[DRAFT] Illegal content Codes of Practice for user-to-user services

DRAFT FOR CONSULTATION

Draft published for consultation: 25 November 2026

Closing date for responses: 20 January 2027

Contents

1. Introduction 4

The Illegal content Codes of Practice for user-to-user services 4

The recommended measures 4

Data protection 5

2. Application and scope 6

3. Index of recommended measures 8

4. Recommended measures 20

A. Governance and accountability 20

B. [Not used] 24

C. Content moderation 27

D. Reporting and complaints 51

E. Recommender systems 66

F. Settings, functionalities and user support 71

G. Terms of service 75

H. User access 77

I. [Not used] 83

J. User controls 84

5. Definitions and interpretation 91

Risks of illegal harm 114

User numbers 121

DRAFT FOR CONSULTATION

1. Introduction

The Illegal content Codes of Practice for user-to-user services

1.1 Under the Online Safety Act 2023 (the 'Act'), Ofcom is required to prepare and issue Codes of Practice ('Codes') for providers of Part 3 services, describing measures recommended for the purpose of compliance with specified duties imposed on those providers by the Act.

1.2 This document contains Codes relevant to providers of regulated user-to-user services (including providers of combined services, so far as the duties applicable to user-to-user services apply in relation to those services) for the purpose of compliance with the following duties:

a) the illegal content safety duties set out in section 10(2) to (9) of the Act;
b) the duty about content reporting set out in section 20 of the Act, so far as it relates to illegal content; and
c) the duties about complaints procedures set out in section 21 of the Act, so far as relating to the complaints set out in section 21(4).

1.3 Recommended measures for search services are set out separately in the Illegal content Codes of Practice for search services.

1.4 The Act requires Ofcom to prepare and issue separate Codes for terrorism (arising from the offences set out in Schedule 5 to the Act) and child sexual exploitation and abuse ('CSEA') (arising from the offences set out in Schedule 6 to the Act) and one or more Codes for the purpose of compliance with the relevant duties relating to illegal content and harms (except to the extent measures are included in the Codes for terrorism and CSEA). Many of our recommended measures apply to more than one kind of illegal harm. To minimise duplication and simplify the regime for service providers, we have produced one document containing the main Codes for terrorism, CSEA and other duties. We identify the relevant Code(s) for each measure in the index of recommended measures which can be found at Section 3 of this document. We may from time to time amend these Codes or produce other Codes on specialist topics, which will be published on Ofcom's website (at https://www.ofcom.org.uk).[1] Providers should ensure they check this site.

1.5 Over time Ofcom will update the Codes to take account of technological developments, new evidence, and any other relevant matters.

The recommended measures

1.6 Section 4 of this document sets out the recommended measures and is divided into subsections by thematic area. The meaning of terms in **bold**, terms in ***bold and italics*** and terms which are underlined is explained in Section 5.

[1] https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/online-safety-regulatory-documents. As of 16 November 2026, Ofcom has submitted to the Secretary of State a draft Code in respect of proactive technology measures.

DRAFT FOR CONSULTATION

1.7 The Act provides that service providers which implement measures recommended to them in these Codes will be treated as complying with the relevant duty or duties to which those measures relate.

1.8 Where a service provider implements measures recommended to it in these Codes which include safeguards for the protection of freedom of expression and/or for the protection of the privacy of United Kingdom users, the Act provides that they will also be treated as complying with the duties set out in section 22(2) (in respect of freedom of expression) and section 22(3) (in respect of privacy).

1.9 Service providers may seek to comply with a relevant duty in another way by adopting what the Act refers to as alternative measures. In doing so, service providers would also need to comply with the duty to have particular regard to the importance of protecting United Kingdom users' right to freedom of expression and the privacy of United Kingdom users.

1.10 Where they take alternative measures, service providers must also maintain a record of what they have done and how they consider that it meets the relevant duties, including how they have complied with the duty to have particular regard to the importance of protecting freedom of expression and privacy.

Data protection

1.11 Implementing the recommended measures set out in these Codes will inevitably involve the processing of personal data. The Information Commission ('the ICO') is the statutory regulator for data protection law and has made clear that it expects service providers to comply fully with data protection law when taking measures for the purpose of complying with their online safety duties under the Act.

1.12 The ICO has set out that it expects service providers to take a 'data protection by design and by default' approach when implementing online safety systems and processes. It advises service providers to familiarise themselves with the data protection legislation, the ICO's Children's code and relevant ICO guidance, including the updated opinion published by the ICO in January 2024 setting out the Commission's expectations for age assurance under the Children's code, to understand how to comply with the data protection regime.

External documents

1.13 A hard copy of any external document referred to in the recommended measures can be obtained by writing to Ofcom at Riverside House, 2a Southwark Bridge Road, London, SE1 9HA (tel. 020 7981 3000).

DRAFT FOR CONSULTATION

2. Application and scope

2.1 These Codes apply to a ***provider*** in respect of the **regulated user-to-user service** that it provides.

2.2 If a person is the ***provider*** of more than one **regulated user-to-user service**, the recommended measures in these Codes have effect in relation to each such service (so far as applicable).

2.3 These Codes apply regardless of whether or not the ***provider*** of the service is inside the United Kingdom.

2.4 The services in respect of which each recommended measure in these Codes applies are specified in the "application" section of each measure. An overview can be found in the index of recommended measures in Section 3 of this document.

2.5 Section 5 of this document includes provision about a service's risk and size. The subsection headed 'Risks of illegal harm' (which begins at paragraph 5.4) sets out when a service is at medium or high risk of a kind of illegal harm, and includes a definition of a '**multi-risk service**'.

2.6 The subsection headed 'User numbers' (which begins at paragraph 5.7) explains when a service is to be treated as having more than a particular number of monthly **active United Kingdom users** for those measures which apply in relation to **services** of a certain size, and how to calculate the number of monthly **active United Kingdom users**. The definition of '**large service**' is included in the definitions section in Section 5 of this document.

