Benefits and effectiveness'.
- **The further assessment will relate to the change:** We understand that some stakeholders had concerns about the duty to carry out a risk assessment before implementing a significant change to any aspect of its design or operation, as they felt this has the potential to be more burdensome and incur greater costs compared to the above duties to update the risk assessment. We note that this is a direct requirement of the Act. In addition, updating the risk assessment whenever they make a significant change will ensure that their risk assessment remains up to date. This will help service providers manage risk more effectively. We also highlight that the further risk assessment should assess the proposed change specifically, rather than amount to a new assessment of every aspect of the service. We discuss this in further detail in the sections 'Our decisions' and 'Costs'.
- **Best practice in risk management:** We considered responses which pointed to the burden of carrying out a further risk assessment, and those which had concerns that the threshold for a 'significant change' is too low and risks stifling innovation. As we emphasised at the outset of this chapter, illegal content risk assessments are critical to the online safety regime. The adoption of good practice in risk assessment is a legal obligation for service providers and a key component of delivering the wider industrial and cultural change that will put safety at the heart of services' design and decision making. This Act requires service providers to conduct a risk assessment *before* making any significant changes to any aspect of a service's design or operation. This assessment should relate to the impact of the proposed change. However, we are also conscious that the threshold of change which the duty is trying to capture should not be minor updates or create an unmanageable burden for service providers.
- **Flexibility of our guidance:** Related responses calling for more detail as to the material impact or effect of a proposed change would also undermine the flexible approach we have taken for service providers to judge whether a proposed change amounts to a significant change in the Act. Given the diversity of service providers in scope of the Act and the fast-moving nature of technology, it is not feasible for us to anticipate every

[94] [✂]; [✂].
[95] Google response to November 2023 Consultation, p.8.

eventuality that could constitute a significant change. Therefore, we have decided to frame our guidance on significant change in principles-based terms.

- **Significant change under the Act is distinct:** We considered responses which call for harmonisation with the EU's Digital Services Act, which has a similar requirement regarding "deploying functionalities that are likely to have a critical impact".[96] In general, we recognise the importance of international regulatory coherence and support efforts to work towards this where appropriate. However, in this instance, the requirements in the DSA differ from the underlying legal requirements imposed by the Act. Therefore, it is not possible for us to align our guidance or approach to how service providers should meet this duty. While the DSA and OSA are both risk-based, the implementation is very different and service providers in scope of each piece of legislation will need to respond accordingly. Equally, we are bound to develop guidance for the OSA regime, which differs from the DSA legislation.
- **A standard review period is necessary to capture risk:** While a service may not have made a change big enough to amount to a significant change that requires a new risk assessment, several smaller changes to the service design or operation or environment mean that the assessment could be inaccurate and needs review so that it remains suitable and sufficient. Therefore, updating the assessment through regular reviews will improve a service's ability to identify and mitigate new risks and lead to better safety outcomes. If a service was to change significantly then the service would be caught by the requirement to carry out a new risk assessment relating to a 'significant change', which is sufficient to capture risk to users. The risk assessment duties also sit alongside governance Codes, some of which relate to the tracking and management of risks to users.

Our decisions

3.99 We have taken the following decisions around reviewing and updating risk assessments:

- We have recommended that service providers should keep a written policy explaining who is responsible for the risk assessment, how frequently they intend to review their risk assessment, and that risk assessments should be reviewed at least annually. This explanation is in Part 1 of the Risk Assessment Guidance where we explain the specific duties service providers need to meet under the Act.
- We have retained the information on significant change, broadly replicating what we proposed in our November 2023 consultation. It is included in Part 3 of the Risk Assessment Guidance to emphasise that carrying out a new risk assessment relating to the proposed significant change is a duty in itself. Further, we have clarified that when determining whether a proposed change is likely to be significant, service providers should think about the material impact it could have on users, risk and safety. We also explain that we do not intend to capture very minor or routine system changes. We address these decisions in the section entitled 'Benefits and effectiveness'.

3.100 Taking the responses into account, we have added detail to some parts of the guidance to improve clarity regarding how service providers can meet this duty. These have included adding specific metrics for algorithms and algorithmic changes (in line with the updates made to evidence inputs).

[96] Article 34 of the Digital Services Act.

3.101 We have also further emphasised that the duty is to carry out a new risk assessment relating to the proposed change. In practice, this could be at product level but should also consider the impact of the proposed change on wider systems and functions of the service.[97] It is up to service providers to implement this duty in a way which makes sense for their own approach to risk assessment, provided the new risk assessment relating to the proposed change meets all the requirements set out in the Act.

3.102 We have drawn principles and examples from relevant parts of the risk assessment duty 9(5)(a)-(h) and 26(5)(a)-(d), and we have decided to build on this to add clarity to key areas where understanding has developed since consultation. Specifically, regarding algorithmic assessment, we have reviewed language in how we offer guidance to providers regarding the metrics they can gather to assess the role of algorithms on their service.

3.103 Our proposed clarifications affect how we explain how service providers are expected to keep their risk assessment up to date. Our changes in this part of the guidance are about adding clarity for service providers regarding the kinds of things they should consider when deciding if they need to carry out a new risk assessment outside of their usual risk assessment cycle. We think that these changes are minor and will have limited impact on service providers beyond providing additional support to meet their duties regarding keeping a risk assessment up to date.

3.104 Taking account of the consultation responses we received, we remain of the view that the proposed approach works well given the range of service providers in scope and that the policy intent is to support service providers to come to their own conclusions about the potential impact of a proposed change. However, we have improved the clarity of the expectation for how service providers will meet these duties by restructuring how we present the guidance on different duties to review, update or carry out a new risk assessment.

Our reasoning

Benefits and effectiveness

3.105 The Act establishes that service providers have a legal duty to keep risk assessments up to date.[98] The recommendation to update risk assessments at least every 12 months is aligned with service providers' duty under the Act to carry out a Child Access Assessment "not more than one year apart", and also with comparable regimes internationally, such as the Digital Services Act.[99] [100] In addition, it aligns with common governance cycles and accounting periods. If service providers were to leave a longer gap in between completing risk assessments, we believe that there would be a material probability that their risk assessments would become out of date. This would detract from the service provider's ability to identify risks of harm and implement suitable mitigations on their service, thereby causing detriment to users and their safety. Furthermore, the fact that service providers must review and update their risk assessments before implementing a significant change and when there is a change to our risk profiles gives us confidence that the risk assessments

[97] When we use the word 'product' we are using it as an all-encompassing term that includes any functionality, feature, tool, or policy that you provide to users for them to interact with through your service. This includes but is not limited to terms and conditions (Ts&Cs), content feeds, react buttons or privacy settings.
[98] Section 9(3) of the Act for U2U service providers and Section 26(3) of the Act for Search service providers.
[99] Section 36(3) of the Act.
[100] Article 34(1) of the Digital Services Act.

will stay up to date even if they don't review more often than once every 12 months as a matter of course.

3.106 It is also a legal duty under the Act to carry out a further risk assessment before making a significant change to a service.[101] Also, carrying out a further risk assessment before making a significant change will help providers ensure that their risk assessment remains current, thereby improving their ability to manage risks effectively. As we noted in our November 2023 consultation, we consulted with experts internally and externally to help us understand the circumstances under which a change would be significant enough to cause a risk assessment to become out of date and no longer reflect a suitable and sufficient assessment of risk on the service.[102] We believe that our guidance on significant change will help service providers to identify ways to mitigate the risk of harm on their service, and to meet their legal obligations under the Act.

Costs

3.107 We recognise that service providers will incur costs to keep their risk assessments up to date, and that the costs to meet our recommendations are greater than would be the case if we recommended less frequent review. However, to meet this legal obligation under the Act we consider that a 12-month review period is appropriate to maintain a suitable and sufficient risk assessment in a changing risk environment. Further, if over the 12-month period very little has changed about the service the provider is undertaking the assessment for, then the process should be relatively straightforward to comply with.

3.108 We also recognise that updating a risk assessment due to a significant change will carry costs for service providers. However, we highlight that the updated risk assessment will be one that is focused on the change to the service rather than an entirely new one. We consider this to be proportionate to the benefit of helping service providers meet their legal duties and improving safety outcomes for users. We also note the emphasis on the size of a service in our guidance on significant change and the proportionality of this feature for smaller service providers.

Conclusion

3.109 Upon assessment of the stakeholder feedback and our arguments set out above, we maintain that our guidance on reviewing and updating risk assessments is beneficial and appropriate under the Act. We consider that that our recommended timeline for reviewing risk assessments meets best practice and is well suited to help service providers identify and mitigate emerging risks on their service.

3.110 We believe that the modifications we have made will provide clarity for service providers, and specifically that the further guidance we have provided on specific change is proportionate.

Rights impact

3.111 We do not consider that following the Risk Assessment Guidance will have an impact on users' or service providers' rights to privacy or to freedom of expression.

[101] Section 9(4) of the Act for U2U service providers and Section 26(4) of the Act for Search service providers.
[102] Volume 3, November 2023 Consultation, p.84.

3.112 The objective of the four-step methodology is to guide service providers in carrying out a suitable and sufficient risk assessment to meet their legal duties under the Act. Elements such as the Risk Level Tables and guidance on evidence inputs and reviewing and updating risk assessments have been included to aid service providers in assessing the risk of harm on their services and keeping their risk assessments up to date in order to meet their obligations under the Act.

3.113 Our assessment of the rights impact of the measures that will be applied as part of Step 3 of the four-step methodology can be found in the relevant measures sections of this Statement.

Conclusion

3.114 Having reviewed all consultation responses to our proposed Risk Assessment Guidance for Service Providers, we welcome the broad support we have received for our proposed guidance. As a result, we have concluded to broadly retain our approach to the guidance. This includes:

- The four-step methodology.
- Including the Risk Level Tables to help services come to judgements about risk level for each kind of priority illegal content and other illegal content.
- Guidance on 'core' and 'enhanced' evidence.
- What amounts to a 'suitable and sufficient' illegal content risk assessment.
- Guidance on when to review, update or carry out a new risk assessment relating to a proposed significant change.

3.115 As set out in this chapter, where relevant, we have also made changes to clarify our policy intent and expectations for service providers to help them to meet the illegal content risk assessment duties. This includes:

- Clarification on the expectation that services should assess the level of risk as it actually exists on the service, and not 'inherent risk' posed by functionalities and characteristics.
- Improvements to the Risk Level Tables to make them more useful for services as part of their overall illegal content risk assessment.
- Improved explanations of evidence inputs, particularly relating to algorithmic assessment.
- Clarification regarding reviewing and updating duties, particularly relating to the duty to carry out a new risk assessment relating to a proposed significant change to the design or operation of the service (rather than necessarily to the service in its entirety).

4. 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.

Introduction

4.1 Providers of regulated user-to-user (U2U) services and regulated search services are required to keep written records of their risk assessments and the measures taken to comply with some of the new duties and to review compliance regularly. We are required to produce guidance to assist providers with doing so.

4.2 This chapter summarises the main aspects of our Record-Keeping and Review Guidance (RK&RG).[103] It also explains which record-keeping duties the RK&RG covers. There are other record-keeping duties and measures in our Codes of Practice that also entail record-keeping. These are addressed in other regulatory documents we issue, as we explain in this chapter.

Scope of the RK&RG

4.3 The RK&RG covers the record-keeping and review duties that apply to service providers under sections 23 and 34 of the Online Safety Act 2023 (the Act). These are the duties to:

- **keep written records** of the risk assessments and the measures taken to comply with certain duties specified in sections 23 and 34 of the Act, including the illegal content safety duties in sections 10 and 27 of the Act. The specified duties are referred to as 'relevant duties';[104] and
- **review** regularly, and as soon as reasonably practicable after making a significant change to the service, compliance with the duties specified in sections 23 and 34 of the Act. We refer to these as the 'relevant review duties' and they include the duties in section 18 in respect of news publisher content and the duties in sections 71 and 72 in relation to terms of service, in addition to the duties to which the record-keeping duties apply.[105]

4.4 In the RK&RG, we provide guidance on the form that records should take, the matters that they should cover, and when they should be made. We also provide guidance on the frequency with which providers should review compliance with the relevant review duties and the factors that providers should consider when deciding when and whether to conduct a review.

4.5 The RK&RG does not currently include detail on the requirement to keep a written record of children's risk assessments carried out under section 11 or section 28 of the Act, in line with sections 23(2) and 34(2) of the Act. We consulted on our Children's Risk Assessment

[103] Record-Keeping and Review Guidance.

[104] 'Relevant duties' for regulated U2U services means the duties set out in: section 10 (illegal content); section 12 (children's online safety); section 15 (user empowerment); section 17 (content of democratic importance); section 19 (journalistic content); section 20 (content reporting); and section 21 (complaints procedures). 'Relevant duties' for regulated search services means the duties set out in: section 27 (illegal content); section 29 (children's online safety); section 31 (content reporting); and section 32 (complaints procedures).

[105] 'Relevant review duties' for regulated U2U services means the duties set out in: section 10 (illegal content); section 12 (children's online safety); section 15 (user empowerment); section 17 (content of democratic importance); section 18 (news publisher content); section 19 (journalistic content); section 20 (content reporting); section 21 (complaints procedures); section 71 and section 72 (terms of service); and section 75 (disclosure of information about use of service by deceased child users). 'Relevant review duties' for regulated search services means the duties set out in: section 27 (illegal content); section 29 (children's online safety); section 31 (content reporting); section 32 (complaints procedures); and section 75 (disclosure of information about use of service by deceased child users).

Guidance in our May 2024 Protecting Children from Harms Online Consultation.[106] We expect to publish our Statement and regulatory documents on these matters in spring 2025 and will update the RK&RG as appropriate at that time.

4.6 The RK&RG does not cover the record-keeping duties that apply to providers of an online service on which pornographic content is published or displayed by or on behalf of that provider (Part 5 providers).[107] In December 2023 we issued draft Guidance for Service Providers Publishing Pornographic Content Online and expect to publish our Statement and regulatory documents on these matters in January 2025.[108] Record-keeping for Part 5 providers is covered in that Guidance.

4.7 Nor does the RK&RG provide guidance on written records for children's access assessments conducted under section 36 of the Act.[109] In May 2024 we issued our draft Children's Access Assessments Guidance and expect to publish our Statement and accompanying regulatory documents on this in January 2025.[110] Details on making and keeping a written record of each children's access assessment is covered in that Guidance.

4.8 Finally, we are required to issue guidance in respect of other relevant review duties, including the user empowerment duty in section 15, the news publisher content duty in section 18, the terms of service duties in sections 71 and 72, and the duty in section 75 requiring the disclosure of information about the use of service by a deceased child. If we consider it appropriate, we will revise the RK&RG in respect of these relevant review duties.

Structure of this chapter

4.9 In this chapter we summarise the proposals regarding the record-keeping and review duties that we consulted on in our November 2023 Illegal Harms Consultation, and our rationale for these proposals. Where we received significant stakeholder feedback on specific aspects of our proposals, we summarise the feedback in the relevant sections. Other stakeholder feedback regarding the RK&RG is available in our annex of further stakeholder responses.[111] We then outline the decisions we have made, having considered stakeholder responses to our proposals, in finalising and adopting the RK&RG.

General approach of the RK&RG

Our proposals

[106] See May 2024 Consultation: Protecting children from harms online and draft Children's Risk Assessment Guidance.
[107] Part 5 providers are subject to the duties in Part 5 of the Act. These include duties to ensure that children are not normally able to encounter pornographic content on their online service and to keep a record of the kinds of age verification or age estimation used and how they have been deployed.
[108] See draft Guidance on age assurance and other Part 5 duties for service providers publishing pornographic content on online services published December 2023.
[109] Section 36(7) of the Act requires in-scope providers to make and keep a written record, in an easily understandable form, of every children's access assessment.
[110] See draft Children's Access Assessments guidance published May 2024.
[111] Annex 1: Further stakeholder responses.

4.10 Our approach in the Consultation was to ensure that the RK&RG was self-explanatory and straightforward by explaining, in line with the requirements of the Act, what we consider to be appropriate record-keeping.

Stakeholder responses

4.11 Consultation feedback included various positive comments about the record-keeping and review duties and our approach to such; a number of stakeholders acknowledged the importance of such duties and expressed broad support for the "reasonable" approach taken in the RK&RG.[112]

4.12 We also received responses to the record-keeping and review duties regarding: the provision of support for service providers; the need for flexibility; ensuring that the RK&RG does not result in an unnecessary burden on service providers; and the need to align with other regulatory regimes where possible.

4.13 **Supporting service providers:** Several stakeholders highlighted the importance of us providing support in the form of examples, tools, and other additional resources, particularly for smaller service providers.[113]

4.14 **Taking a flexible approach:** We received several responses requesting that we take a flexible and adaptable approach to the record-keeping and review duties, including after implementation of the RK&RG.[114] On the other hand, Protection Group International said there should be an agreed format for record-keeping to ensure a consistent approach to records.[115]

4.15 **Ensuring no unnecessary burden**: Several respondents made the point that we should continue to consider the potential effect of its expectations regarding record-keeping, particularly in the context of smaller service providers, to ensure that the RK&RG is practical and proportionate, and the burden is reduced where possible.[116]

4.16 **Alignment with other regulatory regimes:** Some respondents highlighted that regulatory regimes that have requirements relating to record-keeping or review duties already exist. Respondents requested that we take into consideration existing relevant regulatory regimes and align our RK&RG where possible.[117] This feedback was received specifically in response to our consultation questions about the RK&RG, as well as being a theme of the consultation responses generally.

---

[112] For example: LinkedIn response to November 2023 Illegal Harms Consultation, p.6; [✂]; Nexus response to November 2023 Illegal Harms Consultation, p.6; We Protect Global Alliance response to November 2023 Illegal Harms Consultation, p.7.

[113] British and Irish Law, Education, and Technology Association (BILETA) response to November 2023 Illegal Harms Consultation, pp.4-5; Federation of Small Businesses response to November 2023 Illegal Harms Consultation, p.2; Mid Size Platform Group response to November 2023 Illegal Harms Consultation, p.11; TechUK response to November 2023 Illegal Harms Consultation, p.15.

[114] Mid Size Platform Group response to November 2023 Consultation, p.12; Oxford Disinformation and Extremism Lab response to November 2023 Illegal Harms Consultation, p.5; TechUK response to November 2023 Consultation, p.2.

[115] Protection Group International response to November 2023 Illegal Harms Consultation, p.4.

[116] ACT The App Association response to November 2023 Illegal Harms Consultation, p.7; Mid Size Platform Group response to November 2023 Consultation, p.12; TechUK response to November 2023 Consultation, p.15; Trustpilot response to November 2023 Illegal Harms Consultation, p.10.

[117] Federation of Small Businesses response to November 2023 Consultation p.2; Skyscanner response to November 2023 Illegal Harms Consultation, p.11; TechUK response to November 2023 Consultation, p.15.

Our decisions

4.17 We received positive feedback regarding our approach and did not receive any comments from stakeholders that suggested taking an alternative approach. Therefore, we have maintained our overall approach for the RK&RG.

4.18 **Supporting service providers:** We have an extensive programme of work to make the regulations accessible, and compliance more easily attainable, for all providers of online services that fall within scope of the Act, which include many small or medium-sized enterprises (SMEs). Specifically, alongside this Statement we plan to launch a new 'Digital Support Service', which consists of interactive digital tools for regulated providers based on their perspectives and feedback.[118] Our first release will provide a four-step process for illegal harms, covering services' risk assessment duties, Codes, and record-keeping obligations; we intend the Digital Support Service to specifically cover record-keeping in relation to illegal content risk assessments.

4.19 We consider that the RK&RG, along with the Digital Support Service we are developing, will provide the necessary support for all in-scope service providers to be able to understand and action the relevant record-keeping duties. We plan to continually develop additional resources to support compliance.

4.20 **Taking a flexible approach:** We agree that taking a flexible approach is important to ensure that the RK&RG has relevance for the diverse range of services that are in-scope of the Act. Therefore, in preparing the RK&RG, we have sought to avoid being unduly prescriptive about aspects of record-keeping, such as by specifying a particular format for records of risk assessments. We consider that our approach with the RK&RG is flexible and note that the feedback emphasised the importance of flexibility, rather than raising a concern that our approach was not flexible enough.

4.21 We are balancing the need for flexibility with the need to provide support through the provision of our Digital Support Service, as explained in paragraph 4.18; we will adapt the Digital Support Service, which will cover record-keeping, to the needs of in-scope service providers over time, based on their feedback. We consider that our flexible approach should enable service providers to use their established record-keeping processes and procedures to fulfil their obligations under the Act with minimal adjustment.

4.22 **Ensuring no unnecessary burden**: We have taken account of the fact that the RK&RG will apply to a wide range of services and, in line with our general duties, the need for it to be proportionate. The comments received on this topic did not say that our proposed RK&RG was unduly burdensome, but that we should continue to take the burden on service providers into account. Further, the responses regarding our RK&RG did not provide any substantive evidence to demonstrate that our approach is disproportionately burdensome. Accordingly, we remain satisfied that the RK&RG is proportionate and should not place undue burden on stakeholders.

4.23 **Alignment with other regulatory regimes:** We are required to produce guidance that sets out the duties specified in the Act. As a result, the RK&RG reflects the requirements of the Act and our assessment of the approach that is needed to meet these duties. We have

[118] The Digital Support Service will be accessible via the Ofcom website. We plan to communicate the launch through our usual channels, so will be able to update service providers then. We will publicise the service to providers when it is available to use.

sought to align with other regulatory regimes where it makes sense to do so, although differences in regimes limit how far we can do this.

4.24 For this reason, we have adjusted the retention period we expect for record-keeping, which we explain in more detail in the next section of this chapter. We consider that there are not sufficient advantages to justify a longer retention period for records made and kept under the Act, when compared to the advantages of simplifying the regulatory burden for service providers by aligning with requirements in other relevant regimes where appropriate, such as the EU's Digital Services Act (DSA). As a result of consultation feedback, we have adjusted our guidance regarding record retention to better align with the requirements of the DSA; namely, recommending a **three-year** retention period for record-keeping.

Over-arching guidance for records

Our proposals

4.25 Well-maintained, clear records, and regular, timely reviews will assist service providers in keeping track of compliance with the relevant duties and ensure that the measures taken are fit for purpose. The records will also provide a useful resource for us to monitor how the relevant duties are being fulfilled. Accordingly, for records made and kept under the Act, we proposed that:

- written records should be durable, accessible, easy to understand, and up to date;
- where reasonably practicable, records should be kept in English (or for providers based in Wales, in English or in Welsh). If this is not reasonably practicable, it must be possible for an English translation of the records to be provided;
- records should be updated to capture changes to a risk assessment or Code or alternative measure, but earlier versions should be retained so the provider is able to produce both current and historic records of how it has complied with the relevant duties; and
- records that are no longer current but have not been provided to Ofcom should be retained for a minimum of five years (or for the duration of the provider's record retention policy if that was the greater length of time).

Stakeholder responses

4.26 Several respondents recommended that we make clear in the RK&RG our expectations regarding the format of documents made and kept as part of the record-keeping and review duties. Specifically: Alliance to Counter Crime Online suggested we emphasise that records should be easily searchable and accessible;[119] We Protect Global Alliance recommended that we specify minimum durability standards for records;[120] and Yoti suggested that we clarify what we mean by 'durable' and 'easy to understand'.[121]

4.27 We proposed that records that had not been provided to Ofcom should be retained for a minimum of five years. We received contradicting feedback that this timeframe should be

[119] Alliance to Counter Crime Online response to November 2023 Illegal Harms Consultation, p.4.
[120] We Protect Global Alliance response to November 2023 Consultation, p.7.
[121] Yoti response to November 2023 Illegal Harms Consultation, p.8.

either shortened or lengthened.[122] Snap queried the rationale for the five-year retention period and highlighted that a five-year retention period was not aligned with the requirements of the EU's Digital Services Act (DSA).[123]

Our decisions

4.28 Regarding consultation feedback about the format of documents, we consider that we have made it sufficiently clear in the RK&RG that records should be durable, accessible, easy to understand, and up to date; this includes providing several examples of possible durable mediums and explaining what we mean by 'easy to understand' and 'up to date'.[124]

4.29 As set out in paragraph 4.24 of this chapter, we have decided that a **three-year** retention period for records is appropriate. This should still allow for records to be available if retrospective problems are identified. It should also ensure that records are available for providers to show how they have responded to the evolution of risks over time. As such, the RK&RG sets out that written records which are no longer current but have not already been provided to Ofcom should be retained for a minimum of three years (or in accordance with the service provider's record retention policy, where this is longer).

