Protecting children from harms online

Annexes 10-15

Consultation

Published 8 May 2024
Closing date for responses: 17 July 2024

Contents

Annex

A10. Draft guidance on highly effective age assurance

A10.1 This annex provides draft guidance for service providers to assist in implementing highly effective age assurance in accordance with Measures AA1 to AA6 of the Children's Safety Codes. It sets out:

i) An overview of our recommendations for the implementation of highly effective age assurance as set out in the Codes; and,
ii) Accompanying draft guidance, including additional technical detail and examples to assist service providers in complying with the measures.

A10.2 Our proposed approach to highly effective age assurance aligns with the approach taken in our draft Guidance for service providers publishing pornographic content under Part 5 of the Act. This is to ensure consistency so that service providers in scope of both Part 5 and our Part 3 Codes of Practice are clear what they need to do to prevent children from encountering the most harmful forms of content.

Implementing highly effective age assurance in accordance with the Codes

A10.3 We are making the following recommendations in the Codes relating to the implementation of highly effective age assurance.

Recommendations in the Codes

A10.4 For the use of age assurance to be highly effective at correctly determining the age of users, service providers should choose an appropriate method (or methods) of age assurance that is of such a kind that could be highly effective at correctly determining whether a user is a child.

A10.5 Service providers should ensure that their chosen age assurance process as a whole fulfils each of the criteria of technical accuracy, robustness, reliability and fairness, to ensure it is highly effective in practice.

A10.6 The technical accuracy criterion is fulfilled if:

a) the provider has ensured that the measures1 forming part of the age assurance process for the service have been evaluated against appropriate metrics to assess the extent to which they can correctly determine the age or age range of a person under test lab conditions;
b) where the age assurance process used on the service involves the use of age estimation, the provider uses a challenge age approach; and

[1] We acknowledge the Draft Code Children Safety Codes refers to 'age assurance methods' as 'age assurance measures.' This is to reflect the statutory language of the Act as per Section 41(3), Schedule 4, in particular Sch 4 para 12. In sub-section 'Highly Effective Age Assurance' of the Draft Children Safety Codes, 'age assurance measures' has the same meaning as 'age assurance methods' in the Age Assurance Section.

c) the provider periodically reviews whether the technical accuracy of the age assurance process for the service could be improved by making use of new technology and, where appropriate, makes changes to the age assurance process.

A10.7 The robustness criterion is fulfilled if:

a) The provider has:

i) taken steps to identify methods children use to circumvent the age assurance process used on the service to determine that the relevant individual is not a child; and
ii) taken feasible and proportionate steps to prevent children using those methods; and

b) the provider has ensured that the age assurance measures forming part of the age assurance process for the service have been tested in multiple different environments during the development of the age assurance process.

A10.8 The reliability criterion is fulfilled if:

a) where age assurance measures forming part of the age assurance process rely on artificial intelligence or machine learning, the provider has taken steps to ensure that:

i) the artificial intelligence or machine learning has been suitably tested during the development of the age assurance process to ensure it produces reproducible results;
ii) the artificial intelligence or machine learning is regularly tested to ensure it produces reproducible results;
iii) the outputs of the artificial intelligence or machine learning used are monitored and assessed against key performance indicators designed to identify whether the artificial intelligence or machine learning produces reproducible results;
iv) in circumstances where the artificial intelligence or machine learning used are observed to be producing unreliable or unexpected results, the root cause of the issue is identified and rectified.

b) The provider has taken steps to ensure that any data relied upon as part of the age assurance process comes from a reliable source.

A10.9 The fairness criterion is fulfilled if:

a) The provider has ensured that any elements of the age assurance process for a service, which rely on artificial intelligence or machine learning have been tested and trained on data sets which reflect the diversity in the target population.

A10.10 Service providers should not publish content that directs or encourages United Kingdom users to circumvent the age assurance process or access controls used on the service.

A10.11 When implementing the age assurance process, service providers should have regard to the following principles:

a) the principle that age assurance should be easy to use, including by children of different ages and with different needs;
b) the principle that age assurance should work effectively for all users regardless of their characteristics or whether they are members of a certain group;
c) the desirability of ensuring interoperability between different kinds of age assurance;
d) the latest version of the age appropriate design code and the Information Commissioner's opinion entitled "Age Assurance for the Children's code" published on 18 January 2024.

A10.12 The provider should ensure that users are able to easily access information about what a provider's age assurance process is intended to do and how the provider's age assurance process works prior to commencing the age assurance process for the service.

A10.13 When implementing age assurance, service providers should have regard to the ICO's Children's code, and the Commissioner's Opinion on Age Assurance for the Children's code.[2]

Draft guidance on highly effective age assurance

A10.14 We expect to publish accompanying guidance to assist services in implementing highly effective age assurance in accordance with age assurance Measures 1-6.

A10.15 The remainder of this annex sets out our proposed guidance, including additional technical detail and examples.

Age assurance methods

Box A10.1: What is an age assurance method?

An **age assurance method** refers to the particular system or technology that underpins an age assurance process.

Box A10.2: What is an age assurance process?

An **age assurance process** refers a system or process designed to determine whether a particular user is, or is not, a child that is comprised of one or more age assurance methods[3]. The effectiveness of an age assurance method will depend on how it is implemented, including whether by itself or in combination with other methods.

A10.16 No age assurance method is inherently highly effective at correctly determining whether or not a particular user is a child. Instead, effectiveness depends on how the age assurance method is implemented by the service provider.

A10.17 Below, we set out a non-exhaustive list of kinds of age assurance that we consider **could** be highly effective at correctly determining whether or not a user is a child. We recognise that age assurance methods are developing at pace and this list may expand in time.

A10.18 We have also proposed to specify age assurance methods which we consider are not capable of being highly effective. Services should not rely on these methods to determine whether a user is a child in the absence of other measures.

A10.19 All age assurance methods involve the processing of personal data. As such, they are subject to the requirements of the UK's data protection regime. We discuss how services can have regard to data protection law in the 'Privacy and data protection' sub-section below.

[2] See ICO, Age assurance for the Children's code. [accessed 19 April 2024].

[3] We acknowledge the Draft Code Children Safety Codes refers to 'age assurance methods' as 'age assurance measures.' This is to reflect the statutory language of the Act as per Section 41(3), Schedule 4, in particular Sch 4 para 12. In sub-section 'Highly Effective Age Assurance' of the Draft Children Safety Codes, 'age assurance measures' has the same meaning as 'age assurance methods' in the Age Assurance Section.

Kinds of age assurance that could be highly effective

- **Open banking.** This works by accessing the information a bank has on record regarding a user's age, with the user's consent. Confirmation of whether or not the user is over 18 is shared with the relying party.[4] The user's date of birth is not shared with the relying party, nor is any other information.
- **Photo-identification (photo-ID) matching.** This works by capturing relevant information from an uploaded photo-ID document and comparing it to an image of the user at the point of ID upload to verify that they are the same person.
- **Facial age estimation**. This works by analysing the features of a user's face to estimate their age.
- **Mobile-network operator (MNO) age checks**. Each of the UK's MNOs have agreed to a code of practice whereby they automatically apply a content restriction filter (CRF), which prevents children from accessing age-restricted websites over mobile internet on pay-as-you-go and contract SIMs. Users can remove the CRF by proving they are an adult.[5] MNO age checks rely on checking whether the CRF on a user's mobile phone has been removed. If the CRF has been removed, this indicates that the recorded user of the device is over 18. Confirmation of whether or not the recorded user is over 18, based on the status of the CRF, is shared with the relying party.
- **Credit card checks**. In the UK, you must be 18 or over to obtain a credit card, therefore, credit card issuers are obliged to verify the age of applicants before providing them with a credit card. Credit-card based age checks work by asking a user to input their credit card details, after which a payment processor sends a request to check the card is valid by the issuing bank. Approval by the issuing bank can be taken as evidence that the user is over 18.[6]
- **Reusable Digital ID services.** A digital identity is a digital representation of a person which enables them to prove who they are during interactions and transactions online and in person. Reusable digital identities are those which can be used multiple times for different interactions and transactions.[7] This includes digital identity wallets which enable users to verify and securely store their attributes (such as age) in a digital format. This verification may take place using a variety of methods, including those listed above. Once their identity or an attribute of their identity has been verified and stored in the wallet, a user may choose to share individual attributes, such as their age, or their status as an adult, with a relying party.

A10.20 We note that the Part 5 consultation referred to 'digital identity wallets' as an example of an age assurance method which could be highly effective. In this draft guidance, we have proposed to broaden this to 'reusable digital ID services' to reflect that digital wallets are just one example of how reusable digital identities can be used in the age assurance context. Subject to considering and completing the Part 5 consultation process, we propose to reflect this change in the final guidance for service providers publishing pornographic content.

---

[4] 'Relying party' refers to the service that is trying to establish the age of the user. In this context, the relying party is likely to be the regulated service.

[5] There are several ways to remove a CRF, depending on the MNO.

[6] Possession of credit card details is not guaranteed to be evidence that the user with the details is the person who took out the credit card.

[7] As defined in Department for Science, Innovation and Technology (DSIT), 2023. UK digital identity and attributes trust framework beta version (0.3). [accessed 22 March 2024].

A10.21 Our ongoing analysis of responses to our Part 5 consultation may result in further changes to our approach to highly effective age assurance. We will seek to maintain consistency in our approach to highly effective age assurance in developing our final guidance and Codes across both Part 3 and Part 5 of the Act.

A10.22 It is for the service provider to determine which age assurance method(s) is appropriate to meet its duties under the Act. Implementing one of the example methods is not a guarantee that the service is acting in accordance with the Code measures – service providers need to be able to demonstrate that the method(s) has been implemented in such a way that ensures the overall process as a whole is highly effective.

Kinds of age assurance that are not capable of being highly effective

- **Self-declaration of age:** The Act states that measures which require users to self-declare their age are not to be regarded as age assurance.[8] There is evidence to support this, and such methods are therefore not appropriate for the purposes of compliance with this measure. These include:

  i) asking a user to input their date of birth without any further evidence to confirm this information; or
  ii) asking a user to tick a box to confirm that they are 18 years of age or over.

- **Age verification through online payment methods which do not require a user to be over the age of 18.** For example, Debit, Solo or Electron cards, or any other card where the card holder is not required to be 18.
- **General contractual restrictions** on the use of the regulated service by children. For example:

  iii) including as part of the terms of service a condition that prohibits users who are under 18 years old from using the service, without any additional age assurance;
  iv) general disclaimers asserting that all users should be 18 years of age or over; or
  v) warnings on specific content that the content is suitable for adults.

Criteria to ensure an age assurance process is highly effective

A10.23 We recommend that service providers should ensure that the age assurance process as a whole fulfils each of the criteria of technical accuracy, robustness, reliability and fairness.

A10.24 We recognise that there may be trade-offs in how well an age assurance method performs against each of the criteria, and it is for the service provider to determine which trade-offs are appropriate to ensure that the overall process is highly effective at correctly determining whether or not a particular user is a child.

A10.25 Below, we provide additional detail on how a service provider can implement the practical steps recommended in the measure.

A10.26 We recognise that, as well as building an in-house age assurance method, or purchasing a method from an age assurance provider, there may be wider system-level age assurance processes service providers can use to distinguish between children and adults on their service. Regardless of where the age assurance occurs in the ecosystem, it is the responsibility of the provider of the regulated U2U service to ensure that age assurance is

[8] Section 230(4) of the Act.

used in such a way that it is highly effective at determining whether or not a user is a child, in accordance with the proposed age assurance measures.

Figure A10.1: Summary of our approach to implementing highly effective age assurance

Criteria that the age assurance should fulfil to be highly effective
• **Technically accurate**
• **Robust**
• **Reliable**
• **Fair**
Examples of age assurance methods that could be highly effective:
✓ **Open banking**
✓ **Photo-ID matching**
✓ **Facial age estimation**
✓ **MNO age checks**
✓ **Credit cards**
✓ **Reusable digital identity services**
✓ **Oher methods that fulfil each of the criteria**
Examples of age assurance methods that are not capable of being highly effective
✗ **Self-declaration**
✗ **Debit, Solo, or Electron cards**
✗ **Other payment methods which do not require the user to be over 18**
✗ **General contractual restrictions on the use of the service by children**
Principles that service providers should have regard to
• **Accessibility**
• **Interoperability**
• **Transparency**

Technical accuracy

Box A10.3: What is technical accuracy?

Technical accuracy in this context refers to the degree to which an age assurance method can correctly determine the age of a user under test lab conditions.

A10.27 Technical accuracy is important as an age assurance method which performs poorly in test conditions will perform worse in a real-world deployment and is therefore very unlikely to

be highly effective at correctly determining the age of users when deployed. This would indicate that an alternative or additional age assurance method is likely to be required.

A10.28 Age assurance methods either produce:

a) A binary result (for example, categorising users as either over or under the age of 18).
b) A continuous result (for example, providing an estimation of the user's age).[9]

A10.29 We have recommended in the Codes that service providers should take the following steps to fulfil the criterion of technical accuracy:

a) the provider has ensured that the measures forming part of the age assurance process for the service have been evaluated against appropriate metrics to assess the extent to which they can correctly determine the age or age range of a person under test lab conditions;
b) where the age assurance process used on the service involves the use of age estimation, the provider uses a challenge age approach; and
c) the provider periodically reviews whether the technical accuracy of the age assurance process for the service could be improved by making use of new technology and, where appropriate, makes changes to the age assurance process.

A10.30 When evaluating the age assurance method(s) against appropriate metrics, these metrics could be derived from the providers' own internal testing (if feasible); testing by third party providers; or, testing by an independent third party. Where testing has been carried out by third parties, providers should understand what tests have been conducted and the metrics which have been used to measure the results.

A10.31 In the case of methods that produce a binary result, examples of appropriate metrics include but are not limited to; the True Positive Rate (TPR), False Positive Rate (FPR), and False Negative Rate (FNR).

Box A10.4: What is the True Positive Rate (TPR)?

True positives (TP) refer to an outcome where a model correctly predicts a positive class i.e., a user is under 18 and the model predicts their age as under 18.

For the purpose of age assurance, the **true positive rate (TPR)** measures the proportion of TP predictions out of all actual positive instances (i.e., TP and false negative (FN)). This metric highlights the model's performance in correctly identifying positive cases. The formula is TPR = TP / (TP + FN).

Box A10.5: What is the False Positive Rate (FPR)?

False positives (FP) for the purpose of age assurance refer to an outcome where a model incorrectly predicts a positive class i.e., a user is 18 or over and the model predicts their age as under 18.

[9] The estimation of the user's age will usually be accompanied by a confidence interval or range, which conveys the algorithm's level of uncertainty regarding the prediction. For example, where an age estimation method predicts that a user is 25 years old with a confidence interval of ±2 years, this means that the method estimates the user's age to fall within the range of 23 to 27 years.

The **false positive rate (FPR)** measures the proportion of FP against all positive predictions (i.e., FP and TN). FPR highlights the performance of the model in yielding FP results, and this should be minimised. The formula is FPR = FP / (FP + TN).

Box A10.6: What is the False Negative Rate (FNR)?

False negatives (FN) refer to an outcome where a model incorrectly predicts a negative class i.e., a user is under 18 and the model predicts their age 18 or over.

The **false negative rate (FNR),** also known as the 'miss rate,' measures the proportion of FN against all negative predictions (i.e., FN and TP). FPR highlights the performance of the model in yielding FP results, and this should be minimised. The formula is FNR = FN / (FN + TP).

A10.32 In the case of methods that produce a continuous result, examples of appropriate metrics include but are not limited to the Standard Deviation, Mean Absolute Percentage Error (MAPE), and Cumulative Score (CS).

Box A10.7: What is the Standard Deviation?

Standard deviation (SD) is a measure of variation or dispersion of the dataset relative to the mean. A low SD suggests datapoints closer to the mean, whereas a high SD suggests datapoints are more dispersed.

s = ∑((X − MAE)^2/(n – 1)) where X = is the i*th* point in the dataset, MAE = is the mean absolute error, and n = the number of datapoints in the dataset.

Error refers to the user's age determined by the technology minus the user's actual age. An overestimation yields a positive value, whereas an underestimation yields a negative value.

Absolute error (AE) is the same as the 'error,' but disregards the sign (i.e., positive or negative) thus focusing only on the magnitude (size) of the difference between the technologically determined age and actual age.

Mean absolute error (MAE) is the central value of the absolute error. It describes the average discrepancy between a user's technology determined age and their actual age, ignoring whether it is an over- or under-estimation. It is calculated by summing the absolute errors for a given number of absolute errors, then dividing this by the number of absolute errors. The formula is MAE = (1/n) Σ(i=1 to n) |y – x| where n = number of observations in the dataset, y = is the true value, x = is the predicted value.

Box A10.8: What is the Mean Absolute Percentage Error (MAPE)?

The **mean absolute percentage error (MAPE)** is a metric that used to measure the accuracy in a regression analysis. This is useful where relative errors (age range

estimations) are more meaningful than absolute errors. M = (1/n) Σ(t=1 to n)|(At – Ft) / At) |* 100 Where n = number of times the summation iteration happens, At = actual value and Ft = forecast value.

Box A10.9: What is the Cumulative Score? (CS)

The **cumulative score (CS)** is an aggregated score that is calculated by summing the individual score across over a period of time/category etc.

Challenge age

A10.33 A 'challenge age' approach is widely used offline when selling age-restricted products in retail environments, for instance, through the retailing strategy 'Challenge 25.' In this approach, anyone who appears to the provider of restricted products to be under the age of 25 should be challenged to provide acceptable ID proving that they are over the age of 18 if they wish to buy alcohol. The 'challenge age' in this scenario would be 25. [10]

A10.34 In an age assurance process, a challenge age approach refers to where a user who is estimated as being under a given challenge age must then undergo a second age assurance step (for example, a different age assurance method) to confirm that they are over the age required.[11]

A10.35 In the Codes, we specifically recommend that a challenge age should be used where a service uses age estimation. The challenge age should be set according to the limits of the technical accuracy of that method, for example, where system testing suggests that there is a significant risk of incorrectly estimating a 17-year-old's age by 7 years above or below. To manage this risk a buffer can be set above the age by 8 years, so if the age of interest is 18 then the Challenge Age would be 25. For users estimated to be over the age of 25, no additional verification will be required. Where the method estimates that the user's age is under the challenge age, the user could be required to undergo another age check by a second method that is more technically accurate for that age group.

A10.36 Using a 'challenge age' approach can help to improve the overall effectiveness of the age assurance process by preventing or minimising borderline cases where the age estimation method produces an output that could have been subject to error.

Robustness

Box A10.10: What is robustness?

Robustness describes the degree to which an age assurance method can correctly determine the age of a user in unexpected or real-world conditions.

A10.37 Conditions in the real world will vary considerably to those in a test scenario. If the age assurance method is not robust, there are likely to be discrepancies in how it performs

[10] Drink Aware, Challenge 25. [accessed 22 March 2024].
[11] ACCS, 2022. Measurement of Age Assurance Technologies. [accessed 22 April 2024].

across varying conditions. Conditions that change the quality of characteristics of the input include poor lighting, blurring, brightness, contrast, or positioning of the user in the image.

A10.38 In addition, there may be circumvention techniques which are easily accessible to children and where it is reasonable to assume that children may use them. These might require limited cost in terms of time, money or materials. For example, a child uploading an image of an ID that does not belong to them. If the age assurance process is not robust, it will be more vulnerable to circumvention.

A10.39 It is therefore important that the age assurance process fulfils the criterion of robustness. We have recommended in the Codes that service providers should take the following steps to fulfil this criterion:

a) The provider has:

i) taken steps to identify methods children use to circumvent the age assurance process used on the service to determine that the relevant individual is not a child; and
ii) taken feasible and proportionate steps to prevent children using those methods; and

b) the provider has ensured that the age assurance measures[12] forming part of the age assurance process for the service have been tested in multiple different environments during the development of the age assurance process.

A10.40 Testing the age assurance measure in multiple environments during its development will help to minimise any discrepancy in the performance of the method in unexpected or real-world conditions.

A10.41 With regard to taking steps against circumvention, this recommended step relates to circumvention techniques which are specific to certain age assurance methods. For example, when using photo-ID matching, a provider might:

- require a photo of the user at the point of ID upload to verify that the photo ID belongs to that user;
- implement liveness detection to ensure that the user undergoing the age assurance process is present at the time the check is carried out; and,
- ensure the method can detect basic levels of falsified documentation or manipulation, for instance, by using the Government-issued guidance on how to prove and verify someone's identity ('GPG45') which provides some useful indicators on how a document can be scored to detect certain levels of faked documentation. To prevent the most basic levels of fake documentation getting through, this could align to a photo-ID method meeting at least level 2 checks from GPG45.

[12] We acknowledge the Draft Code Children Safety Codes refers to 'age assurance methods' as 'age assurance measures.' This is to reflect the statutory language of the Act as per Section 41(3), Schedule 4, in particular Sch 4 para 12. In sub-section 'Highly Effective Age Assurance' of the Draft Children Safety Codes, 'age assurance measures' has the same meaning as 'age assurance methods' in the Age Assurance Section.

Reliability

Box A10.11: What is reliability?

Reliability describes the degree to which the age output from an age assurance method is reproducible and derived from trustworthy evidence.

Box A10.12: What is reproducibility?

Reproducibility describes the ability for an age assurance method to perform in a consistent manner, producing the same or similar outputs when given the same or similar inputs.[13]

Box A10.13: What do we mean by strength of evidence?