2.7 The measures in these Codes are recommended for the purpose of compliance with the i**llegal content safety duties** and the **reporting and complaints duties** and their scope and application should be construed accordingly. In particular, the recommended measures should be construed in light of section 8 of the **Act** which provides that:

a) the duties set out in Chapter 2 of Part 3 of the **Act** which must be complied with in relation to a ***user-to-user service*** that includes ***regulated provider pornographic content*** does not extend to—

   i) the ***regulated provider pornographic content***, or

   ii) the design, operation or use of the service so far as relating to that content;

b) the duties set out in Chapter 2 of Part 3 of the **Act** which must be complied with in relation to a ***combined service*** do not extend to:

   i) the ***search content*** of the service,

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

   iii) anything relating to the design, operation or use of the ***search engine***; and

c) the duties set out in Chapter 2 of Part 3 of the **Act** which must be complied with in relation to a ***user-to-user service*** extend only to:

   i) the design, operation and use of the service in the United Kingdom, and

DRAFT FOR CONSULTATION

ii) 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.

DRAFT FOR CONSULTATION

3. Index of recommended measures

Recommended measure		Application	Code(s)	Relevant duties
Governance and accountability				
ICU A1	Annual review of risk management activities	**Large services**.	CSEA Terrorism Other duties	Section 10(2), (3), and (5) to (9) Section 20(2)* Section 21(2)† and (3)†
ICU A2	Individual accountable for illegal content safety duties and reporting and complaints duties	All **services**.		Section 10(2), (3), and (5) to (9) Section 20(2)* Section 21(2)† and (3)†
ICU A3	Written statements of responsibilities	**Large** or **multi-risk services**.		Section 10(2), (3), and (5) to (9) Section 20(2)* Section 21(2)† and (3)†
ICU A4	Internal monitoring and assurance	**Large services** that are **multi-risk services**.		Section 10(2), (3), and (5) to (9) Section 20(2)* Section 21(2)† and (3)†
ICU A5	Tracking evidence of new and increasing illegal harm	**Large** or **multi-risk services**.		Section 10(2) and (3)
ICU A6	Code of conduct regarding protection of users from illegal harm			Section 10(2), (3), and (5) to (9) Section 20(2)* Section 21(2)† and (3)†
ICU A7	Compliance training			Section 10(2), (3), and (5) to (9) Section 20(2)* Section 21(2)† and (3)†

CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
Age assurance				
ICU B1	Age assurance process	**Services** that use an **age assurance process** to determine which ***United Kingdom users*** of the service are ***children***.	CSEA Other duties	Section 10(2) and (3)
Content moderation				
ICU C1	Having a content moderation function to review and assess suspected illegal content	All **services**.	CSEA Terrorism Other duties	Section 10(2) and (3) Section 21(2)(b)†
ICU C2	Having a content moderation function that allows for the swift take down of illegal content	All **services**.	CSEA Terrorism Other duties	Section 10(2) and (3) Section 21(2)(b)†
ICU C2A	Having a content moderation function that allows for the take down of intimate image abuse content within 48 hours	All **services**.	Other duties	Section 10(3A) and (3B)
ICU C3	Setting internal content policies	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 10(2) and (3)
ICU C4	Performance targets	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 10(2) and (3)
ICU C4A	Steps to take on reversal of content decisions	All **services**.	CSEA Terrorism Other duties	Section 10(12) Section 22(2) and 22(3)
ICU C5	Prioritisation	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 10(2) and (3)
ICU C6	Resourcing	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 10(2) and (3)

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU C7	Provision of training and materials to individuals working in content moderation (non-volunteers)	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 10(2) and (3)
ICU C8	Provision of materials to volunteers	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 10(2) and (3)
ICU C9	Using hash matching to detect and remove CSAM	**Large services** that are at medium or high **risk** of image-based CSAM. **Services** that are at high **risk** of image-based CSAM and: (a) have more than 700,000 monthly **active United Kingdom users**, (b) are **file-storage and file-sharing services**, or (c) whose principal purpose is the hosting or dissemination of **regulated pornographic content**.	CSEA	Section 10(2) and (3)
ICU C10	Detecting and removing content matching listed CSAM URLs	**Large services** that are at medium or high **risk** of CSAM URLs. **Services** that have more than 700,000 monthly **active United Kingdom users** and are at high risk of CSAM URLs.	CSEA	Section 10(2) and (3)
ICU C11	[Intentionally blank]	[Intentionally blank]	[Intentionally blank]	[Intentionally blank]
ICU C12	[Intentionally blank]	[Intentionally blank]	[Intentionally blank]	[Intentionally blank]

FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU C13	Using hash matching to detect terrorism content	**Services** that enable ***regulated user-generated content*** in the form of photographs, videos or visual material to be generated, uploaded or shared, and are: a) **services** at high risk of ***terrorism content***, and (i) have more than 700,000 monthly **active United Kingdom users** or (ii) are **file-storage and file-sharing services**, or b) **large services** that are at medium or high risk of ***terrorism content***.	Terrorism	Section 10(2) and (3)

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU C14	Using hash matching to detect intimate image abuse content	Services that enable ***regulated user-generated content*** in the form of photographs, videos or visual images (whether or not combined with written material) to be generated, uploaded or shared, and: a) are at high **risk** of intimate image abuse, and: i) the principal purpose of the **service** is the hosting or dissemination of **regulated pornographic content**; or ii) have more than 700,000 monthly **active UK users**; or iii) are **file-storage and file-sharing services**; or b) are **large services** at medium or high **risk** of intimate image abuse.	Other duties	Section 10(2) and (3)
ICU C15	Crisis response	**Large services** that are at medium **risk** of one or more specified **kinds of illegal harm**; or **services** that are at high **risk** of one or more specified **kinds of illegal harm**.	Terrorism Other duties	Section 10(2) and 10(3)