Record-keeping: risk assessments

Our proposals

4.30 Service providers are required to make and keep a written record of all aspects of every risk assessment they carry out, including details about how the assessment was carried out and its findings. The record needs to include how the provider has considered the required elements in section 9 or section 26 (as applicable) of the Act, and the evidence the provider has relied on to assess the risks posed by the provider's service.

4.31 At Consultation, we aligned our guidance on making and keeping a written record of all risk assessments with our draft Service Risk Assessment Guidance.[125] We considered that the record of a risk assessment should help service providers to demonstrate how their risk assessments are suitable and sufficient, as required by the Act. To achieve this, we set out a list of information that should be included in any risk assessment record.[126]

4.32 We also proposed that written records of risk assessments (or revisions to the assessment) should be made at the same time as the assessment is conducted, to ensure that the record is up to date and accurate.

Stakeholder responses

4.33 Meta queried how confidentiality would apply to records of risk assessments.[127] We understood this to be primarily in the context of providers of Category 1 U2U services and

[122] Centre for Competition Policy response to November 2023 Illegal Harms Consultation, p14; Evri response to November 2023 Illegal Harms Consultation, p.3.
[123] Snap response to November 2023 Illegal Harms Consultation, p.8.
[124] See paragraphs 2.1 to 2.7 of the Record-Keeping and Review Guidance.
[125] See Annex 5: Draft Risk Assessment Guidance published November 2023.
[126] See section 9(5) or section 26(5) of the Act (as applicable) for the list of matters that need to be assessed for an illegal content risk assessment.
[127] Meta response to November 2023 Illegal Harms Consultation, p.13.

Category 2A search services being required to provide written records of their risk assessments, in full, to Ofcom.[128]

4.34 We also received feedback emphasising the need for transparency regarding records made and kept as part of the record-keeping duties, which could be achieved by service providers making their records available to regulators or the public. One comment was specific to records of risk assessments, whereas other feedback was also applicable to other records made and kept as part of such duties:

- CELE said that documents resulting from the record-keeping and review duties should be made public.[129]
- An individual respondent and Yoti stated that there should be transparency regarding records made to fulfil the record-keeping duties, particularly so that regulators could review such records.[130]
- National Trading Standards eCrime Team recommended that service providers should publish a summary of their risk assessments and measures taken to demonstrate to the public how online harms are being dealt with.[131]

Our decisions

4.35 The RK&RG includes specific guidance about the records that service providers must make and keep in relation to risk assessments. As we explain in paragraph 4.5 of this chapter, the RK&RG does not currently contain specific information on record-keeping in relation to children's risk assessments at this time; we will update it when we publish the Protecting Children from Harms Statement in spring 2025.

4.36 For illegal content risk assessments, the RK&RG should be read along with our guidance on illegal content risk assessments.[132] Our decisions specific to our guidance on illegal content risk assessments are covered in Chapter 3: Risk Assessment Guidance for Service Providers.[133]

4.37 Regarding record-keeping of risk assessments, we have not substantially altered our proposals. Therefore, in relation to a record of an illegal content risk assessment made under section 9 or 26, we consider that it should cover the matters set out at paragraphs 3.2, 3.10, and 3.11 of the RK&RG, which are the same matters as those set out in our Risk Assessment Guidance and Risk Profiles document. We set out our rationale for a record of an illegal content risk assessment to include such areas as part of our November 2023 Illegal Harms Consultation.[134]

---

[128] Section 23(10) and section 34(9) of the Act as applicable. Category 1 U2U services and Category 2A search services are services that Ofcom considers meet the applicable threshold conditions set out in regulations to be made by the Secretary of State under Schedule 11 of the Act and that are entered into a public register to be kept by Ofcom under section 95 of the Act.

[129] Centro de Estudios en Libertad de Expresion y Acceso a la Informacion (CELE) response to November 2023 Illegal Harms Consultation, p.6.

[130] Are, C response to November 2023 Illegal Harms Consultation, p.5; Yoti response to November 2023 Consultation, p.8.

[131] National Trading Standards eCrime team response to November 2023 Illegal Harms Consultation 2023, p.4.

[132] Risk Assessment Guidance and Risk Profiles.

[133] Chapter 3: Risk Assessment Guidance for Service Providers.

[134] See Table 9.2, Section 9 of Volume 3: How should services assess the risk of online harm? pages 52 to 56.

4.38 We have maintained our expectation that the written record of a risk assessment, or any revision to it, should be made contemporaneously, to ensure that it is accurate and up to date.

4.39 Regarding the confidentiality of records of risk assessments, section 393(1) of the Communications Act prohibits the disclosure of information relating to a business, such as records of risk assessments, that Ofcom has obtained as a result of its powers under the Online Safety Act. However, this prohibition is subject to the gateways in section 393(2) of the Communications Act, which enable the disclosure of such information in certain circumstances. This includes a disclosure that Ofcom makes for the purpose of carrying out its functions. Information covered by the disclosure provisions of section 393 includes information relating to a business that Ofcom has obtained as a result of its powers.

4.40 Where we consider it necessary to disclose information contained in a risk assessment to carry out our functions, we will generally redact confidential information or exclude it from the disclosure. Occasionally, disclosure of confidential information may be appropriate. Where this is the case, we will take reasonable steps to inform the provider and give it a reasonable opportunity to make representations before making a final decision on whether to disclose the information.

4.41 In response to consultation feedback that service providers should be transparent about their records, particularly by providing them to regulators or making records publicly available, we note that we can obtain such records through our information gathering powers if necessary.[135] In the RK&RG, we highlight that the Act requires providers of Category 1 U2U services and Category 2A search services to provide Ofcom with a copy of written records of risk assessments, in full, as soon as is reasonably practicable.[136] These should be sent to Ofcom's dedicated email address (as published on our website at the time of submission) as soon as the risk assessment, or any revision to it, is concluded.

4.42 We have also added to the RK&RG a reference to the following duties:

- providers of Category 1 U2U services have a duty to summarise in their terms of service the findings of the most recent illegal content risk assessment of a service (including as to levels of risk and as to nature, and severity, of potential harm to individuals);[137]
- Category 2A search services have a duty to publish a summary of the findings of their most recent illegal content risk assessment (including as to levels of risk and as to nature, and severity, of potential harm to individuals).[138]

4.43 We consider that explaining in the RK&RG what must be made public according to the Act is helpful for stakeholders to understand what aspects of service providers' records will be publicly available. We also consider that Category 1 and 2A service providers may find it useful, when reviewing the RK&RG on record-keeping in relation to risk assessments, to be

[135] See our July 2024 Consultation on Online Safety Information Guidance for more detail on our information gathering powers. Our final Online Safety Information Guidance will be published in early 2025.
[136] Section 23(10) and section 34(9) of the Act as applicable. Category 1 U2U services and Category 2A search services are services that Ofcom considers meet the applicable threshold conditions set out in regulations to be made by the Secretary of State under Schedule 11 of the Act and that are entered into a public register to be kept by Ofcom under section 95 of the Act.
[137] Section 10(9) of the Act.
[138] Section 27(9) of the Act.

reminded of this associated duty. We anticipate publishing draft proposals regarding the additional duties on categorised services no later than early 2026.[139]

Record-keeping: Code measures

Our proposals

4.44 Providers are required to keep a written record of any measures taken to comply with the relevant duties described in a Code of Practice issued by Ofcom under section 41 of the Act.[140]

4.45 In the draft Guidance, we set out the information that we proposed should be included in addition to the record of each Code measure taken or in use, including the relevant Code of Practice and the date on which the measure takes effect. We also proposed that the record should be made promptly, to ensure that it is up to date and accurate.

Stakeholder responses

4.46 We did not receive significant feedback regarding our guidance on making and keeping records of Code measures.

Our decisions

4.47 Service providers must make and keep a written record of each measure that is taken or in use as described in the Code of Practice. As set out in the RK&RG, this should:

a) provide a description of the measure in question;
b) identify the relevant Code of Practice; and
c) give the date that the measure takes effect.

4.48 Where a measure says a document should be made, or information recorded, it should be kept and maintained in accordance with our RK&RG, as part of the record made for the purposes of the record-keeping duty under section 23(3) or section 34(3) of the Act.[141] We consider that this is a practical recommendation that ensures all relevant information is kept and maintained in a consistent manner.

4.49 To ensure that records are accurate and up to date, the written record of a Code measure should be made promptly after the measure has been taken or, where the measure is already in effect prior to the relevant duty coming into force, promptly after this date.

[139] See our October 2024 approach to implementing the Act update for information on expected timings regarding the additional requirements that fall on categorised services.
[140] 'Relevant duties' for regulated U2U services means the duties set out in: section 10 (illegal content); section 12 (children's online safety); section 15 (user empowerment); section 17 (content of democratic importance); section 19 (journalistic content); section 20 (content reporting); and section 21 (complaints procedures). 'Relevant duties' for regulated search services means the duties set out in: section 27 (illegal content); section 29 (children's online safety); section 31 (content reporting; and section 32 (complaints procedures).
[141] The record of the Codes measures should be durable, accessible, easy to understand, and up to date.

Record-keeping: alternative measures

Our proposals

4.50 It is open to a service provider to take alternative measures to those described in a Code of Practice, to comply with a relevant duty. In line with the Act, our draft RK&RG set out that, where a provider opts to take an alternative measure, its record of the measure in question must cover the matters specified in section 23(4) and (5) and section 34(4) and (5) of the Act (as applicable).

4.51 We consider that the matters specified in the Act as needing to be covered are clearly articulated. As such, the draft RK&RG specified what the record of alternative measures should include as set out in the Act. We also proposed that the record should be made promptly, to ensure that the record is accurate and up to date.

Stakeholder responses

4.52 We did not receive significant feedback regarding our guidance on making and keeping records of alternative measures.

Our decisions

4.53 The RK&RG provides that a written record of each adopted alternative measure should include, as set out in the Act:

a) the applicable measures in a Code of Practice that have not been taken or are not in use;[142]
b) the alternative measures that have been taken or are in use;
c) how those alternative measures amount to compliance with the duty in question; and
d) how the provider has complied with section 49(5) (freedom of expression and privacy).

4.54 For evidencing or reviewing compliance, service providers should record the date that the alternative measure(s) came into effect.

4.55 To ensure that such records are accurate and up to date, we have maintained our position that the written record of alternative measure should be made promptly after the measure has been taken or, where the measure is already in effect prior to the relevant duty coming into force, promptly after this date.

Review duties

Our proposals

[142] These are measures set out by Ofcom in a Code of Practice which apply to the relevant service provider. There is no obligation on a service provider to keep a written record of a measure (from the Code of Practice) that does not apply to it (for example, where particular measures only apply to a subset of services based on size or risk of a particular harm).

4.56 Service providers are required to review compliance with the relevant review duties regularly, and as soon as reasonably practicable after making any significant change to any aspect of the design or operation of the service.[143]

4.57 We proposed general guidance on matters that service providers should consider when conducting a review, such as reviewing whether the measures in place are sufficient to secure compliance with the relevant online safety duties. The aim of our proposals was to ensure the effectiveness of compliance reviews conducted in line with the RK&RG.

4.58 We proposed that, as a minimum, service providers should conduct a compliance review at least once a year. We considered that this aligned with the frequency of the annual and financial reporting cycle and the guidance we issued on the frequency with which providers should undertake their risk assessments.[144]

4.59 We provided guidance on what constitutes a significant change by reference to our draft Service Risk Assessment Guidance.[145] We considered that this latter Guidance would likely also be of assistance to service providers when considering whether there is a significant change that may affect a provider's compliance with other relevant review duties.

Stakeholder responses

4.60 We proposed that compliance reviews should be undertaken on an annual basis as a minimum. We received contradictory feedback that this may be too frequent or not frequent enough.[146] Both Google and Mid Size Platform Group expressed concern that it could be disproportionate for all service providers to have to review their risk assessments every 12 months.[147]

4.61 We received requests for further clarity regarding our definition of 'significant change', as well as comments challenging the criteria we used to define when a change is 'significant', particularly in the context of our expectation that a provider should review its compliance whenever there is a significant change to its service.[148]

Our decisions

4.62 We have maintained our decision that providers should conduct annual compliance reviews. As explained at Consultation, conducting reviews on a yearly basis is in line with the financial and annual reporting cycle for UK companies, as well as being in line with our

[143] Section 23(6) or 34(6) as applicable.
[144] See Annex 5: Draft Service Risk Assessment Guidance published November 2023, page 34, paragraph A5.96. Service providers required to complete a Children's Access Assessment must also redo these at least every 12 months (see section 36(3) of the Act).
[145] See Annex 5: Draft Service Risk Assessment Guidance published November 2023, pages 48 to 51, paragraphs A5.132 to A5.135 and Table 13 for guidance on significant change.
[146] Are, C response to November 2023 Consultation, p.5; Match Group response to November 2023 Illegal Harms Consultation, pp.6-7; SkyScanner response to November 2023 Consultation, p.11; TechUK response to November 2023 Consultation, p.15.
[147] Google response to November 2023 Illegal Harms Consultation, p.25; Mid Size Platform Group response to November 2023 Consultation, p.12.
[148] BILETA response to November 2023 Consultation, pp.4-5; Booking.com response to November 2023 Illegal Harms Consultation, p.5; DuckDuckGo response to November 2023 Illegal Harms Consultation, p.6; Mid Size Platform Group response to November 2023 Consultation, p.12.

guidance on the frequency that a provider should review its risk assessments.[149] As such, we consider that it is reasonable to expect providers to conduct reviews on an annual basis. We acknowledge that this will require specific resource, however, we do not consider that this is unduly burdensome, particularly given that conducting reviews is a duty under the Act.

4.63 We make it clear in the RK&RG that a yearly review is our minimum expectation. Where a service provider becomes aware of compliance concerns, or where a provider implements new measures to comply with a relevant review duty, it may be appropriate to conduct earlier or more frequent reviews. We therefore consider that the RK&RG also addresses the concern that a year may be too long between compliance reviews as we make it clear that, in practice, reviews may need to be completed on a more regular basis.

4.64 Our Risk Assessment Guidance and Risk Profiles regulatory document sets out our expectation that service providers should review and update their risk assessment at least every 12 months, as well as setting out our definition of what constitutes a significant change.[150] We provide an explanation for our decisions regarding how often risk assessments should be reviewed and the definition of significant change in chapter 3: Risk Assessment Guidance for Service Providers.[151]

4.65 As noted, we will be producing guidance on other relevant review duties, such as the user empowerment duty and the news publisher content duty, and therefore may issue further guidance on the question of when there is a significant change for the purposes of the review duties in due course.

Exemption from relevant duties

Our proposals

4.66 Ofcom may (under sections 23(7) and (8) and sections 34(7) and (8)) exempt certain types of service providers from any or all the record-keeping and review duties which are the subject of the RK&RG.

4.67 At consultation, we proposed not to exempt any types of service from the record-keeping and review duties and asked respondents to confirm whether or not they agreed with our position.[152]

Stakeholder responses

[149] See 'Part 1: Duties and carrying out an illegal content risk assessment' in the Risk Assessment Guidance and Risk Profiles document, specifically the section titled 'Review and update at least every 12 months'. Service providers required to complete a Children's Access Assessment must also redo these at least every 12 months (see section 36(3) of the Act).

[150] See 'Part 3: Supporting documents' of the Risk Assessment Guidance and Risk Profiles, specifically '4. Making a significant change to your service'.

[151] See the 'Reviewing and updating risk assessment' section in Chapter 3: Risk Assessment Guidance for Service Providers, particularly the 'Our consideration of stakeholder responses' and 'Our decisions' sub-sections.

[152] See Volume 3: How should services assess the risk of online harm? published November 2023, page 93, paragraphs 10.27 to 10.29, for detail on our reasoning.

4.68 We asked respondents whether they agreed with our proposal not to exercise our power to exempt specified descriptions of services from the record-keeping and review duties for the time being. Of those who responded to this question, the majority agreed with our proposal.[153] However, a minority of respondents disagreed.[154] We also received a suggestion that there should be a staggered approach to implementation, with larger and high or multi risk service providers being required to implement the record-keeping and review duties before smaller service providers.[155]

Our decisions

4.69 We note that the majority of respondents agreed with our proposal not to exercise our exemption powers regarding record-keeping or review duties. We remain of the view that we should not exempt any types of service from the record-keeping and review duties at this time, notwithstanding the feedback from a minority of stakeholders that they disagreed with our decision. This is because we do not have a sufficiently strong evidence base to justify any such exemptions.[156] Our analysis of risk across the sector indicates a wide variation in levels of risk across services, independent of type, which we consider would make an exemption by type of service difficult to implement in a reasonable and proportionate way.

4.70 We have considered whether there may be an undue burden arising from the record-keeping and review duties, particularly for smaller or lower risk service providers. However, we do not consider that the record-keeping and review duties are onerous, given the proportionate approach we have set out in the RK&RG. We are also mindful of the importance of the risk assessment duties to the regulatory regime and hence the importance of having a record to demonstrate that a service provider's risk assessment is suitable and sufficient.

4.71 Finally, we note that the underlying duties to conduct risk assessments and take measures to comply with the relevant duties would not be removed by any exemption. We consider

---

[153] 5Rights Foundation response to November 2023 Illegal Harms Consultation, p.17; ACT The App Association response to November 2023 Consultation, p.7; Are, C response to November 2023 Consultation, p.5; Betting and Gaming Council response to November 2023 Illegal Harms Consultation, p.4; Center for Countering Digital Hate (CCDH) response to November 2023 Illegal Harms Consultation, p.7; Centre for Competition Policy response to November 2023 Consultation, p.14; [✂]; Institute for Strategic Dialogue response to November 2023 Illegal Harms Consultation, p.7; Logically response to November 2023 Illegal Harms Consultation, p.14; Match Group response to November 2023 Consultation, p.7; Mencap response to November 2023 Consultation, p.5; National Trading Standards eCrime team response to November 2023 Consultation, p.5; Nexus response to November 2023 Consultation, p.7; OnlyFans response to November 2023 Illegal Harms Consultation, p.3; Protection Group International response to November 2023 Consultation, p.4; RSPCA response to November 2023 Illegal Harms Consultation, p.3; Safe Space One response to November 2023 Illegal Harms Consultation, p.6; SafeCast (1) response to November 2023 Illegal Harms Consultation, p.4; Segregated Payments Ltd response to November 2023 Illegal Harms Consultation, p.4; Snap response to November 2023 Consultation, p.8; The Board of Deputies of British Jews response to November 2023 Illegal Harms Consultation, p.5; The Cyber Helpline response to November 2023 Illegal Harms Consultation, p.6.

[154] Bolton, C response to November 2023 Illegal Harms Consultation, p.3; Dwyer, D response to November 2023 Illegal Harms Consultation, p.3; Evri response to November 2023 Consultation, p.3; Name withheld 3 response to November 2023 Illegal Harms Consultation, p.6; Mid Size Platform Group response to November 2023 Consultation, p.12; SkyScanner response to November 2023 Consultation, p.11; Wikimedia Foundation response to November 2023 Illegal Harms Consultation, pp.13-14.

[155] Federation of Small Businesses response to November 2023 Consultation, p.2.

[156] As part of this, we considered exempting small or micro businesses but did not have sufficiently strong evidence to understand the potential effects if such businesses provide higher risk services.

that it is good practice for all service providers to keep written 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. We will keep our position on exemptions under review.

Failure to comply with relevant duties

Our proposals

4.72 Our draft RK&RG highlighted that compliance with the record-keeping and review duties is an enforceable requirement under the Act, so that providers would understand the importance of compliance and the possible consequences of failing to do so.

Stakeholder responses

4.73 We received feedback that setting out the consequences of non-compliance may incentivise adherence.[157] We also received a request for more detail on how we intend to regulate the record-keeping and review duties.[158]

Our decisions

4.74 Our draft RK&RG stated that the record-keeping and review duties are enforceable by Ofcom. In response to respondent feedback, we have included in the RK&RG more detail about what the possible consequences of non-compliance could be, to incentivise compliance with the relevant duties.

4.75 In response to requests for information regarding how we will enforce under the Online Safety Act, we have signposted our Online Safety Enforcement Guidelines, which is one of the regulatory documents published with this Statement and which provides detail on how we will enforce under the Act.[159]

Implementation

4.76 The record-keeping and review duties take effect at the same time as the obligations to conduct illegal content risk assessments and the obligations to comply with the relevant duties (for the purpose of the record-keeping duties) and the relevant review duties (for the purpose of the review duties).[160]

4.77 Regarding Category 1 U2U services and Category 2A search services, as soon as reasonably practicable after making a written record of the first illegal content risk assessment, such regulated providers are required to provide this written record (in full), and in an electronic format, to Ofcom.[161] The additional requirements for categorised services will only come

[157] BILETA response to November 2023 Consultation, pp.4-5.
[158] The Cyber Helpline response to November 2023 Consultation, p.6.
[159] See Online Safety Enforcement Guidance.
[160] Note that these duties are subject to different contingencies and so are likely to take effect on different dates.
[161] This also includes any subsequent changes to the service provider's written records.

into effect after the thresholds for categorisation have been confirmed in regulations to be made by the Secretary of State, under Schedule 11 of the Act.

Impact assessment

4.78 Ofcom is required by legislation to provide guidance on the record-keeping and review duties. As we have discretion over the nature of this guidance, we have carried out an impact assessment, as defined in Section 7 of the Communications Act (2003).

4.79 To the extent that the RK&RG results in additional costs to those necessarily incurred by service providers in fulfilling their statutory duties and ensuring ongoing compliance, we consider that such costs are minimal and outweighed by the regulatory benefits of ensuring the availability of clear, well-maintained records, and timely reviews of compliance. For the same reason, we consider that, to the extent that our decision not to grant any exemptions from the record-keeping and review duties creates burdens, these are outweighed by the benefits of maintaining the application of the duties. In any event, as set out in paragraph 4.69, at this stage of the new regime, we do not have a sufficient evidential basis for exercising our power to grant an exemption.

5. 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	Service 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	Service 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	Service 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	Service 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	Service 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 and multi-risk services.
ICU A7/ ICS A7	Service 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.

Introduction

5.1 Under the Online Safety Act 2023 ('the Act'), service providers have duties to carry out risk assessments, to keep a written record of how risk assessments are carried out, and to

regularly review compliance with their safety duties and reporting and complaints duties.[162] Strong governance processes help service providers meet these duties by helping ensure that their organisational structure facilitates effective risk monitoring, identification, and management.

5.2 Effective governance and accountability structures provide the foundation for service providers to identify, manage, and review illegal harms risks to their users. By embedding principles like accountability, oversight, independence, transparency and clarity of purpose into their operations, we expect providers to have 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 strategic decision making. This will assist providers in implementing appropriate risk mitigations. These governance and organisational design processes should also help providers in preparing to deal with changes in the online landscape that may increase risks to users, including sudden spikes in illegal content and sensitive events, as well as monitoring and reviewing the effectiveness of measures designed to reduce risk. In this way, governance and organisational design should be seen as a fundamental part of ongoing risk management.

5.3 This chapter sets out our decisions on the governance and accountability measures ('governance measures') to be included in the Illegal Content Codes of Practice ('Codes'). It details the responses received to our November 2023 Illegal Harms Consultation ('November 2023 Consultation') and outlines how we have reached our final decisions.[163] We first summarise and consider the responses we received on our overall approach to governance and then the responses received on individual governance measures. We expect the governance measures presented in this chapter to lead to effective risk management processes that set company-wide standards for high-quality risk assessments.[164]

5.4 In our May 2024 Consultation on Protecting Children from Harms Online ('May 2024 Consultation'), we consulted on separate proposed governance measures. In this chapter, we have included and considered some responses to our May 2024 Consultation, where they are also relevant to the governance measures from the November 2023 Consultation.

Our approach

5.5 In the November 2023 Consultation, we explained that good practice in governance processes is multi-faceted. This means that there is no single governance and accountability intervention that can manage the possible complexities of risk mitigation and management. As such, we considered that it would be appropriate to recommend more than one governance measure for some service providers. This is to ensure these providers can provide the more complex oversight needed to address the harms and risks on their services, and to ensure that illegal content risks are mitigated and managed effectively.