Strength of evidence describes the relative weight that should be afforded to the underlying data or documents used as evidence for a user's age.[14] It concerns how trustworthy the documents or data are and therefore is indicative of how much reliance, or doubt, a service should place on the output of an age assurance method derived from this evidence.

A10.42 Without reproducibility, an age assurance method might correctly determine the same user to be a child in some instances, but not in others. Demonstrating that a method can account for variance and create reproducible outputs is therefore an important element of ensuring that children are prevented or protected from encountering harmful content online.

A10.43 In addition, where age assurance does not rely on trustworthy age evidence, there is a risk that a service incorrectly determines a child to be an adult based on evidence that wrongly suggests they are over 18 in some instances.

A10.44 We have recommended in the Codes that service providers should take the following steps to fulfil the criterion of reliability:

a) where age assurance measures[15] forming part of the age assurance process rely on artificial intelligence or machine learning, the provider should take steps to ensure that:

i) the artificial intelligence or machine learning has been suitably tested during the development of the age assurance process to ensure it produces reproducible results;

ii) the artificial intelligence or machine learning is regularly tested to ensure it produces reproducible results;

[13] Gundersen OE, Kjensmo S, 2018. State of the Art: Reproducibility in Artificial Intelligence, Proceedings of the AAAI Conference on Artificial Intelligence. [accessed 22 April 2024].

[14] 'Strength' refers to evidence being harder to forge or counterfeit, as defined in GPG45.

[15] We acknowledge the Draft Code Children Safety Codes refers to 'age assurance methods' as 'age assurance measures.' This is to reflect the statutory language of the Act as per Section 41(3), Schedule 4, in particular Sch 4 para 12. In sub-section 'Highly Effective Age Assurance' of the Draft Children Safety Codes, 'age assurance measures' has the same meaning as 'age assurance methods' in the Age Assurance Section.

iii) the outputs of the artificial intelligence or machine learning used are monitored and assessed against key performance indicators designed to identify whether the artificial intelligence or machine learning produces reproducible results;

iv) in circumstances where the artificial intelligence or machine learning used are observed to be producing unreliable or unexpected results, the root cause of the issue is identified and rectified.

b) The provider has taken steps to ensure that any data relied upon as part of the age assurance process comes from a reliable source.

Reproducibility

A10.45 Age assurance methods relying on statistical modelling or artificial intelligence, such as facial age estimation and photo-ID matching, are likely to produce outputs with a degree of variance. This can be due to several reasons, including data variability, model complexity, and 'model drift.'

Box A10.14: What is model drift?

Model drift is where the data the method has been trained on becomes less representative of the population using the age assurance method. For example, population demographics may shift over time resulting in a greater degree of variance.

A10.46 Therefore, services using these methods should follow the recommendation to ensure the method(s) used has undergone regular testing. This will ensure that the method produces consistent outputs when presented with the same inputs.

A10.47 This testing should be accompanied by regular monitoring and measurement of key performance indicators of the system. For example:

- Age Verification Accuracy Rate (AVAR): the percentage of users correctly identified as belonging to the appropriate age group.
- Age Verification Efficiency (AVE): the time taken to complete the age verification process.
- Drift Threshold: establish predefined thresholds for AVAR and AVE beyond which significant model drifting is considered to have occurred.

A10.48 Where a service identifies issues through conducting root analysis under step A1.43 (c) above, appropriate steps to rectify this might include retraining the relevant machine learning model.

A10.49 For other kinds of age assurance methods, including credit card age checks, open banking, and MNO age checks, outputs do not generally exhibit any variance of the type described above.

Strength of evidence

A10.50 We expect service providers to have confidence in the evidence that the age assurance method is relying on by considering, for example:

- the nature and properties of any identity documents, profiles, accounts, data, etc. used as part of the age assurance process; and,
- the source of the underlying data or documents.

A10.51 In assessing the nature and properties of the relevant evidence, service providers should identify features that they would expect to see in a reliable source. When deploying photo-ID matching, for example, these features might include that:

- the evidence has originated from a country or organisation that is recognised as trustworthy;
- the positioning of the photographs on the evidence does not suggest they have been edited or replaced;
- the layout or any logos look as expected; and/or,
- the visible security features are genuine.[16]

Fairness

Box A10.15: What is fairness?

Fairness describes the extent to which an age assurance method avoids or minimises bias and discriminatory outcomes. It refers here to the internal operation of an age assurance method, as opposed to external factors, such as a lack of access to a particular form of identification required by the age assurance method.[17]

A10.52 Implementing a fair age assurance process is important to avoid discriminatory outcomes for certain groups. For example, where an age assurance method provides outputs with a lower degree of technical accuracy for users of certain ethnicities when relying on facial estimation. Such an outcome might lead to children being incorrectly determined to be adult users, or adult users being incorrectly determined to be children.

A10.53 Fairness is also important to ensure services abide with duties under the Equality Act 2010 ('the EA 2010'), which prohibits discrimination against persons sharing protected characteristics (including race, age, disability, sex, and gender assignment).[18]

A10.54 To fulfil the criterion of fairness, we have recommended in the Codes that service providers should ensure that any elements of the age assurance process which rely on artificial intelligence or machine learning have been tested and trained on data sets which reflect the diversity in the target population.

Record keeping

A10.55 Services are required to make and keep a written record of any measures taken or in use to comply with a relevant duty which:

a) Are described in a code of practice and recommended for the purpose of compliance with the duty in question, and
b) Apply in relation to the provider and the service in question.[19]

---

[16] Further examples and information on checking that evidence is genuine or valid can be found in GPG45.
[17] The technical criterion of fairness is distinct from the principle of fairness in the UK General Data Protection Regulation (GDPR) which concerns how a user's personal data is processed. For more information, see ICO, Principle (a): Lawfulness, fairness and transparency. [accessed 22 March 2024] and ICO, 2023. Guidance on AI and data protection. [accessed 22 March 2024].
[18] Section 4 of the Equality Act 2010.
[19] Section 23(3) of the Act.

A10.56 Guidance on how to keep a written record can be found in our draft Guidance on Record keeping and Review.

Principles to consider

A10.57 Service providers should ultimately ensure that the age assurance process used is highly effective at correctly determining whether a particular user is a child.

A10.58 Alongside fulfilling the criteria, we also consider it is important that the age assurance process is easy to use and works for all users, so adult users are not unduly prevented from accessing legal content.

A10.59 In recognition of this, we recommend additional principles in the Codes that services should have regard to when implementing the age assurance process. These are:

a) The principle that age assurance should be accessible;
b) The principle of interoperability (defined below); and,
c) The principle of transparency

Accessibility

A10.60 The Act sets out that in recommending the use of age assurance, or which kinds of age assurance to recommend, Ofcom must have regard to the principles that age assurance should:

a) be easy to use, including by children of different ages and with different needs;[20] and,
b) work effectively for all users regardless of their characteristics or whether they are members of a certain group.[21]

A10.61 We refer to these principles collectively using the term **accessibility.** We propose that services should also have regard to these principles to ensure that, as far as possible, adults are not unduly excluded from accessing legal content. An inaccessible age assurance process might be one which is too difficult to use, leading users to abandon the process. Alternatively, the requirements of an age assurance process might make it inaccessible to certain groups of users, thereby excluding them from the process.

A10.62 Accessibility will also help to ensure that adult users are not excluded based on them holding certain characteristics or being part of a certain group. This will assist regulated services in complying with their duties under the EA 2010, as set out above under the criterion of fairness.

A10.63 To have regard to accessibility, service providers may wish to assess the potential impact that age assurance might have on users with different characteristics. In doing so, service providers may find it helpful to:

- Consider the potential impacts on users with "protected characteristics" under equalities legislation (including age, disability, gender reassignment, race, and sex, as well as users of different nationalities that may speak different languages).
- Challenge themselves to think as broadly as possible, including indirect or cumulative impacts (intersectionality). For example, you may not consider that a particular age assurance method will discriminate against or adversely affect users of a particular race, but it is possible that this assessment may change when, for

[20] Schedule 4, paragraph 12(2)(e) to the Act.
[21] Schedule 4, paragraph 12(2)(f) to the Act.

example, you think about users of that race that are also of a particular sex and/or within a particular age bracket (specifically in this case, under or above 18).
- Consider collecting evidence to properly assess potential impacts on particular groups e.g., through focus groups, surveys or reaching out to representative bodies, charities or communities.
- Continually consider and review potential impacts and, where appropriate, revise the assessment as thinking progresses.
- Record the assessment of potential impacts and mitigating steps so that it is clear how they determined what would be the most appropriate age assurance process for the service.

A10.64 It is for service providers to consider what steps are most appropriate for their service to have regard for accessibility. Examples of practical steps to improve accessibility include:

- Considering whether to offer a variety of age assurance methods to assist users who may be unable to, or may find it more difficult to, use certain kinds of age assurance;[22] and,
- Designing the user journey through the age assurance process to be accessible for a wide range of abilities.

A10.65 The Web Accessibility Initiative's Web Content Accessibility Guidelines provide recommendations for how service providers can make content more accessible to users with a wide range of disabilities, including blindness, deafness, limited movement, and learning disabilities.

Interoperability

Box A10.16: What is interoperability?

Interoperability describes the ability for technological systems to communicate with each other using common and standardised formats. It relies on consistent technological approaches being adopted across different methods.

In the context of age assurance, interoperability may involve re-using the result of an age check across multiple services allowing different providers of age assurance methods to share the information in line with data protection laws.

A10.66 Interoperability offers a potential benefit to the user experience, as it limits the amount of information that users need to provide when accessing a new service if they have already proved their age elsewhere. This could reduce the time and effort required by users to understand, and input into, different age assurance processes. It could also reduce the data protection risks that might otherwise occur.

A10.67 We recognise that the development of interoperable solutions is still at an early stage. We are therefore not making any specific recommendations relating to interoperability, beyond recommending that services should have regard to interoperability as a principle.

A10.68 Service providers can have regard to interoperability by staying up to date with developments in this area, and considering whether to implement interoperable solutions to age assurance where they exist and are appropriate for the service.

[22] For example, those without credit cards will be unable to complete a credit card check. Those without a driving licence or passport will be unable to undergo a photo-ID check that relies on these documents.

A10.69 Current efforts at enabling interoperable age assurance include:

- The UK Government's Digital Identity and Attributes Trust Framework (DIATF) may enable interoperability between providers of digital identity and attribute services by encouraging the consistent adoption of common rules and standards. Certain digital identity and attribute services may provide or specialise in age assurance methods. The DIATF will come into full effect once the Data Protection and Digital Information Bill receives Royal Assent.
- The euCONSENT project is a non-profit non-governmental organisation that has been established with the intention of designing, testing, and implementing extensions to the eIDAS infrastructure to enable open-system, secure and certified interoperable age verification.
- The Open Wallet Foundation (OWF) is a consortium of companies and non-profit organisations collaborating to drive the global adoption of open, secure and interoperable digital wallet solutions.

Transparency

Box A10.17: What is transparency?

Transparency refers to the practice of disclosing relevant information so that others can make informed decisions.

A10.70 We consider it important that users are informed about the age assurance process before completing an age check. This includes explaining to users what the age assurance process is designed to do and how it works, so that users can understand why it is necessary and how to complete the process.

A10.71 Setting this information out clearly and accessibly in the terms of service will help services comply with the duties to include provisions in their terms of service specifying how children are to be prevented from encountering PPC and protected from encountering PC/NDC on their service.[23] We set out more information on this in our Terms of Service Measures 1 and 2 protection of Children Codes of Practice, Section 18.

A10.72 As well as in the terms of service, it may be helpful for services to make information on the age assurance available in the form of a pop up prior to completing the age check, for example, as a smaller, new window that appears overlayed on top of the webpage, drawing the user's attention. The text could be included in this window, or the pop up could feature a button prompting users to click for more information.

A10.73 Services likely to be accessed by children also have a duty to operate a complaints procedure in relation to complaints by a user who is unable to access content because measures used to comply with a duty set out in section 12(2) or (3) have resulted in an incorrect assessment of the user's age. This complaints procedure must be easy to access, easy to use and transparent, and services should take appropriate action in response to these complaints.[24] We set out recommendations to support these duties in User Reporting Section 17.

A10.74 Transparency is also an important data protection principle under the UK's data protection regime and refers to being open and honest with people about how and why you use their personal data, for instance through publishing privacy notices. It is also the fourth standard

[23] Section 12(9) and 12(13) of the Act.
[24] Section 21(2) and 21(5)(e) of the Act.

in the ICO's Children's code (also known as the Age Appropriate Design code). Service providers may wish to consult the relevant ICO guidance when having regard to the principle of transparency when implementing age assurance.[25]

Privacy and data protection

A10.75 All age assurance methods involve the processing of personal data. As such, they are subject to the requirements of the UK's data protection regime.

The Data Protection Regime

A10.76 The UK data protection regime is made up of several pieces of legislation, including the Data Protection Act (DPA) 2018, the UK GDPR, and the Privacy and Electronic Communications Regulations (PECR) 2003.

A10.77 Together, this legislation provides a risk-based framework for making sure the processing of personal data respects the fundamental risks and freedom of individuals. The Information Commissioner's Office (ICO) is responsible for upholding information rights through its oversight and enforcement of the legislation.

A10.78 Services providers should consult ICO guidance when implementing age assurance to understand how to comply with the data protection regime, including its guides to the data protection principles, identifying an appropriate lawful basis, and how to respond to users exercising their individual rights afforded by the UK GDPR.[26]

ICO guidance on data protection and age assurance

A10.79 The data protection principles are the cornerstone of the UK GDPR.[27] The ICO guidance includes the data protection principles for UK GDPR which are:

a) Lawfulness, fairness and transparency[28]
b) Purpose limitation[29]
c) Data minimisation[30]
d) Accuracy[31]
e) Storage limitation[32]
f) Security[33]
g) Accountability.[34]

A10.80 To assist in implementing age assurance while protecting user privacy, we have recommended in the Codes that service providers should familiarise themselves with ICO's Children's code, and the Commissioner's Opinion on Age Assurance for the Children's code.

---

[25] Transparency is discussed in section 6.1.3 of the Commissioner's Opinion on Age Assurance for the Children's code. Additional guidance can be found at ICO, 'Transparency,' in the ICO's Accountability Framework. [accessed 22 March 2024].
[26] ICO, 2023. A guide to the data protection principles [accessed 22 March 2024]; ICO, A guide to lawful basis [accessed 9 January 2024]; and ICO, Individual rights – guidance and resources [accessed 22 March 2024].
[27] For an overview of each principle, see the ICO's guide to the data protection principles.
[28] ICO, Principle (a): Lawfulness, fairness and transparency. [accessed 22 March 2024].
[29] ICO, Principle (b): Purpose limitation. [accessed 22 March 2024].
[30] ICO, Principle (c): Data minimisation. [accessed 22 March 2024].
[31] ICO, Principle (d): Accuracy. [accessed 22 March 2024].
[32] ICO, Principle (e): Storage limitation. [accessed 22 March 2024].
[33] ICO, Principle (f): Integrity and confidentiality (security). [accessed 22 March 2024].
[34] ICO, Accountability and governance. [accessed 22 March 2024].

A10.81 The ICO's Children's code is a statutory code of practice which sets out 15 standards that internet society services likely to be accessed by children must conform with to protect children's information rights online. The standards include that the best interests of the child should be a primary consideration when designing and developing online services likely to be accessed by children. Services should take the standards of the Children's code into account when implementing highly effective age assurance.[35]

A10.82 The Opinion outlines how the data protection principles and other requirements can be considered in the context of age assurance. In particular, the Opinion explains how age assurance can form part of an appropriate and proportionate approach to reducing or eliminating the personal information risks children face online and enabling conformance with the Children's code. The considerations set out in the Opinion are technology neutral, making them applicable to any kind of age assurance.[36]

Having regard to privacy under the Act

A10.83 Under section 22 of the Act, when deciding on, and implementing, safety measures, services have a duty to have particular regard to the importance of protecting users from a breach of any statutory provision or rule of law concerning privacy.[37] Where we have concerns that a provider has not complied with its obligations under data protection laws, we may refer the matter to the ICO.

A10.84 To demonstrate compliance with this duty, service providers may find it helpful to keep a written record of how they have taken privacy into account when implementing highly effective age assurance.

A10.85 The examples listed below, which reflect relevant principles set out in the ICO's Children's code, are ways to demonstrate consideration of data protection law, which service providers may wish to provide details on in the written record.

- **Conducting a Data Protection Impact Assessment (DPIA).** These are required by data protection law where processing is likely to result in a high risk to the rights and freedoms of individuals. DPIAs will assist service providers in identifying and mitigating the risks arising from their processing of personal data, which can help demonstrate that they have had regard to the importance of protecting users from a breach of any statutory provision or rule of law concerning privacy. As set out in Standard 2 of the Children's code, a DPIA can also help services to minimise and identify the specific risks to children who are likely to access the service which arise from the processing of their personal data.[38] Detailed guidance on how to carry out a DPIA, and a sample template, can be found on the ICO website.
- **Providing privacy information to users.** Service providers should give users information about why they need to provide any personal data, how it will be processed, how long it will be retained, and if it will be shared with anyone else. Doing so in a child-friendly way will also help services to meet Standard 4 of the

[35] A summary of the 15 standards can be found at ICO, 'Code standards' in Age appropriate design: a code of practice for online services. [accessed 20 March 2024].
[36] ICO, Children's code guidance and resources. [accessed 22 March 2024].
[37] Section 22(3) of the Act.
[38] ICO, '2. Data protection impact assessments' in Age appropriate design: a code of practice for online services. [accessed 22 March 2024].

Children's code: transparency.[39] More information on privacy notices can be found on the ICO website.[40]

- **Keeping written records of processing activities.** Most organisations that process personal data must document their processing activities to some extent.[41]
- Having up to date data protection policies along with a record of how providers make staff aware of them. This provides staff with clarity and consistency around their data protection obligations.[42]
- **Having a record of which staff have completed any data protection training programme that is in place.** This helps to ensure all staff have adequate knowledge of data protection, as appropriate for their role.[43]
- **Clearly documenting technical and organisational security measures**.[44]

[39] ICO, '4: Transparency' in Age appropriate design: a code of practice for online services. [accessed 20 March 2024].
[40] See ICO, Transparency (cookies and privacy notices). [accessed 22 April 2024] and ICO, How to write a privacy notice and what goes in it. [accessed 22 April 2024].
[41] ICO, Records of processing and lawful basis. [accessed 22 March 2024]. Also see 'Governance and Accountability' in Age appropriate design: a code of practice for online services. [accessed 20 March 2024].
[42] ICO, Policies and procedures. [accessed 23 November 2023]. Also see 'Governance and Accountability' in Age appropriate design: a code of practice for online services. [accessed 20 March 2024].
[43] ICO, Training and awareness. [accessed 23 November 2023]. Also see 'Governance and Accountability' in in Age appropriate design: a code of practice for online services. [accessed 20 March 2024].
[44] ICO, A guide to data security. [accessed 22 March 2024].

A11. Summary of Child Safety Measures Across Platforms Most Children Use

Methodology	This research was conducted via desk research by Ofcom researchers. It did not involve researchers creating accounts on platforms, and all data is based on publicly available information from either a platform's terms of service, their community guidelines, information on their website, or a platform's own press releases. To identify the platforms most children in the UK use for this work, we relied on Ipsos iris data about the sites and apps most visited by 15-17 year olds and findings from Ofcom's Children's Media Literacy Tracker about the apps and sites most used by 3-17 year olds. We conducted the research in order to have an overview of the safety measures different platforms had in place.								
Category	**Content Moderation**							**Goods & Service Restrictions**	
Safety Measure	**Content Moderation**	**Use of AI to block content and contact harms**	**Keyword detection**	**Permit but not promote certain types of content?**	**Use of warnings & grey-out screens**	**User reporting**	**Additional criteria related to children**[45]	**Goods & Service Restrictions**	**Age-restrict goods or services?**[46] [47]
Number of the 33 platforms assessed that stated they employ the following measures and further details.	All 33 platforms assessed stated that they employ a form of the content moderation measures described here.	22 of the 33 of the platforms assessed stated that they use AI to block types of content or contact harms, and 32 of the 33 assessed provided definitions for content and contact harms.	20 of the 33 platforms assessed stated that they employ a form of keyword detection in their content moderation efforts, and 13 of them stated that they do so in a child safety-specific use case.	9 of the 33 platforms assessed stated that they permit but do not promote certain types of content that may be ineligible for recommendation according to their community guidelines and terms of service.	10 of the 33 platforms assessed stated they employ warnings and grey-out screens for potentially sensitive or violative content, and all 10 published criteria for employing this measure.	All 33 of the platforms assessed allowed users to report or flag content, and 26 published criteria for the evaluation of these reports.	12 of the 33 platforms assessed stated that they have additional criteria for content moderation efforts as they apply to children.	25 of the 33 platforms assessed stated that they restrict goods and services.	12 of the 33 platforms assessed stated that they age-restrict certain goods and services, and 25 published a list of banned goods and services.