Recommended measure		Application	Code(s)	Relevant duties
ICU C16	Crisis response (dedicated channel for law enforcement)	**Large services** that are at medium or high **risk** of one or more specified **kinds of illegal harm**.	Terrorism Other duties	Section 10(2) and 10(3)
Reporting and complaints				
ICU D1	Enabling complaints, including intimate image content reports	All **services**.	CSEA Terrorism Other duties	Section 20(2)* Section 20A Section 21(2)(a)†
ICU D2	Having easy to find, easy to access and easy to use complaints systems and processes	All **services**.	CSEA Terrorism Other duties	Section 20(2)* Section 20A Section 21(2)(c)†
ICU D3	Provision of information prior to the submission of a complaint	**Services *likely to be accessed by children*** that are **large** or at medium or high **risk** of any **kind of illegal harm**.	CSEA Terrorism Other duties	Section 21(2)(c)†
ICU D3A	Provision of information prior to the submission of an intimate image content report	All **services**.	Other duties	Section 20A
ICU D4	Appropriate action – sending indicative timeframes	**Services** that are **large** or at medium or high **risk** of any **kind of illegal harm**.	CSEA Terrorism Other duties	Section 21(2)(b)† and (c)†
ICU D5	Appropriate action – sending further information about how the complaint will be handled	**Services *likely to be accessed by children*** that are **large** or at medium or high **risk** of any **kind of illegal harm**.	CSEA Terrorism Other duties	Section 21(2)(b)† and (c)†
ICU D6	Opt-out from communications following a complaint	**Services** that are **large** or at medium or high **risk** of any **kind of illegal harm**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D7	Appropriate action for relevant complaints about suspected illegal content	All **services**.	CSEA Terrorism Other duties	Section 10(3) Section 21(2)(b)†

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU D8	Appropriate action for relevant complaints which are content appeals – determination (large or multi risk services)	**Large** or **multi-risk services**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D9	Appropriate action for relevant complaints which are content appeals – determination (services that are neither large nor multi risk)	**Services** that are neither **large** nor **multi-risk**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D10	Appropriate action for relevant complaints which are content appeals – action following determination	All **services**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D10A	Appropriate action for relevant complaints which are content appeals – action following determination upholding the appeal, except for CSEA content or CSEA content proxy (large or multi-risk services)	**Large** or **multi-risk services.**	Terrorism Other duties	Section 21(2)(b)†
ICU D11	Appropriate action for relevant complaints about proactive technology, which are not content appeals	All **services**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D12	Appropriate action for all other relevant complaints	All **services**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D12A	Expedited complaints procedure in relation to intimate image abuse content	All **services**.	Other duties	Section 21(2A)

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU D13	Exception: manifestly unfounded complaints	All **services**.	CSEA Terrorism Other duties	Section 21(2)(b)†
ICU D14	Dedicated reporting channel for trusted flaggers to report fraud	**Large services** that are at medium or high **risk** of fraud.	Other duties	Section 10(3)
ICU D15	Appropriate action for age assessment appeals (services that are large or multi-risk)	**Services** that: a) use an **age assurance process** to determine which ***United Kingdom users*** of the service are ***children***; and b) are **large services** or **multi-risk services** (or both).	CSEA Other duties	Section 21(2)(b)†
ICU D16	Appropriate action for age assessment appeals (services that are neither large nor multi-risk)	**Services** that: a) use an **age assurance process** to determine which ***United Kingdom users*** of the service are ***children***; and b) are neither **large services** nor **multi-risk services**.	CSEA Other duties	Section 21(2)(b)†
ICU D17	Livestreaming reporting for imminent physical harm	**Services** that: a) have **one-to-many livestreaming**; and b) are at medium or high **risk** of one or more specified **kinds of illegal harm**.	CSEA Terrorism Other duties	Section 20(2)* Section 21(2)(a)†

DRAFT CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
Recommender systems				
ICU E1	Collection of safety metrics during on-platform testing of content recommender systems	**Services** that conduct **on-platform testing** of **content recommender systems** and are at medium or high **risk** of two or more specified **kinds of illegal harm**.	CSEA Terrorism Other duties	Section 10(2)
ICU E2	Content recommender systems: excluding priority illegal content (terrorism and hate)	**Services likely to be accessed by children** that are (a) at medium or high **risk** of one or more specified **kinds of illegal harm**, and where (b) a **child-accessible part** of the service has a **content recommender system**.	Terrorism Other duties	Section 10(2)
ICU E3	Content recommender systems: excluding priority illegal content (encouraging or assisting suicide (or attempted suicide) and serious self-harm)	**Services likely to be accessed by children** that are (a) at medium or high **risk** of encouraging or assisting suicide (or attempted suicide) and serious self-harm, and where (b) a **child-accessible part** of the service has a **content recommender system.**	Other duties	Section 10(2)
Settings, functionalities and user support				
ICU F1	Safety defaults for child users	**Services** that are at high **risk** of grooming. **Large services** that are at medium **risk** of grooming.	CSEA Other duties	Section 10(2)

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU F2	Support for child users	**Services** that are at high **risk** of grooming. **Large services** that are at medium **risk** of grooming.	CSEA Other duties	Section 10(2)
Terms of service				
ICU G1	Terms of service: substance (all services)	All **services**.	CSEA Terrorism Other duties	Section 10(2), (5) and (7) Section 21(3)†
ICU G1A	Terms of service: substance (large or multi-risk services)	**Large** or **multi-risk services**.	Terrorism Other duties	Section 10(2) and (5)
ICU G2	Terms of service: substance (Category 1 services)	***Category 1 services***.		Section 10(9)
ICU G3	Terms of service: clarity and accessibility	All **services**.		Section 10(8) Section 21(3)†
User access				
ICU H1	Removing accounts of proscribed organisations	All **services**.	Terrorism	Section 10(2) and (3)
ICU H2	User sanctions	**Large** or **multi-risk services**.	Terrorism Other duties	Section 10(2)
ICU H3	Banning users for CSEA	All **services**.	CSEA	Section 10(2)

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
User controls				
ICU J1	User blocking and muting	**Services** that have **user profiles**, at least one of the specified functionalities, and are any of the following: a) a **large service** which is at medium or high **risk** of one or more specified **kinds of illegal harm**; b) a service, other than a **large service**, that is ***likely to be accessed by children*** and is either (i) at high **risk** of one or more specified **kinds of illegal harm**, or (ii) at medium **risk** of one or more specified **kinds of illegal harm** and has more than 700,000 **active United Kingdom users**.	CSEA Other duties	Section 10(2)

DRAFT FOR CONSULTATION

Recommended measure		Application	Code(s)	Relevant duties
ICU J2	Disabling comments	**Services** that have **commenting on content functionality** and are any of the following: a) a **large service** which is at medium or high **risk** of one or more specified **kinds of illegal harm**; b) a service, other than a **large service**, that is ***likely to be accessed by children*** and is either (i) at high **risk** of one or more specified **kinds of illegal harm**, or (ii) at medium **risk** of one or more specified **kinds of illegal harm** and has more than 700,000 **active United Kingdom users**.	CSEA Other duties	Section 10(2)
ICU J3	Notable user and monetised labelling schemes	**Large services** that are at medium or high **risk** of one or more specified **kinds of illegal harm**, and on which **user profiles** are labelled under a **notable user scheme** or a **monetised scheme**.	Other duties	Section 10(2)

* So far as it relates to illegal content.