---

[162] Sections 9, 26, 23(2), 23(6) and 34(6) of the Act.
[163] Ofcom, 2024. Consultation: Protecting people from illegal harms online.
[164] Assessments should be specific to the service and reflect the risks accurately to ensure that they are "suitable and sufficient". Service providers have a duty to keep a record of this. Implementing such measures will make providers better equipped to design and organise their services in a way that helps to effectively mitigate against illegal harms. See 'Risk Assessment Guidance' and 'Record-Keeping and Review Guidance'.

5.6 Our measures are supported by a review of good practice standards and principles in risk management and corporate governance across a range of different industries.[165]

5.7 In our Consultations, we structured our proposed recommendations into thematic areas, ranging from the overall governance structure to policies for individual staff members. The measures we proposed in the November 2023 Consultation were organised into four thematic areas.

a) Annual review of risk management activities.

   i) Boards or overall governance bodies should carry out an annual review and record how the provider has assessed risk management activities in relation to illegal harms, and how developing risks are being monitored and managed. This would apply to providers of large user-to-user ('U2U') services and providers of large general search services.[166]

b) Senior accountability and responsibility.

   i) A named person should be accountable to the most senior governance body for compliance with illegal content safety duties, and reporting and complaints duties. This would apply to providers of all U2U and search services.
   ii) Providers should have written statements of responsibilities for senior members of staff. This would apply to providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk).

c) Internal assurance and compliance functions.

   i) Providers should have an internal monitoring and assurance function to provide independent assurance that measures taken to mitigate illegal harms are effective. This would apply to providers of services that are both large and multi-risk.
   ii) Providers should track evidence of new and increasing illegal harm. This would apply to providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk).

d) Staff incentives, policies, and processes.

   i) Providers should have a code of conduct regarding protection of users from illegal harm. This would apply to providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk).
   ii) Providers should offer compliance training for staff involved in the design and operation of a service. This would apply to providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk).

Feedback on our overall approach[167]

[165] Alongside work by Milliman commissioned by Ofcom (2023) this includes, for example, relevant ISO standards, UK Government Orange Book, guidance from the National Cyber Security Centre (NCSC), and case studies such as: OECD, 2012. Corporate Governance for Process Safety OECD Guidance for Senior Leaders in High Hazard Industries, and Consolidated complaint regarding Boeing accessed via the Washington Post, 2021

[166] In this context, we define a large service as with more than seven million monthly active United Kingdom users. Further information can be found in 'Our approach to developing Codes measures'.

[167] Note that this list is not exhaustive and further responses can be found in Annex 1.

5.8 We received a range of responses regarding this overall approach to governance. In this section, we summarise and address the following themes:

- support for the package of measures,
- feedback on who these measures apply to,
- design of measures,
- oversight and management of all risks identified,
- feedback on rights impact and risks, and
- feedback on a single governance process.

5.9 Responses which focussed on specific measures are outlined in the 'Summary of stakeholder feedback' sections and addressed in the 'Our reasoning' sections for each individual measure.

Support for the package of measures

5.10 Several respondents gave broad support for the proposed governance measures as a package in both the November 2023 and May 2024 Consultations, though in some cases qualified this support with some caveats.[168] [169] Where relevant, these have been drawn out in the rest of this chapter.

Feedback on who these measures apply to

Proportionality for small and large low-risk services

5.11 Many respondents argued that recommending some of our measures to providers of **small services** and providers of **large services that are low-risk** was not proportionate. Several respondents argued that the proposed governance measures were disproportionately

[168] Association of British Insurers response to November 2023 Illegal Harms Consultation, p.3; Association of Police and Crime Commissioners response to May 2024 Consultation on Protecting Children from Harms Online, p.8; Association of Police and Crime Commissioners response to November 2023 Illegal Harms Consultation, p.3; Bolton, C. response to November 2023 Illegal Harms Consultation, p.1; [✂]; Born Free Foundation response to November 2023 Illegal Harms Consultation, p.4; Canadian Centre for Child Protection (C3P) response to November 2023 Illegal Harms Consultation, p.4; Centre for Excellence for Children's Care and Protection (CELCIS) response to May 2024 Consultation on Protecting Children from Harms Online, p.6; East Riding of Yorkshire Council response to May 2024 Consultation on Protecting Children from Harms Online, p.2; Evri response to November 2023 Illegal Harms Consultation, p.2; Kooth response to May 2024 Consultation on Protecting Children from Harms Online, p.6; Match Group response to November 2023 Illegal Harms Consultation, pp.2-3; Meta (WhatsApp) response to May 2024 Consultation on Protecting Children from Harms Online, p.11; Meta (WhatsApp) response to November 2023 Illegal Harms Consultation, p.6; Molly Rose Foundation response to November 2023 Illegal Harms Consultation, pp.39-40; National Trading Standards eCrime team response to November 2023 Illegal Harms Consultation, p.2; OnlyFans response to November 2023 Illegal Harms Consultation, p.1; [✂]; Scottish Society for the Prevention of Cruelty to Animals response to November 2023 Illegal Harms Consultation, p.14; Skyscanner response to May 2024 Consultation on Protecting Children from Harms Online, p.8; Skyscanner response to November 2023 Illegal Harms Consultation, p.4; Snap response to May 2024 Consultation on Protecting Children from Harms Online, p.14; Snap response to November 2023 Illegal Harms Consultation, p.5; South East Fermanagh Foundation response to November 2023 Illegal Harms consultation, p.3; Stop Scams UK response to November 2023 Illegal Harms Consultation, p.5; techUK response to November 2023 Illegal Harms Consultation, p.8; The Cyber Helpline response to November 2023 Illegal Harms Consultation, p.3; The Lego Group response to May 2024 Consultation on Protecting Children from Harms Online, p.2; Trust Alliance Group response to November 2023 Illegal Harms Consultation, p.2; TSB Bank response to November 2023 Illegal Harms Consultation, p.1; [✂].

[169] In response to our equivalent measures in the May 2024 Consultation, [✂].

burdensome for smaller services.[170] Other respondents argued that the costs of these measures were disproportionately high for providers of **small and low-risk services**. Some stakeholders emphasised that any measures proposed must be directly related to, and targeted at, the specific risks present on a service.[171] There was also a general concern over the potentially disproportionate cost of the governance measures.[172]

5.12 As to feedback on both **large low-risk** and **small (including low-risk) services**, we agree that where we have recommended measures, they must be proportionate to service risk level and size. We have impact assessed each measure and, based on this, have only recommended measures to service providers when in our judgement they would deliver benefits to users and are proportionate.[173] We explore this in greater detail throughout the sections under 'Who this measure applies to for each measure.

5.13 We have also recommended applying a greater number of measures specifically to **large services** (including large low risk). While we acknowledge the views of respondents regarding the application of measures to **large low-risk services**, our assessment is that these services are likely to be able to access necessary resources to implement them, and applying these governance measures will bring sufficient benefits to justify it. As set out in our chapter 'Our approach to developing Codes measures', these measures help to ensure that risks are properly identified in a timely manner. This is especially important in larger services given the potential larger number of users that could be impacted by the harm. This is because governance measures can improve the accuracy and comprehensiveness of a service provider's risk assessment and management of the risks it identifies. For instance, levels of risk can change over time, so the recommended measures can help service providers better identify emerging risks that can impact their users. An emerging new risk on a large low risk service could impact many users before it is identified if appropriate governance is not in place. In addition, larger providers are likely to have more members of staff and/or a more complex organisational structure than smaller services, increasing the risk that new information or emerging issues may not be identified as relevant to safety or escalated appropriately. Measures to help the flow of information through levels of seniority and governance are likely to have more benefit here and are likely to improve the risk assessment.

Broadening the scope of who the measures apply to

5.14 Some other respondents, including service providers, argued that the scope of governance measures should be broadened, such as by being recommended for all services, or for services that are at a medium or high risk of just one kind of illegal harm.[174] The Association

---

[170] Bolton, C. response to November 2023 Consultation, p.1; Global Partners Digital response to November 2023 Illegal Harms Consultation, p.10; Mega response to November 2023 Illegal Harms Consultation, p.5; Mid Size Platform Group response to November 2023 Illegal Harms Consultation, pp.7-8; Name Withheld 3 response to November 2023 Illegal Harms Consultation, p.3; Online Dating and Discovery Association (ODDA) response to November 2023 Illegal Harms Consultation, pp.1-2; techUK response to November Consultation, pp.3, 8.

[171] Airbnb response to November 2023 Illegal Harms Consultation, p.3; [✂]; Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE) response to November 2023 Illegal Harms Consultation, p.4; Evri response to November 2023 Consultation, p.2.

[172] Spotify response to November 2023 Illegal Harms Consultation, p.6.

[173] Please see 'Our approach to developing Codes measures' for details on this framework.

[174] Alliance to Counter Crime Online (ACCO) response to November 2023 Illegal Harms Consultation, p.2; Internet Watch Foundation (IWF) response to November 2023 Illegal Harms Consultation, p.22; Logically

of Police and Crime Commissioners stated that operational experience shows that the harms caused by smaller platforms with less moderation can be higher in potential and risk compared to larger organisations.[175] The Cyber Helpline highlighted it was vital that as many services as possible are in scope of the measures as there is a risk of "driving some perpetrator behaviour currently being encountered on major platforms onto smaller, less visible services".[176]

5.15 Google commented that high user reach alone is not indicative of being high risk, suggesting that some of the more onerous of our proposed measures apply only to large services at high risk of harm.[177] [✂][178]

5.16 As set out in our 'Our approach to developing Codes measures', we have decided not to extend the services in scope of these measures at this time. In Spring 2025, we intend to consult further on the case for applying some or all of the measures currently recommended only to multi-risk services to some or all single-risk services.[179]

Design of measures

5.17 Some stakeholders saw our measures as overly prescriptive and encouraged us to allow for more flexibility in how service providers implement these measures.[180] Stakeholders also requested additional flexibility for the equivalent measures we consulted on in the May 2024 Consultation.[181]

5.18 Other stakeholders raised concerns about the applicability of our governance measures to decentralised, non-profit and community-moderated services – this can capture a diverse group of services whose architecture or operational practices are performed in non-centralised ways, and thus rely more on users or communities to manage and operate.[182] They argued that our proposals were modelled on services with a "centralised approach" to moderation practices or wider service architecture.[183] These responses suggested that we had not properly considered the governance and accountability frameworks of

---

response to November 2023 Illegal Harms Consultation, p.12; Molly Rose Foundation response to November 2023 Consultation, pp.39-40; National Society for the Prevention of Cruelty to Children (NSPCC) response to November 2023 Consultation, p.6; NSPCC response to May 2024 Consultation on Protecting Children from Harms Online, p.19; Online Safety Act Network (OSA Network) response to November 2023 Illegal Harms Consultation, p.95; Snap response to November 2023 Consultation, p.5; The Cyber Helpline response to November 2023 Consultation, p.3, Yoti response to November 2023 Illegal Harms Consultation, p.2.

[175] Association of Police and Crime Commissioners (PCC) response to November 2023 Illegal Harms Consultation, p.3.

[176] The Cyber Helpline response to November 2023 Consultation, p.3.

[177] Google response to November 2023 Illegal Harms Consultation, pp.27-28.

[178] [✂].

[179] We note that in this Statement, we refer to services that are at medium or high risk for exactly one kind of illegal harm as 'single-risk' services. For further information, please refer to Volume 2: chapter 13: 'Combined impact assessment'.

[180] Meta (WhatsApp) response to November 2023 Consultation, p.6; Microsoft response to November 2023 Illegal Harms Consultation, p.2; Reddit response to November 2023 Illegal Harms Consultation, pp.8, 20; techUK response to November 2023 Consultation, p.8; X response to November 2023 Illegal Harms Consultation, p.3.

[181] Meta (WhatsApp) response to May 2024 Consultation, p.3.

[182] Center for Data Innovation response to November 2023 Illegal Harms Consultation, p.3; Global Partners Digital response to November 2023 Consultation, p.10; Wikimedia Foundation response to November 2023 Illegal Harms Consultation, pp.7, 20.

[183] Wikimedia Foundation response to November 2023 Consultation, p.7.

decentralised or community-led mechanisms, and that our draft measures were, therefore, ill-suited for such services.[184]

5.19 We recognise that a wide variety of service types will fall in scope of the Act, including services with a decentralised and community-moderated organisational model. However, we have intentionally designed each measure with flexibility in mind as we recognise the importance for providers to have some flexibility in how they will implement these measures. We encourage providers to consider the safety outcome expected and to implement the measures in a way that is appropriate and effective for their own service and organisational structure.

5.20 Having considered stakeholder feedback, we therefore maintain that regardless of the organisational structure in question, having appropriate governance measures in place will reduce risks to users from illegal content for the reasons set out in the 'Introduction' section of this chapter. Our final measures therefore continue to apply for all services that fall in scope, regardless of their organisational model; this includes decentralised, non-profit and community-moderated services.

5.21 We address any other measure specific comments on flexibility in the sections titled 'How this measure works' and 'Who this measure applies to' for each specific measure.

Oversight and management of all risks identified

5.22 As set out in the 'Our approach to developing Codes measures' chapter, some respondents to our November 2023 and May 2024 Consultations stated that there was an unaccounted 'gap' between the risks identified in a service's own risk assessment and our draft Illegal Content and Children's Safety Codes of Practice. Some respondents raised concerns that this would result in high levels of unmanaged risk even after compliance with our draft Codes measures - examples provided ranged from risks from livestreaming and addictive design to those arising from certain business models.[185]

---

[184] Center for Data Innovation response to November 2023 Consultation, p.3; Wikimedia Foundation response to November 2023 Consultation, pp.7, 20.

[185] 5Rights Foundation response to May 2024 Consultation on Protecting Children from Harms Online, p.1-2; 5Rights Foundation response to November 2023 Illegal Harms Consultation, pp.2, 19-20; Barnardo's response to May 2024 Consultation on Protecting Children from Harms Online, p.2; Barnardo's response to November 2023 Illegal Harms Consultation, p.10; Center for Countering Digital Hate (CCDH) response to May 2024 Consultation on Protecting Children from Harms Online, pp.2, 8, 9; CCDH response to November 2023 Illegal Harms Consultation, p.7; Children's Coalition for Online Safety response to May 2024 Consultation on Protecting Children from Harms Online, p.2; Children's Commissioner for England response to May 2024 Consultation on Protecting Children from Harms Online, p.29; Global Action Plan response to May 2024 Consultation on Protecting Children from Harms Online, p.1; Internet Matters response to May 2024 Consultation on Protecting Children from Harms Online, p.2; Internet Matters response to November 2023 Illegal Harms Consultation, p.2; Internet Watch Foundation (IWF) response to November 2023 Illegal Harms Consultation, pp.10-11; IWF response to May 2024 Consultation on Protecting Children from Harms Online, p.4; Marie Collins Foundation response May 2024 Consultation on Protecting Children from Harms Online, p.2; Molly Rose Foundation response to May 2024 Consultation on Protecting Children from Harms Online, p.41; NSPCC response to May 2024 Consultation, pp.23-24; OSA Network response to May 2024 Consultation on Protecting Children from Harms Online, pp.43-51; OSA Network response to November 2023 Consultation, pp.61-62; UK Safer Internet Centre (UKSIC) response to November 2023 Illegal Harms Consultation, p.33; UKSIC response to May 2024 Consultation on Protecting Children from Harms Online, p.22; Violence Against Women and Girls (VAWG) Sector Experts response to May 2024 Consultation on Protecting Children from Harms Online, pp.11, 15; Vodafone response to May 2024 Consultation on Protecting Children from Harms Online, p.2.

5.23 We note that some respondents would like to see our Codes recommend providers to mitigate all risks from a service provider's risk assessment. As explained in 'Our approach to developing Codes measures', we cannot do this. The safety duties in the Act only require providers to take proportionate steps and we can only make recommendations we are satisfied are proportionate, having impact assessed them. We cannot assess the impact of a proposal if we do not know what compliance with it would entail.

5.24 Our policy intent is to provide adequate oversight of risk management practices through effective governance. In light of these objectives, and taking into account stakeholder feedback, we consider that two of the measures on which we consulted (ICU A1/ICS A1 and ICU A2/ICS A2) should be amended.

5.25 These two measures are intended to ensure there are formalised accountability, reporting and audit processes in place for activities related to managing risks (including risks remaining after implementing Codes of Practice), as identified in a service's risk assessment. We provide an explanation in 'How this measure works' section for each of these measures.

5.26 The Governance measures should be understood as complementing the risk assessment practice outlined in our Risk Assessment Guidance. Service providers should consider how to develop and operationalise the structures and processes for illegal harms risk assessments that help them to identify, monitor and manage risk of harm as appropriate. A service's risk assessment should provide a robust understanding of all its risks. This should enable the provider to identify proportionate measures it can take to manage and mitigate its risks.

5.27 To reinforce this point, and taking into account stakeholder feedback received on our Risk Assessment Guidance, we have also revised its Step 4 to report, review and update the risk assessment. Here, we advise services to monitor the effectiveness of safety measures at reducing the risk of harm to users (including taking into account the role of any controls in place), as well as the level of risk exposure after appropriate measures are implemented.[186]

Feedback on rights impact and risks

5.28 We received some responses that asked us to consider potential negative impacts from our proposed measures on freedom of expression, mishandling of personal data, and other harms not solely related to exposure to illegal content.[187]

5.29 In our November 2023 Consultation, we set out our view that our proposals in this area would not have implications for freedom of expression or privacy. This is because governance and accountability are wholly concerned with the organisation, internal structure, and processes of regulated services as businesses. We also considered that good

[186] See our Risk Assessment Guidance.

[187] CELE response to November 2023 Consultation, p.3; Centre for Competition Policy response to November 2023 Illegal Harms Consultation, p.5; Global Partners Digital response to November 2023 Consultation, p.10; X response to November 2023 Consultation, pp.3-4.

governance can positively reinforce compliance, be this with internal processes and procedures, or other legal requirements. We considered that a well-managed business is, in general, more likely to comply with its obligations under privacy and data protection laws. We have now included a rights assessment for each measure within the sections titled 'Rights' to clarify our position further.

Feedback on a single governance process

5.30 In May 2024, we consulted on a separate set of draft governance measures which we proposed to include in the Children's Safety Codes.[188] We proposed adopting a single consistent approach to governance and accountability (as outlined in our November 2023 Consultation) across our work on illegal harms and the protection of children. In practice, this means that providers of services likely to be accessed by children may choose to adopt a single process to meet both their illegal content duties and child safety duties. However, we clarified that services must still ensure this process effectively addresses both risks of illegal harms and risks to children's safety from other harmful content.

5.31 Several stakeholders, including civil society organisations and service providers, agreed with our expectation that the governance measures set out in both the November 2023 and May 2024 Consultations could be implemented through a single process.[189]

5.32 While we have not yet reached a final decision on the governance measures to be included in the Children's Safety Codes, we are working to ensure consistency in our approach to equivalent measures in the Illegal Content and Children's Safety Codes. This is to ensure that providers of services can discharge both their illegal harms and protection of children duties if they are in scope of both. We consider this to be important for services to embed an appropriate culture of good governance.

Measure on annual review of risk management activities

5.33 In our November 2023 Consultation, we proposed that providers of large user-to-user ('U2U') services and providers of large general search services conduct an annual review of their risk management processes in relation to online safety and examine how developing risks are being monitored and managed. We recommended that this be undertaken by a service provider's most senior governance body.

5.34 Our expectation was that the implementation of annual reviews would result in better risk management, reducing the level of harm users are exposed to.

---

[188] Ofcom, May 2024 Consultation on Protecting Children from Harms Online.

[189] Canadian Centre for Child Protection (C3P) response to May 2024 Consultation on Protecting Children from Harms Online, p.13; CELCIS response to May 2024 Consultation, p.7; Derbyshire OPPC response to May 2024 Consultation on Protecting Children from Harms Online, p.7; Federation of Small Businesses response to May 2024 Consultation on Protecting Children from Harms Online, p.4; Dean, J. response to May 2024 Consultation on Protecting Children from Harms Online, p.9; Kooth response to May 2024 Consultation on Protecting Children from Harms Online, p.6; Meta (WhatsApp) response to May 2024 Consultation, p.11; Microsoft response to May 2024 Consultation on Protecting Children from Harms Online, p.5; [✂]; Nexus response to May 2024 Consultation on Protecting Children from Harms Online, p.9; NSPCC response to May 2024 Consultation, pp.19, 43; OneID response to May 2024 Consultation 2024 on Protecting Children from Harms Online, p.2; Scottish Government response to May 2024 Consultation on Protecting Children from Harms Online, p.9; Snap response to May 2024 Consultation, p.14; [✂].

5.35 This was also in line with the annual review cycle we were suggesting in our Risk Assessment Guidance for service providers.

Summary of stakeholder feedback[190]

5.36 Several stakeholders expressed support for this measure, though some caveated this with reservations.[191] We have identified several other themes that emerged from stakeholder feedback which we discuss in the following paragraphs:

- Frequency of reviews.
- Feedback on the suitability of existing corporate functions.
- Feedback on who this measure applies to.

Frequency of reviews

5.37 Lloyds Banking Group welcomed the focus of this proposal of annual review of risk management activities. However, it indicated that review periods of at least 12 months could mean that there are substantial periods of new risks not being factored into risk assessments, a particular challenge for fraud where new schemes can arise quickly.[192] In relation to the equivalent measure included in our May 2024 Consultation, one respondent said they found an iterative process of continuously reviewing and updating trust and safety policies and risk management activities helpful to respond to the quickly evolving nature of risks, rather than an annual scheduled review.[193] We address this in the 'Benefits and effectiveness' section.

Feedback on the suitability of existing corporate functions

5.38 The Children's Commissioner for England shared concerns in its response to our May 2024 Consultation that Ofcom "has largely recommended existing industry models of governance that, while effective in protecting company interests, abstract from the objective of the Act". It recommended that Ofcom reiterates that "improving online safety might require a change in the overall structure and dynamic of the way these services are run".[194] We address this feedback in the 'How this measure works' and 'Benefits and Effectiveness' sections.

Feedback on who this measure applies to

5.39 We received a range of responses regarding which services this measure would apply to, and the equivalent measure included in our May 2024 Consultation.[195] [196] Some respondents commented on the size of the services this measure would apply to, in

---

[190] Note that this list is not exhaustive and further responses can be found in Annex 1.
[191] Age Verification Providers Association response to November 2023 Illegal Harms Consultation, p.2; C3P response to November 2023 Consultation, pp.5-7; LinkedIn response to November 2023 Illegal Harms Consultation, p.3; Lloyds Banking Group response to November 2023 Illegal Harms Consultation, pp.3-4; Meta (WhatsApp) response to November 2023 Consultation, p.7; Microsoft response to November 2023 Consultation, p.2, Yoti response to November 2023 Consultation, p.2.
[192] Lloyds Banking Group response to November 2023 Consultation, pp.6-7.
[193] Match Group response to May 2024 Consultation on Protecting Children from Harms Online, p.6.
[194] Children's Commissioner for England response to May 2024 Consultation, p.21-22.
[195] Airbnb response to November 2023 Consultation, p.3; [✂]; Logically response to November 2023 Consultation, p.12; NSPCC response to November 2023 Consultation, p.7.
[196] C3P response to May 2024 Consultation, pp.10-11; Children's Commissioner for England response to May 2024 Consultation, pp.21-22; East Riding of Yorkshire Council response to May 2024 Consultation, p.3; Match Group response to May 2024 Consultation on Protecting Children from Harms Online, pp.5-6.

particular that providers of smaller services should be in scope of this measure. The National Society for the Prevention of Cruelty to Children (NSPCC) argued that this measure should also be applied to small services. It stated that it should be reasonable to expect all services to adapt their risk management approach year-on-year in response to an "online threat landscape [that] is constantly shifting".[197] Logically argued that small services should be in scope of this measure due to the threat of foreign interference.[198] East Riding of Yorkshire Council stated that the scope of the measure could produce a gap in the effectiveness of the changes achieved through implementation of the Act.[199]

5.40 Others said that the measure should be more risk-based, and applied regardless of service size. One stakeholder argued that the measure of annual review of risk management activities should not apply to providers solely on the basis of the size of their service, but to "multi-risk" services or large services which are multi-risk, where there is a high risk of at least one priority offence.[200] One stakeholder stated that there should be senior management oversight of how potential harms are being addressed within a high-risk organisation regardless of its size.[201]

5.41 The Canadian Centre for Child Protection ('C3P') argued against measures based on size, stating that it is illogical for a platform aimed at or developed for children to be exempt from annual reviews if it does not meet the threshold to be classed as a large service. It suggested that some form of scaled annual review of risk management activities should be required for all services.[202] C3P noted in their response to the November 2023 Consultation that smaller companies can be targeted by those sharing CSAM and that a lower monthly user threshold or a scaled down version of this measure for such providers could help mitigate risks.[203]

5.42 We address this feedback in the section 'Who this measure applies to'.

Our decision

5.43 We have decided to broadly confirm the measure we proposed in the November 2023 Consultation with one minor change to reinforce the role of the governance body in reviewing risk management pertaining to all illegal harm risks identified at risk assessment, including the risk services are exposed to after the implementation of appropriate Codes measures.