Category	**User Age Policies**[48]				**User Support**		
Safety Measure	**User Age Policies**	**Forms of self-declaration for user age policies**	**Minimum age policy in place?**[49]	**Restrict content for those self-declared under 18?**[50]	**User Support**	**Provision of advice regarding content and contact harms**[51]	**Reactive intervention for suicide, self-harm, eating disorders, and substance abuse**
Number of the 33 platforms assessed that stated they employ the following measures and further details.	All 33 platforms assessed stated that that they employ a form of the user age policy mitigations described here.	29 of the 33 platforms assessed relied on a user self-declaration process for their user age policies, 2 analysed user photos to estimate age, and 1 stated that they require parental certification or permission.	32 of the 33 platforms assessed had a minimum age policy in their terms of service and community guidelines.	17 of the 33 platforms assessed stated that they restrict content for those who have self-declared that they are under 18.	28 of the 33 platforms assessed stated that they employ a form of the user support measures described here.	27 of the 33 platforms assessed provided advice to their users regarding content and contact harms.	11 of the 33 platforms assessed stated that they employ reactive interventions for those at risk of suicide, self-harm, eating disorders, or substance abuse, and 6 stated that they use AI to support vulnerable users.

[45] This refers to any child-related content moderation measure not already covered by the other categories.
[46] Note that this figure includes platforms that do not have an explicit 'banned goods and services' list but still mention banned goods and/or services in other parts of their guidelines.
[47] Note that the age-restricted figure also includes restrictions applied to advertisements or events.
[48] Note that this is limited only to statements described in platforms' terms of service and community guidelines, and measures enforced through user self-declaration.
[49] Note that this figure also includes platforms targeted at children that may instead have a maximum age policy.
[50] Note that the figure also includes restrictions that apply to branded content only as well as 'mature content' restrictions.
[51] Note that this also includes resources based outside of the UK.

A12. Further detail on economic assumptions and analysis

A12.1 This annex provides further information related to economic analysis used to support our provisional conclusions for some of the measures we propose to include in our Children's Online Safety Codes of Practice ('Codes'). We outline:

a) General assumptions we have used to develop quantified cost estimates across several of the measures; and
a) More detail on specific assumptions and analysis related to our proposed measures for Age Assurance (discussed in section 15).

General cost assumptions

A12.2 This annex describes some of the general assumptions we have made on costs. where these assumptions apply to our analysis of many of the proposed measures. These general assumptions are usually combined with other assumptions that are specific to each measure to determine the costs of measure in the chapters in the main body of the report. Any additional assumptions that are used in the cost analysis are described in the costs section of the relevant chapters.

Price Level

A12.3 All quantified estimates of costs or benefits are provided in 2023 prices, unless otherwise stated. We have used 2023 prices, as that is the year of the most recent Annual Survey of Hours and Earnings ('ASHE'), which we use to develop estimates for the labour cost required to implement some code measures.[52]

A12.4 Note that our previous Illegal Harms Consultation, which includes cost analysis for some similar measures to those we propose in the current consultation, used 2022 ASHE data as this was the most recent available data when that analysis was conducted.

Labour Costs

A12.5 To develop estimates for labour costs, we have estimated a salary range for employees across four types of professions, who are likely to develop and/or manage the systems and processes that in-scope services will need to have to comply with the regime. For the lower end of the range, we have used the ASHE 2023 gross median full-time earnings for the relevant occupation, which includes both base and incentive pay.[53]

A12.6 The four professions we have determined to be most relevant for developing our proposed measures, and their relevant Standard Occupational Classification ('SOC') 2020 references are as follows:

[52] Office for National Statistics ('ONS'), 2023. *Annual Survey of Hours and Earnings ('ASHE')*, Table 14, 2023 provisional estimates.
[53] ASHE documentation does not explicitly state that gross salaries include bonuses, but our understanding is that the gross pay includes bonuses, tips and other payments.

a) We use the Programmers and software development professionals salary (2134) to estimate the cost of 'software engineer' time used when developing our cost estimates.
b) We use the Database administrators and web content technicians (3133)[54] salary to estimate the cost of 'content moderator' time used when developing our cost estimates.
c) We use the Professional Occupations (2) estimate to cover a range of professions that are employed at various online services and might be required to implement code measures. This could be legal employees, operations, product managers and so forth.
d) We use Graphic and multimedia designers salary (2142) to estimate the cost of 'Graphic and multimedia designer' time used for creating audio-visual support materials.

A12.7 We recognise that for some services, median UK wage rates may differ from actual salary rates. This may be especially the case for larger services based in the US, who may have higher salary levels. We also appreciate that the salary costs of some types of staff, such as software engineers with certain specialisms, may vary and may be considerably higher in some cases. To take account of this, we also include a higher estimate, which we have assumed is double the value of our lower estimate.

A12.8 Conversely, we are aware that some services may outsource some relevant work to locations where average pay is lower than the UK, which may reduce these costs. To the extent this is the case, our salary range may tend to overstate costs.

A12.9 Table A12.1 shows the resulting low and high estimates we use for the four occupations.

Table A12.1: Gross Annual Wages Estimates

Occupation	Gross Annual Wage Estimates (ASHE 2023)	
	Low	High
Software Engineer	£49,430	£98,860
Content Moderator	£31,500	£63,000
Professional Occupations	£43,191	£86,382
Graphic and multimedia designer	£29,104	£58,208

A12.10 We also assume **a 22% uplift** to the gross wage costs to account for non-wage labour costs, such as employers' National Insurance contributions.[55]

A12.11 When producing cost estimates for our measures, we have used resourcing estimates based on different time periods (e.g. days/weeks/months) suitable for the particular measure. To help understand more clearly the unit labour costs used in each of these situations, the

[54] This four-digit SOC 2020 code (unit group code 3133) includes occupations such as content, chat, web, and website moderators as well as other occupations such as database administrators and web content. technicians. ONS, SOC 2020 Volume 2: the coding index and coding rules and conventions [accessed 25 March 2024]. The associated ONS spreadsheet can be found here: SOC 2020 Volume 2: the coding index.
[55] This is the non-wage uplift recommended by the Regulatory Policy Committee ('RPC'). Source: RPC, 2019. *RPC guidance note on 'implementation costs'*. It is also the uplift used by DSIT in its Impact Assessment for the Online Safety Bill.

following tables illustrate these costs on a daily,[56] weekly and monthly[57] basis. These figures include the 22% uplift mentioned above.

Table A12.2: Estimated daily labour cost

Occupation	Estimated Daily Labour Cost	
	Low	High
Software Engineer	£265	£530
Content Moderator	£169	£338
Professional Occupations	£231	£463
Graphic and multimedia designer	£156	£312

Table A12.3: Estimated weekly labour cost

Occupation	Estimated Weekly Labour Cost	
	Low	High
Software Engineer	£1,177	£2,354
Content Moderator	£736	£1,472
Professional Occupations	£1,047	£2,093
Graphic and multimedia designer	£696	£1,393

Table A12.4: Estimated monthly labour cost

Occupation	Estimated Monthly Labour Cost	
	Low	High
Software Engineer	£5,025	£10,051
Content Moderator	£3,203	£6,405
Professional Occupations	£4,391	£8,782
Graphic and multimedia designer	£2,959	£5,918

[56] The daily labour cost is estimated by increasing the annual salary by 22% and dividing by the number of working days in a year. We assume on average there are 228 working days in a year. This assumes people work 5 days a week and that there are 8 bank holidays and on average people take 25 days of leave a year.
[57] The monthly labour cost is estimated by increasing the annual salary by 22% and dividing by the number of months in a year (12).

A12.12 In the Governance and Accountability section, we have used assumptions of salary for a Senior Leader (£100,000 per year), a Senior Director (£150,000 per year), and a S&P 500 Independent Director (£250,000 per year)[58] to approximate the costs to businesses of measures which would be conducted at least in part by senior individuals.

Non-engineering Costs for System Changes

A12.13 Where system or other software changes associated with a proposed measure involve a software engineering cost, we typically match the amount of engineering time with an equivalent amount of non-engineering time for work carried out by people in professional occupations. This is to account for non-engineering labour time that a business might need to spend on a system change, for instance legal or project management associated with the change.

Maintenance Costs for System Changes

A12.14 Where system or other software changes associated with a proposed measure involve an initial cost, we typically assume there is also an ongoing annual maintenance cost of 25% of the initial cost. These ongoing costs reflect work likely required to ensure the system continues to operate as intended. We apply this assumption unless we have more specific information about the ongoing maintenance costs.

Further detail on age assurance cost analysis

A12.15 This sub-section provides further analysis of costs which has been used to support our provisional conclusions on age assurance measures, as set out in section 15. We discuss:

- Our general cost assumptions.
- Direct costs to services. We consider that all direct costs are likely to depend on how a service provider approaches its implementation of the measures, but in all cases we consider that the main costs are likely to relate to:
  - preparing to implement age assurance; and
  - implementing and operating a third-party age assurance method; or
  - building and operating an in-house age assurance method.
- Indirect costs to services due to our requirements to implement age assurance.

Our general cost assumptions

A12.16 We adopt several general assumptions to estimate a range of costs services may incur, with the higher end of the range representing a conservative upper bound for costs. Where these assumptions do not hold, the costs may be materially lower but also higher in some cases. We also make some further assumptions in relation to specific cost elements, as explained separately.

A12.17 **We have assumed that users will have to confirm their age for each service separately.** We recognise that where an online service provider manages multiple services it may be

[58] Spencer Stuart, 2023. 2023 S&P 500 Compensation Snapshot [Accessed April 2024], converted to GBP.

possible a user is only required to prove their age once (e.g., across multiple pornography services provided by the same company), which may reduce direct costs for the service and also friction on users. Reusability of age checks and/or interoperability of age assurance methods may become more widely available in future, for example, where a user can complete an age check that is valid for many service providers. This could reduce costs and make implementing and operating age checks more cost effective for more services. We understand that the age assurance industry expects interoperability to increase over time. We expect that services will be incentivised to facilitate this process, as this could reduce user friction and increase user numbers. We also anticipate that some services will have an incentive to sell access to their developed age assurance methods in order to bring in additional revenue.

A12.18 **We have assumed that service providers have no existing systems in place that can facilitate age assurance.** Where services already have systems to gate access for users in some way (e.g., a payment system for subscription charges), the costs of implementing age assurance may be lower than our estimates suggest.

A12.19 **We assume that age checks are one-off.** Services can decide to reverify users if they think it is appropriate, for example, depending on risks, but we have not recommended this in our first set of Codes. We expect this to mean that adults will need to have an account with a service to access adult appropriate experiences, which would facilitate a one-off age check.

A12.20 **We have assumed that services apply age assurance to all users**. In practice, services may be able to implement the measures and only age assure a subset of users. For example, as explained in age assurance Section 15, for proposed Measures AA3 and AA4 a service may only conduct age checks for users who are specifically seeking access to identified PPC and PC. Depending on the specific context of a service, this may significantly reduce costs compared to the estimates we present. For instance, if a service conducted age checks for 50% of its users, then we would estimate its ongoing costs related to conducting age checks to be up to 50% lower.

A12.21 We recognise that our cost estimates are dependent on the assumptions we have made and in practice costs could be higher or lower, depending on how service providers have decided to comply with their online safety duties and implement age assurance.

Preparatory costs relating to the introduction of age assurance

A12.22 All U2U services are likely to incur some one-off preparatory labour costs relating to the preparation of adopting age assurance. These may include staff familiarising themselves with our proposed measures and requirements, researching and assessing the suitability of different age assurance options for their service, considering how to implement age assurance in a way that is highly effective, meeting the relevant criteria and having regard to the other principles (such as accessibility).

A12.23 Where a service provider decides to use a third-party age assurance provider, the procurement process is likely to involve some time and effort related to governance and budget processes, senior management engagement, which can take weeks. Going through an existing supplier or a digital experience platform can reduce the procurement time and costs, while a formal tendering process could tie up internal staff's time and take significantly longer. All these preparatory processes may take longer in large organisations with more complex procurement processes.

A12.24 Overall, these preparatory costs are likely to depend on the size and type of service and are expected to be larger for large services because of different governance processes but also the number of employees likely to be involved.

Costs associated with third-party age assurance methods

A12.25 **There may be upfront costs linked to the age assurance provider setting up a client account to prepare the age assurance method for use**, or in some cases, this charge may be part of an ongoing maintenance support service.[59] We understand that some third-party solutions are developed with ease of integration in mind, meaning that connecting to a services' existing systems may be relatively easy and cheap. We recognise that in some cases upfront set-up costs could be more material. For example, if the existing service infrastructure needs adjusting or there are other complexities with linking up the third-party technology with the services' systems or data.

A12.26 **The service provider may have to also introduce access or content controls** as part of implementing the proposed Age Assurance measures and the related Content Moderation and Recommender System measures. For example, this may involve tagging users so that children can get a more age-appropriate experience or are prevented from accessing parts of the services not suited for them. This may require changes to the existing ICT infrastructure or building of a new user interface to integrate age assurance with the service, which could mean costs could be material, although requirements and costs are likely to vary by service and approach to implementation.

A12.27 The service provider may also need to train some of its staff who work closely with the age assurance process (e.g., software engineers maintaining the running of the age assurance software) when the process becomes operational.

A12.28 **We expect that the main cost component relating to third-party age assurance methods is the per-check cost**, covering both a one-off cost to check existing users and an ongoing cost to check new users. These costs are likely to vary depending on the solution and age assurance provider, as underlying costs and pricing approaches vary. According to DSIT's impact assessment of the Online Safety Bill, some age assurance providers offer volume discounts to services requiring a large number of checks and discounted fees for small clients and start-ups in some cases,[60] while subscription-based verification packages often include a fixed number of checks for users.[61] The DSIT data also suggest that price per check ranges from less than 1p to £1, depending on the provider and method used.[62]

A12.29 To illustrate what these per check costs may mean for a service, we consider stylised cost examples for hypothetical services with a different number of users in Table A12.5. According to the Government's impact assessment on the Online Safety Act, most per-check costs provided were 10p or lower.[63] For our analysis, we have chosen to use a low estimate of 5p per check, and a high estimate of 20p.

---

[59] Based on Yoti's price list data from May 2022, setting up an organisational account is £750 per organisation. GC-13 Yoti Age Verification Pricing (digitalmarketplace.service.gov.uk) [accessed 23 February 2024].
[60] DSIT, 2022. Online Safety Bill impact assessment, paragraph 185 [accessed 7 February 2024].
[61] DSIT, 2022. Online Safety Bill impact assessment, paragraph 183 [accessed 7 February 2024].
[62] It is possible that due to inflation in 2022 and 2023 these examples are now out of date. Publicly available per check prices are greater than the bottom end of this range, and in these cases, it is not clear who these prices would apply to. DSIT, 2022. Online Safety Bill impact assessment, paragraph 182 [accessed 5 February 2024].
[63] DSIT, 2022. Online Safety Bill impact assessment, paragraph 182 [accessed 12 April 2024].

Table A12.5: Illustrative cost estimates of age checks via third-party age assurance providers*

	Existing UK user base	New users each year	Age assurance for existing users	Age assurance for new users (annual ongoing cost)
Smaller services	100,000	10,000	£5,000 - £20,000	£1,000 - £2,000
	350,000	35,000	£18,000 - £70,000	£2,000 - £7,000
	700,000	35,000	£35,000 - £140,000	£2,000 - £7,000
Larger services	1,000,000	50,000	£50,000 - £200,000	£3,000 - £10,000
	7,000,000	70,000	£350,000 - £1,400,000	£4,000 - £14,000
	20,000,000	200,000	£1,000,000 - £4,000,000	£10,000 - £40,000

Source: Ofcom analysis

**Note: All cost estimates have been rounded up to the nearest thousand. These stylised examples assume a faster rate of user base growth, in proportionate terms, for the smallest services (10% growth rate) and a lower rate for the largest services (1% growth rate).*

A12.30 If our proposed code measures come into force, **our cost estimates assume that services will incur a one-off cost of checking the age of their entire existing user base.** We multiply the number of existing users by the per-check cost (for example, 100,000 existing users x 5p = £5,000). We estimate that this one-off cost may be between £5,000 and £20,000 initially for a service with 100,000 users, or between £18,000 and £70,000 for a service with 350,000 users. For a service with 700,000 users, we estimate the upfront age check cost to be between £35,000 and £140,000, and between £50,000-£200,000 for a service with 1 million users. A service with 7 million users could incur a cost of between £350,000 and £1.4 million upfront, and between £1 million and £4 million if the service has 20 million users.

A12.31 As noted in our general cost assumptions, in practice we expect services would be able to implement some of the proposed measures while only conducting age checks on a subset of all users (e.g., those who want access to restricted content). Costs would be lower in those cases.

A12.32 **We also estimate the annual ongoing cost of carrying out age checks for new users.** After the first year, we consider that the ongoing age check costs will depend on the size of the user base at the start of regime, the annual growth in new users and the price of age checks. We have assumed a higher growth rate for smaller services (10%) because of their relative ease to grow compared to larger services (1%).[64]

A12.33 We expect that ongoing age checks on new users will continue annually, and that: (a) the cost per check remains unchanged over time; (b) all checks for a service cost the same; and (c) the nature of the service does not influence the per-check cost. Table A12.1 captures what these ongoing checks could cost. We estimate that a service with 10,000 new users per year could incur age check costs of between £1,000-£2,000 annually, or between £2,000 and £7,000 if user numbers grow by 35,000 annually. Where a service has 700,000 users initially

[64] Small absolute increases in user numbers reflect higher growth rates on smaller services. Despite slower growth of larger services, we recognise that this can add tens of thousands of new users to check per year, which can mean tens of thousands of pounds in costs on an ongoing basis.

and adds 35,000 new users annually could incur ongoing costs of £2,000-£7,000, while a service with 1 million users to start with could incur annual costs of £3,000 to £10,000 if user numbers grow by 50,000 every year. A larger service with 70,000 new users annually could incur between £4,000 and £14,000 in costs, while 200,000 new users could cost a larger service between £10,000 and £40,000 yearly.

A12.34 In addition to ongoing costs relating to checking users' age, services may incur other annual costs including licensing of age assurance software if not captured by the ongoing age checks, training costs although given the age checks are managed externally this is likely to be limited, other software used to support the age assurance process, and data storage costs, although in most cases we assume these would be captured by the ongoing age check costs.

A12.35 We note that various testing and evaluation activities are recommended under our highly effective age assurance criteria. Where services use third-party age assurance providers, we have assumed that those third parties would carry out the bulk of these activities, which may limit further costs incurred by services. However, first-party service providers would still be expected to maintain due oversight and understanding of any third-party testing and evaluation, as it is the service providers in scope of our age assurance measures who are ultimately responsible for ensuring that their approach to age assurance is highly effective.

A12.36 Due to the fast-developing age assurance industry and emerging new verification tools, we consider that future costs of third-party age methods are uncertain. We think there is a significant likelihood that costs of age assurance will fall over time, as well as the possibility of interoperability of different solutions to increase in the future. We therefore consider that our estimates of costs are likely to be sufficiently conservative.

Costs of developing an age assurance method in-house

A12.37 For illustrative purposes, **we have also considered what an age estimation method could cost to develop and run**. We assume that the overall development phase takes six months, which includes the development, testing and deployment of age assurance software. We recognise that development time and costs are likely to vary by the approach taken and the estimates we present below are intended to provide an illustrative example of the broad magnitude of costs associated with developing a single in-house age assurance method.

A12.38 The main costs are likely to be:

a) One-off labour costs relating to the upfront expense of developing, testing and deploying the software. This would include meeting recommendations related to technical accuracy (evaluating methods against appropriate metrics) and fairness (testing and training the method on diverse datasets).

b) Ongoing staff costs of monitoring, supporting, and maintaining of the age assurance model. This would include meeting recommendations related to reliability, including monitoring key performance indicators and rectifying issues related to unexpected or unreliable predictions.

A12.40 Our high-level indicative analysis in the context of a very large business (which we consider the more likely scenario[65]), suggests that **the upfront staff costs could be in the region of many hundreds of thousands and potentially up to £1 million.**[66] In addition to these quantified costs, a provider may incur substantial one-off costs relating to acquiring relevant datasets for developing its age assurance method and one-off software/hardware costs relating to additional computational resources to develop and train its age assurance method, which may include cloud infrastructure and data security.[67] While a large service may be able to use existing infrastructure to encompass its new age assurance processes, and this way optimise resource utilisation and not incur additional costs because of the method development, there is an opportunity cost to this because these resources are not available for other uses.

A12.41 **There would also be ongoing staff costs relating to the method monitoring and maintenance**, and there could be additional ongoing data costs if the method requires significant improvements and/or changes in the future. We estimate that these ongoing staff costs could reach £1 million annually or potentially more, depending on a service's approach. Our estimates are based on the same salary assumptions for upfront and ongoing costs. In practice, it is possible that some ongoing activities could be conducted by more junior staff on lower salaries, such that ongoing costs could be lower than suggested here.[68]

A12.42 As with our examples on third-party methods, these cost estimates are only intended to be illustrative and depend on the different assumptions we have made.

A12.43 Any services seeking to develop age assurance methods in-house are likely to be relatively large, due to the substantial upfront costs relating to software development and testing. This may still be more cost effective if a service predicts a high number of new service users over time, while expects the ongoing engineering costs to be lower than what ongoing age checks by a third-party would be.

A12.44 To the extent that smaller services have the relevant capabilities to pursue an in-house approach, it is possible that they may be able to do so in a more cost-effective way than suggested by our indicative cost estimates (e.g. due to having simpler organisational processes and lower overheads in relation to the relevant activities).

A12.45 The service may also incur some one-off staff training costs after age assurance is deployed to users, but these are likely to be relatively small in comparison to the one-off and ongoing costs relating to developing and deploying age assurance approach in-house and will depend primarily on the number of people that need to be trained and how much training is required.