† So far as relating to the complaints set out in section 21(4).

DRAFT FOR CO

4. Recommended measures

A. Governance and accountability

ICU A1 Annual review of risk management activities

Application

ICU A1.1 This measure applies to a ***provider*** in respect of each **large service** it provides.

Recommendation

ICU A1.2 The provider's most senior **governance body** in relation to the service should carry out and record an annual review of risk management activities having to do with **illegal harm** as it relates to individuals in the UK, including as to risk that is remaining after the implementation of appropriate Codes of Practice measures. The review should include how developing risks are being monitored and managed.

ICU A2 Individual accountable for illegal content safety duties and reporting and complaints duties

Application

ICU A2.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU A2.2 The provider should name an individual accountable to the most senior **governance body** for compliance with the **illegal content safety duties** and the **reporting and complaints duties**.

ICU A2.3 Being accountable means being required to explain and justify actions or decisions regarding:

a) **illegal harm** risk management and mitigation (including as to risks remaining after the implementation of appropriate Codes of Practice measures), and
b) compliance with the relevant duties,

to the most senior **governance body**.

DRAFT FOR CONSULTATION

ICU A3 Written statements of responsibilities

Application

ICU A3.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU A3.2 The provider should have written statements of responsibilities for senior managers who make decisions about the management of risks having to do with **illegal harm** in relation to individuals in the UK.

ICU A3.3 A statement of responsibilities is a document which clearly shows the responsibilities that the senior manager performs in relation to the management of risks having to do with **illegal harm** in relation to individuals in the UK and how those responsibilities fit in with the provider's overall governance and management arrangements in relation to the service.

ICU A4 Internal monitoring and assurance

Application

ICU A4.1 This measure applies to a ***provider*** in respect of each **service** it provides that is both a **large service** and a **multi-risk service**.

Recommendation

ICU A4.2 The provider should have an internal monitoring and assurance function to provide independent assurance that measures taken to mitigate and manage the risks of harm to individuals identified in the **risk assessment** are effective on an ongoing basis. This function should report to, and its findings should be considered by, either:

a) the body that is responsible for overall governance and strategic direction of a service; or
b) an audit committee.

ICU A4.3 This independent assurance may be provided by an existing internal audit function.

DRAFT FOR CONSULTATION

ICU A5 Tracking evidence of new and increasing illegal harm

Application

ICU A5.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU A5.2 The provider should track evidence of new kinds of illegal content on the service, and unusual increases in particular kinds of illegal content or illegal content proxy, or equivalent changes in the use of the service for the commission or facilitation of ***priority offences***. Relevant evidence may include, but is not limited to, that derived from:

a) complaints processes;
b) content moderation processes;
c) referrals from law enforcement; and
d) information from **trusted flaggers** and any other expert group or body the provider considers appropriate.

ICU A5.3 The provider should ensure that any new kinds of illegal content or unusual increases in particular kinds of illegal content or illegal content proxy, or equivalent changes in the use of the service for the commission of facilitation of ***priority offences***, are regularly reported through relevant governance channels to the most senior **governance body**.

ICU A5.4 To understand this, the provider should establish a baseline understanding of how frequently particular kinds of illegal content, illegal content proxy, or the commission or facilitation of ***priority offences*** occur on the service to the extent possible based on its internal data and evidence. The provider should use this baseline to identify unusual increases in the relevant data.

ICU A5.5 References in this Recommendation ICU A5 to "illegal content" or "illegal content proxy" are to be read as references to ***illegal content*** or ***illegal content proxy*** that may be encountered by ***United Kingdom users***.

DRAFT FOR CONSULTATION

ICU A6 Code of conduct regarding protection of users from illegal harm

Application

ICU A6.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU A6.2 The provider should have a code of conduct that sets standards and expectations for individuals working for the provider around protecting ***United Kingdom users*** from risks of **illegal harm**.

ICU A7 Compliance training

Application

ICU A7.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU A7.2 The provider should secure that individuals working for the provider who are involved in the design and operational management of the service are trained in the service's approach to compliance with the **illegal content safety duties** and the **reporting and complaints duties**, sufficiently to give effect to them. This measure does not apply in relation to **volunteers**.

ICU A7.3 This does not affect Recommendations ICU C7 (provision of training and materials to individuals working in content moderation (non-volunteers)), ICU C8 (provision of materials to volunteers) and ICU J3 (notable user and monetised labelling schemes) (see ICU J3.3(f)).

DRAFT FOR CONSULTATION

B. Age assurance

ICU B1 Age assurance process

Application

ICU B1.1 This measure applies to a ***provider*** in respect of each **service** that uses an **age assurance process** to determine which ***United Kingdom users*** of the service are ***children***.

Recommendation

ICU B1.2 In order to ensure that its **age assurance process** is highly effective at correctly determining whether or not a ***user*** is a ***child***, the provider should ensure that it meets each of the criteria listed under:

a) ICU B1.3 (technical accuracy),
b) ICU B1.4 (robustness),
c) ICU B1.5 (reliability), and
d) ICU B1.6 (fairness).

ICU B1.3 The technical accuracy criterion is fulfilled if:

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 and the results indicate that the measures are able to correctly determine whether or not a particular ***user*** is a ***child***;
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.**

ICU B1.4 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;
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**; and

DRAFT FOR CONSULTATION

c) the provider has identified, and taken appropriate steps to mitigate against, methods of circumvention that are easily accessible to **children in the United Kingdom** and where it is reasonable to assume that **children in the United Kingdom** may use them.

ICU B1.5 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**; and
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; and

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

ICU B1.6 The fairness criterion is fulfilled if:

a) the provider has ensured that any elements of the **age assurance process** for the **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; and
b) the provider has ensured that any elements of the **age assurance process** which rely on artificial intelligence or machine learning have been evaluated against either the outcome or the error parity and the results indicate that the **age assurance process** does not produce significant bias or discriminatory outcomes.