5.44 This adjustment reinforces our position that service providers need to do a thorough review of the risks on their services, as explained in the 'Oversight and management of all risks identified'. The measure now reads:

5.45 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. The full text of the measure can be

[197] NSPCC response to November 2023 Consultation, p.7.
[198] Logically response to November 2023 Consultation, p.12.
[199] East Riding of Yorkshire Council response to May 2024 Consultation, p.3.
[200] Airbnb response to November 2023 Consultation, p.3.
[201] [✂].
[202] C3P response to May 2024 Consultation, pp.10-11.
[203] C3P response to November 2023 Consultation, pp.5-7.

found in our Illegal Content Codes of Practice for U2U services and search services on terrorism, child sexual exploitation and abuse (CSEA) and other duties. We refer to this measure as ICU A1 for U2U services and ICS A1 for search services.

Our reasoning

How this measure works

5.46 Providers of large services should conduct an annual review of their risk management in relation to all risks relating to illegal harm identified in their risk assessment, including 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.

5.47 The description, size, complexity, or name given to a governance body of a provider will vary. Service providers may define what they consider to be their most senior governance body. In our November 2023 Consultation, we referred to it as the body responsible for the overall governance and strategic direction of a service.

5.48 The review may form part of existing governance processes for annually reviewing strategic risks. A good review might include:

- all online safety risks identified in risk assessments,
- risk management processes, policies, and procedures,
- the effectiveness of mitigation measures in place,
- the monitoring and management of risk trends,
- the monitoring and management of residual risk levels, and
- lessons learned from past mistakes.

5.49 We have recommended that this be undertaken by the service provider's most senior governance body to ensure online safety risk management is embedded in decision making and becomes part of an organisation-wide approach to risk management. In response to the Children's Commissioner for England, we would expect that this approach would make online safety a "company interest".[204] In essence, we want regulated entities' bodies to pay attention to tracking and addressing online safety risks as they do to commercial risks.

Benefits and effectiveness

5.50 Regular review of risk management and regulatory compliance by a governance body is required for appropriate oversight over internal controls. Evidence supporting this principle can be found in good governance practice principles and codes.[205] We note that two

[204] Childrens' Commissioner for England response to May 2024 Consultation, p.21.

[205] Under the UK Corporate Code, companies with a premium listing on the London Stock Exchange are already required to follow principles related to board oversight. This includes provision 29 which states that boards "should monitor the company's risk management and internal control systems and, at least annually, carry out a review of their effectiveness and report on that review in the annual report". Monitoring and review activities are intended to cover all material controls including financial, operational and compliance controls. Source: Financial Reporting Council, 2018. The UK Corporate Governance Code, pp.10. The OECD's Principles of Corporate Governance similarly suggests that a key function of the Boards should be "reviewing and guiding corporate strategy, major plans of action [and] risk management policies and procedures". The Principles suggest that while committees or other sub-bodies may have specific responsibilities of different areas of risk, "the board should retain final responsibility for oversight of the company's risk management system and for ensuring the integrity of the reporting systems. Source: OECD, 2015.

respondents suggested these review periods should be different, as outlined in paragraph 5.37. However, we maintain that an annual cycle remains appropriate. Based on our analysis of best practice, annual reviews of risk management activities can be tied in with financial and company results reporting or public disclosure. Our Risk Assessment Guidance for service providers also suggests an annual review cycle.

5.51 In our November 2023 Consultation, we cited evidence showing that effective use of data and information to report on risks to boards is associated with good risk management.[206] [207] The communication of this information to the board is important for assessing the effectiveness of internal controls and for deciding whether any changes are required to improve risk management.[208] We also noted that best practice guidance and codes for governance bodies and boards points to the importance of setting a regular schedule for the review of risk management activities.[209] We also noted a study about the importance of independence of governance body members in ensuring that a Board or governance body discharges its duties effectively.[210]

5.52 In light of this information, we consider that this measure will enhance in-scope service providers' ability to identify and manage online safety risks effectively. This should in turn reduce the risk of illegal harms taking place on the services they operate, thereby delivering significant benefits.

5.53 As detailed in paragraph 5.38, the Children's Commissioner for England questioned our approach. We have not based our proposals solely on what service providers already do, but have used these examples to demonstrate how these measures may be practically implemented. Our recommendations have been informed by a review of good practice standards and principles in risk management and corporate governance across a range of different industries which have mature and effective cultures of risk management. This has been key to giving confidence that the measures will work effectively in application. It has also enabled us to impact assess measures. The safety duties in the Act only require providers to take proportionate steps and we can only make recommendations we are satisfied are proportionate, having impact assessed them. We cannot assess the impact of a proposal if we do not know what compliance with it would entail.

5.54 The evidence set out above leads us to conclude that where senior governance bodies review risk management activities regularly safety outcomes will be better. Consequently, we consider that the measure under discussion will deliver important benefits.

Costs

[206] This study by the Financial Reporting Council with participants from over 40 listed companies concluded that it was important for Boards to have a whole view of risk (including "gross" or inherent risks) to engage in meaningful discussion. Several organisations specified that they reported on emerging risks as well as more conventional risk registers to improve Boards' oversight across risk areas. Source: FRC 2011. Boards and Risk A summary of discussions with companies, investors and advisers. [accessed 29 November 2024]

[207] The Board Imperative: How can data and tech turn risk into confidence?, Ernst & Young [accessed 4 May 2023].

[208] UK Government Finance Function (GFF), 2023. Good Practice Guide: Risk Reporting. [accessed 29 November 2024].

[209] Best practice for overall governance bodies is to maintain an annual cycle of planned activity, to ensure that there is time for full consideration of specific exposures. Source Milliman, 2023.

[210] Guluma, T. F., 2021. The impact of corporate governance measures on firm performance: the influences of managerial overconfidence, Future Business Journal, 7 (50). [accessed 29 November 2024].

5.55 We estimate the extra cost for the main board of a large organisation to review and scrutinise an annual risk management paper to be approximately £16,000 to £37,000 per year.[211] Costs will be higher for service providers with larger and more highly paid boards. We would expect most providers of large services to already have governance bodies in place for the overall management of the business.[212] In these instances, it will be a case of additional ongoing costs associated with conducting this review process.

5.56 Providers of large services that have assessed low risks will also tend to have lower costs because reporting the annual review of risk management activities related to online safety to the governance body will be simpler if the risks are low.

5.57 We did not receive any specific stakeholder responses regarding direct costs for this measure, although there were general concerns over the potentially disproportionate cost of the governance measures (paragraph 5.11).[213] While our cost assumptions and estimates are largely unchanged from the November 2023 Consultation, we consider that the flexibility given to service providers regarding how they implement this measure will enable them to choose approaches with costs appropriate to their business.[214]

5.58 This measure may result in other costs – for example, if the service provider needs to make changes as a result of the annual review. We are not counting these in our estimates for this measure, and we assume if an effective annual review results in the service provider taking action, then the costs of those actions will be reasonable in order to prevent harm.

---

[211] This is derived from the assumptions set out in Annex 5 which has been updated with the latest wage data released by the Office for National Statistics (ONS) and combined with the following specific assumptions. We assume it takes 10 to 20 days for a person in a "professional occupation" to prepare the paper for the board and that on average each director on the board spends one to two hours in total to read, consider, and discuss the paper. We assume on average directors spend 250 hours a year on board related activities for each company they are a director for, based on PwC's 2022 Annual Corporate Directors Survey. For total remuneration per board member, we assume $321,220 per year, based on the average for 2023 of S&P 500 independent board directors (from 2023 U.S. Spencer Stuart Board Index – a report by Spencer Stuart, a leadership consultancy). Many of the largest services are owned by US companies. For the number of board members, we assume boards have on average 11 members, based on the average S&P 500 board size, from Diversity, Experience, and Effectiveness in Board Composition, Harvard Law School Forum on Corporate Governance, Merel Spierings, 14 June, 2022. [accessed 25 October 2024].

[212] Of the services we are aware of that are large, most are ultimately owned by listed companies. Companies listed on the New York Stock Exchange are required to have an audit committee which is required to discuss "policies with respect to risk assessment and risk management". Source: NYSE 2009. NYSE Audit Committee Responsibilities [accessed 25 October 2024].

[213] In response to the November 2023 Illegal Harms Consultation, several stakeholders raised concerns about the proportionality for small organisations. These included: Evri response to November 2023 Consultation, p.2; Global Partners Digital response to November 2023 Consultation, p.10; [✂]; Name Withheld 3 response to November 2023 Consultation, p.3; ODDA response to November 2023 Consultation, pp.1-2; techUK response to November 2023 Consultation, p.8. Some stakeholders raised concerns about the proportionality of measures for large low risk or low risk service providers. These included: Airbnb response to November 2023 Consultation, p.3; CELE response to November 2023 Consultation, p.4; Spotify response to November 2023 Consultation, p.6; techUK response to November 2023 Consultation, p.8.

[214] We have updated the estimates since the November 2023 Consultation in line with the latest wage data released by the Office of National Statistics (ONS) and Spencer Stuart, as outlined in paragraph 5.55. We received some feedback on the general cost assumptions (such as salary assumptions) that are fed into these costs. We consider that feedback in Annex 5.

Rights impact

Freedom of expression

5.59 As explained in 'Introduction, our duties, and navigating the Statement', as well as in Volume 2: chapter 14: 'Statutory tests', Article 10 of the ECHR sets out the right to freedom of expression, which encompasses the right to hold opinions and to receive and impart information and ideas without unnecessary interference by a public authority. Article 11 of the ECHR sets out the right to associate with others. We must exercise our duties under the Act in light of users' and services' Article 10 and 11 rights and not interfere with these rights unless we are satisfied that to do so is prescribed by law, pursues a legitimate aim, is proportionate to the legitimate aim and corresponds to a pressing social need.

5.60 The objective of this measure is to ensure that the quality of risk management activities (in relation to illegal content risks) is regularly reviewed. Providers may choose to take actions based on the findings of the annual review of risk management. However, the measure itself does not recommend any steps to be taken with respect to particular kinds of content and, therefore, we consider that this measure would not constitute an interference with users' or providers' freedom of expression or association rights.

5.61 To the extent that it helps to reduce harm on the service and make users feel safer, this could also positively impact on their human rights.

Privacy and Data Protection

5.62 As explained in 'Introduction, our duties, and navigating the Statement', as well as in chapter 14 of this Volume: 'Statutory tests', Article 8 of the ECHR sets out the right to respect for individuals' private and family life. An interference with this right must be in accordance with the law, pursue a legitimate aim, be proportionate to the legitimate aim and correspond to a pressing social need. We consider the privacy and data protection impacts of this measure to be inextricably linked.

5.63 We do not consider this measure will have a negative impact on users' rights to privacy. Service providers will have to evaluate information and review trends in order to implement the measure. Where a provider elects to collect any metrics or is already collecting metrics, it will need to ensure that any personal data is processed in accordance with the relevant data protection legislation. Providers should refer to relevance guidance from the Information Commissioner Office ('ICO').[215]

5.64 We also consider that a well-managed business is, in general, more likely to comply with its obligations under privacy and data protection laws (as well as, for that matter, other laws such as those relating to consumer protection and equality). As such, our recommended measure may help to safeguard these.

Who this measure applies to

5.65 In our November 2023 Consultation, we proposed applying this measure to all providers of large U2U services and large general search services. We have considered stakeholder arguments for both expanding and reducing the scope of this measure (see paragraphs 5.39-5.41).

5.66 We remain of the view that it is proportionate to apply this measure to all large U2U and general search services, even if they have not identified any risk any kind of illegal harm in

[215] ICO, UK GDPR guidance and resources. [accessed 4 November 2024].

their latest risk assessment. The governance bodies in large providers with complex organisations have a greater role to play in a provider's approach to identifying and addressing illegal content because of the need for high-level oversight of coordination and consistency in risk management. We expect the direct costs of the measure to be manageable for providers of large services which are already likely to have a suitable established governance body responsible for oversight of risk management.

5.67 We acknowledge that this measure may have fewer benefits for providers of large services that has assessed low risk of all kinds of illegal harms since there may be less scope for reducing harm to users from illegal content. However, there are still important benefits of governance measures that help to ensure that risks are properly identified in a timely manner. We consider that a failure in oversight of risk management by the provider of a large service can affect their ability to identify risks that would impact a large number of users and could have a significant adverse effect. Furthermore, a large service with a low risk of harm is likely to incur lower costs when implementing this measure because reporting the annual review of risk management activities for online safety will be simpler if the risks are low.

5.68 As with some of the other governance measures, we currently do not consider it proportionate to recommend this measure to large vertical search services as they typically present very limited risks, if any.[216]

5.69 We do not consider there to be new evidence demonstrating that it would be proportionate to recommend annual reviews to providers of smaller services. Many providers of smaller services may not have a fully developed governance body, and to establish one would entail significant staff and resource costs relative to the size of the organisation. Furthermore, the benefits of this measure will likely be lower for providers of smaller services because they have less complex organisational structures, and it is easier for management to ensure coordination and consistency in approach.

5.70 These governance measures are also not being recommended in isolation, and it is critical for providers of smaller services to have a good understanding of risks. They can achieve this through their duty to carry out and keep risk assessments up to date at least once a year.[217] [218] Given the costs in question and the fact that the benefits of applying the measure to smaller services are lower, we are not at this point persuaded that it is necessary and proportionate to supplement this recommendation with a recommendation that providers of smaller services conduct an annual review of risk management activities.

5.71 Therefore, we have decided not to recommend that providers of smaller services create new governance structures to follow this measure. We are maintaining our position of recommending this measure for all providers of large U2U services and providers of large general search services.

Conclusion

[216] By their nature vertical search services are unlikely to have content that is as rapidly changing as U2U services and search results are more under a service's control than for U2U content. We are also not aware of evidence of such services showing illegal content. Any benefits of applying this measure would therefore be low for vertical search services. See a more detailed explanation in our Register of Risks chapter titled 'Search'.

[217] Section 9(3) and (4) and 26(3) and (4) of the Act.

[218] See Risk Assessment Guidance.

5.72 Based on our analysis, we consider that this measure will offer significant benefits to users by ensuring sufficient oversight and scrutiny of risk management activities. Whilst there are costs associated with implementing this measure, these should be manageable for providers of large services.

5.73 We have decided to recommend this measure to all providers of large U2U and providers of general search services, regardless of level of risk, for the reasons outlined in paragraphs 5.69-5.71. We have decided to make a small adjustment to the wording of this measure in response to stakeholder feedback to clarify that the governance board should review risk management pertaining to the illegal harm risks identified at risk assessment, including in relation 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.

5.74 This measure is included in our Codes for U2U services and search services on terrorism, CSEA and other duties, and is referred to as ICU A1 for U2U services and ICS A1 for search services.

Measures on senior accountability and on written statements of responsibilities

5.75 In our November 2023 Consultation, we proposed two measures to support strong governance related to senior accountability and responsibility:

- A measure recommending that all service providers name an individual accountable to the most senior governance body for compliance with illegal content duties and reporting and complaints duties.
- A measure recommending providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk) implement written statements of responsibilities for senior members of staff who make decisions related to the management of online safety risks.

5.76 We considered that senior accountability for online safety was critical in building a culture that prioritises safety for users. We explained that users are more likely to be exposed to illegal content risks if there is a lack of accountability at senior management level for compliance with illegal content safety duties. This would be indicative of an absence of senior oversight and responsibility for decisions which have a material impact on user safety. We identified that risks could also arise if services did not provide clarity on roles or responsibilities for managing illegal content risks. This lack of clarity could contribute to inconsistent application of risk management measures.

5.77 We have set out our assessment and decisions related to these measures together as we consider them to be closely linked in their intended outcome.

Summary of stakeholder feedback[219]

Measure on senior accountability

5.78 Several respondents welcomed or expressed support for this measure.[220] In response to both our November 2023 and May 2024 Consultations, Meta (WhatsApp) broadly supported this measure as part of building a culture that prioritises safety.[221] The measure was also supported in responses to the May 2024 Consultation by the Scottish Government.[222]

5.79 We have identified several other themes that emerged from stakeholder feedback, which we discuss in the following paragraphs:

- feedback on the accountable person,
- burden on small services, and
- potential unintended outcomes.

Feedback on the accountable person

5.80 Some respondents raised concerns or requested clarifications about our proposed recommendation that service providers name an accountable person:

- The NSPCC highlighted that this measure must be distinct from the enforcement power which enables a senior manager to be held liable for compliance with a confirmation decision.[223]
- Snap asked for further clarification on who we envisage would be the "person accountable to the most senior governance body for compliance", asking if this was equivalent to the EU's DSA Head of Compliance role and if it could be performed by the same individual.[224]
- Google argued that the measure should recognise the complexity of large and multi service providers and allow them to name more than one accountable person.[225]
- Protection Group International queried how Senior Management could be held accountable if based outside the UK.[226]

[219] Note this list in not exhaustive, and further responses can be found in Annex 1.
[220] Association of Police and Crime Commissioners response to November 2023 Consultation, p.3; Molly Rose Foundation response to November 2023 Consultation, pp.39-40; NSPCC response to November 2023 Consultation, p.6; OneID response to November 2023 Illegal Harms Consultation, p.1; Stop Scams UK response to November 2023 Consultation, p.5.
[221] Meta (WhatsApp) response to May 2024 Consultation, p.11; Meta (WhatsApp) response to November 2023 Consultation, p.6.
[222] Scottish Government response to May 2024 Consultation, p.8.
[223] NSPCC response to November 2023 Consultation, p.6.
[224] Snap response to November 2023 Consultation, p.5. We note that Snap made the same point in response to our May 2024 Consultation. Source: Snap response to May 2024 Consultation, p.14.
[225] Google response to November 2023 Consultation, p.10. We note that Google made the same point in response to our May 2024 Consultation. Source: Google response to May 2024 Consultation on Protecting Children from Harms Online, p.12.
[226] Protection Group International response to November 2023 Illegal Harms Consultation, pp.1-2.

- TikTok and Inkbunny asked for clarification on whether the accountable person must be named publicly or internally, and to Ofcom, and suggested that there could be risks to naming such a person publicly.[227]
- techUK suggested that in practice, responsibilities will be spread across a team, and that rather than have a single accountable person, services should provide Ofcom with a single point of contact.[228]
- The Lego Group also suggested such a role could be fulfilled by multiple people, and asked for clarity on how the accountable person's name would be shared with Ofcom.[229]
- Safe Space One queried the feasibility of holding a specific person in a company accountable for "illegal content duties" by services using end-to-end encryption.[230]

5.81 We have addressed these points in the section 'How these measures work'.

5.82 An individual respondent disagreed with this measure and noted that services in scope of the regime may be run by people outside of the UK who may wish to run their service anonymously to protect themselves, such as activists. Requiring providers to name a person accountable could endanger that anonymity and result in severe consequences for them. This respondent also expressed concerns that small services will shut down due to fear of prosecution and fines.[231] We have addressed these points in the sections titled 'How these measures work', 'Risks' and 'Rights impact'.

Burden on small and overseas services

5.83 Global Partners Digital highlighted the possible challenge for smaller companies to hire or designate somebody to this role and said that it could be the case that only the largest services could provide sufficient remuneration for someone to assume such a high level of responsibility.[232]

5.84 One respondent argued that the measure is unclear and would be disproportionately costly for smaller service providers, particularly those operating personal non-commercial services.[233]

5.85 techUK said it would be disproportionate to apply this measure to smaller services. When discussing the potential "governance burden" that complying with measures could inadvertently place on SMEs, it asked us to explore whether a named person accountable for compliance is essential for smaller services .[234]

5.86 Open Rights Group questioned how realistic it is to expect small overseas operators run by volunteers to comply with the measure.[235]

5.87 We address this feedback in the sections titled 'Costs' and 'Who these measures apply to'.

---

[227] Inkbunny response to May 2024 Consultation on Protecting Children from Online Harms, p.7; TikTok response to May 2024 Consultation on Protecting Children from Harms Online, p.2.
[228] techUK response to May 2024 Consultation on Protecting Children from Harms Online, p.10.
[229] The Lego Group response to May 2024 Consultation on Protecting Children from Harms Online, p.2.
[230] Safe Space One response to November 2023 Illegal Harms Consultation, p.3.
[231] Name Withheld 3 response to May 2024 Consultation on Protecting Children from Harms Online, p.5.
[232] Global Partners Digital response to November 2023 Consultation, p.9.
[233] Name Withheld 2 response to November 2023 Illegal Harms Consultation, p.3.
[234] techUK response to November 2023 Consultation, p.3.
[235] Open Rights Group response to May 2024 Consultation on Protecting Children from Harms Online, p.5.

Potential unintended outcomes

5.88 Global Network Initiative expressed concern about the liability of the accountable person for compliance failures and said that the measure may, without sufficient safeguards, "make it less likely that companies will push back on overbroad government demands or restrictions" given the growing trend of potential liability for company personnel in various jurisdictions.[236] Big Brother Watch expressed concern that making one individual responsible for having to justify the entire service's compliance decisions, coupled with broad definitions and a low threshold of acceptable expression, would result in platforms unscrupulously removing lawful content on their sites, risking widespread censorship online.[237] We address this feedback in the section 'Rights Impact'.

Measure on written statements of responsibility

5.89 Several respondents welcomed or expressed support for this measure. The Molly Rose Foundation and Meta (WhatsApp) agreed with the proposal.[238]

5.90 We have also identified several other themes that emerged from stakeholder feedback:

- **Staff captured by the measure:** Snap requested further clarification regarding who in its organisation should be considered as the "staff who make decisions" related to the management of online safety risks. It queried whether this would be the most senior Product/Engineering and Operations managers responsible for deciding on and implementing the risk mitigation measures.[239] We address this feedback in the section 'How these measures work'.
- **Design of the measure:** Meta (WhatsApp) requested sufficient flexibility to adapt to a variety of organisational designs and structures that may change over time, advising against a prescriptive nature to allow for such changes. It also asked for clarity from Ofcom on retaining the confidentiality of statements and related names.[240] We address this feedback in the section 'How these measures work'.
- **Feedback on costs:** Meta (WhatsApp) further noted that the implementation and maintenance of a statement of responsibilities would require significant resources and disagreed with the time estimate of this measure stated in our November 2023 Consultation.[241] We address this feedback in the section 'Costs'.
- **Feedback on risks:** 5Rights Foundation agreed that it is key to place accountability on senior staff but highlighted the importance of implementing safety standards throughout all levels of an organisation, stating that decisions affecting the safety of children should not be the sole responsibility of upper management.[242] We address this feedback in the section 'Risks'.

---

[236] Global Network Initiative response to November 2023 Illegal Harms Consultation, p.3. We note that Global Network Initiative made the same point in response to our May 2024 Consultation. Source: Global Network Initiative response to May 2024 Consultation on Protecting Children from Harms Online, p.7-8.
[237] Big Brother Watch response to May 2024 Consultation on Protecting Children from Harms Online, p.16.
[238] Meta (WhatsApp) response to November 2023 Consultation, p.6.; Molly Rose Foundation response to November 2023 Consultation, p.40.
[239] Snap response to November 2023 Consultation, p.5.
[240] Meta (WhatsApp) response to November 2023 Consultation, p.6.
[241] Meta (WhatsApp) response to November 2023 Consultation, p.6.
[242] 5Rights Foundation response to November 2023 Consultation, p.9.