---

[65] For example, Google has appeared in a registry of providers approved by the Age Check Certification Scheme (ACCS), the UK's program for age verification systems. https://www.biometricupdate.com/202312/google-receives-certificate-for-facial-age-estimation-in-the-uk [accessed 7 February 2024].

[66] We assume that the upfront labour input would involve 16 full-time equivalent employees. The cost range is based on an annual software engineer pay of £49,430 (low) and £98,860 (high), uplifted by 22% to account for non-wage labour costs, such as employers' National Insurance contributions. This may be an overestimate given that we expect services could use more junior staff for some model monitoring, maintenance and support functions.

[67] A service developing an age assurance method is likely to require a cloud security solution that runs all the time and scans information regularly. Securing the data and systems is needed from the development phase but the service will continue to incur this as the systems and data need to be secured on an ongoing basis.

[68] The ongoing labour costs we assume require 14 FTEs annually.

Indirect costs on services

A12.46 Our research suggests that some users may be reluctant to prove their age due to not wanting to share personal information with a service or worries about data privacy.[69] This may result in some users leaving services that are required to implement highly effective age assurance. As a result, user numbers and engagement on these services could fall which in turn is likely to reduce advertising and/or subscription revenues. [70] For example, Aylo suggested that after implementing the state-recommended age assurance method in Louisiana in 2022 it lost 80% of its traffic[71], while a video-sharing platform noted an approximate 20% drop in new UK registrants after implementing Yoti on its services in 2021.[72] However, we do not know whether the drop in usage in this case was due to people not wanting to complete an age check, or because these users were children and unable to do so, and whether this drop was ongoing or temporary as users become more familiar with age assurance. Some of these users may have also switched to services which do not age check.

A12.47 Most service providers already aim to maximise revenues from subscriptions and advertising and should have incentives to minimise the loss of users because of our age assurance recommendations. In some cases, the service may be able to offer users access to a child appropriate version of the service where users do not have to confirm their age, which may limit the extent of any negative impacts on user engagement and revenue. We also consider it plausible that light users of a service may, on average, be less likely to be willing to invest time in undertaking an age check, whereas heavy users who value access to the service highly – and who tend to generate more revenue – may be on average more willing to conduct age checks (although we lack specific evidence to this effect). Where this is the case, it would tend to limit any adverse impacts on overall usage and engagement levels.

[69] Ofcom, 2022, Adult Users' Attitudes to Age Verification on Adult Sites. p. 10. [accessed 2 May 2024].
[70] We recognise that a service dedicated to PPC or PC content may require users to verify their age when entering the site which may reduce advertising and subscription revenues from adults who are not prepared to confirm they are 18 or over increasing indirect costs.
[71] BBC (2023) UK porn watchers could have their faces scanned. [accessed 2 March 2024].
[72] [**CONFIDENTIAL ✂**] response dated 4 August 2023 to the RFI dated 23 June 2023.

A13. Legal framework: duties of providers and Ofcom in relation to the protection of children

This annex sets out the duties relating to the protection of children, as they apply to providers of user-to-user services ('U2U services'); providers of search services; and to Ofcom, and which are relevant to this consultation.

This annex does not cover other duties set out in the Online Safety Act 2023 ('the Act')[73], except where relevant to the protection of children. We have not referred to aspects of the legal and regulatory framework which relate to illegal content, which were covered in our consultation entitled *Protecting people from illegal harms online,* (Illegal Harms Consultation), which was published on 9 November 2023.[74] We have also not referred to aspects of the legal framework which relate to phase three of our implementation roadmap, such as transparency, user empowerment and other duties on categorised services.

Provider duties in relation to children's access assessments (U2U and Search)

A13.1 The Act places providers of regulated U2U services; and providers of regulated search services, under a duty to conduct a suitable and sufficient children's access assessment and to keep a written record of the same, in an easily understandable form.[75]

A13.2 A children's access assessment first involves determining whether it is possible for children in the UK to access all or part of the service.[76] [77] The Act provides that a service can only conclude that it is *not* possible for children in the UK to access the service[78] if age verification or age estimation is used on the service with the result that children are ordinarily prevented from accessing the service.[79]

A13.1 If a provider determines that it is possible for children in the UK to access all or part of the service, the provider must go on to consider whether the child user condition is met in relation to all or the relevant part of that service.[80] That will be the case where:

a) there are a significant number of children in the UK who are users of the service or of the relevant part of it, or

[73] Online Safety Act 2023.
[74] Ofcom, 2023. Protecting people from illegal harms online, see Annex 12 Part B.
[75] Section 36 of the Act.
[76] Sections 35(1)(a) and 35(5)(a) of the Act.
[77] Services do not need to assess whether parts of the service which are not, or are not included in, the U2U part of the service or a search engine can be accessed by children in the UK. See section 35(5)(b) of the Act.
[78] Or the relevant part of the service, as applicable.
[79] Section 35(2) of the Act.
[80] Section 35(1)(b) of the Act.

b) the service, or the relevant part of it, is of a kind likely to attract a significant number of users who are children in the UK.[81]

A13.2 In relation to limb (a), the Act provides that whether or not the test is met should be assessed using evidence about actual users (and not who the intended users are).[82] If the number of users that are children in the UK is significant in proportion to the total number of UK users of the service (or the relevant part of it), then the number of children in the UK who are users is significant.[83]

A13.3 Providers who provide more than one U2U or search service must carry out a separate children's access assessment for each service.[84]

A13.4 Part 1 of Schedule 3 to the Act specifies the deadline by which providers must complete their first children's access assessment. Providers of services in operation immediately before the publication of Ofcom's Children's Access Assessments Guidance (see paragraph 2.21 of the draft guidance published as Annex 5 to this consultation) are required to complete the first children's access assessment relating to the service within three months of the date on which that guidance is published. Providers of services that start up or otherwise become Part 3 services after the publication of Ofcom's Children's Access Assessments Guidance must complete their first children's access assessment within three months of becoming a Part 3 service.[85]

A13.5 If, having conducted a children's access assessment, a provider determines that a service (or the relevant part of it) is *not* likely to be accessed by children, then it must carry out a further children's access assessment no more than one year later.[86] Such a provider is also required to carry out a further assessment:

a) before making any significant change to any aspect of the service's design or operation to which such an assessment is relevant;
b) in response to evidence about reduced effectiveness of age verification or age estimation that is used on the service in order to achieve the result that children are not normally able to access the service;[87] or
c) in response to evidence about a significant increase in the number of children using the service.[88]

A13.6 Ofcom is required to issue guidance for U2U and search services to assist with completing the children's access assessment.[89]

[81] Section 35(3) of the Act.
[82] Section 35(4)(b) of the Act.
[83] Section 35(4)(a) of the Act.
[84] Section 36(5) of the Act.
[85] Different provisions apply to providers of video-sharing platform (VSP) services currently regulated by Part 4B of the Communications Act 2003. These providers must complete the first children's access assessment relating to those services by the deadline specified in Part 3 of Schedule 3.
[86] Section 36(3) of the Act.
[87] See section 35(2) of the Act.
[88] Section 36(4) of the Act.
[89] Section 52(3)(b) of the Act.

When services will be likely to be accessed by children

A13.7 Section 37 of the Act sets out when a service will be treated as likely to be accessed by children for the purposes of the Act.

a) First, this will be the case where a children's access assessment carried out by the provider of the service concludes that it is possible for children in the UK to access all or part of the service and the child user condition is met (see paragraphs A13.2 to A13.7 above).[90] In that case, the service will be treated as likely to be accessed by children from the date on which the children's access assessment is completed.[91]

b) Second, this will be the case where the provider of the service fails to carry out the first children's access assessment by the deadline specified in Schedule 3 to the Act.[92] In that case, the service will be treated as likely to be accessed by children from the date by which the assessment should have been completed until the first children's access assessment has been completed.[93] [94]

c) Third, the Act provides that in specific circumstances Ofcom can take action which will result in a service being treated as likely to be accessed by children for the purposes of the Act. This will be the case where, following an investigation into the failure to complete a children's access assessment in accordance with the relevant requirements,[95] Ofcom determine that it is possible for children in the UK to access the service (or the relevant part of it) and the child user condition is met in relation to the service (or the relevant part of it)[96] [97] and, as such mandate that the children's safety duties must be complied with by the service. In that case, the service will be treated as likely to be accessed by children from the date specified by Ofcom.[98] Ofcom has the power to specify the circumstances in which the service will cease to be treated as likely to be accessed by children.[99] [100]

[90] Section 37(2) of the Act.
[91] Section 37(3) of the Act.
[92] Section 37(4) of the Act.
[93] Section 37(5) of the Act.
[94] If the conclusion of that assessment is that it is possible for children in the UK to access all or part of the service and the child user condition is met then the service will continue to be treated as likely to be accessed by children by virtue of section 37(2) of the Act.
[95] Such a failure may arise either in circumstances in which no children's access assessment has been completed at all or in circumstances in which an assessment has been completed but the relevant requirements have not been complied with, for example because the assessment that has been completed is not suitable and sufficient.
[96] Sections 135(4) and 135(5) of the Act give Ofcom the power to make such a determination.
[97] See paragraphs A13.2-A13.4 above for further detail on the meaning of "possible for children in the UK to access the service" and the "child user condition".
[98] The date will be specified in a confirmation decision given to the provider of the service under sections 132 and 135 of the Act.
[99] Section 135(5)(b) of the Act.
[100] The circumstances will be specified in a confirmation decision given to the provider of the service under sections 132 and 135 of the Act.

Duties of providers of U2U services likely to be accessed by children

A13.8 Providers of U2U services are given specific duties under the Act in relation to the protection of children. These include the "children's risk assessment duties"[101] and the "safety duties protecting children".[102]

A13.9 Providers of U2U services are also subject to "additional duties" which are relevant, among other things, to the protection of children. These additional duties are as follows:

a) "Duties about content reporting and complaints procedures", which include:
   i) "Duties about content reporting",[103] and
   ii) "Duties about complaints procedures";[104] and
b) so-called "Cross-cutting duties", which include:
   iii) "Duties about freedom of expression and privacy"[105] and
   iv) "Record-keeping and review duties".[106]

A13.10 These are set out in more detail below. Section 7 of the Act states that all providers of regulated U2U services must comply with these duties (and the other duties set out under section 7(2)).

Connection with the United Kingdom

A13.11 These duties only apply to:

a) the design, operation and use of the service in the United Kingdom, and
b) in the case of a duty that is expressed to apply in relation to users of a service, the design, operation and use of the service as it affects United Kingdom users of the service.[107]

Combined services

A13.12 Where the U2U service is a combined service (i.e. providing both a regulated U2U and regulated search service), these duties will not apply to:

a) the search content of the service,
b) any other content that, following a search request, may be encountered as a result of subsequent interactions with internet services, or
c) anything relating to the design, operation or use of the search engine.[108]

A13.13 However, the duties that apply to regulated search services in relation to the protection of children (see paragraphs A13.44-A13.70 below) will still apply to those aspects of a combined service.

[101] Section 11 of the Act.
[102] Section 12 of the Act.
[103] Section 20 of the Act.
[104] Section 21 of the Act.
[105] Section 22 of the Act.
[106] Section 23 of the Act.
[107] Section 8(3) of the Act.
[108] Section 8(2) of the Act.

Safety duties for services likely to be accessed by children

Protection of children risk assessment duties

A13.14 Providers of regulated U2U services that are likely to be accessed by children have a duty to carry out a suitable and sufficient children's risk assessment[109] at the specific times set out in Schedule 3 to the Act.[110]

A13.15 A children's risk assessment means an assessment of the following matters, taking into account the risk profiles that relate to the services of that kind:[111]

a) the user base, including the number of users who are children in different age groups;
b) the level of risk of children who are users of the service encountering the following by means of the service—
   i) each kind of primary priority content that is harmful to children (with each kind separately assessed),
   ii) each kind of priority content that is harmful to children (with each kind separately assessed), and
   iii) non-designated content that is harmful to children,
   giving separate consideration to children in different age groups, and taking into account (in particular) algorithms used by the service and how easily, quickly and widely content may be disseminated by means of the service;
c) the level of risk of harm to children presented by different kinds of content that is harmful to children, giving separate consideration to children in different age groups;
d) the level of risk of harm to children presented by content that is harmful to children which particularly affects individuals with a certain characteristic or members of a certain group;
e) the extent to which the design of the service, in particular its functionalities, affects the level of risk of harm that might be suffered by children, identifying and assessing those functionalities that present higher levels of risk, including functionalities—
   i) enabling adults to search for other users of the service (including children), or
   ii) enabling adults to contact other users (including children) by means of the service;
f) the different ways in which the service is used, including functionalities or other features of the service that affect how much children use the service (for example a feature that enables content to play automatically), and the impact of such use on the level of risk of harm that might be suffered by children;

---

[109] Section 11(2) of the Act.

[110] The deadline for completing the first risk assessment depends on the day on which a provider of U2U services starts its operations. In particular:

i. U2U services that are already in operation at the outset of this regime, must complete their first children's risk assessment within a period of three months from the day on which Ofcom's risk assessment guidance ('RAG') is published;
ii. new U2U services that start operations after the RAG is published must complete their first children's risk assessment within a period of three months from the day on which they begin their new services; and
iii. existing services that become U2U services (having previously provided a different type of service) after the RAG is published must complete their first children's risk assessment within a period of three months from the day on which their services become a U2U service. See Schedule 3 to the Act.

[111] Section 11(6) of the Act.

g) the nature, and severity, of the harm that might be suffered by children from the matters identified in accordance with paragraphs (b) to (f), giving separate consideration to children in different age groups;
h) how the design and operation of the service (including the business model, governance, use of proactive technology, measures to promote users' media literacy and safe use of the service, and other systems and processes) may reduce or increase the risks identified.

A13.16 The provider of a U2U service that is likely to be accessed by children must take appropriate steps to keep a children's risk assessment up to date, including when Ofcom makes a significant change to a relevant risk profile (see paragraphs A13.74-A13.76).[112]

A13.17 The provider of a U2U service that is likely to be accessed by children is under an obligation to carry out a further suitable and sufficient children's risk assessment, before making any significant changes to any aspect of a service's design or operation. This further children's risk assessment must relate to the impact of that proposed change.[113]

A13.18 Where a children's risk assessment of a service identifies the presence of non-designated content that is harmful to children, the provider of the U2U service is under duty to notify Ofcom of—

a) the kinds of such content identified, and
b) the incidence of those kinds of content on the service.[114]

Safety duties relating to the protection of children

A13.19 Providers of regulated U2U services likely to be accessed by children have specific safety duties in relation to children's online safety as set out under section 12 of the Act. These duties extend to such parts of a service as it is possible for children to access.[115] [116] The duties are as follows:

a) A duty, in relation to a service, to take or use proportionate measures relating to the design or operation of the service to effectively—
   i) mitigate and manage the risks of harm to children in different age groups, as identified in the most recent children's risk assessment of the service (see paragraph A13.17 and section 11(6)(g)) of the Act), and
   ii) mitigate the impact of harm to children in different age groups presented by content that is harmful to children present on the service.[117]
b) A duty to operate a service using proportionate systems and processes designed to—
   i) prevent children of any age from encountering, by means of the service, primary priority content that is harmful to children; [118]
   ii) protect children in age groups judged to be at risk of harm from other content that is harmful to children[119] (or from a particular kind of such content), as assessed by the

[112] Section 11(3) of the Act.
[113] Section 11(4) of the Act.
[114] Section 11(5) of the Act.
[115] A provider is only entitled to conclude that it is not possible for children to access a service, or a part of it, if age verification or age estimation is used on the service with the result that children are not normally able to access the service or that part of it – see section 13(6) of the Act.
[116] The Act, section 13(5).
[117] Section 12(2) of the Act.
[118] With the harm arising by virtue of the nature of the content rather than the fact of its dissemination, see section 13(4) of the Act.
[119] With the harm arising by virtue of the nature of the content rather than the fact of its dissemination, see section 13(4) of the Act.

provider of a service in the most recent children's risk assessment of the service,[120] from encountering it by means of the service.[121]

The duty in paragraph i) above requires a provider to use age verification or age estimation (or both) that is of such a kind, and used in such a way, that it is highly effective at correctly determining whether or not a particular user is a child, to prevent children of any age from encountering primary priority content that is harmful to children which the provider identifies on the service, except where:

(A) a term of service indicates (in whatever words) that the presence of that kind of primary priority content that is harmful to children is prohibited on the service, and
(B) that policy applies in relation to all users of the service.[122]

c) A duty to include provisions in the terms of service specifying—
   i) how children of any age are to be prevented from encountering primary priority content that is harmful to children (with each kind of primary priority content separately covered);
   ii) how children in age groups judged to be at risk of harm from priority content that is harmful to children (or from a particular kind of such content), as assessed by the provider of a service in the most recent children's risk assessment of the service,[123] are to be protected from encountering it, where they are not prevented from doing so (with each kind of priority content separately covered);
   iii) how children in age groups judged to be at risk of harm from non-designated content that is harmful to children (or from a particular kind of such content), as assessed by the provider of a service in the most recent children's risk assessment of the service,[124] are to be protected from encountering it, where they are not prevented from doing so.[125]
d) A duty to apply the provisions of the terms of service referred to in paragraph c) above consistently.[126]

e) If a provider takes or uses a measure designed to prevent access to the whole of the service or a part of the service by children under a certain age, a duty to—
   i) include provisions in the terms of service specifying details about the operation of the measure, and
   ii) apply those provisions consistently.
f) A duty to include provisions in the terms of service giving information about any proactive technology used by a service for the purpose of compliance with a duty set out in paragraph a) or b) above[127] (including the kind of technology, when it is used, and how it works).[128]
g) A duty to ensure that the provisions of the terms of service referred to in paragraphs c), e) and f) above[129] are clear and accessible.[130]

---

[120] Section 13(3) of the Act.
[121] Section 12(3) of the Act.
[122] Sections 12(4)-(6) of the Act.
[123] Section 13(3) of the Act.
[124] Section 13(3) of the Act.
[125] Section 12(9) of the Act.
[126] Section 12(10) of the Act.
[127] Those paragraphs refer to the provisions in sections 12(2) and 12(3) of the Act.
[128] Section 12(12) of the Act.
[129] Those paragraphs refer to the provisions in sections 12(9), 12(11) and 12(12) of the Act.
[130] Section 12(13) of the Act.

A13.20 So far as the above duties relate to non-designated content that is harmful to children, the relevant duty is to be taken to extend only to addressing risks of harm from the kinds of such content that have been identified in the most recent children's risk assessment (if any have been identified).[131]

A13.21 The duties set out in paragraphs A13.21a) and A13.21b)[132] apply across all areas of a service, including the way it is designed, operated and used as well as content present on the service, and (among other things) require the provider of a service to take or use measures in the following areas, if it is proportionate to do so—

a) regulatory compliance and risk management arrangements,
b) design of functionalities, algorithms and other features,
c) policies on terms of use,
d) policies on user access to the service or to particular content present on the service, including blocking users from accessing the service or particular content,
e) content moderation, including taking down content,
f) functionalities allowing for control over content that is encountered, especially by children,
g) user support measures, and
h) staff policies and practices.[133]

A13.22 Age verification or age estimation to identify who is or is not a child user or which age group a child user is in are examples of measures which (if not required by section 12(4) of the Act (see paragraph A13.21b)) may be taken or used (among others) for the purpose of compliance with a duty set out in paragraph A13.21a) or A13.21b).[134]

A13.23 Providers of Category 1 Services likely to be accessed by children are also subject to a duty to summarise in the terms of service the findings of the most recent children's risk assessment of a service (including as to levels of risk and as to nature, and severity, of potential harm to children).[135]

A13.24 In determining what is "proportionate" for the purposes of the safety duties set out above, the following factors, in particular, are relevant:

a) all the findings of the most recent children's risk assessment (including as to levels of risk and as to nature, and severity, of potential harm to children), and
b) the size and capacity of the provider of the service.[136]

Providers' judgements about whether content is harmful to children

A13.25 Section 192 of the Act sets out the approach to be taken by providers to judgements about the status of content in the following circumstances:

a) a system or process operated or used by a service provider for the purpose of compliance with relevant requirements,
b) a risk assessment required to be carried out by Part 3, or

[131] Section 13(2) of the Act.
[132] Those paragraphs refer to the provisions in sections 12(2) and 12(3) of the Act.
[133] Section 12(8) of the Act.
[134] Those paragraphs refer to the provisions in sections 12(2) and 12(3) of the Act.
[135] Section 12(14) of the Act.
[136] Section 13(1) of the Act.

c) an assessment required to be carried out by section 14, involves a judgement by a provider about whether content is content of a particular kind.[137]

A13.26 Such judgements are to be made on the basis of all relevant information that is reasonably available to a provider, where the following factors, in particular, are relevant:

a) the size and capacity of the provider, and
b) whether a judgement is made by human moderators, by means of automated systems or processes or by means of automated systems or processes together with human moderators.[138]

A13.27 In considering a provider's compliance with section 192 requirements, Ofcom may take into account whether providers' judgements follow the approaches set out in this section (including judgements made by means of automated systems or processes, alone or together with human moderators).[139]

Duties about content reporting and complaints procedures

A13.28 The duties about content reporting and complaints procedures for providers of U2U services are contained in sections 20 and 21 of the Act.