ICU B1.7 The provider should have regard to the following when implementing **highly effective age assurance** on the service:

a) the principle that **age assurance** should be easy to use, including by **children in the United Kingdom** 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**;

DRAFT FOR CONSULTATION

d) the **age appropriate design code** and the Information Commission's Opinion titled "Age Assurance for the Children's code" published on 18 January 2024.

ICU B1.8 The provider should ensure that ***users*** are able to easily access information about what a provider's **highly effective age assurance** is intended to do and how the provider's **highly effective age assurance** works prior to commencing the **highly effective age assurance** for the service.

ICU B1.9 The provider should not publish ***content*** that directs or encourages ***United Kingdom users*** to circumvent the **highly effective age assurance**.

DRAFT FOR CONSULTATION

C. Content moderation

ICU C1 Having a content moderation function to review and assess suspected illegal content

Application

ICU C1.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU C1.2 The provider should, as part of its **content moderation function**, have ***systems and processes*** designed to review and assess ***content*** the provider has reason to suspect may be ***illegal content***.

ICU C1.3 For this purpose, when the provider has reason to suspect that ***content*** may be ***illegal content***, the provider should review the content and either:

a) make an **illegal content judgement** in relation to the content; or
b) where the provider is satisfied that its ***terms of service*** prohibit the type of ***illegal content*** which it has reason to suspect exist, consider whether the content is in breach of those ***terms of service***.

ICU C1.4 This does not affect Recommendations ICU C9 (using hash matching to detect and remove CSAM), ICU C10 (detecting and removing content matching listed CSAM URLs), ICU C13 (using hash matching to detect terrorism content) and ICU C14 (using hash matching to detect intimate image abuse content) (see ICU C9.4, ICU C10.5, ICU C13.5(a) and ICU C14(5)(a), respectively).

Safeguards for freedom of expression and privacy

ICU C1.5 The following measures are safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***:

za) Recommendation ICU C4A;
a) where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8 (in relation to content moderation);
b) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;
c) Recommendations ICU D8 or ICU D9 (whichever is applicable) and ICU D10 (in relation to **content appeals**); and

DRAFT FOR CONSULTATION

d) Recommendation ICU D10A (in relation to **content appeals**), where applicable.

ICU C2 Having a content moderation function that allows for the swift take down of illegal content

Application

ICU C2.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU C2.2 The provider should, as part of its **content moderation function,** have systems and processes designed to swiftly take down ***illegal content*** and/or **illegal content proxy** of which it is aware, (see also ICU C1.2), unless it is currently not technically feasible for them to achieve this outcome.

ICU C2.3 For this purpose, when the provider determines that:

a) the content is ***illegal content*** (pursuant to ICU C1.3(a)); or
b) the content is in breach of its ***terms of service*** (pursuant to ICU C1.3(b),

the provider should swiftly take the content down.

ICU C2.4 This does not affect Recommendations ICU C9 (using hash matching to detect and remove CSAM), ICU C10 (detecting and removing content matching listed CSAM URLs), ICU C13 (using hash matching to detect terrorism content) and ICU C14 (using hash matching to detect intimate image abuse content) (see ICU C9.4, ICU C10.5, ICU C13.5(a) and ICU C14.5(a) respectively).

Safeguards for freedom of expression and privacy

ICU C2.5 The following measures are safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***:

za) Recommendation ICU C4A;
a) where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8 (in relation to content moderation);
b) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;
c) Recommendations ICU D8 or ICU D9 (whichever is applicable) and ICU D10 (in relation to **content appeals**); and

DRAFT FOR CONSULTATION

d) Recommendation ICU D10A (in relation to **content appeals**), where applicable.

ICU C2A Having a content moderation function that allows for the take down of intimate image abuse content within 48 hours

Application

ICU C2A.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU C2A.2 The provider should, as part of its **content moderation function**, have ***systems and processes*** designed to ***take down***:

a) content in relation to which an ***intimate image content report*** is made to the provider, and
b) any other content identified by the provider as the same or substantially the same as the content referred to in (a),

as soon as reasonably practicable and no later than 48 hours after the provider receives the ***intimate image content report***, unless it is currently not technically feasible for them to ***take down content***.

ICU C2A.3 ICU C2A.2 does not apply if the provider considers:

a) the content is not **intimate image abuse content**, or
b) the person making the ***intimate image content report*** is not:
   (i) the subject of the content, or
   (ii) a person acting on that person's behalf.

Safeguards for freedom of expression and privacy

ICU C2A.4 The following measures are safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendation ICU C4A;
b) where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8 (in relation to content moderation);
c) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;
d) Recommendations ICU D8 or ICU D9 (whichever is applicable) and ICU D10 (in relation to **content appeals**); and
e) Recommendation ICU D10A (in relation to **content appeals**), where applicable.

DRAFT FOR CONSULTATION

ICU C3 Setting internal content policies

Application

ICU C3.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU C3.2 The provider should set and record (but need not publish) internal content policies setting out rules, standards and guidelines around:

a) what ***regulated user-generated content*** is allowed on the service and what is not; and
b) how policies should be operationalised and enforced.

ICU C3.3 The policies should be drafted in such a way that ***illegal content*** (where it is identifiable as such) is not permitted.

ICU C3.4 The provider should:

a) have regard to the **risk assessment** of the service in setting these policies; and
b) have processes in place for updating these policies in response to evidence of new and increasing **illegal harm** on the service (as tracked in accordance with ICU A5.2).

Safeguard for freedom of expression and privacy

ICU C3.5 Recommendation ICU C4A is a safeguard to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***.

ICU C4 Performance targets

Application

ICU C4.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

DRAFT FOR CONSULTATION

ICU C4.2 The provider should set and record performance targets for its **content moderation function**, covering at least:

a) the time period for taking **relevant content moderation action;** and
b) the accuracy of decision making.

ICU C4.3 In setting its targets, the provider should balance the need to take **relevant content moderation action** swiftly against the importance of making accurate moderation decisions.

ICU C4.4 The provider should effectively measure and monitor its performance against its performance targets.

ICU C4.5 For the purpose of ICU C4.2 and ICU C4.3, "**relevant content moderation action**" refers to:

a) the action recommended by ICU C1.3 and ICU C2.3; or
b) to the extent that it is currently not technically feasible for the provider to take down content, the action recommended by ICU C1.3, so far as the suspected ***illegal content*** is of a kind in respect of which the provider may impose a **sanction** under its policy on **sanctions** (see Recommendation ICU H2).