- **Feedback on who the measure applies to:** BT Group stated that this measure should apply to smaller services with specific risks to increase the individual accountability of the named person and senior individuals, encourage risks to be taken seriously, and help set the right culture.[243] In response to the May 2024 Consultation, C3P and the Centre for Countering Digital Hate (CCDH) said this measure should be expanded to include smaller and single-risk service providers.[244] The Children's Commissioner for England recommended that the measure should apply to all services.[245] One stakeholder argued that this measure should not apply to providers solely on the basis of the size of their service, but to multi-risk services or large services which are multi-risk, where there is a high risk of at least one priority offence.[246] There were also general comments stating that this measure could be disproportionate for providers of some types of services,[247] that a more proportionate or risk-based approach should be adopted,[248] or that the measure could be better tailed by provider capacity.[249] Airbnb argued that the measure on senior accountability and written statements of responsibilities should not apply to providers solely on the basis of the size of their service, but to multi-risk services or large services which are multi-risk, where there is a high risk of at least one priority offence.[250] We address this feedback in the section 'Who these measures apply to'.

Our decision

Measure on senior accountability

5.91 We have decided to broadly confirm the measure we proposed in the November 2023 Consultation.

5.92 We have adjusted this measure to clarify that accountability of the individual entails being able to explain and justify actions or decisions regarding illegal harm risk management and mitigation (including risks remaining after implementing appropriate Codes of Practice measures) and compliance with the relevant duties. This is in response in particular to the concerns raised about what accountability means, and also to draw out, as for measure ICU A1, that the role relates to all illegal harms risks. The measure now reads:

- The provider 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**.
- Being accountable means being required to explain and justify actions or decisions regarding:
  - **illegal harm** risk management and mitigation (including as to risks remaining after the implementation of appropriate Codes of Practice measures), and

[243] BT Group response to November 2023 Illegal Harms Consultation, pp.1-2.
[244] C3P response to May 2024 Consultation, p.12; CCDH response to May 2024 Consultation, p.9.
[245] Children's Commissioner for England response to May 2024 Consultation, p.23.
[246] Airbnb response to November 2023 Consultation, p.3.
[247] [✂]; Evri response to November 2023 Consultation, p.2; Global Partners Digital response to November 2023 Consultation, p.10; ODDA response to November 2023 Consultation, pp.1-2; [✂].
[248] Bolton, C. response to November 2023 Consultation, p.1; Mega response to November 2023 Consultation, p.5; [✂]; Name Withheld 3 response to November 2023 Consultation, p.3; [✂]; techUK response to November Consultation, pp.3, 8.
[249] CELE response to November 2023 Consultation, p.4.
[250] Airbnb response to November 2023 Consultation, p.3.

- compliance with the relevant duties.

5.93 The full text of the measure can be found in our Illegal Content Codes of Practice on terrorism, CSEA and other duties. We refer to this measure as ICU A2 for U2U services and ICS A2 for search services.

Measure on written statements of responsibility

5.94 We have decided to proceed with the measure as proposed in our November 2023 Consultation, with minor amendments. We have changed the wording from 'staff' to 'managers' because we recognise that the individuals concerned may not be employees, and we have made it clear that the measure is about illegal harm as it relates to individuals in the UK.

5.95 The full text of the measure can be found in our Illegal Content Codes of Practice on terrorism, CSEA and other duties. We refer to this measure as ICU A3 for U2U services and ICS A3 for search services.

Our reasoning

How these measures work

Measure on senior accountability

5.96 Service providers should name an individual accountable to the most senior governance body for compliance with the illegal content safety duties, and reporting and complaints duties.

5.97 We do not recommend any specific qualifications for the accountable individual, as we consider service providers to be best placed to determine this. There is no recommendation for the individual to be UK based, as they may be more effective in their role if they are located somewhere else. The service provider is best placed to determine this. We would generally expect the individual to be a senior manager or director with responsibility for overseeing compliance with the illegal content safety duties and the reporting and complaints duties.

5.98 The accountable individual will explain and justify actions or decisions regarding online safety risk management and mitigation (and compliance with the relevant duties) to the most senior governance body. However, just as an individual can be both an employee and a director of a company, if a provider is run by a single individual nothing in our measure prevents that individual from being both the individual accountable and the only member of the senior governance body.

5.99 We have amended the measure to make it explicit that senior accountability covers all risks related to illegal harms identified in risk assessments, including those remaining after appropriate Codes measures have been implemented. This is because we want it to be absolutely clear that good risk management requires proper oversight of remaining risks, even after Codes measures have been adopted.

5.100 Snap requested clarification regarding who the accountable individual for this measure should be. Given the organisational differences between services, we consider it is best for providers to determine this (as long as the named individual is accountable to the most senior governance body). We consider it essential that the measure is flexible enough to be implemented by a broad range of services. This is particularly important for this measure, which is recommended for all services in scope of the Act.

5.101 In response to the request for clarity from several respondents on the confidentiality of this appointment, we do not expect service providers to publish the name of the accountable individual or require that service providers routinely notify it to Ofcom.[251]

5.102 In response to the points about naming more than one individual, only one person should be named. In the November 2023 Consultation, we pointed to work commissioned by Ofcom from Milliman which puts individual accountability as the first principle of good governance, drawing on the Institute of Internal Auditors Three Lines Model.[252] Having multiple persons in the accountability role would dilute its effectiveness. We view service providers as best placed to decide how this role works most effectively within their structures. For example, providers of larger services may decide to have multiple risk owners and subject matter experts responsible for risk management controls reporting into the accountable individual.

5.103 We note that Safe Space One queried how feasible this measure would be for providers operating services with end-to-end encryption. This measure recommends that there be an individual who is accountable for compliance. We see no reason why an end-to-end encrypted provider should be unable to identify such an individual. Providers of end-to-end encrypted services are subject to the safety duty and the reporting and complaints duties. In our Codes we have made a number of recommendations applicable to them. In particular, in Volume 2: chapter 2: 'Content moderation', we speak about our position on content moderation of complaints for such services.

5.104 To clarify, in response to stakeholders' concerns (outlined in paragraph 5.80) about the liability of the accountable individual, this measure is not associated with individual liability to comply with information notices issued by Ofcom (unless the provider wishes to give these roles to the same individual).[253] The role relates to internal accountability for compliance with the Act to ensure effective corporate governance. Please see the 'Rights Impact' section below.

Measure on written statement of responsibilities

5.105 The measure recommends that service providers have written statements of responsibilities for senior managers who make decisions related to the management of risks relating to illegal harm. We have changed the wording from 'staff' to 'managers' because we recognise that the individuals concerned may not be employees.

5.106 A statement of responsibilities is a document which clearly shows the responsibilities that the senior manager performs in relation to illegal harm risk management and how these fit in with the provider's overall governance and management arrangements in relation to the service.

5.107 We expect providers to draw up statements to specify areas of responsibility for senior managers, including all key responsibilities for decision-making in illegal harm risk management. This would include, for example, decisions related to the design of products, data-related user safety, and the implementation of trust and safety policies.

---

[251] We have powers to obtain the information if we need it.
[252] Milliman, 2023. Report on principles-based best practices for online safety Governance and Risk Management. This report was commissioned by Ofcom.
[253] Section 103 of the Act

5.108 We acknowledge the requests for clarification on which senior managers would be covered by this measure and for flexibility as organisations change over time. As with our response to Snap in paragraph 5.100, we recognise the differences in how providers operate and have set out this measure with some flexibility to implement it as appropriate for their service. This is to ensure our measure works for the diverse set of services for which it is recommended. Therefore, we have not made any recommendations regarding precisely which managers should be included in the written statement of responsibilities.

5.109 In response to the request for clarity from Meta (WhatsApp) on the confidentiality of this statement, we do not expect service providers to publish the written statement of responsibilities (unless they want to). Nor do we suggest that service providers should routinely notify their statement to Ofcom.

Benefits and effectiveness

Measure on senior accountability

5.110 In our November 2023 Consultation, we highlighted evidence indicating that risk management practices among services improve following the implementation of senior accountability requirements. Senior accountability is a cornerstone of other regulatory regimes, including the Senior Managers and Certification Regime ('SM&CR') jointly regulated by the Financial Conduct Authority ('FCA') and the Prudential Regulation Authority ('PRA').[254] Findings from a 2020 review by the PRA reported positive behavioural change and improvement in risk management practices among services which have implemented SM&CR.[255] These findings were corroborated by our commissioned research on best practice, including a report by Milliman which highlighted individual accountability as the first principle of good governance.[256]

5.111 In the context of online safety risk management, we expect that users are more likely to be exposed to illegal content and risk of harm where there is a lack of senior accountability, oversight, and responsibility. There is the potential for risks to not be adequately considered during decision-making in cases where there is no ownership of this responsibility within organisations.

5.112 By recommending that providers name an accountable individual, we expect that the added scrutiny to the design and operations of a service will lead to an overall material improvement to user safety on the service.

Measure on written statement of responsibilities

5.113 Specifying responsibilities for senior decision-makers is an important feature of other regulatory regimes and has been found to be effective in improving outcomes. Findings from a 2020 review of the FCA's SM&CR reported that many firms surveyed said the requirements of the regime had resulted in clearer articulation of authority and had

[254] The SM&CR is directly underpinned by legislation and serves different outcomes related to compliance with financial regulation, we consider the broad lessons and findings from the FCA's implementation of the regime as instructive for other areas of risk management and regulatory compliance. Source: FCA, 2023. Senior Managers and Certification Regime. [accessed 29 November 2024].
[255] Bank of England, Prudential Regulation Authority, December 2020. Evaluation of the Senior Managers and Certification Regime. [accessed 29 November 2024]. Subsequent references are to this document throughout.
[256] Milliman, 2023. Report on principles-based best practices for online safety Governance and Risk Management.

improved focus on accountability and responsibility.[257] These findings mirrored a 2014 cost benefit analysis of the SM&CR, where large banks surveyed anticipated that statements of responsibility would positively impact behaviour around decision-making and risk.[258]

5.114 From an online safety perspective, the purpose of statements of responsibilities is to ensure that all key responsibilities for decision-making in online safety risk management are assigned to senior management, and that there is clarity in how these responsibilities are owned within a provider.

5.115 International corporate governance principles support ensuring that senior decision-makers have clear responsibilities as part of good risk management. This includes the G20/Organisation for Economic Co-operation and Development's (OECD) Principles of Corporate Governance, which suggests that the specification of accountabilities and responsibilities for managing risk is a "crucial guideline for management" within organisations.[259] We found corroboration of this principle among several good practice models for governance and risk management, including the Committee of Sponsoring Organisations of the Treadway Commission ('COSO') Enterprise Risk Management ('ERM') framework and the Institute of Internal Auditors' ('IIA') Three Lines Model.[260] [261]

5.116 The Milliman report similarly highlights the importance of "having clearly defined roles and responsibilities for all senior managers" and individual accountability in forward-looking risk management systems.[262]

5.117 Based on this research, we expect that senior managers having clearly defined responsibilities over daily operations will contribute to better quality risk assessments. This should lead to a safer design of services and more effective safety mitigations, thereby delivering material benefits. We also would expect this to help embed online safety across the organisation.

Costs and risks

Costs

Measure on senior accountability

5.118 We anticipate that most services will choose to add accountability for compliance to the current portfolio of a senior manager or director who already oversees an online safety, compliance, or risk function. The costs of selecting and naming such an individual are likely

---

[257] PRA, 2020.

[258] "Large banks and investment firms did consider it likely that the policies would result in behavioural changes
as senior managers sought to ensure they would be protected in the event that misconduct or a regulatory breach was discovered, driven by the statement of responsibilities and the presumption of senior responsibility. Such behaviour includes increased due diligence, monitoring and sign-off processes, as well as more formalised and considered decision-making. These actions are all likely to contribute to an increased likelihood that potential and actual regulatory breaches are identified and prevented." Source: Europe Economics, 2014. Cost Benefit Analysis of the New Regime for Individual Accountability and Renumeration. [accessed 29 November 2024].

[259] OECD, 2015.

[260] This model focuses on how clarity of responsibilities among managers supports the proper functioning of internal controls. Compliance Risk Management: Applying the COSO ERM Framework. Source: COSO, November 2020. [accessed 29 November 2024]. Subsequent references are to this document throughout.

[261] The Three Lines model specifies that second line management roles require expertise, support, monitoring and challenge on risk-related matters including risk management. Source: IIA, July 2020.

[262] Milliman, 2023.

to be negligible for such services. Any providers that have not already appointed a suitable individual will incur greater costs in following this measure as they would need to make changes to their internal structure. However, in order to comply with the Act, service providers would already need someone in a suitably senior role who understands the service provider's legal duties. Therefore, we consider the direct incremental costs of implementing this measure are small. In the May 2024 Consultation, we estimated a cost for identifying and training the accountable individual as less than £2,000 and will use that same estimate here.[263] We expect the actual cost to vary depending on the complexity of the organisation, and the regulatory requirements the individual will be accountable for.

5.119 Additional costs associated with the individual spending extra time overseeing online safety (such as backfilling other roles) will depend in part on the risk level of the service and its existing organisational setup. In the November 2023 Consultation, we estimated that if a senior leader who is named the accountable individual in a larger, high-risk service spends ten days more a year considering online safety than they otherwise would, this would result in an increase in costs in the region of £8,000 a year.[264]

5.120 tech UK and Global Partners Digital expressed concerns about costs for providers of smaller services.[265] One individual respondent argued that the measure did not make sense for individuals operating personal non-commercial services. As set out above, in cases where a service is provided by a single individual, that individual would need to be the accountable individual (as well as constituting the most senior governance body of the provider) and thus the direct incremental costs of implementing this measure would be negligible. That individual would anyway need to incur costs in familiarising themselves with their online safety duties under the Act and take any actions to meet their obligations. Such costs are not a direct result of this measure. More generally, low-risk services and smaller services will require less work from an accountable individual than larger and higher-risk services. The costs might be higher for smaller services with substantial risks, but the likely benefits of this measure would also be higher in such cases.

5.121 In response to concerns about individual liability and also the potential impact on companies struggling to recruit people for such a position (as outlined in paragraph 5.83), and as explained in paragraph 5.104, this measure does not make any individual criminally liable, and so salaries would not need to rise to account for this.[266] The costs associated with the measure arise from the time spent by the named accountable individual considering online safety and not from additional risk to that individual.

5.122 There may also be other costs flowing from the change in behaviour of the accountable individual, as they will need to focus more on the provider's illegal content duties. These could be considered indirect costs of the measure, and may be significantly more than the

[263] May 2024 Consultation, Volume 4, paragraph 11.74. We assume training the accountable person takes two days' time of two staff in professional occupations. Labour costs are estimated using the assumptions outlined in Annex 5 which contains the latest wage data released by the Office of National Statistics (ONS) and considers general feedback we have received on cost assumptions.

[264] We assume the "senior leader" has an annual salary of £150,000 and the total cost includes a further uplift as described in Annex 5. Senior leaders staff salaries have not been inflated since the November 2023 Consultation as these are broader estimates and are not derived from specific annual ONS surveys. This also applies to our analysis of other relevant measures in this chapter.

[265] Global Partners Digital response to November 2023 Consultation, pp.9-10; techUK response to November 2023 Consultation, p.3.

[266] Global Partners Digital response to November 2023 Consultation, p.9.

costs discussed in paragraph 5.118. As providers are only likely to make such changes when it helps protect users from illegal content, any such costs would tend to rise in line with benefits for users.

5.123 We are not aware of evidence suggesting that the direct costs of time spent by the accountable individual overseeing online safety would be higher than our estimates set out in the November 2023 Consultation. Therefore, our cost assumptions and estimates are unchanged, other than the addition of the identification and training cost outlined in the May 2024 Consultation.

Measure on written statement of responsibilities

5.124 In our November 2023 Consultation, we explained that this measure would involve one-off costs to develop the statements as well as on-going costs to maintain a centralised reference of responsibilities and review it when necessary. There may be some additional costs to agree on areas of responsibilities. We estimated a first-year cost of approximately £16,000 to produce written statements of responsibilities for ten senior managers.[267]

5.125 These costs will increase with service size and with the number and complexity of risks. Costs will naturally be smaller for providers of services with few senior managers and will increase for providers of larger services (depending on the number of senior managers that make decisions relating to the management of risks). However, the costs will likely be a larger proportion of the annual revenue of providers of smaller services.

5.126 In response to our November 2023 Consultation, Meta (WhatsApp) said the estimated time investment to create the written statement of responsibilities was too short for a global and matrixed company.[268] We also received general concerns over the total cost of the governance measures.

5.127 Costs incurred by providers of services will vary according to service complexity and size. We consider our assumptions for time and resource reasonable and illustrative of a broad range of services, but very large and highly complex services may incur costs that are higher than our estimates.

5.128 For most services, we continue to assume that if there were ten senior managers who needed statement of responsibilities and it took a few days each to develop the statements, this will lead to an approximate implementation cost of £16,000.[269] Using our standard assumption that on-going costs are 25% of implementation costs, the annual costs of reviewing and updating the statements will be up to £4,000.

Risks

Measure on senior accountability

5.129 Some respondents expressed concerns that the named accountable individual could face criminal prosecution. As outlined in paragraph 5.92, this measure relates only to compliance with the specified duties in the Act. It does not make the named individual

[267] We assumed that on average it would take three days to develop and agree each statement, and that the time would mostly be of senior managers. We assumed an annual salary of £100,000 for senior managers of a large service, and used the non-wage uplift assumption in Annex 14 of the November 2023 Consultation.
[268] Meta response to November 2023 Consultation, p.6.
[269] We assume on average it would take three days to develop and agree each statement, and that the time would mostly be of senior managers. As in our November 2023 Consultation, we assume an annual salary of £100,000 and the total cost includes a further uplift as described in Annex 5.

liable for responses to information requests or for any overseas compliance failures on the part of the service provider.[270]

Measure on written statement of responsibilities

5.130 5Rights Foundation argued that decisions impacting safety should be understood and implemented at all organisational levels, noting that such a measure could interfere with wider accountability for mitigating against risk.[271]

5.131 We agree with the importance of establishing a culture of compliance within organisations, and expect that, as a package, the measures included in this statement will contribute to this (including those which relate to training and content moderation).[272] This will help to ensure that better decisions are made about service design and operation in the context of safety risks posed to users. Moreover, this measure does not place total responsibility for compliance with the senior managers who have a written statement of responsibilities. Instead, it provides clarity regarding responsibility for different decisions that affect the management of illegal harm risks and ensures safe outcomes for users.

Rights impact

Freedom of expression

5.132 We have considered whether these measures impact on any the rights to freedom of expression or association. In particular, we have considered the responses which stated that these measures could pose risks of censorship and increased potential for influence from governments, see paragraph 5.88.

5.133 As outlined in paragraphs 5.92, the purpose of these measures is to identify individuals and managers responsible for compliance with the duties in the Act, and to clarify their roles in relation to the provider's internal processes and structures. It does not have any impact on their roles in relation to the laws of other jurisdictions, nor does it require providers to take steps with regard to particular types of content. As such, we do not accept that these measures impact on users' or services' rights to freedom of expression. Nor do we see any impact on rights to freedom of association.

5.134 To the extent that it helps to reduce harm on the service and make users feel safer, this measure could also positively impact on their human rights.

Data Protection and Privacy

5.135 We have also considered whether these measures have an impact on the right to respect for privacy and family life.

5.136 The measures recommend that aspects of individuals' roles be defined in a way which means the provider holds a record of the information. A third party might ask to see the information, and may have the power to compel its provision. Insofar as information which would not otherwise be recorded about an individual would, as a result of the recommendation, be recorded, it could have a small impact on privacy. However, the information relates only to compliance with specified duties in the Act, and would in any event only be codifying the activities the individual was supposed to be carrying out as a

[270] See sections 103 and 110 of the Act.
[271] 5Rights Foundation response to November 2023 Consultation, p.9.
[272] See Volume 2: chapter 2: 'Content moderation' for further details.

part of their role. We would therefore expect the information to exist in some form in any event.

5.137 It can be assumed that a service provider would have to process the personal data of the accountable person and of the senior members of staff named in the statement of responsibilities. However, it is likely that most providers would have already collected and processed this data in some form (for example, through employment contracts, role descriptions and performance management).

5.138 Where a provider processes the personal data of any individual, it should do so in accordance with applicable data protection legislation, which acts as a safeguard for the individual's rights. We recognise that accountable individuals and senior managers may be located in jurisdictions which do not have data protection laws, and to which UK data protection laws do not apply. However, our measure does not recommend that the record be held in those jurisdictions. Ofcom's primary duty is towards citizens and consumers in the UK.[273]

5.139 As with other governance measures, we consider that a well-managed business is, in general, more likely to comply with its obligations under privacy and data protection laws (as well as, for that matter, other important laws such as those relating to consumer protection and equality). As such, our proposals may help to safeguard these.

5.140 Overall, we consider that if there is any interference with the right to privacy, it is proportionate in the interests of public safety, for the prevention of crime, for the protection of morals, and for the protection of the rights and freedoms of others, in the UK and for UK users.

Who these measures apply to

Measure on senior accountability

5.141 In our November 2023 Consultation, we proposed this measure for all U2U and search service providers.

5.142 As set out in paragraph 5.110, clearly defining senior accountability materially improves risk management and associated safety outcomes. This measure is an important complement to the risk assessment process as it ensures that a particular individual is responsible for and must be able to explain the actions a service provider takes to protect users from the risks identified. It helps to ensure that risks to users online are given due weight and consideration within an organisation and reduces the risk of governance failures due to a lack of accountability which would put users at risk. It can also provide indirect benefits for some services. For example, providers of smaller services can manage risk more effectively from an early stage (evolving and expanding their approach as the business grows). They can therefore address any illegal harms issues early and may even save costs overall.

5.143 We consider that it is proportionate to recommend this measure for all U2U and search services, given the significant benefits and relatively low cost of implementing this measure. As set out in paragraph 5.120 in the 'Costs section', the measure imposes minimal costs on providers of small low-risk services, and costs for riskier services may be higher.[274]

[273] Section 3 Communications Act 2003.

[274] For example, services that identify more or higher levels of risks for illegal content will incur higher costs to implement the measure, as the impact on the named person's time (as a result of being formally made accountable) will be greater.

However, we expect the possible benefits for users to be correspondingly greater for riskier services.

Measure on written statement of responsibilities

5.144 In our November 2023 Consultation, we consulted on this measure for all providers of large U2U and large general search services as well as all providers of multi-risk services (including vertical search services which are multi-risk).

5.145 This measure will ensure clear lines of responsibility, supporting good risk management and ensuring managers consider how their decisions contribute to risks having to do with illegal harm in relation to individuals in the UK (see paragraphs 5.113-5.117).

5.146 In their response to the November 2023 Consultation, BT Group stated that this measure should apply to smaller services with specific risks, while three respondents to the May 2024 Consultation suggested this measure should apply to smaller and single-risk, or all, service providers.[275] As set out in paragraph 5.90, some stakeholders argued that governance measures that apply to multi-risk services should be extended to single-risk services, or even to all services.[276] There were also general comments stating that this measures could be disproportionate for providers of some types of services,[277] that a more proportionate or risk-based approach should be adopted,[278] or that the measure could be better tailed by provider capacity.[279]

5.147 We recommend the measure for all multi-risk services. Having statements of responsibilities for senior managers who perform functions relevant to online safety risk management will have considerable benefits for multi-risk services because they pose significant risks and need to co-ordinate activities across multiple harms. Unclear responsibilities in this context could result in gaps in risk management. Furthermore, we consider that costs are likely to be small in relation to these benefits.

5.148 We also recommend this measure for all providers of large U2U and general search services that are not multi-risk (including those that are low-risk or that have a single medium or high risk). The complexity of organisational structures within large organisations means that clarity of responsibilities is important in ensuring that risk management activities are properly scrutinised to ensure their effectiveness. It also reduces the risk of service providers failing to identify risks that may affect a large number of users (particularly in a rapidly changing risk environment). We acknowledge that benefits may be lower for providers of large services with low risk of harm, but we also expect that costs will be lower than for providers of large multi-risk services because the written statement of

---

[275] BT Group response to November 2023 Consultation, pp.1-2; C3P response to May 2024 Consultation, p.12; CCDH response to May 2024 Consultation, p.9; Children's Commissioner for England response to May 2024 Consultation, p.23.

[276] ACCO response to November 2023 Consultation, p.2; IWF response to November 2023 Consultation, p.22; Logically response to November 2023 Consultation, p.12; Molly Rose Foundation response to November 2023 Consultation, pp.39-40; NSPCC response to May 2024 Consultation, p.19; NSPCC response to November 2023 Consultation, p.6; OSA Network response to November 2023 Consultation, p.95; Snap response to November 2023 Consultation, p.5; The Cyber Helpline response to November 2023 Consultation, p.3.