Duties about content reporting

A13.29 All providers of regulated U2U services are required to use systems and processes in the operation of their services which allow users and "affected persons" to easily report certain types of content, depending on the kind of service. For instance, such systems and processes must be put in place to enable users and affected persons to report "Content that is harmful to children, present on a part of a service that it is possible for children to access"[140] on all U2U services likely to be accessed by children.[141]

A13.30 For the purposes of the duties about content reporting and complaints procedures, an "affected person" means a person, other than a user of the service in question, who is in the United Kingdom and who is:

a) the subject of the content,
b) a member of a class or group of people with a certain characteristic targeted by the content,
c) a parent of, or other adult with responsibility for, a child who is a user of the service or is the subject of the content, or
d) an adult providing assistance in using the service to another adult who requires such assistance, where that other adult is a user of the service or is the subject of the content.[142]

A13.31 In applying the content reporting duty, the cross-cutting duties will also be relevant (see paragraphs A13.37-A13.43).

Duties about complaints procedures

A13.32 There are two main duties in respect of complaints procedures which apply in relation to all regulated U2U services. These are:

a) A duty to operate a complaints procedure, in relation to a service, that:

---

[137] Section 192(1) of the Act.
[138] Section 192(2) and (3) of the Act.
[139] Section 192(8) of the Act.
[140] Section 20(6) of the Act states that: "a provider is only entitled to conclude that it is not possible for children to access a service, or a part of it, if age verification or age estimation is used on the service with the result that children are not normally able to access the service or that part of it."
[141] sections 20(2) and (4) of the Act.
[142] Section 20(5) of the Act.

i) allows for relevant kinds of complaint to be made (as set out below),
ii) provides for appropriate action to be taken by the provider of the service in response to complaints of a relevant kind, and
iii) is easy to access, easy to use (including by children) and transparent.[143]

b) A duty to include provisions in the terms of service which are easily accessible (including to children) specifying the policies and processes that govern the handling and resolution of complaints of a relevant kind.[144]

A13.33 For all regulated U2U services, a relevant complaint will be:

a) complaints by users and affected persons about content present on a service which they consider to be illegal content;
b) complaints by users and affected persons (see definition at paragraph A13.32) if they consider that the provider is not complying with: their Illegal content duties, the content reporting duty (paragraph A13.31), or the freedom of expression and privacy duties (paragraphs A13.37-A13.38);
c) complaints by a user who has generated, uploaded or shared content on a service if that content is taken down on the basis that it is illegal content;
d) complaints by a user of a service if the provider has given a warning to the user, suspended or banned the user from using the service, or in any other way restricted the user's ability to use the service, as a result of content generated, uploaded or shared by the user which the provider considers to be illegal content;
e) complaints by a user who has generated, uploaded or shared content on a service if—
i) the use of proactive technology on the service results in that content being taken down or access to it being restricted, or given a lower priority or otherwise becoming less likely to be encountered by other users, and
ii) the user considers that the proactive technology has been used in a way not contemplated by, or in breach of, the terms of service (for example, by affecting content not of a kind specified in the terms of service as a kind of content in relation to which the technology would operate).[145]

A13.34 For services that are likely to be accessed by children the following will also be a relevant complaint:

a) complaints by users and affected persons about content, present on a part of a service that it is possible for children to access, which they consider to be content that is harmful to children;
b) complaints by users and affected persons if they consider that the provider is not complying with the children's safety duties (see paragraphs A13.21-A13.26);
c) complaints by a user who has generated, uploaded or shared content on a service if that content is taken down, or access to it is restricted, on the basis that it is content that is harmful to children;
d) complaints by a user of a service if the provider has given a warning to the user, suspended or banned the user from using the service, or in any other way restricted the user's ability to use the service, as a result of content generated, uploaded or shared by the user which the provider considers to be content that is harmful to children;
e) complaints by a user who is unable to access content because measures used to comply with the children's safety duties have resulted in an incorrect assessment of the user's age.[146]

---

[143] Section 21(2) of the Act.
[144] Section 21(3) of the Act.
[145] Sections 21(4)(a)-(e) of the Act.
[146] Section 21(5) of the Act.

Cross-cutting duties

A13.35 The Act also creates so-called "cross-cutting duties", which apply to regulated U2U services in relation to the performance of *other* duties under the Act. For instance, the freedom of expression and privacy duties are concerned with how "safety measures and policies" are introduced in relation to a regulated U2U service. These "safety measures and policies" refer to any measures or policies designed to secure compliance with the safety duties relating to the protection of children (section 12 of the Act, paragraphs A13.21-A13.26), and the duties about content reporting (paragraphs A13.31-A13.33) and complaints procedures (section 21 of the Act, paragraphs A13.34-A13.36), as well as other duties in relation to illegal content (section 10 of the Act), and user empowerment (section 15 of the Act).[147]

A13.36 In a similar vein, the record-keeping and review duties apply to the performance of the children's risk assessment duties (section 11 of the Act, paragraphs A13.16-A13.20); and other "relevant duties", including the children's safety duties (section 12, paragraphs A13.21-A13.26), and content reporting (section 20 of the Act, paragraphs A13.31-A13.33) and complaints procedures (section 21 of the Act, paragraphs A13.34-A13.36).[148]

Duties about freedom of expression and privacy

A13.37 All regulated U2U services will have the following duties when deciding on, and implementing, "safety measures and policies":

a) a duty to have particular regard to the importance of protecting users' right to freedom of expression within the law;[149] and
b) a duty to have particular regard to the importance of protecting users from a breach of any statutory provision or rule of law concerning privacy that is relevant to the use or operation of a U2U service (including, but not limited to, any such provision or rule concerning the processing of personal data).[150]

A13.38 In addition, regulated U2U services which are also Category 1 services will have the following duties (although these are not covered in this consultation and will be covered at a later stage of Ofcom's work):

a) A duty to carry out impact assessments:
   i) when deciding on safety measures and policies, to determine the impact that such measures or policies have on (i) users' right to freedom of expression within the law, and (ii) the privacy of users;[151] and
   ii) to determine the impact that any adopted safety measures and policies have on (i) users' right to freedom of expression within the law, and (ii) the privacy of users.[152]

   An impact assessment relating to a service must include a section which considers the impact of the safety measures and policies on the availability and treatment on the service of content which is news publisher content or journalistic content in relation to the service.

b) A duty to keep an impact assessment up to date, and to publish impact assessments.[153]

[147] Section 22 of the Act.
[148] Section 23 of the Act.
[149] Section 22(2) of the Act.
[150] Section 22(3) of the Act.
[151] Section 22(4)(a) of the Act.
[152] Section 22(4)(b) of the Act.
[153] Section 22(6) of the Act.

c) A duty to specify in a publicly available statement the positive steps that the provider has taken in response to an impact assessment to— (i) protect users' right to freedom of expression within the law, and (ii) protect the privacy of users.[154]

Record-keeping and review duties

A13.39 All regulated U2U services will have the following duties:

a) A duty to make and keep a written record, in an easily understandable form, of every children's risk assessment under section 11 (see paragraphs A13.16 to A13.20).[155] [156]

b) A duty to make and keep a written record of any measures taken or in use to comply with a relevant duty which—
   i) are described in a Code of Practice and recommended for the purpose of compliance with the duty in question, and
   ii) apply in relation to the provider and the service in question. Such measures are referred to as "applicable measures in a Code of Practice".[157]

c) If alternative measures (see paragraph A13.42 below) have been taken or are in use to comply with a relevant duty, a duty to make and keep a written record containing the following information—
   i) the applicable measures in a Code of Practice that have not been taken or are not in use,
   ii) the alternative measures that have been taken or are in use,
   iii) how those alternative measures amount to compliance with the duty in question, and
   iv) how the provider has had regard to the importance of protecting the right of users to freedom of expression within the law, and protecting the privacy of users in taking or using alternative measures.[158] [159]

d) A duty to review compliance with the relevant duties in relation to a service—
   v) regularly, and
   vi) as soon as reasonably practicable after making any significant change to any aspect of the design or operation of the service.[160]

A13.40 'Alternative measures' means measures other than measures which are (in relation to the provider and the service in question) applicable measures in a Code of Practice.[161] If alternative measures have been taken or are in use to comply with any of the safety duties relating to the protection of children set out in section 12(2) or (3)) of the Act (see paragraphs A13.21a and A13.21b), [162] these records must also indicate whether such measures have been taken or are in use in every area listed at paragraph A13.23[163] in relation to which there are applicable measures in a Code of Practice (see paragraphs A13.80-A13.104).[164]

---

[154] Section 22(7) of the Act.
[155] Or section 9 (Illegal Content Risk assessment duties)
[156] Section 23(2) of the Act.
[157] Section 23(3) of the Act.
[158] Section 23(4) of the Act.
[159] Section 49(5) of the Act.
[160] Section 23(6) of the Act.
[161] Section 23(11) of the Act.
[162] or with the safety duties about illegal content (the Act, section 10(2) or 10(3)).
[163] These are the areas listed in section 12(8) of the Act.
[164] Similarly, if alternative measures have been taken or are in use to comply with the safety duties about illegal content in section 10(2) or 10(3) of the Act, these records must also indicate whether such measures have been taken or are in use in every area listed in section 10(4) of the Act in relation to which there are applicable measures in a Code.

A13.41 We consulted on our draft guidance on record keeping and review as part of our Illegal Harms Consultation.[165] As relevant, we have suggested amendments to our draft guidance to reflect the duties as relevant to services likely to be accessed by children. See Volume 4, Section 12, from paragraph 12.77.

Duties of providers of search services likely to be accessed by children

A13.42 Providers of regulated search services are also given specific duties under the Act in relation to the protection of children. These include: "children's risk assessment duties";[166] and "safety duties protecting children".[167]

A13.43 Providers of regulated search services are also subject to additional duties which are relevant to the protection of children, but also apply to other types of content and in respect of other regulatory requirements as set out under the Act. These are:

a) "Duties about content reporting and complaints procedures", which include:
   i) The "duty about content reporting",[168] and
   ii) "Duties about complaints procedures";[169] and
b) the "Cross-cutting duties", which include:
   iii) "Duties about freedom of expression and privacy";[170] and
   iv) "Record-keeping and review duties".[171]

A13.44 These are set out in more detail below. Section 24 of the Act states that all providers of regulated search services must comply with these duties (and the other duties set out under section 24(2)).

Connection with the United Kingdom

A13.45 These duties only apply to:

a) the design, operation and use of the search engine in the United Kingdom, and
b) in the case of a duty that is expressed to apply in relation to users of a service, the design, operation and use of the search engine as it affects United Kingdom users of the service.[172]

Combined services

A13.46 Where a service is a combined service (i.e. providing both a regulated U2U and regulated search service), the duties applying to U2U services likely to be access by children will apply save for in relation to:

a) the search content of the service,
b) any other content that, following a search request, may be encountered as a result of subsequent interactions with internet services, or

[165] See Volume 3, Chapter 10 and Annex 6.
[166] Section 28 of the Act.
[167] Section 29 of the Act.
[168] Section 31 of the Act.
[169] Section 32 of the Act.
[170] Section 33 of the Act.
[171] Section 34 of the Act.
[172] Sections 25(2) and (3) of the Act.

c) anything relating to the design, operation or use of the search engine.[173]

A13.47 The duties that apply to regulated search services will apply only to the search content of the combined service.[174]

Safety duties for services likely to be accessed by children

Protection of children risk assessment duties

A13.48 Providers of regulated search services that are likely to be accessed by children have a duty to carry out a suitable and sufficient children's risk assessment[175] at the specific times set out in Schedule 3 to the Act.[176].

A13.49 A children's risk assessment means an assessment of the following matters, taking into account the risk profiles that relate to the services of that kind:[177]

a) the level of risk of children who are users of the service encountering search content of the following kinds—
   i) each kind of primary priority content that is harmful to children (with each kind separately assessed),
   ii) each kind of priority content that is harmful to children (with each kind separately assessed), and
   iii) non-designated content that is harmful to children,

   giving separate consideration to children in different age groups, and taking into account (in particular) risks presented by algorithms used by the service and the way that the service indexes, organises and presents search results;

b) the level of risk of children who are users of the service encountering search content that is harmful to children which particularly affects individuals with a certain characteristic or members of a certain group;

c) the extent to which the design of the service, in particular its functionalities, affects the level of risk of harm that might be suffered by children, identifying and assessing those functionalities that present higher levels of risk, including a functionality that makes suggestions relating to users' search requests (predictive search functionality);

---

[173] Section 8(2) of the Act.

[174] Section 25(1) of the Act.

[175] Section 28(2) of the Act.

[176] The deadline for completing the first risk assessment depends on the day on which a search service's provider starts its operations. In particular:

i. search services that are already in operation at the outset of this regime, must complete their first children's risk assessment within a period of three months from the day on which Ofcom's risk assessment guidance ('RAG') is published;

ii. new search services that start operations after the RAG is published must complete their first children's risk assessment within a period of three months from the day on which they begin their new services; and

iii. existing services that become search services (having previously provided a different type of service) after the RAG is published must complete their first children's risk assessment within a period of three months from the day on which their services become a search service. See Schedule 3 to the Act.

[177] Section 11(6) of the Act.

d) the different ways in which the service is used, including functionalities or other features of the service that affect how much children use the service, and the impact of such use on the level of risk of harm that might be suffered by children;
e) the nature, and severity, of the harm that might be suffered by children from the matters identified in accordance with paragraphs (a) to (d), giving separate consideration to children in different age groups;
f) how the design and operation of the service (including the business model, governance, use of proactive technology, measures to promote users' media literacy and safe use of the service, and other systems and processes) may reduce or increase the risks identified.

A13.50 The provider of a search service that is likely to be accessed by children must take appropriate steps to keep a children's risk assessment up to date, including when Ofcom makes a significant change to a relevant risk profile (see paragraphs A13.74-A13.76).[178]

A13.51 The provider of a search service that is likely to be accessed by children is under an obligation to carry out a further suitable and sufficient children's risk assessment, before making any significant changes to any aspect of a service's design or operation. This further children's risk assessment must relate to the impact of that proposed change.[179]

Safety duties relating to the protection of children

A13.52 Providers of regulated search services likely to be accessed by children have specific safety duties in relation to children's online safety as set out under section 29 of the Act. These duties extend to such parts of a service as it is possible for children to access.[180] [181] The duties are as follows:

a) A duty, in relation to a service, to take or use proportionate measures relating to the design or operation of the service to effectively—
   i) mitigate and manage the risks of harm to children in different age groups, as identified in the most recent children's risk assessment of the service (section 28(5)(e) of the Act), and
   ii) mitigate the impact of harm to children in different age groups presented by content that is harmful to children present on the service.[182]
b) A duty to operate a service using proportionate systems and processes designed to—
   i) minimise the risk of children of any age encountering search content that is primary priority content that is harmful to children[183] [184];
   ii) minimise the risk of children in age groups judged to be at risk of harm from other content that is harmful to children[185] (or from a particular kind of such content), as assessed by the provider of a service in the most recent children's risk assessment of the service,[186] encountering search content of that kind.[187]

---

[178] Section 28(3) of the Act.
[179] Section 28(4) of the Act.
[180] A provider is only entitled to conclude that it is not possible for children to access a service, or a part of it, if age verification or age estimation is used on the service with the result that children are not normally able to access the service or that part of it – see section 30(6) of the Act.
[181] Section 30(5) of the Act.
[182] Section 29(2) of the Act.
[183] With the harm arising by virtue of the nature of the content rather than the fact of its dissemination, see section 30(4) of the Act.
[184] Section 30(4) of the Act.
[185] With the harm arising by virtue of the nature of the content rather than the fact of its dissemination, see section 30(4) of the Act.
[186] Section 30(3) of the Act.
[187] Section 29(3) of the Act.

c) A duty to include provisions in a publicly available statement specifying how children are to be protected from search content of the following kinds—
   i) primary priority content that is harmful to children (with each kind of primary priority content separately covered),
   ii) priority content that is harmful to children (with each kind of priority content separately covered), and
   iii) non-designated content that is harmful to children.[188]
d) A duty to—
   i) include provisions in a publicly available statement giving information about any proactive technology (see paragraphs A13.96 to A13.100) used by a service for the purpose of compliance with a duty set out in paragraph a) or b) (including the kind of technology, when it is used, and how it works);[189] and
   ii) ensure that the provisions of that public statement are clear and accessible.[190]

A13.53 So far as the above duties relate to non-designated content that is harmful to children, the relevant duty is to be taken to extend only to addressing risks of harm from the kinds of such content that have been identified in the most recent children's risk assessment (if any have been identified).[191]

A13.54 The duties set out in paragraphs A13.54a) and A13.54b) apply across all areas of a service, including the way the search engine is designed, operated and used as well as search content of the service, and (among other things) require the provider of a service to take or use measures in the following areas, if it is proportionate to do so—

a) regulatory compliance and risk management arrangements,
b) design of functionalities, algorithms and other features relating to the search engine,
c) functionalities allowing for control over content that is encountered in search results, especially by children,
d) content prioritisation,
e) user support measures, and
f) staff policies and practices.[192]

A13.55 Providers of Category 2A services likely to be accessed by children are additionally subject to a duty to summarise in a publicly available statement the findings of the most recent children's risk assessment of a service (including as to levels of risk and as to nature, and severity, of potential harm to children).[193]

A13.56 In determining what is 'proportionate' for the purposes of the safety duties for search services likely to be accessed by children, the following factors, in particular, are relevant:

a) all the findings of the most recent children's risk assessment (including as to levels of risk and as to nature, and severity, of potential harm to children), and
b) the size and capacity of the provider of a service.[194]

---

[188] Section 29(5) of the Act.
[189] Section 29(7) of the Act.
[190] Section 29(8) of the Act.
[191] Section 30(2) of the Act.
[192] Section 29(4) of the Act.
[193] Section 29(9) of the Act.
[194] Section 30(1) of the Act.

Duties about content reporting and complaints procedures

Duty about content reporting

A13.57 All providers of regulated search services are required to operate a service using systems and processes that allow users and 'affected persons' to easily report certain types of search content, depending on the type of service.[195] For instance, such systems and processes must be put in place to enable users and affected persons to report 'illegal content' on *all* of the search service.[196]

A13.58 For services that are likely to be accessed by children, the duty also applies in respect of content that is harmful to children.[197]

A13.59 For the purposes of the duties about content reporting and complaints procedures, an "affected person" has the same definition as for U2U services (see paragraph A13.32 above).

Duties about complaints procedures

A13.60 There are two main duties in respect of complaints procedures which apply in relation to all regulated search services. These are as follows:

a) A duty to operate a complaints procedure in relation to a service that—
   i) allows for relevant kinds of complaint to be made (as set out below),
   ii) provides for appropriate action to be taken by the provider of the service in response to complaints of a relevant kind, and
   iii) is easy to access, easy to use (including by children) and transparent.[198]
b) A duty to make the policies and processes that govern the handling and resolution of complaints of a relevant kind publicly available and easily accessible (including to children).[199]

A13.61 Relevant complaints in relation to a regulated search service are:

a) complaints by users and affected persons (see paragraphs A13.61 and A13.32 above) about search content which they consider to be illegal content;
b) complaints by users and affected persons if they consider that the provider is not complying with their illegal content duties, content reporting duties (paragraphs A13.59-A13.61)), or freedom of expression and privacy (see paragraph A13.69);
c) complaints by an interested person if the provider of a search service takes or uses measures in order to comply with their illegal content safety duties that result in content relating to that interested person no longer appearing in search results or being given a lower priority in search results;
d) complaints by an interested person if—
   i) the use of proactive technology (see paragraphs A13.96-A13.100 below) on a search service results in content relating to that interested person no longer appearing in search results or being given a lower priority in search results; and
   ii) the interested person considers that the proactive technology has been used in a way not contemplated by, or in breach of, the provider's policies on its use (for example, by affecting content not of a kind specified in those policies as a kind of content in relation to which the technology would operate).[200]

---

[195] Section 31(2) of the Act.
[196] Section 31(3) of the Act.
[197] Section 31(4) of the Act.
[198] Section 32(2)(a)-(c) of the Act.
[199] Section 32(3) of the Act.
[200] Sections 32(4)(a)-(d) of the Act.

A13.62 For services that are likely to be accessed by children the following will also be a relevant complaint:

a) complaints by users and affected persons about search content which they consider to be content that is harmful to children;
b) complaints by users and affected persons if they consider that the provider is not complying with the children's safety duties (see paragraphs A13.54-A13.58 above);
c) complaints by an interested person if the provider of a search service takes or uses measures in order to comply with the children's safety duties that result in content relating to that interested person no longer appearing in search results or being given a lower priority in search results;
d) complaints by a user who is unable to access content because measures used to comply with the children's safety duties described in paragraphs A13.54a) and A13.54b) above[201] have resulted in an incorrect assessment of the user's age.

A13.63 For the purposes of the duties about complaints procedures for regulated search services, an 'interested person' means a person that is responsible for a website or database capable of being searched by the search engine, provided that—

a) in the case of an individual, the individual is in the United Kingdom;
b) in the case of an entity, the entity is incorporated or formed under the law of any part of the United Kingdom.[202]

Cross-cutting duties

A13.64 The Act also creates 'cross-cutting' duties which apply to regulated search services in relation to the performance of other duties under the Act. For instance, the duties about freedom of expression and privacy are concerned with how "safety measures and policies" are introduced in relation to a regulated search service. These "safety measures and policies" refer to any measures or policies designed to secure compliance with the safety duties relating to the protection of children (section 29, paragraphs A13.54-A13.58), and the duties about content reporting (section 31, paragraphs A13.59-A13.61) and complaints procedures (section 32, paragraphs A13.62-A13.64), as well as other duties in relation to illegal content (section 27).[203]

A13.65 In a similar vein, the record-keeping and review duties apply to the performance of the risk assessment duties under section 28 and other "relevant duties", including the children's safety duties , and content reporting and complaints procedures.[204]

A13.66 The cross-cutting duties for regulated search services are set out in sections 33 and 34 of the Act.