ICU C4A Steps to take on reversal of content decisions

Application

ICU C4A.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU C4A.2 If a ***provider*** reverses a previous decision that ***content*** is ***illegal content*** or **illegal content proxy**, and this reversal is not in relation to a **relevant complaint** that is a **content appeal** (for which, see ICU D10), it should take the steps listed in ICU D10.2(a), (b) and (c).

ICU C5 Prioritisation

Application

ICU C5.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

DRAFT FOR CONSULTATION

Recommendation

ICU C5.2 The provider should prepare and apply a policy in respect of the prioritisation of ***content*** for review. In setting the policy, the provider should have regard to at least the following:

a) the desirability of minimising the number of ***United Kingdom users*** encountering a particular item of ***illegal content***;
b) the severity of potential harm to ***United Kingdom users*** if they ***encounter illegal content*** on the service, including whether the ***content*** is suspected to be ***priority illegal content***, the **risk assessment** of the service, and the potential harm to **children in the United Kingdom**; and
c) the likelihood that ***content*** is ***illegal content***, including whether it has been reported by a **trusted flagger**.

ICU C6 Resourcing

Application

ICU C6.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU C6.2 The provider should resource its **content moderation function** so as to give effect to its internal content policies and performance targets having regard to at least:

a) the propensity for external events to lead to a significant increase in demand for content moderation on the service; and
b) the particular needs of its ***United Kingdom user*** base as identified in its **risk assessment**, in relation to languages.

ICU C7 Provision of training and materials to individuals working in content moderation (non-volunteers)

Application

ICU C7.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

DRAFT FOR CONSULTATION

Recommendation

ICU C7.2 The provider should ensure individuals working in content moderation receive training and materials that enable them to fulfil their role in moderating ***content*** including in relation to Recommendations ICU C1 and ICU C2 and the internal content policies set in accordance with Recommendation ICU C3. This measure does not apply in relation to **volunteers**.

ICU C7.3 The provider should ensure that in doing so:

a) it has regard to at least the **risk assessment** of the service and evidence of new and increasing **illegal harm** on the service (as tracked in accordance with ICU A5.2); and
b) where the provider identifies a gap in the understanding of individuals working in content moderation in relation to a specific kind of **illegal harm**, it gives training and materials to remedy this.

ICU C8 Provision of materials to volunteers

Application

ICU C8.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU C8.2 The provider should ensure **volunteers** in its **content moderation function** have access to materials that enable them to fulfil their role in moderating ***content*** including in relation to Recommendations ICU C1 and ICU C2 and the internal content policies set in accordance with Recommendation ICU C3.

ICU C8.3 The provider should ensure that in doing so:

a) it has regard to at least the **risk assessment** of the service and evidence of new and increasing **illegal harm** on the service (as tracked in accordance with ICU A5.2); and
b) where the provider identifies a gap in such volunteers' understanding of a specific kind of **illegal harm**, it gives materials to remedy this.

ICU C9 Using hash matching to detect and remove CSAM

DRAFT FOR CONSULTATION

Application

ICU C9.1 This measure applies to a ***provider*** in respect of each **service** it provides that:

a) is at high **risk** of image-based CSAM, and:
   i) has more than 700,000 monthly **active United Kingdom users** (see paragraphs 5.7 to 5.10);
   ii) is a **file-storage and file-sharing service**; or
   iii) the principal purpose of the **service** is the hosting or dissemination of **regulated pornographic content**; or
b) is a **large service** and is at medium or high **risk** of image-based CSAM.

Key definition

ICU C9.2 In this Recommendation ICU C9 "relevant content" means:

a) any ***regulated user-generated content*** in the form of photographs, videos or visual material that:
   i) may be ***encountered*** by ***United Kingdom users*** of the service by means of the service, and
   ii) is communicated publicly[2] by means of the service; or
b) any material which, if it were present on the service, would be content within sub-paragraph (a).

Recommendation

ICU C9.3 The provider should ensure that, where technically feasible, **perceptual hash matching technology** is used effectively to analyse relevant content to assess whether it is **CSAM**.

ICU C9.4 The provider should ensure that appropriate measures are taken to swiftly ***take down*** (or prevent from being generated, uploaded or shared) **detected content** that is **CSAM** (but see ICU C9.15 to ICU C9.18 in relation to the use of human moderators).

ICU C9.5 For the purposes of ICU C9.3, the provider should ensure that:

a) all relevant content present on the service at the time the technology is implemented is analysed within a reasonable time; and
b) relevant content that is generated on, uploaded to or shared on the service (or that a user seeks to so generate, upload or share) after

[2] Ofcom has published **guidance on content communicated 'publicly' and 'privately' under the Online Safety Act** for this purpose.

DRAFT FOR CONSULTATION

the technology is implemented is analysed before or as soon as practicable after it can be ***encountered*** by ***United Kingdom users*** of the service.

ICU C9.6 For the use of the **perceptual hash matching technology** to be effective, it should:

a) use a suitable perceptual hash function to compare relevant content to an appropriate set of hashes (see ICU C9.7 to ICU C9.11); and
b) be configured so that its performance strikes an appropriate balance between **precision** and **recall** (see ICU C9.12 to ICU C9.14).

The set of hashes

ICU C9.7 For the set of hashes to be appropriate, it should include hashes of **CSAM**:

a) sourced from one or more persons with expertise in the identification of **CSAM** and who meet (in relation to the hashes in question) the other requirements set out in ICU C9.8 (but may also include other hashes of **CSAM**, such as **CSAM** identified by the service's **content moderation function**);
b) that, taken together, reflects the range of material that is illegal under the criminal law of any part of the United Kingdom (including images of ***children*** that are indecent but may not show sexual activity).

ICU C9.8 The requirements are that the person has arrangements in place:

a) to identify suspected **CSAM**;
b) to secure (so far as possible) that **CSAM** is correctly identified before hashes of that material are added to its database (such as assessment by persons with expertise in making such judgements);
c) which, in relation to identifying or assessing suspected **CSAM**, do not plainly discriminate on the basis of protected characteristics (within the meaning of Part 2 of the Equality Act 2010);[3]
d) to regularly update its database with hashes of **CSAM**; and
e) to review cases where material is suspected to have been incorrectly identified as **CSAM**, and remove such hashes from the database where appropriate;
f) to secure its database from unauthorised access, interference or exploitation (whether by persons who work for that person or are providing a service to that person, or any other person).