[277] [✂]; Evri response to November 2023 Consultation, p.2; Global Partners Digital response to November 2023 Consultation, p.10; ODDA response to November 2023 Consultation, pp.1-2; [✂].

[278] Bolton, C. response to November 2023 Consultation, p.1; Mega response to November 2023 Consultation, p.5; [✂]; Name Withheld 3 response to November 2023 Consultation, p.3; [✂]; techUK response to November Consultation, pp.3, 8.

[279] CELE response to November 2023 Consultation, p.4.

responsibilities will be simpler. There will also be flexibility over how it is implemented for service providers with fewer resources. We expect the costs to be low relative to the benefits to user safety for such services.

5.149 We do not currently consider it proportionate to recommend this measure to large vertical search services that are not multi-risk because they typically pose very limited risks (if any).[280]

5.150 We have considered the responses from certain stakeholders calling on us to extend governance measures to smaller providers of single-risk services or small low risk services.

5.151 Our current assessment is that it would not be proportionate to recommend this measure for smaller services that are low risk. We consider the additional benefits of formal written responsibilities would be more limited compared to smaller multi-risk services (all else being equal) as they are likely to have fewer and less complex relevant risk management activities. For such services we rely on our measure on having an accountable person and give providers the flexibility to choose how to ensure clarity of responsibilities without recommending formal written arrangements.

5.152 As set out in 'Our approach to developing Codes measures', we have decided not to extend the services in scope of these measures at this time. In Spring 2025, we intend to consult further on the case for applying some or all of the measures currently recommended only to multi-risk services to some or all single-risk services

5.153 Based on our analysis, we are maintaining our position of recommending this measure for all providers of large U2U and large general search services as well as all providers of multi-risk services.

Conclusion

5.154 We have decided to recommend both of the measures essentially as proposed in our November 2023 Consultation with the minor clarification that accountability of the individual entails being required to explain and justify actions or decisions regarding illegal harm risk management and mitigation (including risks remaining after the implementation of appropriate Codes of Practice measures), and compliance with the relevant duties, to the most senior governance body. We have also made some changes to the drafting to make it clearer that individuals need not be employed by the provider, and to focus the measure on illegal harms.

5.155 Based on our analysis, we consider that the measures in question will confer significant benefits and that the costs associated with them are relatively limited.

5.156 Our measure on senior accountability is recommended to all regulated service providers, whereas our measure on written statement of responsibilities is recommended to providers of large services (excluding large vertical search services) and all providers of multi-risk services.

5.157 These measures will be included in our Codes of Practice on terrorism, CSEA and other duties. They are referred to within these Codes as ICU A2 and ICS A2 for our senior

[280] By their nature vertical search services are unlikely to have content that is as rapidly changing as U2U services and search results are more under a service's control than for U2U content. We are also not aware of evidence of such services showing illegal content. Any benefits of applying this measure would therefore be low for vertical search services. See a more detailed explanation in our Register of Risks chapter titled 'Search'.

accountability measure and ICU A3 and ICS A3 for our written statements of responsibilities measure.

Measures on internal assurance and compliance functions, and on tracking evidence of new and increasing illegal harm

5.158 In our November 2023 Consultation, we proposed two measures to support providers in developing appropriate internal assurance and compliance functions:

- A measure recommending providers of services that are both large and multi-risk to have an internal monitoring and assurance function to provide independent assurance that measures are effective.
- A measure recommending providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk) to track evidence of new kinds of illegal content on a service, or increases in particular kinds of illegal content or illegal content proxy, and report this to the most senior governance body. For U2U services, the proposal also recommended that the provider track equivalent changes in the use of the service for the commission or facilitation of priority offences.

5.159 We stated that for risk management activities to be effective, organisations need to establish policies and processes to confirm that internal controls are appropriate and effective to address identified risks of harm. This ensures that risks identified in risk assessments are properly managed and mitigated on an ongoing basis, and that compliance requirements are met.

5.160 We also highlighted that as well as assuring the effectiveness of the measures taken, this may also enable services to identify and take action to address new forms of illegal content.

5.161 In this section, as in our November 2023 Consultation, we set out our assessment and decisions relating to these two measures together as they both relate to internal assurance and compliance functions, and we consider them to be closely linked in their intended outcome.

5.162 In our November 2023 Consultation, we also considered several other options regarding how services can assure their measures to mitigate and manage illegal content are effective.[281] These could be alternatives to the measure proposed or could supplement them. However, we do not consider there is currently enough information on the effectiveness of other possible measures to be able to recommend them in our Codes at this stage. Responses to all of these other options are outlined in Annex 1.

[281] These included external audits, requiring services to undertake due diligence on third-party providers of online safety services to assure that their approaches to mitigating and managing risks to users are effective and requiring metrics to measure the effectiveness of measures.

Summary of stakeholder feedback[282]

Internal assurance and compliance functions

5.163 We received a broad range of support from some stakeholders on this measure, including some agreeing on who this measure should be applied to.[283] Others provided more details on what they supported specifically:

- The Centre for Competition Policy supported the decision to extend the obligations of large and multi-risk service providers to include the duties under this measure.[284]
- The Molly Rose Foundation welcomed the independent assurance on risk mitigation processes that this measure will provide.[285]

5.164 We have identified several other themes that emerged from stakeholder feedback which we discuss in the following paragraphs.

Feedback on how the measure works

5.165 In response to both the November 2023 Consultation, and the May 2024 Consultation, some stakeholders asked about aspects of the measure:

- Protection Group International queried who would be suitably qualified carry out the independent assurance checks, and whether an internal audit mechanism by a platform can be relied on.[286]
- UK Safer Internet Centre (UKSIC) told us there should be independent mechanisms to assess the effectiveness of measures, in addition to internal oversight.[287]
- Health Professionals for Safer Screens asked who will check if this and other governance measures are being effectively implemented.[288]
- The Association of Police and Crime Commissioners suggested we specify a maximum time between reviews.[289]

5.166 We address this feedback in the section 'How these measures work'.

Feedback on who measure applies to

5.167 We also received several responses regarding who this measure should apply to:

- Roblox raised the concern that this measure would put a disproportionate burden on providers that are not the largest, have nimble resourcing, and/or have existing effective solutions in place. It shared the view that a dedicated assurance team is not always necessary, especially when there are less burdensome solutions available that are proven effective. Examples of available solutions that could be used instead of a

---

[282] Note that this list in not exhaustive and further responses can be found in Annex 1.
[283] Booking.com response to November 2023 Consultation, p.4; Centre for Competition Policy response to November 2023 Consultation, p.8; LinkedIn response to November 2023 Consultation, p.3; Meta (WhatsApp) response to November 2023 Consultation, p.7; Microsoft response to November 2023 Consultation, p.2; Molly Rose Foundation response to November 2023 Consultation, p.41.
[284] Centre for Competition Policy response to November 2023 Consultation, p.8.
[285] Molly Rose Foundation response to November 2023 Consultation, p.41.
[286] Protection Group International response to November 2023 Consultation, p.2.
[287] UKSIC response to May 2024 Consultation, p.24.
[288] Health Professionals for Safer Screens response to May 2024 Protecting Children from Harms Online Consultation, p.8.
[289] Association of Police and Crime Commissioners response to May 2024 Consultation, p.8.

separate assurance function were existing trust and safety teams, or a combination of multiple teams in the online safety process.[290] We address this feedback in the section titled 'Who this measure applies to'.

- Age Verification Providers Association stated that this measure should apply to large services with specific risks. It said that it does not seem logically defensible that the number of risks would affect the degree of governance for risk, noting that large services will generally have these in place already (so an additional focus would not have a disproportionate impact).[291] We address this feedback in the section 'Who this measure applies to.'

5.168 We address this feedback in the section titled 'Who these measures apply to'.

Tracking evidence of new and increasing illegal harm

5.169 Several respondents to both our November 2023 and May 2024 Consultations supported this measure, with some of them providing additional suggestions.[292] Some shared concerns about our approach, or stated that we should go further in areas.

Design of measure

5.170 Respondents queried how our measure may work:

- The Cyber Helpline highlighted the need for adequate resources and capacity for the providers in scope of the measure. It also encouraged us to use this as an opportunity to work with non-governmental organisations (NGOs) and charities to provide insight based on users experiencing issues relating to illegal harm. It said that this would allow us to make informed decisions on the accountability measures taken when our Codes are breached.[293]

- Two respondents expressed concerns that content tracking was too limited.[294] One of these respondents highlighted that it would miss "harms like fraud and some ephemeral interactions like in gaming or VR/AR", allowing them to go unnoticed, that it "puts the focus on finding and destroying content, not systemic approaches that could have an impact on reducing the production of, and/or incentives to create, harmful content". It also noted that the measure could misrepresent content removed as a metric of success and subsequently detract focus from "dealing with root causes of harms or the scale of harms caused, such as total number of users exposed to harms".[295] The other respondent suggested that "tracking and reporting of content should ideally be broadened to encapsulate reporting detected and self-reported harms and safety events on these services (for example, user usage of safety/moderation

[290] Roblox response to November 2023 Consultation, p.3. We note Roblox made similar comments in response to the May 2024 Consultation. Source: Roblox response to May 2024 Consultation on Protecting Children from Harms Online, p.7.
[291] Age Verification Providers Association response to November 2023 Consultation, p.2.
[292] BT Group response to November 2023 Consultation, p.2; [✂]; The Cyber Helpline response to November 2023 Consultation, p.3; TSB Bank response to November 2023 Consultation, p.1; Welsh Government response to May 2024 Consultation on Protecting Children from Harms Online, p.6.
[293] The Cyber Helpline response to November 2023 Consultation, p.3.
[294] SPRITE+ (University of Glasgow) response to November 2023 Illegal Harms Consultation, p.3; Integrity Institute response to November 2023 Illegal Harms Consultation, p.3.
[295] Integrity Institute response to November 2023 Illegal Harms Consultation, p.3.

tools)".[296] They added that transparency from providers of 'social VR' services would be important in analysing the extent of harms which may manifest on them.

- 5Rights Foundation supported the measure but noted that it does not include an obligation to take the insights gathered from tracking new and unusual illegal content to the board, or for the board to take action upon receiving such information and argued that this should be the case.[297]
- [✂] and White Ribbon Canada wanted us to ensure that any rise in detection of illegality was met with a strong and proactive approach by service providers.[298]
- In addition, the OSA Network said the equivalent measure in our May 2024 Consultation is not forward-looking, as horizon scanning means identifying risks before they occur, not after evidence of them is available. It also expressed concern that there was no related governance responsibility to act on the findings from this measure.[299]

5.171 We address this feedback in the section 'How this measure works'.

Feedback on who this measure applies to

5.172 We received a number of responses relating to who this measure should apply to:

- Roblox expressed concerns that this measure may not be proportionate for some small services due to costs. It stated that this measure seems to be based on an assumption that smaller service providers would be able to absorb the measures, when it could be difficult for services that do not have a head-start on such measures to implement them.[300] NSPCC and C3P also argued that the measure should apply to smaller services to ensure a proactive approach in tackling illegal harms.[301]
- One stakeholder argued that this measure should not apply to providers solely on the basis of the size of their service, but to large services which are multi-risk or multi-risk services, where there is a high risk of at least one priority offence.[302]
- One stakeholder argued that downstream search services should be exempt from this measure. This is because downstream search services may not have the means to track increases in, or new kinds of, harmful content appearing in search results because such services obtain their results from upstream search services.[303]
- Mega queried the technical feasibility of tracking and monitoring encrypted content (such as messages) by services using end-to-end encryption, specifically for small service providers.[304]
- Logically argued that smaller services should be in scope of this measure due to foreign interference risks.[305]

---

[296] SPRITE+ (University of Glasgow) response to November 2023 Illegal Harms Consultation, p.3.
[297] 5Rights Foundation response to November 2023 Consultation, p.10.
[298] [✂]; White Ribbon Canada response to November 2023 Illegal Harms Consultation, p.2.
[299] OSA Network response to May 2024 Consultation, p.22.
[300] Roblox response to November 2023 Consultation, p.16. We note Roblox made a similar point in response to the Ofcom May 2024 Consultation. Source: Roblox response to May 2024 Consultation, p.7.
[301] NSPCC response to November 2023 Consultation, p.7; C3P response to November 2023 Consultation, p.5.
[302] Airbnb response to November 2023 Consultation, p.3.
[303] [✂].
[304] Mega response to November 2023 Illegal Harms Consultation, pp.3-5.
[305] Logically response to November 2023 Consultation, p.12.

- BT Group stated that this measure should apply to smaller services, especially those with particular or multi-risks, as providers could continue to categorise themselves as low risk without it.[306]
- In response to an equivalent measure in our May 2024 Consultation, the Scottish Government, and OSA Network said this measure should be expanded to all services.[307]

5.173 We address this feedback in the section 'Who this measure applies to'.

Feedback on rights

5.174 The ICO noted that the tracking of evidence of illegal harms as part of this measure is likely to involve the processing of personal data and that service providers will need to comply with data protection law when doing so.[308] We address this feedback in the section 'Rights impact'.

Our decision

Internal assurance and compliance functions

5.175 We have decided to confirm the measure broadly as we proposed in the November 2023 Consultation.

5.176 We have adjusted this measure and clarified that the function should report its findings to a governance body or audit committee, which should consider the report. This is to reiterate the importance that the findings inform decision making as appropriate, which we consider likely ultimately to result in a more effective approach to risk management. See section 'Oversight and management of all risks identified' at paragraph 5.22 for further details. The amended version of this measure now reads:

- The provider 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:
  - the body that is responsible for overall governance and strategic direction of a service; or
  - an audit committee.
- This independent assurance may be provided by an existing internal audit function.

5.177 This measure will apply to providers of services that are both large and multi-risk.

5.178 The full text of the measure can be found in our Illegal Content Codes of Practice for U2U and search services on terrorism, CSEA and other duties. We refer to this as ICU A4 for U2U services and ICS A4 for search services.

Tracking evidence of new and increased harm

5.179 Having analysed responses, we have decided to recommend the measure essentially as proposed in our November 2023 Consultation, although we have made some small amendments to be clear that it relates to the UK.

---

[306] BT Group response to November 2023 Consultation, pp.1-2.
[307] Online Safety Act Network response to May 2024 Consultation, p.22; Scottish Government response to May 2023 Consultation, p.9.
[308] Information Commissioner's Office ('ICO') response to November 2023 Illegal Harms Consultation, pp.9-10.

5.180 The full text of the measure can be found in our Codes of Practice for U2U and search services on terrorism, CSEA and other duties. We refer to this as ICU A5 for U2U services and IC5 A5 for search services.

Our reasoning

How these measures work

Internal assurance and compliance functions

5.181 Service providers should have an internal monitoring and assurance function in place to provide independent assurance on an ongoing basis on how effective measures in place are at mitigating and managing the risks of harm identified in the risk assessment.[309] This function should report to an overall governance body or audit committee, which should consider the findings presented to it. We do not recommend any specific qualifications for this function, as we consider service providers best placed to determine this.

5.182 As set out in paragraph 5.165, a respondent requested we provide maximum time limits between reviews. At this stage, we consider that the frequency and timing of independent assurance should be part of the service provider's risk management strategy. We expect that a service provider's leadership is best placed to determine the review and reporting requirements that will ensure measures are effective on an ongoing basis.

5.183 As set out in paragraph 5.165, some respondents questioned the reliability of internal assurance functions and suggested that this measure could be strengthened by having regular checks on the effectiveness of the internal assurance and compliance functions.[310] Another recommended putting in place assurance checks by an independent third party in addition to internal assurance and compliance functions to avoid relying solely on internal oversight.[311]

5.184 We considered the role of independent third-party audits in our November 2023 Consultation. Based on the evidence we have reviewed, we expect that the internal assurance and compliance functions measure outlined in this chapter will bring sufficient material benefits in mitigating risks. However, this does not mean that providers are not expected to consider third-party information. Our Risk Assessment Guidance for Service Providers includes third-party expertise as a kind of enhanced input in risk assessments. Our Register of Risks ('Register') can also help providers consider third-party information about how harms manifest online.

Tracking evidence of new and increased harm

5.185 With this recommended measure, service providers should track evidence of new kinds of illegal content, as well as unusual increases in particular kinds of illegal content.[312] This

[309] 'Assurance' refers to the verification of risks and mitigation and internal controls, including activities around effectively identifying, measuring, and managing risks.
[310] Health Professionals for Safer Screens response to May 2024 Consultation, p.8; Protection Group International response to November 2023 Consultation, p.2; UKSIC response to May 2024 Consultation, p.24.
[311] UKSIC response to May 2024 Consultation, p.24.
[312] To determine this, the provider should establish a baseline understanding of how frequently particular kinds of illegal content, illegal content proxy, or the commission of facilitation of priority offences occurs on the service to the extent possible based on its internal data and evidence. The service provider should use this baseline to identify unusually high spikes in the relevant data.

monitoring and evaluation can ensure effective risk management, including obtaining relevant evidence to make accurate judgements of risk on the service.

5.186 Given that illegal harm is highly likely to change over time, monitoring how harms are manifesting on services will be necessary to ensure that existing mitigations are effective and adequate to prevent all kinds of illegal harm. Any new kinds of illegal content, unusual increases in illegal content or illegal content proxy, or (for U2U services) equivalent changes in the use of the service for the commission of facilitation of priority offences, should be reported through a provider's relevant governance channels to the most senior governance body.

5.187 To track these changes, providers may use evidence derived from reporting and complaints processes, content moderation processes, referrals from law enforcement and information from trusted flaggers and third parties, proxies, sampling, market research, and any other expert groups. The measure allows service providers the flexibility to track proxies of illegal content (provided these are suitable) instead of tracking changes in illegal content.[313]

5.188 There will be differences in how this applies to U2U and search services. U2U services need to track for new and increasing illegal harm, including evidence of their service being used to commit or facilitate an offence, as well as the presence of illegal content. Search services only need to track new and increasing kinds of illegal content.

5.189 We recognise that the way in which this works in practice is likely to be different for different services, depending in particular on how closely their complaints and content moderation processes use definitions which precisely track the UK's definition of illegal content and (for U2U services) the facilitation and commission of offences. A provider may need to use proxies, sampling, market research among users, and/or information from third parties such as non-governmental organisations (NGOs) and charities to provide insight based on users experiencing issues relating to illegal harm where it does not have precise information. Providers of end-to-end encrypted services, for whom activity on their service is less visible, may also need to look to alternative sources of information to understand their risks fully.

5.190 Service providers should establish a baseline of how frequently illegal harms occur on their service and monitor unusual trends against this. We recognise stakeholder concerns, outlined in paragraph 5.170, around the lack of an obligation to take action on these. However, we have explained in paragraph 1.39-1.43 of 'Our approach to developing Codes measures' why we cannot recommend this in Codes. Providers should report findings on new kinds of illegal content or increases in illegal content through relevant governance channels to the most senior governance body, for them to make appropriate and proportionate risk-based decisions.

5.191 As outlined in paragraph 5.170, two respondents expressed concerns about content tracking missing a number of harms, as well as ephemeral interactions. The Online Safety Act includes risk assessment duties for all in-scope providers to assess the risk of illegal content, taking into consideration how the service is used and its relevant functionalities

[313] As set out in Volume 2: chapter 2: 'Content moderation', we recognise that many service providers design their terms of service and community guidelines both to comply with existing laws in multiple jurisdictions and to meet their own commercial needs. We therefore consider that service providers should have a choice. They may set about making illegal content judgements in relation to individual pieces of content for the express purpose of complying with the safety duties. The alternative is that they moderate illegal content by reference to provisions in their terms of service which would be cast broadly enough to necessarily cover illegal content.

such as ephemeral messaging. Ofcom's Risk Profiles will also help services identify risks associated to specific features or functionalities. We expect that a service should be able to use its governance process to consider whether the measures it has in place continue to be appropriate and proportionate to changes in risks, or whether action is required. We also recognise the role that transparency reporting can play in increasing knowledge about safety online. We have separately consulted on the transparency powers conferred on us by the Online Safety Act and will publish our decisions in due course. Further, service providers are in any case required to have proportionate systems and processes to remove illegal content of which they become aware, and the Act does not give us powers to require providers to preserve evidence for criminal enforcement purposes.[314] As our powers to address illegal harms are therefore related specifically to the removal of illegal content, the decision to retain data is a matter for the service provider.

Benefits and effectiveness

Internal assurance and compliance functions

5.192 Most service providers will have to implement some mitigations to protect their users from risks of harm on their service, but these mitigations will only work if they are adequately and independently scrutinised for effectiveness.

5.193 As outlined in the November 2023 Consultation, evidence demonstrates that strong internal controls are instrumental to the effectiveness of risk mitigations across industries.[315] For example, consultations by both the Department for Business, Energy, and Industrial Strategy (BEIS) and the European Commission found that strengthening internal controls can improve the quality and reliability of corporate reporting and ensure effective management of risk, including compliance risk.[316] This is corroborated by best practice guidelines and controls on governance and internal assurance and audit. This includes references to monitoring and review of the effectiveness of risk controls in ISO 31000 on risk management.[317]

5.194 Ensuring that roles which provide objective assurance and advice on the adequacy and effectiveness of governance and risk management are independent is usually necessary to ensure objectivity, authority, and credibility. Independence in these functions can be established by having direct accountability between the function and the overall governing body, having unfettered access to people, resources, and data necessary to complete work,

[314] Section 10(3)(b) of the Act.
[315] European Commission, 2022. Corporate reporting – improving its quality and enforcement [accessed 29 November 2024]; Department for Business, Energy, and Industrial Strategy, 2022. Restoring trust in audit and corporate governance. [accessed 29 November 2024].
[316] The European Commission's consultation on corporate reporting found overall support from respondents in favour of ensuring effective internal controls to improve the effectiveness and efficiency of corporate governance mechanisms. Notable responses to this consultation included comments from professional services organisations, which pointed to evidence that establishing and embedding a system for monitoring and reporting of internal controls improves the quality of financial reporting (PwC) and reduces the risk of corporate failure and fraud (Deloitte). We also found support for stronger internal control frameworks reflected in response to a 2022 BEIS consultation, which cites improved reporting and audit and better corporate governance as key outcomes. Sources: European Commission, 2022. Corporate reporting – improving its quality and enforcement [accessed 25 October 2024]; Department for Business, Energy, & Industrial Strategy, 2022. Restoring trust in audit and corporate governance. [accessed 25 October 2024].
[317] International Organization for Standardization, 2020. ISO 31000 Risk Management. [accessed 25 October 2024].

and having freedom from bias or interference in the delivery of findings on effectiveness of controls.[318]

5.195 We do not envisage independence as requiring providers to engage an independent third party (such as an external auditor) to confirm effectiveness of mitigations, although providers may choose to do so. In our November 2023 Consultation we sought additional evidence of the efficacy, costs and risks associated with a potential future measure to requiring services to have measures to mitigate and manage illegal content risks audited by an independent third-party. We want to thank the large number of stakeholders that have taken the time to engage with this, detailed further in Annex 1. At this time, we have decided not to add any additional measures to what we consulted on in November 2023. We will continue to assess the proportionality of potential measures as we plan and work towards future iterations of our Codes.

5.196 The overall objective for independence of the monitoring and compliance function is to ensure that providers find a way to achieve as much independent oversight and challenge as possible for each task. For services where having dedicated members of staff in a monitoring and assurance function is not possible, there may be an option to structure the organisation to try to ensure that oversight of tasks within the monitoring and assurance function can be done by another individual in the firm who is not directly involved with that task.

5.197 Analyses of serious corporate governance failures have focused remediation and prevention recommendations on improving the effectiveness of internal controls and oversight processes – including assurance and compliance functions. Bolstering the independence of assurance functions, and having heads of functions report directly to the overall governance body or supervisory board have been found to support more robust internal oversight.[319]

5.198 However, we also note evidence that questions the efficacy of internal assurance measures if risk management policies and processes are not properly implemented. The proper implementation of controls is a prerequisite for the effectiveness of risk management policies and processes, as demonstrated by case studies where scrutiny of internal controls failed despite the presence of internal assurance and audit functions.[320]

5.199 By implementing a function that is independent and reporting to the overall governance body, service providers will be better informed of how effective measures are in mitigating risks and the levels of risk the service remains exposed to after implementing measures to address the illegal harms identified in the provider's risk assessment. As a result, providers will be better able to check that mitigations and controls are adequate, and address any safety concerns effectively. We consider that this will make providers better placed to make

[318] IIA, 2020.