Duties about freedom of expression and privacy

A13.67 All regulated search services will have the following duties when deciding on, and implementing, "safety measures and policies" (see above):

a) a duty to have particular regard to the importance of protecting the rights of users and interested persons to freedom of expression within the law;[205] and

[201] These are the duties in sections 29(2) and 29(3) of the Act.
[202] Sections 32(6) and 227(7) of the Act.
[203] Section 33 of the Act.
[204] Section 34 of the Act.
[205] Section 33(2) of the Act.

b) a duty to have particular regard to the importance of protecting users from a breach of any statutory provision or rule of law concerning privacy that is relevant to the use or operation of a search service (including, but not limited to, any such provision or rule concerning the processing of personal data).[206]

Record-keeping and review duties

A13.68 All regulated search services will have the following duties:

a) A duty to make and keep a written record, in an easily understandable form, of every risk assessment made under section 28 or section 26.[207]
b) A duty to make and keep a written record of any measures taken or in use to comply with a relevant duty which—
   i) are described in a Code of Practice and recommended for the purpose of compliance with the duty in question, and
   ii) apply in relation to the provider and the service in question. In this section such measures are referred to as "applicable measures in a code of practice".[208]
c) If alternative measures have been taken or are in use to comply with a relevant duty**,** a duty to make and keep a written record containing the following information—
   iii) the applicable measures in a Code of Practice that have not been taken or are not in use,
   iv) the alternative measures that have been taken or are in use,
   v) how those alternative measures amount to compliance with the duty in question, and
   vi) how the provider has had regard to the importance of protecting the right of users and interested persons to freedom of expression within the law, and protecting the privacy of users in taking or using alternative measures (i.e. under section 49(5)).[209]
d) If alternative measures have been taken or are in use to comply with the Safety duties about the protection of children (specifically sections 29(2) or (3), this record must also indicate whether such measures have been taken or are in use in every area listed at section 29(4) (or section 27(4) of the Act as the case may be) in relation to which there are applicable measures in a Code of Practice.[210]
e) A duty to review compliance with the relevant duties in relation to a service— regularly, and as soon as reasonably practicable after making any significant change to any aspect of the design or operation of the service.[211]
f) Ofcom may provide that particular descriptions of providers of search services are exempt from any or all of the record-keeping and review duties, and must publish details of any exemption.[212]

Ofcom's duties in relation to the protection of children

A13.69 The Act gives specific duties to Ofcom in relation to the protection of children. These are set out below.

[206] Section 33(3) of the Act.
[207] Section 34(2) of the Act.
[208] Section 34(3) of the Act.
[209] Section 34(4)(a)-(d) of the Act.
[210] Section 34(5) of the Act.
[211] Sections 34(6)(a) and(b) of the Act.
[212] Section 34(7) of the Act.

Ofcom sector risk assessment

A13.70 Ofcom is under a duty to carry out a risk assessment to identify and assess the risks of harm to children in the United Kingdom, in different age groups, presented by content that is harmful to children.[213] [214]

A13.71 Ofcom's risk assessment must, among other things, identify the characteristics of U2U and search services (which include functionalities, user base, business model and governance, and other systems and processes) that are relevant to the risks of harm and assess the impact of these characteristics on the risks of harm.[215]

Register of Risks and Risk Profiles

A13.72 Ofcom must prepare and publish a register of risks that reflects the findings of its sector risk assessment (the 'Register of Risks'). The Register of Risks must be prepared as soon as reasonably practicable after completion of the risk assessment.[216]

A13.73 Further to the Register of Risks, after completing its risk assessments, Ofcom must prepare and publish Risk Profiles for U2U services and search services that relate to each risk of harm, as applicable (the 'Risk Profiles'). In preparing the Risk Profiles, Ofcom can group U2U services and search services as appropriate and having regard to (i) the characteristics of the services and (ii) the risk levels and other matters identified in the risk assessment.[217]

A13.74 Ofcom must review and revise the risk assessments and the Risk Profiles from time to time to keep them up to date.[218]

Risk assessment guidance for services

A13.75 Ofcom must prepare and publish guidance to help U2U services and search services comply with their duties to prepare children's risk assessments under sections 11 and 28 respectively (the 'Children's Risk Assessment Guidance')).[219]

A13.76 Ofcom must prepare the Children's Risk Assessment Guidance as soon as reasonably practicable after having published the risk profiles relating to the risks of harm to children.[220]

A13.77 Ofcom must revise and publish updated Children's Risk Assessment Guidance when it carries out a new risk assessment and/or revises the risk profiles.[221]

"Protection of children" Codes for U2U and search

Ofcom's duty to prepare and issue Codes of Practice in relation to the protection of children

A13.78 Ofcom must issue Codes for regulated U2U and search services containing measures recommended for the purposes of compliance with certain duties including:

---

[213] Section 98(1)(c) of the Act.
[214] Ofcom has discretion in relation to whether to combine the risk assessments with the risk assessments relating to illegal content it is required to carry out under section 98(1)(a) and (b) of the Act. Ofcom may assess regulated U2U services and regulated search services separately or together. Section 98(3) of the Act.
[215] Sections 98(2) and (11) of the Act.
[216] Section 98(4) of the Act.
[217] Sections 98(5)-(7) of the Act.
[218] Section 98(8) of the Act.
[219] Sections 99(3) and (6) of the Act.
[220] Section 99(3) of the Act.
[221] Section 99(5) of the Act.

a) the protection of children safety duties in sections 12 and 29;[222]
b) the content reporting duties in sections 20 and 31;[223]
c) the complaints procedure duties in sections 21 and 32.[224]

A13.79 Schedule 4 to the Act sets out general principles and online safety objectives which the Codes must follow, as well as what content must be included. These are set out below.

General principles

A13.80 In preparing a draft Code, Ofcom must consider the appropriateness of provisions of the Code to different kinds and sizes of U2U and search services, and to providers of differing sizes and capacities (paragraph 1 of Schedule 4). It must also have regard to the following principles:

a) providers of U2U and search services must be able to understand which provisions of the Code of Practice apply in relation to a particular service they provide;
b) the measures described in the Code of Practice must be sufficiently clear, and at a sufficiently detailed level, that providers understand what those measures entail in practice;
c) the measures described in the Code of Practice must be proportionate and technically feasible: measures that are proportionate or technically feasible for providers of a certain size or capacity, or for services of a certain kind or size, may not be proportionate or technically feasible for providers of a different size or capacity or for services of a different kind or size;
d) the measures described in the Code of Practice that apply in relation to U2U and search service providers of various kinds and sizes must be proportionate to Ofcom's assessment of the risk of harm presented by services of that kind or size (see paragraph A13.73 above).[225]

Online safety objectives

A13.81 Ofcom must ensure that measures described in the Codes are compatible with the pursuit of the online safety objectives.[226]

A13.82 For U2U services, these are:

a) That a service should be designed and operated in such a way that—
    i) the systems and processes for regulatory compliance and risk management are effective and proportionate to the kind and size of service,
    ii) the systems and processes are appropriate to deal with the number of users of the service and its user base,
    iii) UK users (including children) are made aware of, and can understand, the terms of service,
    iv) there are adequate systems and processes to support United Kingdom users,
    v) (in the case of a Category 1 service) users are offered options to increase their control over the content they encounter and the users they interact with,
    vi) the service provides a higher standard of protection for children than for adults,
    vii) the different needs of children at different ages are taken into account,
    viii) there are adequate controls over access to the service by adults, and
    ix) there are adequate controls over access to, and use of, the service by children, taking into account use of the service by, and impact on, children in different age groups; and

[222] Sections 41(3) and 41(10)(b) of the Act.
[223] Sections 41(3) and 41(10)(f) of the Act.
[224] Sections 41(3) and 41(10)(g) of the Act.
[225] The Act, Schedule 4, subparagraphs 2(a)-(d).
[226] The Act, Schedule 4, paragraph 3.

b) that a service should be designed and operated so as to protect individual UK users from harm, including with regard to—
   i) algorithms used by the service,
   ii) functionalities of the service, and
   iii) other features relating to the operation of the service.[227]

A13.83 For search services, these are:

a) That a service should be designed and operated in such a way that—
   i) the systems and processes for regulatory compliance and risk management are effective and proportionate to the kind and size of service,
   ii) the systems and processes are appropriate to deal with the number of users of the service and its user base,
   iii) United Kingdom users (including children) are made aware of, and can understand, the publicly available statement referred to in relation to the safety duties protecting children in section 29,[228]
   iv) there are adequate systems and processes to support United Kingdom users,
   v) the service provides a higher standard of protection for children than for adults, and
   vi) the different needs of children at different ages are taken into account; and
b) that a service should be assessed to understand its use by, and impact on, children in different age groups; and
c) that a search engine should be designed and operated so as to protect individuals in the United Kingdom who are users of the service from harm, including with regard to—
   i) algorithms used by the search engine,
   ii) functionalities relating to searches (such as a predictive search functionality), and
   iii) the indexing, organisation and presentation of search results. [229]

A13.84 For combined services:

a) the online safety objectives that apply to U2U services (paragraph A13.84 above) do not apply in relation to the search engine;
b) the online safety objectives that apply to search services apply in relation to the search engine (and, accordingly, in this context, references to a search service are to be read as references to the search engine);
c) the reference to a publicly available statement includes a reference to provisions of the terms of service which relate to the search engine. [230]

A13.85 The Secretary of State may amend these objectives by way of regulations.[231]

Content of Codes of Practice

A13.86 The Act also sets out what type of measures must be included in the content of the Codes, and the principles in light of which such measures should be designed. Such measures may only relate to the design or operation of the relevant service in the United Kingdom, or as it affects United Kingdom users of the service. In particular:

a) Codes that describe measures recommended for the purpose of compliance with the Safety Duties for providers of U2U services (i.e. in relation to taking proportionate measures relating to

[227] The Act, Schedule 4, paragraph 4.
[228] This provision also applies to the statement relating to the illegal content safety duties referred to in section 27 of the Act.
[229] Schedule 4, paragraph (5)(a)-(c) of the Act.
[230] Schedule 4, paragraph 6(a)-(c) of the Act.
[231] Schedule 4, paragraph 7 of the Act.

the design or operation of the service, or to operate a service using proportionate systems and processes),[232] must include measures in each of the areas of a service listed at paragraph A13.23.[233] [234]

b) Codes that describe measures recommended for the purpose of compliance with the Safety Duties for providers of search services set out at paragraphs A13.54a) and A13.54b) (i.e. in relation to taking proportionate measures relating to the design or operation of the service, or to operate a service using proportionate systems and processes)[235] must include measures in each of the areas of a service listed at paragraph A13.56 above.[236] [237]

A13.87 Any measures described in a Code of Practice which are recommended for the purpose of compliance with any of the relevant duties must be designed in the light of the following principles:

a) the importance of protecting the right of users and (in the case of search services or combined services) interested persons to freedom of expression within the law, and

b) the importance of protecting the privacy of users.[238]

A13.88 Where appropriate, such measures must also incorporate safeguards for the protection of the matters mentioned in those principles.

Age assurance

A13.89 In deciding whether to recommend the use of age assurance, or which kinds of age assurance to recommend, in a code of practice as a measure recommended for the purpose of compliance with any of the duties set out in paragraphs A13.21a) or A13.21b) [239] (these apply to U2U services) or paragraphs A13.54a) or A13.54b) [240] (these apply to search services) Ofcom must, in addition to the general principles set out above,[241] have regard to the following:

a) the principle that age assurance should be effective at correctly identifying the age or age-range of users;

b) relevant standards set out in the latest version of the code of practice under section 123 of the Data Protection Act 2018 (age-appropriate design code);

c) the need to strike the right balance between:

   i) the levels of risk and the nature, and severity, of potential harm to children which the age assurance is designed to guard against, and

   ii) protecting the right of users and (in the case of search services or the search engine of combined services) interested persons to freedom of expression within the law;

d) the principle that more effective kinds of age assurance should be used to deal with higher levels of risk of harm to children;

---

[232] These are the measures in sections 12(2) and 12(3) of the Act.
[233] These are the areas listed in section 12(8) of the Act.
[234] Schedule 4, paragraph 9(2) of the Act.
[235] These are the measures in sections 29(2) and 29(3) of the Act.
[236] These are the measures in section 29(4) of the Act.
[237] Schedule 4, paragraph 9(3) of the Act.
[238] This refers to protecting the privacy of users from a breach of any statutory provision or rule of law concerning privacy that is relevant to the use or operation of a U2U or search service (including any provisions concerning the processing of personal data), Schedule 4, paragraph 10(3).
[239] These are the duties in section 12(2) or (3) of the Act.
[240] These are the duties in section 29(2) or (3) of the Act.
[241] Schedule 4, paragraph 12(1) of the Act.

e) the principle that age assurance should be easy to use, including by children of different ages and with different needs;
f) the principle that age assurance should work effectively for all users regardless of their characteristics or whether they are members of a certain group;
g) the principle of interoperability between different kinds of age assurance.[242]

A13.90 If a code of practice does recommend age assurance for the purpose of complying with the duties set out paragraphs A13.21a) or A13.21b)[243] (these apply to U2U services) then it must also describe measures for the purpose of complying with the following duties:

a) the duties regarding the inclusion of clear information in the terms of service described in paragraphs A13.21c)-f)[244]; and
b) the duties regarding complaints about age assurance described in paragraph A13.36. [245]

A13.91 If a code of practice does recommend age assurance for the purpose of complying with the duties set out in paragraphs A13.54a) or A13.54b)[246] (these apply to search services) then it must also describe measures for the purpose of complying with the following duties:

a) the duties regarding the inclusion of clear information in the publicly available statement described in paragraphs A13.54c) and d)[247]; and
b) the duties regarding complaints about age assurance.[248]

A13.92 A provider of a U2U service likely to be accessed by children is required to use age verification or age estimation (or both) to prevent children of any age from encountering primary priority content that is harmful to children which the provider identifies on the service.[249] The age verification or age estimation must be of such a kind, and used in such a way, that it is highly effective at correctly determining whether or not a particular user is a child.[250] Part 5 of the Act imposes specific duties on service providers that display or publish pornographic content on their online services. Further detail is set out below.

A13.93 The Act makes clear that a code of practice may:

a) refer to industry or technical standards for age assurance (where they exist); and/or
b) elaborate on the principles mentioned in paragraphs (a) and (c) to (g) of paragraph A13.91.[251]

Proactive technology

A13.94 If Ofcom considers it appropriate to do so, and in accordance with the general principles set out at paragraphs 1 and 2 of Schedule 4 and the principles set out at paragraph 10(2) of Schedule 4, it may include in a Code of Practice a measure describing the use of a kind of technology. However, there are constraints on Ofcom's power to include a measure describing the use of "proactive technology" (a "proactive technology measure"). Section 231 defines "proactive technology" as consisting of three types of technology: content identification technology, user profiling technology, and behaviour identification technology (subject to certain exceptions). These are explained in greater detail below.

---

[242] Schedule 4, paragraphs 12(1) and (2) of the Act.
[243] These are the duties in section 12(2) or (3) of the Act.
[244] These are the duties in sections 12(9), 12(11) and 12(13) of the Act.
[245] These are the duties in sections 21(2) and 21(3) of the Act.
[246] These are the duties in sections 29(2) or (3) of the Act.
[247] These are the duties in sections 29(5) and (8) of the Act.
[248] These are the duties in sections 32(2) and (3) of the Act.
[249] Sections 12(3)(a) and 12(4) of the Act.
[250] Section 12(6) of the Act.
[251] The Act, Schedule 4, paragraph 12(8).

A13.95 Content identification technology refers to technology, such as algorithms, keyword matching, image matching or image classification, which analyses content to assess whether it is content of a particular kind (for example, illegal content). Content identification technology is not regarded as proactive technology if it is used in response to a report from a user or other person about particular content.

A13.96 User profiling technology means technology which analyses (any or all of) relevant content (as defined in section 231(8)), user data, or metadata relating to relevant content or user data, for the purposes of building a profile of a user to assess characteristics such as age. However, technology which analyses data specifically provided by a user for the purposes of the provider verifying or estimating the user's age in order to decide whether to allow the user to access a service (or part of a service) or particular content, but which does not analyse any other data or content, is not regarded as user profiling technology.

A13.97 Behaviour identification technology means technology which analyses (any or all of) relevant content (as defined in section 231(8)), user data, or metadata relating to relevant content or user data, to assess a user's online behaviour or patterns of online behaviour (for example, to assess whether a user may be involved in, or be the victim of, illegal activity). But behaviour identification technology is not regarded as proactive technology if it is used in response to concerns identified by another person or an automated tool about a particular user.

A13.98 Ofcom has power to include a proactive technology measure in a Code of Practice for the purpose of compliance with the safety duties in relation to the protection of children described in paragraphs A13.21a) or A13.21b)[252] (these apply to U2U services) or paragraphs A13.54a) or A13.54b) [253] (these apply to search services).[254] However, that power is subject to the following constraints:

a) A proactive technology measure may not recommend the use of technology which operates (or may operate) by analysing user-generated content communicated privately, or metadata relating to such content.[255]
b) A proactive technology measure may be included in a Code of Practice in relation to services of a particular kind or size only if Ofcom is satisfied that the use of the technology by such services would be proportionate to the risk of harm that the measure is designed to safeguard against (taking into account, in particular, Ofcom's risk profile relating to such services published under section 98).[256]
c) In deciding whether to include a proactive technology measure in a Code of Practice, Ofcom must have regard to the degree of accuracy, effectiveness and lack of bias achieved by the technology in question. Ofcom may also refer in the Code of Practice to existing industry or technical standards for the technology (where they exist), or set out principles in the Code of

[252] These are the duties in section 12(2) or (3) of the Act.
[253] These are the duties in section 29(2) or (3) of the Act.
[254] Paragraph 13(3) of Schedule 4. A proactive technology measure may also be recommended for the purpose of compliance with the illegal content safety duties set out in section 10(2) or (3) of the Act (in relation to U2U services) or section 27(2) or (3) of the Act (in relation to search services), or for the purpose of compliance with the fraudulent advertising duties set out in section 38(1) or 39(1) of the Act.
[255] See paragraph 13(4) of Schedule 4. For factors which Ofcom must particularly consider when deciding whether content is communicated "publicly" or "privately" by means of a U2U service for these purposes, see section 232.
[256] See paragraph 13(5) of Schedule 4.

Practice designed to ensure that the technology or its use is (so far as possible) accurate, effective and free of bias.[257]

Relationship between provider duties and Ofcom's Codes of Practice

A13.99 Providers of a regulated U2U or search service who take or use the measures described in a Code of Practice which are recommended for the purpose of complying with a relevant duty will be treated as having complied with that relevant duty.[258] Further, providers who take or use the relevant recommended measures that incorporate safeguards to protect users' rights to freedom of expression within the law, and to protect the privacy of users, respectively, will be treated as having complied with the freedom of expression and privacy duties set out in paragraph A13.39 (for U2U services)[259] and paragraph A13.69 (for search services).[260] [261]

A13.100 Where a provider adopts an alternative measure to those described in a Code of Practice in order to comply with a relevant duty, it must have particular regard to the importance of protecting the right of users and (in the case of search services) interested persons to freedom of expression within the law, and protecting the privacy of users.[262]

A13.101 When assessing whether a provider of a service that has adopted alternative measures is compliant with a duty to protect children's online safety, Ofcom must consider the extent to which the alternative measures taken or in use by the provider extend across all areas of a service listed in sections 12(8) or 29(4), and, where appropriate, incorporate safeguards for the protection of the right of users and (in the case of search services) interested persons to freedom of expression within the law, and protection of the privacy of users.[263]

Effect of the Codes of Practice

A13.102 Failure to comply with a provision of a Code of Practice does not in itself make the provider liable to legal proceedings in a court or tribunal,[264] although the Code will be admissible in evidence in legal proceedings,[265] and any such court or tribunal must take a provision of the Code into account when determining a question which is relevant to that provision, as long as the question relates to a time when the provision was in force.[266] Similarly, Ofcom must take into account a provision of a Code of Practice when determining a question which is relevant to that provision, as long as the question relates to a time when the provision was in force.[267]

---

[257] See paragraph 13(6) of Schedule 4. This requirement does not apply to proactive technology which is a kind of age verification or age estimation technology: see paragraph 13(7) of Schedule 4.
[258] Section 49(1) of the Act.
[259] These are the duties in sections 22(2) and (3) of the Act.
[260] These are the duties in sections 33(2) and (3) of the Act.
[261] Section 49(2)-(3) of the Act.
[262] Section 49(5) of the Act.
[263] Section 49(6) of the Act.
[264] Section 50(1) of the Act.
[265] Section 50(2) of the Act.
[266] Section 50(3) of the Act.
[267] Section 50(4) of the Act.

Children's Access Assessments Guidance

A13.103 Ofcom is required to issue guidance for U2U and search services to assist with determining whether their services are likely to be accessed by children (i.e. completing the children's access assessment).[268]

Guidance on Content Harmful to Children

A13.104 Ofcom must produce guidance which gives examples of content that Ofcom considers to be (or not to be) primary priority and priority content that is harmful to children.[269]

Record keeping Guidance

A13.105 Ofcom must produce guidance for providers of regulated U2U and search services to assist them in complying with their record-keeping and review duties (sections 23 (U2U) and 34 (search)) – paragraphs A13.41-A13.43 (U2U), and A13.70 (search) above.[270] Ofcom produced a draft of this guidance for consultation on 9 November 2023,[271]and have proposed some minor updates as part of this consultation – see Volume 4, paragraph 12.77 onwards.