ICU C9.9 The provider should ensure that the latest versions of any databases sourced from a person in accordance with ICU C9.7 and ICU C9.8 are regularly obtained and then used for the purposes of ICU C9.3.

[3] 2010 c. 15.

DRAFT FOR CONSULTATION

ICU C9.10 Where the set of hashes includes hashes of **CSAM** not sourced from a person in accordance with ICU C9.7 and ICU C9.8, the provider should also ensure that arrangements are in place in relation to those hashes:

a) to secure (so far as possible) that **CSAM** is correctly identified before hashes of that material are added; and
b) to review cases where material is suspected to have been incorrectly identified as **CSAM** and remove such hashes where appropriate.

ICU C9.11 The provider should ensure an appropriate policy is put in place, and that measures are taken in accordance with that policy, to secure any hashes of **CSAM** held for the purposes of this Recommendation ICU C9 from unauthorised access, interference or exploitation (whether by persons who work for the provider or are providing a service to the provider, or any other person).

Technical configuration

ICU C9.12 In configuring the technology so that its performance strikes an appropriate balance between **precision** and **recall**, the provider should ensure that the following matters are taken into account:

a) the service's risk of harm relating to image-based CSAM, reflecting the **risk assessment** of the service and any information reasonably available to the provider about the prevalence of relevant content that is **CSAM** on the service;
b) the proportion of **detected content** that is a **false positive**;
c) the effectiveness of the ***systems and/or processes*** used to identify **false positives**; and
d) the importance of minimising the reporting of **false positives** to the National Crime Agency or a foreign agency (within the meaning of Chapter 2 of Part 4 of the **Act**).

ICU C9.13 The provider should ensure that the performance of the technology, and whether the balance between **precision** and **recall** continues to be appropriate, is reviewed at least every six months.

ICU C9.14 The provider should ensure that a written record is made of how this balance has been struck in configuring the technology, including what information has been considered, and information about reviews and steps taken in response.

Use of human moderators

ICU C9.15 The provider should ensure that a policy is put in place for review of **detected content**, and action is taken in accordance with that policy, which secures that human moderators review an appropriate proportion of content **detected** as **CSAM**.

ICU C9.16 When deciding the policy for review of **detected content**, the provider should ensure that the following things are taken into account:

DRAFT FOR CONSULTATION

a) the level of assurance the provider has in the detection outcomes achieved by the **perceptual hash matching technology** in use and any associated ***systems and/or processes*** (as indicated by the ongoing monitoring, evaluation and quality assurance (including human quality assurance) of the performance of the technology); and
b) to the extent that **detected content** is subject to review for the purpose of Recommendation ICU C1:

i) the potential severity of the harm to those depicted in or ***encountering*** **CSAM**; and
ii) the overall impact of an incorrect decision that the relevant content is ***illegal content*** on a ***United Kingdom user*** who generated, uploaded or shared the content, including the importance of minimising the reporting of **false positives** to the National Crime Agency (NCA) or a foreign agency (within the meaning of Chapter 2 of Part 4 of the **Act**).

ICU C9.17 The provider should ensure that a written record is made of its policy for review of **detected content**, which sets out:

a) the proportion of **detected content** which is intended to be reviewed; and
b) information about how the things mentioned in ICU C9.16 were taken into account in deciding that policy.

ICU C9.18 The provider should keep statistical records about content reviewed in accordance with that policy (including the number of reviews carried out, the proportion of **detected content** that this represents, and the number of **false positives** identified).

Safeguards for freedom of expression and privacy

ICU C9.19 Paragraphs ICU C9.6 to ICU C9.18 of this Recommendation ICU C9 are safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***.

ICU C9.20 The following measures are also safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

za) Recommendation ICU C4A;
a) where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8 (in relation to content moderation);
b) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;
c) Recommendations ICU D8 or ICU D9 (whichever is applicable) and ICU D10 (in relation to **content appeals**); and

DRAFT FOR CONSULTATION

d) Recommendation ICU G1 (terms of service: substance (all services)).

ICU C10 Detecting and removing content matching listed CSAM URLs

Application

ICU C10.1 This measure applies to a ***provider*** in respect of each **service** it provides that:

a) has more than 700,000 monthly **active United Kingdom users** (see paragraphs 5.7 to 5.10) and is at high **risk** of CSAM URLs; or
b) is a **large service** and is at medium or high **risk** of CSAM URLs.

Key definitions

ICU C10.2 In this Recommendation ICU C10:

"CSAM URL" means a **URL** at which **CSAM** is present, or a domain which is entirely or predominantly dedicated to **CSAM**;

"relevant content" means:

a) any ***regulated user-generated content*** in the form of written material or messages (including hyperlinks) that:
   i) may be ***encountered*** by ***United Kingdom users*** of the service by means of the service, and
   ii) is communicated publicly[4] by means of the service; or
b) any material which, if it were present on the service, would be content within sub-paragraph (a).

ICU C10.3 For the purpose of ICU C10.2, a domain is "entirely or predominantly dedicated" to **CSAM** if the ***content*** present at the domain, taken overall, entirely or predominantly comprises **CSAM** (such as indecent images of children) or ***content*** related to ***CSEA content***).

Recommendation

ICU C10.4 The provider should ensure that, where technically feasible, technology is used effectively to analyse relevant content to assess whether it consists of or includes content matching a listed CSAM URL.

[4] Ofcom has published **guidance on content communicated 'publicly' and 'privately' under the Online Safety Act** for this purpose.

DRAFT FOR CONSULTATION

ICU C10.5 The provider should ensure that content **detected** as matching a listed CSAM URL is swiftly ***taken down*** (or prevented from being generated, uploaded or shared).

ICU C10.6 For the purposes of ICU C10.4, the provider should ensure that:

a) all relevant content present on the service at the time the technology is implemented is analysed within a reasonable time; and
b) relevant content that is generated on, uploaded to or shared on the service (or that a user seeks to so generate, upload or share) after the technology is implemented is analysed before or as soon as practicable after it can be ***encountered*** by ***United Kingdom users*** of the service.