[319] Krahnen, P.K., Langenbucher, K., Leuz, C., Pelizzon, L. 2020. Wirecard Scandal: When All Lines of Defense Against Corporate Fraud Fail Oxford Business Law Blog, 23 November. [accessed 25 October 2024]. What are the wider supervisory implications of the Wirecard case, study requested by the European Parliament ECON committee, Katja LANGENBUCHER Christian LEUZ Jan Pieter KRAHNEN Loriana PELIZZON

[320] This includes the case study of India's Yes Bank, which faced charges of money laundering. Despite referring to a clear risk management framework based on the IIA's Three Lines Model and having an internal audit department, ineffective implementation of policies in Yes Bank meant that serious financial risks were not managed or mitigated. Source: Teen, M.Y. (ed.), 2021. Yes Bank, No Governance. Corporate Governance Case Studies 10 [accessed 29 November 2024].

good decisions about what steps to take to protect people from illegal content, thereby delivering significant benefits.

Tracking evidence of new and increased harm

5.200 The online risk environment is dynamic and constantly subject to change. User behaviour can also change over time meaning that users could become more exposed to the risk of encountering illegal content or experiencing harm. Examples may include users becoming desensitised to, or fatigued by, warning messages.[321] In addition, new risks may emerge over time in response to changes in technology, society, or user behaviour.

5.201 As we noted in the November 2023 Consultation, changes in the external environment can often lead to a greater prominence of certain kinds of illegal harm on services, such as foreign interference campaigns organised around conflicts, public health emergencies, and political processes.[322]

5.202 If providers monitor changes on their service, they will be better able to identify and respond to new risks quickly and effectively. We expect that this will make services safer than they otherwise would have been, thus delivering significant benefits. Real-time monitoring will also help to safeguard against any changes to the risk environment that occur after risk assessments have been made. If providers do not implement this measure, changes to the external environment may not trigger risk assessments or necessary mitigation actions. We expect that evidence gathered through tracking will also assist in future risk assessments.

Costs and risks

Internal assurance and compliance functions

5.203 In our November 2023 Consultation, we explained that this measure would have significant costs, largely for staffing and (if necessary) training.

5.204 In considering how many staff service providers might typically need for their internal monitoring and assurance function, an important reference point is the number of staff employed in internal audit functions. These functions evaluate an organisation's internal controls, especially its corporate governance and accounting processes. They may also oversee the organisation's risk management processes and may focus on a specific area of a business, such as cybersecurity. Internal audit functions tend to have fewer staff in non-profit organisations and privately held companies, and more staff in public sector organisations, publicly traded companies, and financial services.[323] Smaller organisations

---

[321] Burgess, M., 2018. The tyranny of GDPR popups and the websites failing to adapt, Wired, 29 August. [accessed 29 November 2024]

[322] Nimmo, B., Torrey, M., 2022. Taking down coordinated inauthentic behavior from Russia and China , Meta. [accessed 29 November 2024]; UK Foreign Commonwealth & Development Office 2022. UK exposes sick Russian troll factory plaguing social media with Kremlin propaganda [accessed 29 November 2024]; DiResta, R., Shaffer, K., Rippel, B., Sullivan, D., Matney, R., Fox, R.,; Albright, J.,; and Johnson, Ben, 2019 The Tactics and Tropes of the Internet Research Agency, New Knowledge [accessed 29 November 2024];
Schliebs, M.,; Bailey, H., Bright, J., and Howard, P. J. 2021 GEC Special Report: The Kremlin's Chemical Weapons Disinformation Campaigns, U.S. Department of State, Global Engagement Center. [accessed 29 November 2024].

[323] The 2022 Internal Audit: A Global View found that 51% of audit functions had 5 or fewer people. At the other extreme, 10% had 51 of more staff. This was based on 3,600 responses to the global survey. Another study, the 2019 Internal Audit Survey Insurance by PwC found that 48% of internal audit functions had 0-10

are less likely to have an internal audit team, although they may have reporting and review functions within operational teams.

5.205 The size of the internal monitoring and assurance function needed will vary by service. If a provider of a larger, riskier, more complex service requires ten additional people to fulfil this function, we estimate the costs to be approximately £500,000 to £1,000,000 per annum. A provider of a smaller service that has a limited range of online safety risks and measures may only require a single person to fulfil its monitoring and assurance function, in which case the estimated annual costs would be approximately £50,000 to £100,000.[324] Because costs are likely to be higher for providers of larger services that face more risks, we would expect the costs to increase with the potential benefits to some extent. We expect that the costs of this measure for providers of smaller services would tend to represent a higher proportion of their annual revenue.

5.206 As demonstrated by responses to our 2022 Illegal Harms Call for Evidence, some large services already have internal assurance processes in place that deal with risks related to online safety.[325] Provided these processes are sufficient and services retain them for online safety purposes, they would not incur any additional costs from this measure.

5.207 We received no specific feedback on the costs of this measure. Our cost assumptions and estimates are largely unchanged from the November 2023 Consultation.[326] Given the benefits we expect this measure will bring to users, we judge it is proportionate.

Tracking evidence of new and increasing illegal harm

5.208 This measure will result in significant costs for services, including both one-off and ongoing costs. It is not possible to meaningfully estimate these costs as they are likely to vary considerably from service to service depending on the kinds of illegal harm they face, how they are able to gather information, and how much content they have.

5.209 As outlined in our November 2023 Consultation, all providers will be required under the online safety regime to establish complaints processes. They will therefore have information from those processes, but they may choose to run complaints processes in a way that does not distinguish between complaints regarding illegal content and complaints regarding content in violation of their terms of service. Providers may need to rely on other sources of information – including but not limited to outcomes of content moderation processes, Trust and Safety activities such as monitoring user activity, investigating policy

---

members, but this was based on responses by only 25 organisations. Source: PWC, 2019. Internal Audit Survey Insurance and Asset Management [accessed 29 November 2024]; Internal Audit Foundation, 2022. 2022 Premier Global Research. [accessed 29 November 2024].

[324] These estimates are based on our assumptions in Annex 5. We use the salaries for the "professional occupations" for the staff of the internal monitoring and assurance function. We recognise that salaries will vary very considerably both between different organisations and also within the internal monitoring and assurance function at any organisation. This is the case for salaries within internal audit function, as shown by the benchmarking of different internal audit roles in 2022 Barclay Simpson Salary & Recruitment Trends Guide: Internal Audit. [accessed 29 November 2024].

[325] For example, Meta indicated that its audit arrangements already cover how the service can be used to facilitate harm or undermine public safety or the public interest. Source: Meta response to 2022 Illegal Harms Call for Evidence.

[326] We have updated the estimates since the November 2023 Consultation in line with the latest wage data released by ONS. Since our estimates are rounded, the cost estimates for this measure have not changed. We received some feedback on the general cost assumptions (such as salary assumptions) that are fed into these costs. We consider that feedback in Annex 5.

violations and evaluating user behaviour, referrals from law enforcement, or flags from expert groups – to have a sufficient understanding of trends in illegal content risks on their service.

5.210 There will be costs associated with monitoring, reporting, and analysing evidence of new and increasing kinds of illegal content, including both one-off costs for establishing processes or automated collection systems and ongoing costs of employing staff to run these systems. Costs will broadly increase with the size and riskiness of the service. They will be higher where services have a higher total volume of material, number of users, volume of potentially illegal content, complexity of monitoring systems, and number of evidence sources. Service providers will have the flexibility to implement this measure in a way that suits their existing monitoring and/or governance systems.

5.211 Although many providers of larger services will also already have teams or individuals in place who are tasked with analysis of data for general online safety purposes, this may not extend to the analysis of trends in illegal harm specifically. The measure allows providers flexibility to track proxies of illegal content (rather than purely illegal content) provided these proxies are suitable. While some providers may already have suitable proxies, we anticipate that others will face additional costs related to tracking illegal content or suitable proxies (even if they currently monitor trends relevant to online safety and would have chosen to continue doing so in the future).

5.212 Providers of smaller services are less likely to have existing teams or systems in place for the ongoing analysis of information related to illegal harm. They may face challenges in accessing and analysing information in a systematic way if, for example, they outsource their content moderation operations to a third party or they do not have appropriate data collection infrastructure or in-house expertise in data analysis.

5.213 We did not receive any specific feedback from stakeholders on the costs relating to this measure. Therefore, our cost assumptions remain unchanged from the November 2023 Consultation.

Rights impact

5.214 We have considered whether these measures will have any impact on the rights to freedom of expression, association, or privacy.

Internal assurance and compliance functions

5.215 As explained at paragraph 5.159, the objective of the measure on internal assurance and compliance functions is to ensure there is a function in place to independently assure and monitor the effectiveness of processes to mitigate and manage illegal content risks on a service. The measure only relates to the internal governance processes of the service provider, and it does not require particular steps to be taken in relation to certain types of content, nor does it require personal data to be processed. Therefore, we consider that this measure does not constitute an interference with users' or providers' rights to freedom of expression or association, nor does it interfere with users' right to privacy.

5.216 To the extent that it helps to reduce harm on the service and make users feel safer, this could also positively impact on their human rights.

Tracking evidence of new and increasing illegal harm

5.217 As explained in paragraph 5.160 the objective of the measure on tracking evidence of new and increasing illegal harm is to ensure that any new or increasing risks that emerge on a

service are continuously monitored by the provider, enabling them to mitigate and manage these risks as and when they arise. The measure does not necessarily recommend the provider to review content it would not otherwise have reviewed, nor does it specify actions to be taken. As such, it will have no impact on rights to freedom of expression or association.

5.218 Any impact of the measure on users' right to privacy will be minimal. We acknowledge that service providers will have to collect some information and review trends in order to carry out the measure, but this process may not need to include user information or particular instances of content. Where a provider elects to use personal data, it will, as the ICO has pointed out (see paragraph 5.174) need to ensure that it is processed in accordance with applicable data protection legislation, including following data minimisation principles.

Who these measures apply to

Internal assurance and compliance functions

5.219 In our November 2023 Consultation, we proposed to apply this measure to all providers of services that are both large and multi-risk.[327]

5.220 Our analysis has shown that there are benefits in ensuring that organisations have independent oversight over internal controls to ensure that governance and risk management are effective. This oversight helps organisations make objective, authoritative, and credible judgments of the efficacy of their approach to risk management. As such, we consider this measure would deliver particular benefits in ensuring that risks of all kinds of illegal harms are mitigated and managed more effectively.

5.221 Roblox expressed concerns that this measure would be a disproportionate burden for some providers of large services that already have other solutions in place.[328] We proposed to apply this measure to providers of large services identifying a medium or high risk for at least two kinds of illegal harm. As providers assess their risk levels after taking account of existing actions, this measure will not apply to any provider already using solutions that effectively reduce risks to the extent that their services are low-risk. In any event, the Codes set out recommendations for how to comply with the safety duty. It is open to a provider to take alternative measures which achieve the same thing.

5.222 We maintain that it is proportionate to apply this measure for providers of large services which are multi-risk. We consider this measure would deliver the greatest benefits for these services as they operate in the most complex risk management environments (in that they typically have more complex organisations and multiple risks to manage) and identify significant risks of harm. Strong assurance processes would increase oversight over risk management processes. While we have identified considerable ongoing costs associated with this measure, we consider that this measure is fundamental to good risk management for large multi-risk services even when added on top of other governance measures. These services are also likely to be able to access necessary resources to implement the measures.

5.223 Age Verification Providers Association advocated extending this measure,[329] and some other stakeholders argued that all governance measures that apply to multi-risk services

[327] This includes large U2U services which are multi-risk, large general search services which are multi-risk, and large vertical search services which are multi-risk.
[328] Roblox response to November 2023 Consultation, p.3. We note Roblox made similar comments in response to the May 2024 Consultation. Source: Roblox response to May 2024 Consultation, p.7.
[329] Age Verification Providers Association response to November 2023 Consultation, p.2.

should be extended to single-risk services.[330] As explained in 'Our approach to developing Codes measures', we expect to consult again on this point in Spring 2025.

5.224 We are not recommending this measure for smaller services. We currently consider that the benefits from having a monitoring and assurance function will be materially lower for smaller services with a lower headcount and less complex governance structures, particularly if they are undertaking the other governance measures recommended to them. While there may be benefits for some smaller services, implementing this measure is likely to entail substantial costs for service providers. The cost burden of this measure could lead to resources and attention being diverted from understanding risks and putting controls in place, or to less innovation and investment in high-quality features. In more extreme cases, it could lead to smaller services exiting the UK market against users' interests.

5.225 We are therefore recommending this measure for all large multi-risk services. We will reconsult on whether to extend this measure to large single-risk services in Spring 2025.

Tracking evidence of new and increasing illegal harm

5.226 In our November 2023 Consultation, we proposed to apply this measure to all providers of large U2U and large general search services as well as all providers of multi-risk services.

5.227 We received feedback from three stakeholders regarding the scope of this measure. Roblox argued that the measure on tracking evidence of new and increased harm could be disproportionate for smaller newer large services which may have other user safety measures in place.[331] In expressing its view on the application of the governance measures more broadly, one stakeholder cited the measure should not apply based on service size alone, but to large services which are multi-risk or multi-risk services, where there is a high risk of at least one priority offence.[332] Logically argued that small services should be included within the scope of the measure due to the threat of foreign interference.[333] BT Group stated that smaller service providers and those that identify specific risks should be required to implement the measure to avoid continuing to categorise themselves as low risk when new risks may materialise.[334] In response to the equivalent measure in the May 2024 Consultation, the Scottish Government and the OSA Network called for the measure to apply to all service providers.[335]

5.228 Our assessment is that this measure is proportionate for multi-risk services. Monitoring and tracking trends in different kinds of illegal content is an important component of risk management. It is complementary to the risk assessment process as it allows changes in risks, which are external to the service, to be identified without triggering a new risk assessment. In addition, the monitoring and collection of up-to-date information is fundamental to ensuring that any mitigations remain effective over time. Managing

---

[330] ACCO response to November 2023 Consultation, p.2; IWF response to November 2023 Consultation, p.22; Logically response to November 2023 Consultation, p.12; Molly Rose Foundation response to November 2023 Consultation, pp.39-40; NSPCC response to May 2024 Consultation, p.19; NSPCC response to November 2023 Consultation, p.6; OSA Network response to November 2023 Consultation, p.95; Snap response to November 2023 Consultation, p.5; The Cyber Helpline response to November 2023 Consultation, p.3.

[331] Roblox response to November 2023 Consultation, p.16. We note Roblox made a similar point in response to the Ofcom May 2024 Consultation. Source: Roblox response to May 2024 Consultation, p.7.

[332] Airbnb response to November 2023 Consultation, p.3.

[333] Logically response to November 2023 Consultation, p.12.

[334] BT Group response to November 2023 Consultation, pp.1-2.

[335] OSA Network response to May 2024 Consultation, p.22; Scottish Government response to May 2024 Consultation.

multiple types of risk is likely to require formal processes to effectively monitor changes in different kinds of illegal content on a service. While the costs may represent a larger proportion of revenue for providers of some smaller multi-risk services and those which have not already implemented similar measures, we consider them to be proportionate (as costs are likely to increase with benefits to some extent). We have given service providers the flexibility to choose the most cost-effective way to implement this measure.

5.229 We maintain that this measure is proportionate for all large U2U and large general search services that are not multi-risk (including those that are low-risk or have a single medium or high risk). We acknowledge that benefits will be lower for providers of large services that have assessed low risk of illegal harm. However, explained in 'Our approach to developing Codes measures', some governance measures may still have benefits for large low-risk services as they will help ensure potential risks are promptly identified (particularly in a changing risk environment). We expect the costs of implementing this measure may be significant for providers of large services because they are likely to have complex operations and a significant volume of content. This will require sophisticated mechanisms to pull together the different sources of evidence required to effectively monitor and track changes in illegal content. However, we consider that these costs are proportionate given the fundamental importance of being able to identify and react to emerging risks that can affect a large number of users. This is particularly the case, given that the costs will increase with the benefits. Furthermore, we expect costs to be lower for providers of large low-risk services than for providers of large multi-risk services, and the measure offers providers flexibility regarding what evidence is gathered and how it is collected in order to achieve the aim of the measure in the least costly manner.

5.230 As with some of the other governance measures, we currently do not consider it proportionate to recommend this measure to large vertical search services that are not multi-risk as they typically present very limited risks (if any).[336]

5.231 BT, Logically, NSPCC, C3P and some of the other more general feedback suggested that this measure should apply to all providers of single-risk services.[337] As explained in 'Our approach to developing Codes measures', we will reconsult on this in Spring 2025.

5.232 We have considered the general responses on extending governance measures to smaller providers which assess themselves as low risk for all kinds of harms. However, we do not consider it proportionate to recommend this measure for such services. The benefits from establishing a system to track all kinds of harm and to report them through governance channels is likely to be more limited for smaller services that did not identify any risks of any kinds of illegal harms in their latest risk assessment.

---

[336] By their nature vertical search services are unlikely to have content that is as rapidly changing as U2U services and search results are more under a service's control than for U2U content. We are also not aware of evidence of such services showing illegal content. Any benefits of applying this measure would therefore be low for vertical search services. See a more detailed explanation in our Register of Risks chapter titled 'Search'.

[337] ACCO response to November 2023 Consultation, p.2; BT Group response to November 2023 Consultation, pp.1-2; C3P response to November 2023 Consultation, p.5; IWF response to November 2023 Consultation, p.22; Logically response to November 2023 Consultation, p.12; Molly Rose Foundation response to November 2023 Consultation, pp.39-40; NSPCC response to May 2024 Consultation, p.19; NSPCC response to November 2023 Consultation, p.6; OSA Network response to May 2024 Consultation, p.22; OSA Network response to November 2023 Consultation, p.95; Scottish Government response to May 2024 Consultation; Snap response to November 2023 Consultation, p.5; The Cyber Helpline response to November 2023 Consultation, p.3.

5.233 We received feedback questioning the technical feasibility of this measure by services using end-to-end encryption.[338] As explained in Volume 2: chapter 2: 'Content moderation', we have evidence that end-to-end encrypted services are still able to make illegal content judgements when receiving complaints, and will need to retain a content moderation function in order to meet their other duties under the Act (specifically in regard to reporting CSEA content). We expect that through such complaints, a service would still be able to apply this measure. Other potential sources of evidence could include information received from law enforcement, civil society groups, academic experts or suspected unusual patterns of user behaviour that may need further investigation.

5.234 Based on our analysis, we are maintaining our position of recommending this measure for all providers of large U2U and large general search services (including downstream general search services, for the reasons set out in 'Our approach to developing Codes measures') as well as all providers of multi-risk services. We will reconsult on whether to extend this measure to single-risk services in Spring 2025.

Conclusion

5.235 Based on our analysis, we consider that the measure on internal assurance and compliance functions will deliver significant benefits to users by ensuring that service providers consider the effectiveness of risk management on an ongoing basis, thereby enabling them to adjust and optimise their risk management strategy where appropriate. We also consider that our measure on tracking evidence of new and increasing illegal harm will deliver significant benefits to users by ensuring that service providers can appropriately consider new risks associated with illegal content which may not have been accounted for, thereby enabling them to respond more effectively.

5.236 Both measures could result in relatively significant costs. However, given the materiality of the benefits of the measures and the important role they will play in ensuring service providers are managing online safety risks appropriately, we consider the imposition of these costs to be proportionate. This is particularly the case given that we are focusing the internal assurance and compliance functions measure on providers of services that are both large and multi-risk and the tracking evidence measure on providers of services that are large or multi-risk.

5.237 We conclude that it is proportionate to include these measures in our Codes for U2U and search services on terrorism, CSEA and other duties. Our internal assurance and compliance function measure is referred to as ICU A4 for U2U services and ICS A4 for search services. Our tracking evidence of new and increasing illegal harm measure is referred to as ICU A5 for U2U services and ICS A5 for search services.

Measures on code of conduct, and compliance training

5.238 In our November 2023 Consultation, we proposed two measures to support providers in developing an appropriate code of conduct and staff compliance training:

- A measure recommending that providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk)

[338] Mega response to November 2023 Consultation, pp.3-5.

have a code of conduct for individuals working for the provider, around protecting users from risks of illegal harm.

- A measure recommending that providers of large U2U services, large general search services, and multi-risk services (including vertical search services which are multi-risk) should ensure that individuals working for the provider who are involved in the design and operational management of a service are sufficiently trained in the service's approach to compliance with online safety duties.

5.239 We set out our assessment and decisions relating to these two measures together as they both contribute to how staff understand and approach user safety within a service provider. Failing to effectively communicate or train staff on a provider's approach to compliance with illegal content safety duties raises the likelihood that risk mitigation and management is not embedded in the day-to-day operation of a service. Where staff do not understand or are not trained on a service's approach to addressing illegal content risks, risk mitigation measures could be applied poorly.

5.240 We said that a failure to embed such a culture of risk management across a service may result in the inconsistent application of measures designed to mitigate and manage risks of all kinds of illegal harm. Alignment of objectives at all levels is important to achieve good safety outcomes for users. This includes having a common understanding and expectation around risk management in relation to illegal content on a service. We noted the possibility that without alignment, staff in different areas of a service will not understand how a service is approaching regulatory compliance, or how it manages and mitigates risks of illegal content to users.

5.241 In our November 2023 Consultation, we also outlined our assessment of tying remuneration for senior managers to positive online safety outcomes. Though we did not recommend this measure in our draft Codes, we have highlighted relevant feedback in Annex 1.

Summary of stakeholder feedback[339]

Measure on code of conduct

5.242 We received several stakeholder responses regarding this measure. Some supported this measure, albeit some with reservations.[340] We have identified the following themes that emerged from stakeholder feedback, which we discuss in the following paragraphs:

- design of the measure; and
- feedback on who the measure applies to.

Design of measure

5.243 Meta (WhatsApp) stated that service providers should have the flexibility to decide how to draft their codes of conduct (to ensure alignment with any existing codes of conduct they have in place).[341] We address this feedback in the section 'How these measures work'.

[339] Note that this list is not exhaustive and further responses can be found in Annex 1.
[340] Association of British Insurers response to November 2023 Consultation, p.3.; Meta (WhatsApp) response to November 2023 Consultation, p.7.
[341] Meta (WhatsApp) response to November 2023 Consultation, p.7.

5.244 C3P suggested we create a sample code for services to adapt.[342] We address this feedback in the section 'How these measures work'.

Feedback on who measure applies to

5.245 We received a number of responses to our November 2023 and equivalent measure in our May 2024 regarding who this measure should apply to:

- Mega questioned the proportionality of this measure, arguing that it placed a burden on smaller services (specifically those that deal with image-based illegal content). It stated in its response that it is unreasonable for us to impose the same set of responsibilities on small and large multi-risk services. It also said that the governance measures as a whole were not suited to the capacity of smaller services and would be costly for those within scope.[343]
- Airbnb argued that the measure should apply to services that are either multi-risk or large and multi-risk. It said that this measure should not be applied solely by virtue of a service being large, but instead should be more risk-based. It also expressed overarching concerns about applying governance measures to low-risk services.[344]
- Mencap argued that this measure should apply to all services in order to help tackle illegal harms.[345]
- Twelve-APP expressed concerns about the documentation burden this measure would create for providers of small and medium services while making it fully operational. Twelve-App suggests focusing more on operational and practical measures to mitigate risks (such as moderation policies and actions, use of qualified safety specialists and/or detection algorithms, etc.), rather than on extensive documentation.[346]
- The Scottish Government said the measure should apply to all services.[347]
- C3P said that the code of conduct should be a basic obligation.[348]
- The Children's Commissioner for England supported this measure but argued that it should not be justified by low costs.[349]

5.246 We address this feedback in the section 'Who these measures apply to'.

Measure on compliance training

5.247 We received several stakeholder responses on compliance training for staff involved in the design and operation of a service.

Request for further detail on how the measure works

[342] C3P response to May 2024 Consultation, p.12. We note that C3P made a similar comment in their response to our November 2023 Consultation. Source: C3P response to November 2023 Consultation, p.5.
[343] Mega response to November 2023 Consultation, p.5.
[344] Airbnb response to November 2023 Consultation, p.3.
[345] Mencap response to November 2023 Illegal Harms Consultation, p.3.
[346] Twelve-App response to May 2024 Consultation on Protecting Children from Harms Online, pp.6-7.
[347] Scottish Government response to May 2024 Consultation, pp.8-9.
[348] C3P response to May 2024 Consultation, p.12. We note that C3P made a similar comment in their response to our November 2023 Consultation. Source: C3P response to November 2023 Consultation, p.5.
[349] Children's Commissioner for England response to May 2024 Consultation, p.25.