Part 5 Guidance

A13.106 The 'Part 5' duties apply where pornographic content is published or displayed by a provider of an internet service (or on behalf of such a provider) on that internet service. These duties include a requirement for service providers to implement highly effective age assurance to ensure that children are not normally able to encounter pornographic content displayed on their service (section 81(2) and (3)) of the Act. As set out in section 79(6) of the Act, pornographic content will be treated as published or displayed on a service where the pornographic content:

i) Is only visible or audible to users as a result of interacting with content that is blurred or obscured (for example, by clicking on the content) where pornographic content is present on the service;
ii) is embedded on the service; and
iii) Is generated on the service by means of an automated tool or algorithm in response to a prompt by a user and is only visible or audible to that user, such as a generative artificial intelligence tool ('Gen AI').

A13.1 Ofcom must produce guidance for providers of internet services which fall within scope of Part 5 to help them comply with the duties outlined above under section 82(1). To that end, Ofcom must include examples of the kinds and uses of age assurance that are, or are not, highly effective at determining whether or not a user is a child. The guidance must also set out the principles that Ofcom proposes to apply when determining if a service provider has

[268] Section 52(3)(b) of the Act.
[269] Section 53 of the Act.
[270] Section 52(3) of the Act.
[271] https://www.ofcom.org.uk/consultations-and-statements/category-1/protecting-people-from-illegal-content-online.

complied with the duties and where we are likely to consider that they have not (section 82(2)).

A13.2 We have sought to ensure a consistent approach to highly effective age assurance across our draft Part 5 guidance and the draft guidance on HEAA published as Annex 10 of this consultation to ensure consistency so that service providers in scope of both Part 5 and Part 3 are clear what they need to do to prevent children from encountering the most harmful forms of content.

A14. Impact assessments

A14.1 This annex outlines our equality impact assessment and Welsh language assessment.

Consultation questions

61. In relation to our equality impact assessment, do you agree that some of our proposals would have a positive impact on certain groups?

62. In relation to our Welsh language assessment, do you agree that our proposals are likely to have positive, or more positive impacts on opportunities to use Welsh and treating Welsh no less favourably than English? If you disagree, please explain why, including how you consider these proposals could be revised to have positive effects or more positive effects, or no adverse effects or fewer adverse effects on opportunities to use Welsh and treating Welsh no less favourably than English.

Equality impact assessment

A14.2 We have given careful consideration as to whether the proposals in this consultation will have a particular impact on persons sharing protected characteristics (including race, age, disability, sex, sexual orientation, gender reassignment, pregnancy and maternity, marriage and civil partnership and religion or belief in the UK and also dependents, and political opinion in Northern Ireland), and in particular whether they may discriminate against such persons or impact on equality of opportunity or good relations. This assessment helps us to comply with our duties under the Equality Act 2010 and the Northern Ireland Act 1998.

A14.3 We consider that some of our proposals would have a positive impact on certain groups. While we have considered the impact of our policy proposals across this consultation, including our draft guidance, we consider that most of these impacts are likely to come from our draft Children's Safety Codes. We also consider that positive impacts will come from our draft Guidance on Content Harmful to Children. We explain the impacts of the different components we are consulting on in turn below.

Guidance on Content Harmful to Children

A14.4 Our draft guidance aims to assist providers when making judgements on content by providing examples of, or kinds of content that we consider to be, or not to be primary priority content and priority content. We consider that our proposed approach to the examples within the guidance will have positive equality impacts for people with protected characteristics.

A14.5 For example, we provide examples of content which we consider to be or not to be content which is abusive or incites hatred and targets any one of a number of characteristics listed in the Act ('listed characteristics'), which are race, religion, sex, sexual orientation, disability and gender reassignment ('abuse and hate content').[272] While it is important to acknowledge that the definitions set out in the Act and what they mean in that context differ from some of the protected characteristics set out within the Equality Act 2010, there is a degree of overlap and similarity. For example, under the Act, 'disability' means any physical or mental impairment[273] whereas under the

[272] Sections 62(2) and 62(3) of the Act.
[273] Section 62(1) of the Act.

Equality Act 2010, a person has a disability if they have a physical or mental impairment, and the impairment has a substantial and long-term adverse effect on the person's ability to carry out normal day-to-day activities. 'Disability' under the Act, is therefore broader than the Equality Act definition.

A14.6 We consider that our proposed approach and examples of such content are appropriately balanced such that the listed characteristics under the Act are equally considered when assessing such content. This is likely to result in positive impacts where providers are able to identify and use appropriate measures (as set out below) against such content where this might disproportionately impact those with the relevant protected characteristics under the Equality Act 2010 to the extent that they may be impacted by abuse and hate content as defined under the Act.

Draft Children's Safety Codes

Terms of service

A14.7 Our proposals that relate to comprehensibility of language may benefit people with protected characteristics that could affect their level of literacy. Benefits could accrue to younger users, people who may not have English as a first language (which can be associated with race) and people with relevant disabilities. We have also made specific proposals for the benefit of users of assistive technologies including keyboard navigation; and screen reading technology which would also provide direct benefits to users (both children and adults) with disabilities.

User support

A14.8 Our user support proposals aim to reduce the risk of children encountering content harmful to children and to minimise the impact of harm to children of such content. In particular we have proposed measures relating to the provision of supportive information when child users restrict their interaction with other users or content, the signposting of children to support at key points of the user journey, and the provision of age-appropriate user support materials.

A14.9 We consider that such proposed measures may have benefits for people with protected characteristics who may be more impacted by content that is harmful to children by providing them with appropriate support. In addition, the provision of age-appropriate user support materials, which includes a consideration of comprehension of such materials, could also have positive impacts on those with protected characteristics which affect their level of literacy.

A14.10 These proposals are intended to ensure that children can understand the safety tools which are available to them on a service when they need them. In particular, the proposed measure related to age-appropriate user support materials recommends that these should include audio-visual elements as well as guidance for parents and carers, which could therefore provide benefits to those with relevant disabilities as noted above, or those who may not have English as a first language.

Governance and Accountability

A14.11 Our proposals in relation to governance and accountability seek to secure robust governance processes as an effective way of ensuring good risk management practices within a service. We consider that our governance and accountability proposals may have benefits for people whose protected characteristics may be more impacted by content that is harmful to children. For example, we have a proposed measure relating to the tracking of unusual increases or new kinds of primary priority content, priority content and non-designated content on a service. We consider that these measures should particularly benefit groups with protected characteristics, who are likely to be targeted by, or particularly at risk of, such content.

Search service design

A14.12 Our proposals in general aim to minimise the risk of children encountering content that is harmful to children. This includes content which is abusive or incites hate and is targeted at people with certain listed characteristics, and suicide, self-harm and eating disorder content. As such, they are likely to have positive impacts for groups that may be disproportionately affected by such content, particularly people that share protected characteristics. For example, girls who have an increased likelihood of encountering content promoting self-harm.

A14.13 More specifically, our proposal in relation to predictive search is, in our view, likely to reduce the likelihood of children being prompted to run searches for abuse and hate content, and content which encourages, promotes or provides instructions for suicide, self-harm or eating disorders. We consider that this measure should particularly benefit groups with protected characteristics, who are likely to be targeted by, or particularly at risk of, such content.

A14.14 Our proposal in relation to crisis prevention information in relation to searches for suicide, self-harm and eating disorder content would also have positive benefits. The requirement to ensure that the information provided is accessible and understandable to users of all ages (particularly children) will also secure that the resources can be easily used, and benefits experienced by younger users and those who may not have English as a first language or with relevant disabilities. To the extent that our measures will apply to all users of search services, the positive impacts outlined above would be experienced by adults in addition to children.

Recommender systems

A14.15 Our proposals aim to reduce the risk of children encountering content that is harmful to them by way of recommender systems, which includes abuse and hate content which targets people with certain listed characteristics as noted above, and suicide, self-harm and eating disorder content. This is likely to have positive impacts for groups that may be disproportionately affected by such content (particularly people that share protected characteristics).

User Reporting and complaints

A14.16 Our proposals aim to reduce barriers children face to reporting and complaints and ensure that processes are transparent, and easy to use and access by users (including children) and affected persons (as defined by the Act). We expect that our proposed measures will result in protecting children from content harmful to them. As with the proposed measures for recommender systems we consider that our proposed reporting and complaints measures are likely to have positive impacts for groups that may be disproportionately affected by content harmful to children (which includes abuse and hate content which targets individuals with certain listed characteristics). Easier reporting may also result in positive impacts for those with protected characteristics which affect their level of literacy. In particular, our proposed measures recommend that services have regard to the particular needs of its UK user base (including children), which includes the needs of people with relevant disabilities.

Age assurance

A14.17 Our proposals aim to help ensure children are protected from harmful content and have an age-appropriate experience online. In developing our proposals, we have sought to ensure accessibility, including by children, and effectiveness for all users regardless of any protected characteristics.

A14.18 We have outlined proposals to minimise the unintended exclusionary effects of age assurance technologies by recommending the use of highly effective age assurance that meet the

requirements of a proposed set of principles, including fairness and technical accuracy to ensure that services use highly effective age assurance that has been tested on diverse data.

A14.19 Our proposed age assurance measures are not intended to apply to all services, and we have limited them to those services that in our view, might present the most risk of content harmful to children being present on that service.

Content moderation for U2U and search services

A14.20 Our proposals aim to reduce the amount of content harmful to children that they encounter on U2U and search services by recommending that services have in place content moderation systems and processes that provide for swift and appropriate action to be taken on this type of content.

A14.21 Our proposals for large and multi-risk services to have content policies and appropriate training should in turn improve awareness of issues affecting groups with protected characteristics and encourage consistency of decision making. We are proposing that services should have regard to the different languages used by UK users when they resource their content moderation functions, which is likely to benefit speakers of languages other than English. This in turn may have positive impacts for those of different races, which may include different nationalities or ethnic backgrounds, as services will be able to more efficiently identify harmful content that may have specific cultural context or content that is from other countries.

A14.22 Services with content moderation functions that are well resourced are able to make considered decisions both about the content and any action taken, which should improve the rate at which decisions are taken fairly and in consideration of protected characteristics of users. The implementation of an effective content moderation function should improve outcomes for any group disproportionately subject to abuse and hate content, or disproportionately at risk of harm from pornography, content that encourages, promotes or provides instructions for suicide, self-harm and eating disorders, which we think is likely to include most groups with protected characteristics, particularly children.

Welsh language

A14.23 The Welsh language has official status in Wales. To give effect to this, certain public bodies, including Ofcom, are required to comply with Welsh language standards.[274] Accordingly, we have considered:

- The potential impact of our policy proposals on opportunities for persons to use the Welsh language;
- The potential impact of our policy proposals on treating the Welsh language no less favourably than the English language; and
- How our proposals could be formulated so as to have, or increase, a positive impact, or not to have adverse effects or to decrease any adverse effects.

A14.24 Ofcom's powers and duties in relation to online safety regulation are set out in the Online Safety Act 2023 and must be exercised in accordance with our general duties under section 3 of the Communications Act 2003. In formulating our proposals in this consultation, where relevant and to the extent we have discretion to do so in the exercise of our functions, we have considered the

[274] The Welsh language standards with which Ofcom is required to comply are available on our website.

potential impacts on opportunities to use the Welsh language and treating the Welsh language no less favourably than English. More generally, we are proposing that services should have regard to the needs of their user base in considering what languages are needed to effectively resource their content moderation, complaints handling, terms of service and publicly available statements. To this extent, we consider our proposals are likely to have positive effects or increased positive effects on opportunities to use the Welsh language and treating the Welsh language no less favourably than English.

A15. Glossary

This glossary defines the terms we have used throughout the consultation.

Term	Definition
2020 Video-Sharing Platform Regulation Call for Evidence	'*Video-sharing platform regulation Call for Evidence*', published by Ofcom on 16 July 2020 , available at Call for evidence: Video-sharing platform regulation (ofcom.org.uk).
2022 Illegal Harms Call for Evidence	'*First phase of online safety regulation Call for Evidence'*, published by Ofcom on 6 July 2022, available at Call for evidence: First phase of online safety regulation (ofcom.org.uk).
2023 Illegal Harms Consultation	*'Consultation: Protecting people from illegal harms online*', published by Ofcom on 9 November 2023, available at Consultation: Protecting people from illegal harms online - Ofcom (ofcom.org.uk).
2023 Protection of Children Call for Evidence	'*Second phase of online safety regulation Call for Evidence*'. Published by Ofcom on 10 January 2023, available at Call for evidence: Second phase of online safety regulation - Ofcom (ofcom.org.uk).
Abuse and hate content	Content which is abusive and which targets any of the following characteristics— (a) race, (b) religion, (c) sex, (d) sexual orientation, (e) disability, or (f) gender reassignment. Content which incites hatred against people— (a) of a particular race, religion, sex or sexual orientation, (b) who have a disability, or (c) who have the characteristic of gender reassignment.
Access controls	mechanisms to determine which users can access online content or spaces.
Act	Online Safety Act 2023.
Advertising-based revenue models	Revenue models that generate income through payments for the display of advertisements promoting a product or service.
Age appropriate user support materials	Materials that are specifically designed to be accessible and understandable to all children permitted to use a service, and to the adults who care for them.
Age assurance	A collective term for age verification and age estimation.
Age assurance method	An age assurance method refers to the particular system or technology that underpins an age assurance process.
Age assurance process	An age assurance process refers the end-to-end process through which the age

	assurance method or combination of methods are implemented to determine whether or not a user is a child. The effectiveness of an age assurance method will depend on how it is implemented, including whether by itself or in combination with other methods.
Age check	An individual instance where a user is required to undergo an age assurance process.
Age estimation	A form of age assurance designed to estimate the age or age-range of the user, for example using facial age estimation.
Age verification	A form of age assurance designed to verify the exact age of the user, for example using a form of identity documentation.
Algorithm	An algorithm is a sequence of computational instructions that help a programme or application achieve a specific goal.[275] Content recommender systems use different kinds of algorithms to learn about content types, user preferences, and match users to content. In addition to personalisation, content recommender systems can be designed to offer content variety, taking into account the diversity and popularity of content on a service.[276]
Algorithm speak	Algorithm speak or 'algospeak' refers to coded language used online in order to circumvent content moderation methods.
Anonymous user profiles	User-to-user service functionality allowing users to create a user profile where their identity[277] is unknown to an extent. This includes instances where a user's identity is unknown to other users; for example through the use of aliases ('pseudonymity'). It also includes where a user's identity may be unknown to a service, for example services that do not require users to register by creating an account.
Artificial intelligence chatbot	An automated software program that uses artificial intelligence and natural language processing to simulate a conversation.
Autoplay features	Feature that allows audiovisual content to continue playing without input from the user.
Avatar research methodology	Research methodology involving accounts or profiles set up on online services by researchers, modelled on the behaviours

[275] Ofcom, 2023. Evaluating recommender systems in relation to illegal and harmful content.
[276] Ofcom, 2023. Evaluating recommender systems in relation to illegal and harmful content.
[277] Identity refers to an individual's formal or officially recognised identity.

	and interests of real users. This method, similar to the 'mystery shopping' market research approach, is often used to understand the experience of a service by a particular group of people.
Block	A U2U functionality where: a) blocked users cannot send direct messages to the blocking user and vice versa; b) the blocking user will not encounter any content posted by blocked users on the service and vice versa; c) the blocking user and blocked user, if they were connected, will no longer be connected.
Blurring	Involves obscuring the view of image-based content. For example, this may be done by a greyscale overlaying the image, accompanied by a content warning.
Bot	An umbrella term that refers to a software application or automated tool which has been programmed by a person to carry out a specific or predefined task without any human intervention.
Business models	The way in which a business operates to achieve its goals. For the purposes of this risk assessment, this includes a service's revenue model and growth strategy.[278]
CA 2003	The Communications Act 2003.
Characteristic	In respect of a regulated service, includes references to its functionalities, user base, business models, governance and other systems and processes.[279]
Child user	A user under the age of 18.
Children	A person under the age of 18.
Children's code	The ICO's Children's code (also known as the Age Appropriate Design code).[280]
Children's safety duties	The safety duties protecting children in section 12 of the Act.
Clear web	Publicly accessible websites that are indexed by search engines.
Codes of practice (Codes)	The set of measures recommended by Ofcom for compliance with the children's safety duties.
Combined Service	A regulated U2U service that includes a public search engine[281].

[278] 'Business model' can be defined more widely to describe the way in which a service creates value to its users (value proposition), how it delivers this value to users, and how it captures value for itself. However, we adopt a narrow definition in the risk assessment to avoid overlap with the other risk characteristics. This does not affect the overall risk assessment as risk factors that would have been identified under the broader definition are captured elsewhere.

[279] Section 98(11) of the Act.

[280] ICO. Age appropriate design: a code of practice for online services | ICO. [accessed 30 April 2024].

[281] Section 4(7) of the Act.

Combining visual media	User-to-user functionality that allows users to join together videos and/or images, often from different sources, into one piece of content that can be shared.
Commenting on content	User-to-user service functionality that allows users to reply to content, or post content in response to another piece of content, visually accessible directly from the original content without navigating away from that content.
Commercial profile	The size of the service in terms of capacity, the stage of service maturity and rate of growth in relation to users or revenue.[282]
Community	Also referred to as "groups" or "forum groups" refer to a user-to-user service functionality allowing users to create online spaces that are often devoted to sharing content on a particular topic. User groups can be open to the public or closed to the public, requiring a registered account and an invitation or approval from existing members to gain access.
Content audience	Refers to whether content is shared on open or closed channels of communication. Open channels are areas of services where content is visible to the general public or any user. Closed channels are areas of a service where content is limited to a smaller audience, and where users can expect more privacy, such as direct messaging or user groups that have controls or restrictions on who can join.
Content controls	Mechanisms to determine the visibility and accessibility of content including its removal or reduction.
Content editing	Functionality type that comprises user-to-user functionalities which allow users to alter user-generated content before or after it is shared.
Content exploring	Functionality type that comprises user-to-user functionalities which allow users to explore and search for user-generated content.
Content format	Refers to the format in which content is made available. This, for instance, includes content in the form of images, video, audio, text and emojis.
Content identification processes	Automated content classifiers (e.g., machine learning and heuristic techniques) and trained moderators can assess whether

[282] In terms of number of employees and/or revenue.

	content is likely to be harmful to children or not and can label content. For example, content identified as likely to be harmful might be labelled as 'violent' meaning that the algorithm can filter this out so that it is not recommended to children.[283] Where content has completed the moderation process and has been found to not be harmful to children, this may re-enter the recommender systems for children.
Content moderation	When a service reviews content to decide whether it is permitted on its platform (either by AI or a human moderator).
Content recommender systems	Type of recommender system that is used to suggest and curate content that users are likely to find engaging, based on, for example, user preferences and/or history, but also content that is popular and trending on the service at a given moment.
Content restriction tools	User tools that allow users to privately (i.e., not visible to any other user of the service, including the creator of the content) restrict their interaction with a piece of content or kind of content, so that less or none of that content appears in their content feed in future. In some cases, the user may still be able to access the content if they search for it directly. These tools have different names on different services. Examples we are aware of include 'see less of this' and 'hide' tools. We would not consider a 'dislike' button to be a content restriction tool, if its primary function is to publicly express an opinion about the content, not to restrict interaction with it. However, a 'not interested' button might be a content restriction tool for the purposes of this measure if its primary function is to allow users to privately restrict interaction with a piece or kind of content.
Content storage and capture	Functionality type that comprises user-to-user functionalities which allow users to record and store user-generated content.
Content tagging	User-to-user service functionality allowing users to assign a keyword or term to content that is shared.
Content	Anything communicated by means of an internet service, whether publicly or

[283] Thorburn, L, Bengani, P, Stray, J., 2020. How platform recommenders work. [accessed 24 April 2024].

	privately, including written material or messages, oral communications, photographs, videos, visual images, music and data of any description.[284]
Crisis Prevention information	Refers to information provided by a search service in search results that typically contains the contact details of helplines and/or hotlines and links to trustworthy and supportive information provided freely by a reputable and reliable organisation.
CSAM (child sexual abuse material)	A category of CSEA content, including in particular indecent or prohibited images of children (including still and animated images, and videos, and including photographs, pseudo-photographs and non-photographic images such as drawings). CSAM also includes other material that includes advice about grooming or abusing a child sexually or which is an obscene article encouraging the commission of other child sexual exploitation and abuse offences. Furthermore, it includes content which links or otherwise directs users to such material, or which advertises the distribution or showing of CSAM.
CSEA (child sexual exploitation and abuse)	Refers to offences specified in Schedule 6 of the Act, including offences related to CSAM and grooming. CSEA includes but is not limited to causing or enticing a child or young person to take part in sexual activities, sexual communication with a child and the possession or distribution of indecent images.
Cumulative harm	Harm that occurs when harmful content (PPC, PC or NDC) is repeatedly encountered by a child, and/or when a child encounters harmful combinations of content. These combinations of content include encountering different types of harmful content (PPC, PC or NDC), or a type of harmful content (PPC, PC, or NDC) alongside a kind of content that increases the risk of harm from PPC, PC or NDC.[285]
Dangerous stunts and challenges content	Content which encourages, promotes or provides instructions for a challenge or stunt highly likely to result in serious injury to the person who does it or to someone else.