5.248 Protection Group International queried how relevant and appropriate training should be carried out and by whom. It also highlighted that each platform currently takes a different approach with no alignment or set agreed format.[350]

5.249 In response to an equivalent measure in our May 2024 Consultation, we received similar requests for guidance. The Association of Police and Crime Commissioners said it would be helpful for us to give direction on what constitutes appropriate training,[351] and one individual respondent suggested that Ofcom provide such training.[352] East Riding of Yorkshire Council asked if safeguarding procedures would be included in the staff compliance training.[353]

5.250 The Marie Collins Foundation commented that staff training should help staff understand why services need to be compliant.[354]

5.251 We address this feedback in the section 'How these measures work'.

Feedback on who the measure applies to

5.252 techUK commented on the resources that will be required to meet training requirements for staff, specifically for providers of smaller services. It argued that we should carry out a more detailed impact assessment in order to understand the potential challenges and costs that this measure would carry for such services.[355] We address this feedback in the section 'Costs and risks' and the section 'Who these measures apply to'.

5.253 Some stakeholders questioned the application of this measure to large and low-risk services. They both expressed the opinion that this measure should not be applied solely on the basis of size but instead on the risk profile of a service.[356]

5.254 BT Group, Mencap, and C3P argued that we should widen the scope of the measure to include all services.[357] Specifically, BT Group and C3P expressed the view that training is essential to ensure compliance and that controls are effective in mitigating harm.[358] UKSIC noted suggested training requirements should be extended to staff in services where they are deemed to be medium to high risk of CSAM.[359]

5.255 We address this feedback in the section 'Who these measures apply to'.

Our decision

5.256 We have decided to proceed with the measures broadly as proposed in our November 2023 Consultation.

5.257 However, we have made a small change to the wording out of concern that the words "staff" and "employees" which we used at consultation might be taken as implying a provider which used contractors for all its work need not train them. The measure now says

---

[350] Protection Group International response to November 2023 Consultation, p.2.
[351] The Association of Police and Crime Commissioners response to May 2024 Consultation, p. 8.
[352] Dean, J. response to May 2024 Consultation, p.9.
[353] East Riding of Yorkshire Council response to May 2024 Consultation, p.2.
[354] Marie Collins Foundation response to November 2023 Illegal Harms Consultation, p.5.
[355] techUK response to November 2023 Consultation, p.3.
[356] Airbnb response to November 2023 Consultation, p.3; [✂].
[357] BT response to November 2023 Consultation, p.2; C3P response to November 2023 Consultation, p.5; Mencap response to November 2023 Consultation, p.3.
[358] BT response to November 2023 Consultation, p.2; C3P response to November 2023 Consultation, p.5.
[359] UKSIC response to November 2023 Consultation, p.31.

that training should be provided to "individuals working for the provider who are involved in the design and operational management of the service" apart from volunteers.[360]

5.258 The full text of the measures can be found in our Illegal Content Codes of Practice for terrorism, CSEA and other duties. We refer to our code of conduct measure as ICU A6 for U2U services and ICS A6 for search services, and our compliance training measure as ICU A7 for U2U services and ICS A7 for search services.

Our reasoning

How these measures work

Measure on code of conduct

5.259 We recommend that service providers have a code of conduct for all individuals working for the provider, which sets standards and expectations around protecting United Kingdom users from risks of illegal harm.

5.260 The objective is that the code be service-specific. However, the code of conduct could include recognition of risks of illegal harm to users of a service, a clear organisational statement around protecting users, and expectations and guidelines for all staff on how to report instances of concern relating to illegal content on the service. A good code would be simple, concise, easy to understand[361], and consistent with other policies and communications. It would also be reviewed by multi-disciplinary teams.

5.261 Due to the importance of the code being service-specific, and the diversity of service providers, we do not think it appropriate, at this stage, to develop a sample code.

5.262 Meta (WhatsApp) said that providers should have discretion in how they draft their code of conduct.[362] The measure is not prescriptive about how a code of conduct should be drafted. We consider service providers to be best placed to decide how to integrate the code of conduct with their existing codes of conduct or people management policies should they choose to do so.

Measure on compliance training

5.263 We recommend that individuals working for the provider and involved in the design and operational management of a service, (other than volunteers), receive sufficient training on the provider's approach to compliance with illegal content safety duties and reporting and complaints duties so that they can apply these effectively in their roles.

5.264 We received stakeholder feedback requesting more guidance on how the training should be carried out and by whom. We consider it essential that the measure is flexible enough to be implemented by the broad range of services in scope and therefore do not think, at this stage, it would be proportionate for us to add limitations to the training for the measure

[360] Volunteers are those who, in relation to the activities in question, are not employed by the provider or anyone else, remunerated, or acting by way of a business. This is a slight broadening of the measure on which we consulted, which related to staff/employees. However, it aligns with our approach on training of content moderators, see Volume 2: chapter 2: 'Content moderation'. We think the change is unlikely to make a practical difference since a provider is unlikely to wish to consider the precise employment status of its paid workers before deciding which of them to train.

[361] Deloitte, 2009. Suggested guidelines for writing a code of ethics/conduct. [accessed 4 September 2023].

[362] Meta (WhatsApp) response to November 2023 Consultation, p.7.

included in the Illegal Content Codes. Each provider is best placed to inform how they can achieve the outcomes set out by this measure.

Benefits and effectiveness

5.265 As we set out in our November 2023 Consultation, how a provider guides and trains its people is relevant to both risk management and regulatory compliance, given the influence these factors have on individual performance, decision-making and risk-taking behaviour.

5.266 Ensuring that the protection of users from illegal content risks is explained as a part of the package of information and training applicable to those working for the provider makes it more likely that opportunities to mitigate those risks will be identified, considered, and adopted.

5.267 In contrast, where staff are not guided in this way, there is a possibility that risk to users will not be appropriately factored into everyday decision-making and operations either because of competing pressures and incentives on staff, or due to ignorance of compliance requirements. It follows that measures may be either inadequate to address risks to users, or improperly implemented. It also could be the case that any measures put in place are not monitored and evaluated for effectiveness over time.

5.268 Failing to effectively communicate or train people working for a provider on a service's approach to compliance with illegal content safety duties raises the likelihood that risk mitigation and management is not embedded in the day-to-day work of operating a service. Where people working for a provider do not understand or are not trained on a service's approach to addressing illegal content risks, risk mitigation measures could be applied poorly.

Measure on code of conduct

5.269 A code of conduct can be an effective way for service providers to communicate expectations regarding behaviour and responsibilities to all who are working for it. Having a code of conduct around protecting users makes it more likely that opportunities to mitigate illegal harms risks will be identified, considered, and adopted.

5.270 Providers who fail to effectively communicate their approach to compliance may risk safety considerations not being fully prioritised or embedded in the day-to-day operations of their services. This may result in safety measures being poorly or inconsistently applied. Opportunities to identify and consider risks to users may also be missed.

5.271 Where individuals working for a provider receive little guidance on managing and mitigating risks for users, there is a risk that principles will not be adequately factored into everyday decision-making and operations. This could result in risks to users being inadequately addressed. Individuals working for a provider will be more likely to have a consistent understanding of the provider's approach to user safety and be able to apply it effectively in their day-to-day work and decision-making. We expect the implementation of a code of conduct to increase the likelihood of safety concerns in all parts of a service being flagged and actioned.

5.272 Responses to our 2022 Illegal Harms Call for Evidence highlighted the importance of providing documentation regarding the values and behaviours expected of staff as part of a broader programme of good corporate governance regarding online safety. This included Google, which mentioned consistent principles, a Code of Conduct and Group Guiding

Principles as part of governance and accountability.[363] Zoom framed this in terms of its standard operating procedures ('SOPs') which govern expectations for analysts dealing directly with content decisions.[364] [✂] highlighted that it had clear policies and operational guidelines for how it governs its approach to user and platform safety and an expectation that responsibility for safe user experiences is shared across the organisation as a core value of the brand.[365]

5.273 In summary, we consider that – together with our other governance measures – the creation of codes of conduct will contribute to the creation of a service-wide culture to support online safety.

Measure on compliance training

5.274 Compliance training is important to achieve good safety outcomes for users. This includes having a common understanding and expectation around risk management in relation to illegal content on a service. There is a possibility that without these efforts for provider-wide understanding, people working in different areas of a service will not understand how the provider is approaching regulatory compliance, or how it manages and mitigates risks of illegal content to users. This is supported by evidence of how the absence of compliance training programmes has contributed to serious corporate scandals. For example, in the case of Siemens – which was subject to regulatory investigations for bribery in 2008 – the failure to embed a programme of compliance and Code of conduct for staff has been cited as playing a "decisive role" in the scandal.[366]

5.275 How people are trained in their roles can inform how they approach user safety considerations within a service. Training is an important way for service providers to communicate compliance requirements and embed risk mitigation and management within their organisations. Given the potential impact of their work on ensuring user safety, those involved in the design and operation of a service are most likely to benefit from training focused on compliance with illegal content safety duties.

5.276 This measure will ensure that service providers communicate the importance of compliance requirements to staff involved in the design and operational management of the service. This will help to embed risk awareness, management, and mitigation into all operations of the service, resulting in more effective decision-making in relation to user safety.

5.277 Several providers referenced the general importance of staff training in the 2022 Illegal Harms Call for Evidence, and demonstrated their commitment to ensuring that their approach to online safety was understood by employees. These providers included Google, which gives its employees specific training on "risk and compliance to raise awareness of requirements from new and emerging regulations which govern online content and behaviours", and [✂], which stated that engineers receive training to ensure they are aware of and accounting for safety concerns while software is being developed.[367]

---

[363] Google response to 2022 Illegal Harms Call for Evidence.
[364] Zoom response to the 2022 Illegal Harms Call for Evidence.
[365] [✂].
[366] Source: Primbs, M., and Wang, C., 2016. Notable Governance Failures: Enron, Siemens and Beyond Comparative Corporate Governance and Financial Regulation. [accessed 25 October 2024].
[367] Dropbox response to 2022 Illegal Harms Call for Evidence; Google response to 2022 Illegal Harms Call for Evidence.

5.278 Evidence from organisations that have faced major governance issues highlights the importance of compliance training. In the November 2023 Consultation, we noted a case study from Siemens relating to redress of governance failings that focused on strengthening compliance programmes through staff training. This study provides supporting evidence to suggest that such changes led to an improvement in perception of how risks were managed.[368]

5.279 Taking the above into consideration, we consider that regular risk management training will result in better safety outcomes and will play a role in creating a healthy culture in relation to the management of online safety risks. We have therefore concluded that this measure will result in significant benefits.

Costs and risks

Measure on code of conduct

5.280 Having considered the stakeholder feedback, our cost assumptions and estimates remain largely unchanged from the November 2023 Consultation.[369]

5.281 We estimate that developing a code of conduct would incur a one-off development cost of less than £10,000 (in cases where service providers do not have an existing code of conduct for individuals working for the provider).[370] We expect there to be some additional lower costs of reviewing and adapting this document over time. We expect costs to be higher for providers of services that have more risks, use more complex reporting systems, or require more time for legal review and for integrating this measure with existing staff policies. We anticipate that implementation costs for providers of smaller lower-risk services would be considerably lower.

5.282 The code of conduct or principles would need to be read by all relevant individuals. We assume that the document would be short and would not take a significant amount of time for them to read and understand.

Measure on compliance training

5.283 The cost of this measure will vary significantly from service to service based on factors such as size and risk level. The important variables in delivering the initial training would be the number of individuals needing to be trained and the length and detail of training. Overall, we expect costs to be higher for larger services as they will need to train more people, but we note that costs are likely to be a larger proportion of revenue for smaller services compared to larger ones. We also expect costs to increase with the number of risks of harms on the service as this is likely to affect the length of the training required.

---

[368] Following a 2008 bribery scandal, Siemens attempted to redress governance failings identified by strengthening its compliance programmes This included ensuring that employees in different levels have been provided with trainings specific to their roles and responsibilities. Source: Institute of Business Ethics (Dietz, G., and Gillespie, N.), 2012. The Recovery of Trust: Case Studies of Organisational Failures and Trust Repair. [accessed 20 November 2024].

[369] We have updated the estimates since the November 2023 Consultation in line with the latest wage data released by ONS. However, since our estimates are rounded, the cost estimates may not have changed. We received some feedback on the general cost assumptions (such as salary assumptions) that are fed into these costs. We consider that feedback in Annex 5.

[370] This is using our cost assumptions set out in Annex 5, assuming it takes less than 20 days of a person in a "professional occupation" to produce the Code of Practice. We have updated these estimates since the November 2023 Consultation in line with the latest wage data released by ONS.

5.284 In the November 2023 Consultation, we estimated that the total cost per person trained would be £2,000 to £4,000 and noted that an important variable would be the number of individuals trained and the length of training. We provided examples of total training cost assuming one week of training for 10 people and for 100 people.

5.285 As noted in the 'Summary of stakeholder feedback' section, techUK expressed concern regarding the governance burden that compliance could put on SMEs.[371] We note that in the May 2024 Consultation we have revised our assessment of cost for this measure to provide more details and better understand the cost variations for different service providers.

- First, we split the cost for the initial creation of material and the ongoing cost of staff[372] time to be trained. A number of variables were identified as causing significant cost variability between service providers in scope. These include the number of staff, their baseline understanding of risk, the number of risks, whether compliance training programmes are already in place, and the risk complexity.
- Second, we revised our assumption around the length of training. The cost estimates provided in our November 2023 Consultation are reasonable if we assume that the training would take a week. We have updated our assumptions, and now assume it would take half a day to train general staff and two days to train senior staff.

5.286 For illustration, if we assume that it would take one person around two to four weeks to create the training materials, we estimate the cost to be approximately £2,000 to £8,000. We estimate the additional cost per person trained per day to be £150 to £250 for general staff involved in the design and operation of the service (such as designers, engineers, and managers) and £500 to £800 for senior staff (such as senior leaders and senior managers).[373]

5.287 As noted in responses to our 2022 Illegal Harms Call for Evidence, some companies are already training those involved in the design and operational management of a service on compliance with online safety responsibilities. Provided this training is adequate and these services intend to continue with it in the future, this measure would not add additional costs.

5.288 We did not receive any additional responses to the May 2024 Consultation regarding these cost estimates.

Rights impact

5.289 We do not consider that these measures have any negative impact on users' or service providers' rights to freedom of expression or with users' rights to privacy.

5.290 To the extent that providing appropriate training and a code of conduct helps to reduce harm on the service and make users feel safer, this could positively impact on their human rights. As with other governance measures, we consider that a well-managed business is, in general, more likely to comply with its obligations under privacy and data protection laws

[371] techUK response to November 2023 Consultation, p.3.
[372] Although the measure applies slightly more widely than just to employees, we think this is a reasonable basis on which to assess costs.
[373] This is derived from assumptions set out in Annex 5 and have been updated in line with the latest wage data released by the ONS. Since estimates are rounded, they may have changed unevenly when using updated wages. We received some general feedback on the cost assumptions (such as salary assumptions) which have been considered in Annex 5.

(as well as, for that matter, other important laws such as those relating to consumer protection and equality). As such, our proposals may help to safeguard these.

Who these measures apply to

Measure on code of conduct

5.291 In our November 2023 Consultation, we proposed to apply this measure to all providers of large U2U and large general search services as well as all providers of multi-risk services.

5.292 In response to the November 2023 and May 2024 Consultations, we received feedback on who this measure should apply to (see paragraph 5.245). This included providers arguing that this measure should apply in relation to all services, and providers highlighting that the measure could pose a relatively greater burden for providers of smaller services.

5.293 As set out in paragraphs 5.269-5.272 in our 'Benefits and effectiveness section', this measure will ensure individuals working for the provider understand its approach to protecting users.

5.294 We consider this measure to be proportionate for providers of multi-risk services as there are clear benefits of implementing a code of conduct and the costs are relatively low. Those working for a provider must understand the service provider's approach to protecting users so that risks are identified, reported, and considered in the design and implementation of products or functions. This is particularly important for multi-risk services where individuals need to understand multiple risks. Given the relatively low cost of this measure, we remain of the view that it is proportionate for all multi-risk services, including smaller services. Providers have the flexibility to implement this measure in a way that is suitable for their needs and service structures. The Children's Commissioner for England stated in response to the May 2024 Consultation that the justification for where the measure is being applied should not be based on costs.[374] We outline in 'Our approach to developing Codes measures' why costs and benefits are important, and inform our assessment of a measure and its proportionality.

5.295 We also consider this measure to be proportionate for all providers of large U2U and large general search services (including those that are low-risk or have a single medium or high risk). The benefits are likely to be material, particularly for large services with potentially more complex reporting lines. If the service provider's obligations under the Act are not understood by those working for it, there is a greater risk that it will not make an accurate risk assessment and may incorrectly regard itself as low risk when it is not. The consequences of this are greater for large services because of their wider reach. We acknowledge that there may be fewer benefits for providers of large services that have assessed low risk any kinds of of illegal harm. However, as set out in 'Our approach to developing Codes measures', there remain important benefits to governance measures which will ensure potential risks are promptly identified (particularly in a changing risk environment). Providers of large services will also tend to have more resources compared to providers of smaller services, making it easier for them to comply with this measure.

[374] Children's Commissioner for England response to May 2024 Consultation, p.25.

5.296 As with some of the other governance measures, we currently do not consider it proportionate to recommend this measure to large vertical search services that are not multi-risk as they typically present very limited risks (if any).[375]

5.297 As set out in 'Our approach to developing Codes measures', we intend to consider extending the measure to single-risk services in a consultation in Spring 2025.

5.298 We have not expanded this measure to apply to providers of all services as we consider the benefits would be materially lower for providers of small low-risk services. In addition, these providers would be less likely to be able to bear the costs of the measure. This is consistent with our general intention to keep the regulatory burden low for such providers, as discussed in Volume 2: chapter 13: 'Combined impact assessment'.

5.299 Based on our analysis, we are confirming our position of recommending this measure for all providers of large U2U and large general search services as well as all providers of multi-risk services.

Measure on compliance training

5.300 In our November 2023 Consultation, we proposed to apply this measure to all providers of large U2U and large general search services as well as all providers of multi-risk services.

5.301 In response to the November 2023 Consultation, techUK raised concerns about the proportionality of this measure for SMEs.[376] One stakeholder argued that the measure is disproportionate for large low-risk services.[377] In expressing its view on the application of the governance measures more broadly, one stakeholder stated that the measure should not apply based on service size alone, but to large services which are multi-risk or multi-risk services, where there is a high risk of at least one priority offence.[378] BT, Mencap, and C3P argued that the measure should apply to all services.[379]

5.302 Training helps to ensure a consistent approach to organisational aims across a service. By embedding risk management and mitigation within their organisational culture and ensuring that staff understand their role in this, service providers can create a strong culture of risk awareness among staff. As we have explained in paragraph 5.276, this will deliver significant benefits.

5.303 We maintain that this measure is proportionate for providers of multi-risk services (including smaller services). Training will be particularly beneficial for multi-risk services because it will help those working there understand interdependencies between the different risks and systems and the processes used to manage them.[380] Where services identify a number of risks, compliance training is likely to be particularly important. While costs may be significant for some providers, we consider them to be justified as costs are likely to increase with service size and the number of risks, thus increasing in line with

---

[375] By their nature vertical search services are unlikely to have content that is as rapidly changing as U2U services and search results are more under a service's control than for U2U content. We are also not aware of evidence of such services showing illegal content. Any benefits of applying this measure would therefore be low for vertical search services. See a more detailed explanation in our Register of Risks chapter titled 'Search'.
[376] techUK response to November 2023 Consultation, p.3.
[377] [✂].
[378] Airbnb response to November 2023 Consultation, p.3.
[379] BT response to November 2023 Consultation, p.2; C3P response to November 2023 Consultation, p.5; Mencap response to November 2023 Consultation, p.3.
[380] Providers of smaller services that manage multiple risks to users will also benefit from having a consistent and well-defined approach to risk management as expressed through staff compliance training.

benefits. Service providers that are SMEs can choose less costly ways of implementing training (provided that those working there are sufficiently trained to implement the service provider's duties to protect users).

5.304 We also maintain that this measure is proportionate for all providers of large U2U services and large general search services. Providers of large services with more complex operations and larger headcount are likely to benefit more from this measure - it will standardise and embed the culture of corporate risk management to all relevant individuals. This reduces the potential for failing to identify risks of illegal harms, particularly in a rapidly changing risk environment or where large enterprises structure around discrete projects and develop their own team culture. This is likely to justify the cost burden of this measure, particularly as we expect costs to increase with the number of harms (as explained in paragraphs 5.283-5.285 in the 'Costs' section).

5.305 We disagree with [✂] and responses from other stakeholders that the measure is disproportionate for large low-risk services. We acknowledge that this measure may have fewer benefits for providers of large services that have assessed low risk of any kinds of illegal harm since there may be less scope for reducing harm to users from illegal content. However, as set out in 'Our approach to developing Codes measures', there are still important benefits from some governance measures (including this measure) as they help to ensure that risks are properly and speedily identified, fostering company-wide safety culture and values. This is important because, given the size of the services, changes in risks could cause significant harm if not identified and dealt with in a timely manner. Furthermore, a large service with a low risk of harm is likely to incur lower costs when implementing this measure as it will need fewer training materials compared to multi-risk services. In general, large services are also more likely to be able to absorb the costs of the measure than smaller services.

5.306 We also consider this measure builds in flexibility for different approaches and capabilities by services for compliance training. For example, from preparing standard internal guidance or a specific learning and development module, to more elaborate solutions such as externally prepared curriculums or assisted learning. Providers ultimately have discretion to decide the best information and educational methods to ensure they achieve the industry culture change needed to put user safety at the heart of service design and decision making. As with some of the other governance measures, we currently do not consider it proportionate to recommend this measure to large vertical search services that are not multi-risk as they typically present very limited risks (if any). In the November 2023 Consultation, we explained that there was a lack of evidence to suggest that vertical search services were used to disseminate priority illegal content. [381]

5.307 In response to UKSIC's point, a service at medium or high risk of CSEA, and another kind of illegal harm will be considered multi-risk, and will fall into scope of this measure. As set out in 'Our approach to developing Codes measures', we do though intend to consider extending the measure to single-risk services in a future consultation in 2025. However, we have not expanded this measure to apply to providers of all services, as we consider the benefits would be limited for smaller and low-risk services. This is consistent with our

[381] By their nature vertical search services are unlikely to have content that is as rapidly changing as U2U services and search results are more under a service's control than for U2U content. We are also not aware of evidence of such services showing illegal content. Any benefits of applying this measure would therefore be low for vertical search services. See a more detailed explanation in our Register of Risks chapter titled 'Search'.

general intention to keep the regulatory burden low for such services, as discussed in Volume 2: chapter 13: 'Combined Impact Assessment'.

5.308 Based on our analysis, we are maintaining our position of recommending this measure for all providers of large U2U and large general search services as well as all providers of multi-risk services.

Conclusion

5.309 Based on our analysis, we consider that the measure on having a code of conduct will deliver benefits to users by helping to communicate the provider's expectations regarding behaviour and responsibilities to all who are working for it. Having a code of conduct around protecting users makes it more likely that opportunities to mitigate illegal harms risks will be identified, considered, and adopted. Together with our other governance measures it will contribute to the creation of a service-wide culture to support online safety.

5.310 For similar reasons, we consider that our measure on staff compliance training will deliver significant benefits to users. It will set a common understanding and expectation around risk management in relation to illegal content on a service.

5.311 Set against this, the cost of the measures is likely to be relatively modest, and we consider it to be justified given the scale of the benefits and the foundational importance of robust governance in ensuring good safety outcomes.

5.312 Both measures apply to all providers of large services (except large vertical search services) and all providers of multi-risk services.

5.313 We conclude it is proportionate to include both of these measures in our Illegal Content Codes of Practice for terrorism, CSEA and other duties. The code of conduct measure is referred to as ICU A6 for U2U services and ICS A6 for search services, and the staff compliance training measure as ICU A7 for U2U services and ICS A7 for search services.