[284] Section 207(1) of the Act.
[285] Section 234(4) of the Act.

Dating services	User-to-user service type describing services that enable users to find and communicate with romantic or sexual partners.
Dedicated Reporting Channel (DRC)	A means for a Trusted Flagger (defined below) to report problems, for example an inbox, a web portal or another relevant mechanism for reporting.
Deepfake	Specific type of media that involves the use of AI algorithms, particularly generative AI models, to modify videos, images or audio to create realistic synthetic content. This is often done by superimposing the face of a person onto the body of another person in a video or image, as well as voice manipulation with lip-syncing. Deepfakes are commonly shared as user-generated content on user-to-user services but could also potentially be created using functionalities present on user-to-user services. Deepfake technology is currently used to create content that can be harmful; however, we acknowledge that it may also have positive use cases.
Deindexing	Involves the removal of URLs (i.e., links to individual webpages) or domains (i.e. entire websites) from a search index. This will prevent the webpage URLs from appearing in search results entirely.
Delisting	Involves adding content to a blacklist to ensure it does not appear in the pool of content returned in search results. Content which has been delisted will still be found in the index.
Direct messaging	User-to-user service functionality allowing a user to send and receive a message to one recipient at a time, and which can only be immediately viewed by that specific recipient.
Discussion forums and chat room services	A user-to-user service type describing general services that generally allow users to send or post messages that can be read by the public or an open group of people.
Downranking	Action taken by a search service which involves altering the ranking algorithm such that a particular piece of search content appears lower in the search results and is therefore less discoverable to users.
Downstream general search service	Search service type describing a subsection of general search services. Downstream general search services provide access to content from across the web, but they are distinct in that they obtain (or supplement)

	their search results from an index created by another general search service which relies solely on its own indexing (the 'upstream search service').[286]
Doxxing	The intentional online exposure of an individual's identity, private information or personal details without their consent.[287]
Early-stage services	Services in the initial phases of their lifecycle, typically encompassing the start-up and early growth stages. These are characterised by their early establishment, limited operational history, and ongoing efforts to establish themselves in the market.
Eating disorder content	Content which encourages, promotes or provides instructions for an eating disorder or behaviours associated with an eating disorder.
Editing visual media	User-to-user service functionality which allows users to alter or manipulate images and videos by means of the service.
Encrypted messaging	User-to-user service functionality that allows users to send and receive messages that are end-to-end encrypted.
Ephemeral messaging	User-to-user service functionality that that allows users to send messages that are automatically deleted after they are viewed by the recipient, or after a prescribed period of time has elapsed.
Explicit Feedback	This refers to direct and intentional actions taken by users to express their preferences and sentiment on content, for example likes/dislikes. Though it can it vary across services; explicit feedback provides recommender systems with clear and unambiguous information about a user's preferences. Depending on the service, reporting/complaints can also be forms of explicit negative feedback.
External content policies	Publicly available documents aimed at users of the service which provide an overview of a service's rules about what content is allowed and what is not. These are often in the form of terms of service and/or community guidelines.
Extreme pornography	An umbrella term to cover several categories of images which are illegal to

[286] Some downstream general search services may not be in control of the operations of the search engine. In such a case, we expect the upstream search service would be the provider of the search service. However, there may be circumstances in which the downstream search service does exercise control, and in those circumstances the downstream service would be the provider.
[287] eSafety Commissioner, 2020. What is doxing or doxxing? [accessed 18 April 2024].

	possess, broadly covering images which are produced principally for sexual arousal, and which depict extreme or obscene behaviours.
File-storage and file-sharing services	User-to-user service type describing services whose primary functionalities involve enabling users to store digital content and share access to that content through links.
Filter bubble	Describes the narrowing of content that is recommended to users, such that content feeds become homogenous and lack variety. Also often referred to as an 'echo chamber.'
Filtering	Involves ensuring that content is not returned in search results based on whether a condition is/isn't met. For example, 'not displaying search results where condition "PPC" is true.'
Functionalities	In relation to a user-to-user service, includes any feature that enables interactions of any description between users of the service by means of the service.[288] In relation to a search service, includes (in particular): (a) a feature that enables users to search websites or databases; (b) a feature that makes suggestions relating to users' search requests (predictive search functionality).[289] In practice, when referring to functionalities in the Register of Risks, 'functionalities' refers to the front-end features of a service. For user-to-user services, 'functionalities' refers to features that enable interaction between users. 'Functionalities for search services' refers to features that enable users to search websites or databases, as well as features that make suggestions relating to users' search requests.
Fundraising services	User-to-user service type describing services that typically enable users to create fundraising campaigns and collect donations from users.
Gaming services	User-to-user service type describing services that allow users to interact within

[288] Section 233(1) of the Act. Please refer to section 233(2) of the Act for a non-comprehensive list of user-to-user functionalities.
[289] Section 233(3) of the Act.

	partially- or fully-simulated virtual environments.
General search services	Search service type describing services that enables users to search the internet and which derives search results from an underlying search index (developed by either the service or a third party).
General search services which rely on their own indexing	Some general search services rely solely on their own indexing, using crawlers ('crawling') to find content across the web, building an index of URLs ('indexing') and using algorithms to rank the content based on relevance of the search request ('ranking'). General search services are also integrating GenAI to support or perform search functions, for example, by integrating a large language model to provide a conversational summary of that search results.
Generative artificial intelligence	Also known as 'GenAI,' generative artificial intelligence is an emerging form of AI that refers to machine-learning models which can create new content in response to a user prompt. These tools can be used to produce text, images, audio, video and code, which closely resemble the broad datasets on which the models are trained.
Governance	Structures that ensure the adequate oversight, accountability, and transparency of decisions within a service which affect user safety. This is in relation to organisational structure as well as product and content governance.
Grooming	An offence under paragraphs 5, 6, 11 or 12 of Schedule 6 to the Act, including but not limited to the act of an abuser communicating with a child.
Group messaging	User-to-user service functionality allowing users to send and receive messages through a closed channel of communication to more than one recipient at a time.
Growth strategy	How the service plans to expand its business, for example, through increasing revenue and number of users.
Harm	Means physical or psychological harm. References to harm presented by content, and any other reference to harm in relation to content, have the same meaning given to it by section 235 of the Act.[290]

[290] Section 201 of the Act.

Harmful substances content	Content which encourages a person to ingest, inject, inhale or in any other way self-administer— (a) a physically harmful substance; (b) a substance in such a quantity as to be physically harmful.
Hate offences	Public order offences relating to stirring up hatred on the grounds of certain protected characteristics.
High-capacity services	Services with a large number of employees and/or revenue.[291]
Highly effective age assurance	Methods of age assurance that are of such a kind and implemented in such a way that is highly effective at correctly determining whether or not a particular user is a child.
Hyperlinking	User-to-user service functionality enabling users to access other internet services by clicking or tapping on content present on the service.
Illegal content	Content that amounts to a relevant offence.
Illegal harm	Harms arising from illegal content and the commission and facilitation of priority offences.
Image or video search	Search service functionality that allows users to search for images and/or videos.
Immersive technology	A technology (most often used in gaming) which creates or enhances a realistic digital environment which users interact with.
Implicit Feedback	This refers to feedback into the recommender systems that the user may not have intended. Implicit feedback can involve the number of times a user clicks on an item, the amount of time they spend interacting with it (e.g., watch time), and how they scroll through content.
Indexing	The process of collecting, parsing, and storing of data by a search engine to facilitate fast and accurate information retrieval.
Infinite scrolling	Feature enables content to be continuously loaded as the user scrolls down.
Information-sharing services	User-to-user service type describing services that are primarily focused on providing user-generated informational resources to other users.
Internal content policies	More detailed versions of external content policies which set out rules, standards or guidelines, including around what content is allowed and what is not, as well as

[291] Our evidence does not currently allow for quantitative thresholds to be drawn for service capacity. Services should nevertheless consider the number of employees and revenue as a risk factor.

	providing a framework for how policies should be operationalised and enforced.
Large service	A service with more than 7 million monthly UK users.
Leet speak	Leet speak or 'l337 speak' refers to an informal online language where numbers or special characters are used to replace vowels or consonants.
Livestreaming	User-to-user service functionality that allows users to simultaneously create and broadcast online streaming media in, or very close to, real time.
Low-capacity services	Services with a small number of employees and/or revenue.[292]
Low-risk service	A service that has not assessed medium or high risk in relation to any kind of content harmful to children in its risk assessment.
Marketplaces and listings services	User-to-user service type describing services that allow users to buy and sell their goods or services.
Meme	An image or video that is spread widely on the internet, often altered by internet users for humorous effect.[293]
Messaging services	User-to-user service type describing services that are typically centred around the sending and receiving of messages that can only be viewed or read by a specific recipient or group of people.
Micro-businesses	Businesses that employ 1-9 full-time employees.
Multi-risk service	A service that assesses itself as being at medium or high risk in relation to at least two or more different kinds of content harmful to children in their latest children's risk assessment.
Muting	A user tool that enables a user to 'mute' another user. The muting user will not encounter any content posted by muted users on the service (unless the muting user visits the user profile of the muted user directly). The muted user is not aware that they have been muted and continues to encounter content posted by the muting user.
Negative sentiment	By negative sentiment, we mean the unfavourable or adverse emotions, or feelings experienced by children when encountering harmful content. This can

[292] Our evidence does not currently allow for quantitative thresholds to be drawn for service capacity. Services should nevertheless consider the number of employees and revenue as a risk factor.
[293] Collins Dictionary. [accessed 18 April 2024].

	include anxiety, sadness, anger, fear, frustration, or any form of distress. In the context of children encountering harmful content, a child user may not always be able to recognise, understand or express distress in a constructive way. It is important for online services to consider this risk when designing their recommender systems and user interaction features.
Non-designated content	A category of content harmful to children defined in the Act, broadly: content, which is not primary priority content or priority content, of a kind which presents a material risk of significant harm to an appreciable number of children in the United Kingdom.[294]
Overlays or interstitials	Elements such as pop-ups or webpages which appear before the target content is displayed, or while navigating between pages. Typically the user will need to take an action, such as clicking through, to reach the target content.
Part 3 or regulated search service	Refers to a search service that falls within the definition of section 4 of the Act.
Part 3 or regulated user-to-user service	A user-to-user service, as defined in section 4 of the Act.
Pile-on	Refers to when a user is criticized or targeted by a large number of other users, often as part of bullying campaigns.
Pornography services	Services whose principal purpose is the hosting or dissemination of pornographic content and who host user-generated pornographic content.
Posting content	User-to-user service functionality allowing users to upload and share content on open channels of communication.
Posting goods or services for sale	User-to-user service functionality allowing users to post content dedicated to offering goods and services for sale. This does not include paid-for advertisements, but may serve the function of allowing users to promote goods or services.[295]
Posting or sending location information	User-to-user service functionality allowing users to share their current or historic location, record a user's movement, or identify which other users of the service are nearby.
Predictive search	Search service functionality that anticipates a search query based on a variety of factors

[294] Section 60(2)(c) of the Act.
[295] See 'advertising-based revenue model' in business models for more information.

	(including those related to the search results' ranking).
Primary Priority Content	A category of content that is harmful to children, as defined in section 61 of the Act.[296]
Priority Content	A category of content that is harmful to children, as defined in section 62 of the Act.[297]
Priority offences	Offences set out in Schedules 5 (Terrorism offences), 6 (CSEA offences) and 7 (Priority offences) to the Act.
Proactive technology	Consisting of three types of technology: content identification technology, user profiling technology, and behaviour identification technology (subject to certain exceptions) as defined in section 231 of the Act.
Protected user characteristics	Age; disability; gender reassignment; marriage and civil partnership; pregnancy and maternity; race; religion or belief; sex; and sexual orientation.[298]
Publicly Available Statement	A statement that search services are required to make available to members of the public in the UK, often detailing various information on how the service operates.
Rabbit hole	The process of recommending ever more extreme content to users over time, which may occur as a result of users engaging with that type of content in the past.[299]
Reacting to content	User-to-user service functionality allowing users to express a reaction, such as approval or disapproval, of content that is shared by other users, through dedicated features that can be clicked or tapped by users.[300]
Recommender systems	An algorithmic system which, by means of a machine learning model, determines the relative ranking of suggestions made to users on a U2U service. The overarching objective of recommender systems is to ensure that users receive suggestions they are likely to find relevant and engaging. This can include suggesting connections, groups, events and content.

[296] We have typically grouped the different kinds of primary priority content as follows: pornographic content, suicide and self-harm content, eating disorder content.
[297] We have typically grouped the different kinds of priority content as follows: abuse and hate content, bullying content, violent content, harmful substances content, dangerous stunts and challenges content.
[298] Section 4 of the Equality Act 2010.
[299] PATTRN.AI, 2023. Evaluating recommender systems in relation to the dissemination of illegal and harmful content in the UK [accessed 22 April 2024].
[300] This for instance includes 'liking' or 'disliking' a post.

Record keeping and review guidance	The guidance that Ofcom is required to produce under section 52(3) of the Act to help services to comply with their record keeping and review duties under sections 23 (U2U) and 32 (search) of the Act. The draft guidance on which we are consulting can be found under annex 6 of this document.
Re-posting or forwarding content	User-to-user service functionality which allows users to re-share content that has already been shared by a user.
Revenue model	How a service generates income or revenue.
Review service	A service which enables users to create and view critical appraisals of people, businesses, products, or services.
Risk assessment	Identifying and assessing the risk of harm to individuals from illegal content and content harmful to children, present on a Part 3 regulated service.
Risk factor	A characteristic associated with the risk of one or more kinds of harm.
Risk of harm	The possibility of individuals encountering harm on a Part 3 service.
Safe search	A feature of several general search services which filters or obscures certain kinds of search content, such as pornographic/sexual or violent content. Safe search features can have levels or can be opted in or out of. In some cases, a safe search feature is enabled by default, for example for children.
Safety by design	Putting user safety at the centre of the design and development of online services and processes.
Screen capturing or recording	User-to-user service functionality that allows users to capture an image or record a video showing the contents of their display.[301]
Search content	Content that may be encountered in or via search results of a search service. It does not include paid-for advertisements, news publisher content, or content that reproduces, links to, or is a recording of, news publisher content.
Search engine	Includes a service or functionality which enables a person to search some websites or databases but does not include a service

[301] While users can often record or capture content using third-party services, screen recordings and captures are often shared on user-to-user services as user-generated content and some user-to-user services have dedicated screen recording and screen capturing functionalities.

	which enables a person to search just one website database.
Search index	A collection of URLs that are obtained by deploying crawlers to find content across the web, which is subsequently stored and organised.
Search prediction and personalisation	Functionality type that comprises search service functionalities, allowing suggestions to be made relating to users' search requests.
Search query inputs	Search service functionality type by means of which users input search queries.
Search result	In relation to a search service, this means content presented to a user of the service by operation of the search engine, in response to a search request made by the user.[302]
Search services	An internet service that is, or includes, a search engine.
Self-declaration	A process where the user is asked to provide their own age. This could be in the form of providing a date of birth to gain entry to a service or by ticking a box to confirm a user is over a minimum age threshold.
Service	A regulated user-to-user or search service, i.e. only the U2U or search part of the service. We also use it as a shorthand way of referring to the provider of the service concerned.
Service design	The design of all the components that shape a user's end-to-end experience of a service. These components can include the business model or decision-making structures, back-end systems and processes, the user interface, and off-platform interventions.
Service type	A characteristic that in general refers to the nature of the service. For example, social media services and messaging services.[303]
Small business	A business that employs 10-49 full-time employees.
Smaller service	A service which is not a large service.
Social media services	User-to-user service type describing services that connect users and enable them to build communities around common interests or connections.

[302] Section 57(3) of the Act.

[303] Certain service types have been selected because our evidence suggests that they play a role in children encountering harmful content.

Specific-risk service	A service which has assessed itself as being at medium or high risk for a specific kind of harm for which we propose a particular measure.
Stories	Feature on some services that allows users to post images and videos that are ephemeral
Stranger pairing	User-to-user functionality that allows users who likely do not know each other into contact, often at random.
Subscription-based revenue models	Revenue models that generate income by selling access (or premium access) to a service for a period of time in return for a fee.
Suggestive search	Search service functionality that recommends search queries that refine or build on the initial search query made by a user.
Suicide and self-harm content	Content which encourages, promotes or provides instructions for suicide or encourages, promotes or provides instructions for an act of deliberate self-injury.
Super-complaint	A complaint made under section 170 of the Act.
Systems and processes	The actions taken by a service, including procedures to mitigate the risk of content harmful to children being encountered, such as human moderators and automated systems or processes.
Takedown duty	The duty under section 10(3)(b) of the Act for a U2U service to use proportionate systems and processes designed to swiftly take down any (priority or non-priority) illegal content when it becomes aware of it.
Targeted safety measures	Measures recommended under Ofcom's Codes that apply to children.
Terms of Service	All documents comprising the contract for use of the service (or of part of it) by United Kingdom users.
The Act	The Online Safety Act 2023.
Trolling	Trolling is when someone post or comments online to deliberately upset others.[304]
Trusted Flagger	Individuals, NGOs, government agencies, and other entities that have demonstrated accuracy and reliability in flagging content that violates a platform's Terms of Service. As a result, they often receive special

[304] eSafety Commissioner, 2024. Trolling | What does trolling mean? [accessed 18 April 2024].

	flagging tools such as the ability to bulk flag content.
U2U services	Shorthand for 'user-to-user' service, which means an internet service by means of which content that is generated directly on the service by a user of the service, or uploaded to or shared on the service by a user of the service, may be encountered by another user, or other users, of the service.
URL (Uniform Resource Locator)	A "uniform resource locator", which is a reference that specifies the location of a resource accessible by means of the internet.
User access	A user's entry into a service and ability to use the functionalities present on that service.
User base demographics	Demographic make-up of the user base, including selected characteristics, intersectional dynamics and other relevant demographic factors.
User base	Users of a service. A user does not need to be registered with a service to be considered a user of that service.[305]
User communication	Functionality type that comprises user-to-user service functionalities which allow users to communicate with one another, either synchronously or asynchronously. Includes communication across open and closed channels.[306]
User connections	User-to-user service functionality that allows users to follow or subscribe to other users. Users must sometimes be connected in order to view all or some of the content that each user shares.
User feedback	Means the various types of data that helps the recommender systems learn about users' preferences, behaviour, and interactions with content.
User groups	User-to-user service functionality allowing users to create online spaces that are often devoted to sharing content on a particular topic. User groups are generally closed to the public and require an invitation or approval from existing members to gain access. However, in some cases they may be open to the public.
User identification	Functionality type that comprises user-to-user service functionalities which allow

[305] Section 195 of the Act makes clear that 'it does not matter whether a person is registered to use a service' for them to be considered a 'user.'

[306] See content audiences for definition of open and closed channels of communication.

	users to identify themselves to other users.
User networking	Functionality type that comprises user-to-user service functionalities which allow users to find or encounter each other, and establish contact.
User profiles	User-to-user service functionality that represents a collection of identifying information about a user, conveyed to other users of the service. This includes information that may be displayed to other users such as images, usernames, and biographies.[307] [308]
User report	User reports are a specific type of complaint about content, submitted through a reporting tool.
User tagging	User-to-user service functionality allowing users to assign other users, typically by their username, to content that is shared.
User-generated content	Content (a) that is: (i) generated directly on the service by a user of the service, or (ii) uploaded to or shared on the service by a user of the service, and (b) which may be encountered by another user, or other users, of the service by means of the service.
User-generated content searching	User-to-user service functionality allowing users to search for user-generated content by means of a user-to-user service.
User-to-user services	An internet service on which users of the service can generate, upload and/or share content, which can then be encountered by other users of the service.
Vent post	Content that is typically posed by a user to express personal problems or challenges.
Vertical search services	Search service type describing services that enable users to search for specific topics, or products or services offered by third-party providers. Unlike general search services, they do not return search results based on an underlying search index. Rather, they may use an API or equivalent technical means to directly query selected websites or databases with which they have a contract, and to return search results to users.

[307] User profiles are distinct from user accounts, which are representations of a user in a service's information system. They may contain information required for registration to a particular service that are often attributes of a user's identity such as name, age, contact details and preferences.

[308] Users can sometimes create fake user profiles, which are not a functionality in themselves, but are user profiles that impersonates another entity or are intentionally misleading.

Video-sharing services	User-to-user service type describing services that allow users to upload and share videos with the public.
Violent content	Content which encourages, promotes or provides instructions for an act of serious violence against a person. Content which— (a) depicts real or realistic serious violence against a person; (b) depicts the real or realistic serious injury of a person in graphic detail. Content which— (a) depicts real or realistic serious violence against an animal; (b) depicts the real or realistic serious injury of an animal in graphic detail; (c) realistically depicts serious violence against a fictional creature or the serious injury of a fictional creature in graphic detail.
Virality	The degree to which online content spreads easily and/or quickly across many online users, alongside how much engagement and/or views a piece of content received (i.e. 'shares', 'likes', and 'view', etc.).
Virtual private network (VPN)	The creation of a private network over a public internet connection.
Volunteer Moderation	Also referred to as "Community-reliant Moderation" and "Distributed moderation" typically refers to a form of moderation that combines formal policy made at the service level with community-specific rules by volunteer moderators at community level. This form of moderation relies on community members moderating content that does not align with community expectations. Volunteer moderation is often used as one type of moderation within a wider system.