ICU C10.7 For the use of the technology to be effective, it should:

a) compare analysed content to one or more lists of CSAM URLs sourced from a person (or persons) with expertise in the identification of **CSAM** and who meets (in relation to the list) the requirements set out in ICU C10.8, and
b) **detect** content as matching a listed CSAM URL where it is a direct match for a listed **URL** or is a **URL** that contains a listed domain (and for these purposes it does not matter whether the content includes an access protocol, such as “https://”).

ICU C10.8 The requirements are that the person has arrangements in place:

a) to identify **URLs** or domains suspected to be CSAM URLs;
b) to secure (so far as possible) that suspected CSAM URLs are correctly identified before they are added to the list;
c) which, in relation to identifying or assessing suspected CSAM URLs, do not plainly discriminate on the basis of protected characteristics (within the meaning of Part 2 of the Equality Act 2010);
d) to regularly update the list with identified CSAM URLs;
e) to regularly review listed CSAM URLs, and remove from the list any which are no longer CSAM URLs; and
f) to secure the list from unauthorised access, interference or exploitation (whether by persons who work for that person, or by any other person).

ICU C10.9 The provider should ensure that the latest version of any list or lists are regularly obtained and then used for the purposes of ICU C10.4.

ICU C10.10 The provider should ensure that an appropriate policy is put in place, and that measures are taken in accordance with that policy, to secure any copy of a list held for the purposes of this Recommendation ICU C10 from unauthorised

DRAFT FOR CONSULTATION

access, interference or exploitation (whether by persons who work for the provider or are providing a service to the provider, or any other person).

Safeguards for freedom of expression and privacy

ICU C10.11 The following elements of this Recommendation ICU C10 are safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***:

a) the arrangements referred to in sub-paragraphs (b), (e) and (f) of ICU C10.8;
b) ICU C10.9 and ICU C10.10.

ICU C10.12 The following measures are also safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;
b) Recommendations ICU D8 or ICU D9 (whichever is applicable) and ICU D10 (in relation to **content appeals**); and
c) Recommendation ICU G1 (terms of service: substance (all services)).

ICU C11 [Intentionally blank]

ICU C12 [Intentionally blank]

ICU C13 Using hash matching to detect terrorism content

Application

ICU C13.1 This measure applies to a ***provider*** in respect of each **service** it provides that enables ***regulated user-generated content*** in the form of photographs, videos or visual material to be generated, uploaded or shared, and:

a) is at high **risk** of ***terrorism content***, and:
   i) has more than 700,000 monthly **active United Kingdom users** (see paragraphs 5.7 to 5.10), or
   ii) is a **file-storage and file-sharing service**, or
b) is a **large service** and is at medium or high **risk** of ***terrorism content***.

Key definition

ICU C13.2 In this Recommendation ICU C13:

"relevant content" means:

a) any ***regulated user-generated content*** in the form of photographs, videos or visual material that:

DRAFT FOR CONSULTATION

i) may be ***encountered*** by ***United Kingdom users*** of the service by means of the service, and
ii) is communicated publicly[5] by means of the service, or

b) any material which, if it were present on the service, would be content within sub-paragraph (a).

Recommendation

ICU C13.3 The provider should ensure that, where technically feasible, **perceptual hash matching technology** is used effectively (see ICU 13.7) to analyse relevant content to assess whether it is ***terrorism content***.

ICU C13.4 In circumstances where relevant content matches with a hash, the provider:

a) may, depending on the level of assurance the provider has in the detection outcomes achieved by the **hash matching technology**, treat the relevant content as ***illegal content*** and swiftly take it down in accordance with ICU C2.3; or
b) should otherwise treat the match as reason to suspect that the relevant content may be ***illegal content*** and review the relevant content in accordance with Recommendation ICU C1.

ICU C13.5 The provider should ensure human moderators review and assess an appropriate proportion of **detected content**, having regard to:

a) the level of assurance the provider has in the detection outcomes achieved by the **hash matching technology** and any associated ***systems and processes*** (as indicated by the ongoing monitoring, evaluation and quality assurance (including human quality assurance) of the performance of the technology), and

b) to the extent that **detected content** is subject to review for the purpose of Recommendation ICU C1:

i) the potential severity of the harm to those depicted in or ***encountering terrorism content***, and
ii) the overall impact of an incorrect decision that the relevant content is ***illegal content*** on a ***United Kingdom user*** who generated, uploaded or shared the content.

ICU C13.6 For the purposes of ICU C13.3, the provider should ensure that:

a) all relevant content present on the service at the time the **perceptual hash matching technology** is implemented is analysed within a reasonable time, and
b) relevant content that is generated on, uploaded to or shared on the service (or that a user seeks to so generate, upload or share) after the **perceptual hash matching technology** is implemented is analysed

[5] Ofcom has published **guidance on content communicated 'publicly' and 'privately' under the Online Safety Act** for this purpose.

DRAFT FOR CONSULTATION

before or as soon as practicable after it can be ***encountered*** by ***United Kingdom users*** of the service.

ICU C13.7 For the use of **perceptual hash matching technology** to be effective, it should:

a) use a suitable perceptual hash function to compare relevant content to an appropriate set of hashes (see ICU C13.8 to ICU C13.11), and
b) be configured so that its performance strikes an appropriate balance between **precision** and **recall** (see ICU C13.12 to ICU C13.14).

The set of hashes

ICU C13.8 For the set of hashes to be appropriate, it should:

a) include hashes of ***terrorism content*** sourced by a person (a "hash database provider") with:
   i) expertise in identification of ***terrorism content***, and
   ii) arrangements in place to identify suspected ***terrorism content*** and to ensure that ***terrorism content*** is correctly identified before hashes of that material are added to the database,
b) include an appropriate range of ***terrorism content***,
c) contain hashes of a significant number of original items of ***terrorism content***,
d) not plainly discriminate, in relation to identifying ***terrorism content***, on the basis of protected characteristics (within the meaning of Part 2 of the Equality Act 2010),
e) be proactively updated with reasonable regularity,
f) be subject to arrangements for identifying and removing hashes that are not capable of amounting to ***terrorism content***, and
g) be secured from unauthorised access or interference (whether by persons who work for the provider or a third-party hash database provider, persons who are providing a service to the provider or third-party hash database provider, or any other person).

ICU C13.9 Where the provider becomes aware of ***terrorism content*** a hash of which is not included in its set of hashes, the provider should secure that a hash of that content is added to its database.

ICU C13.10 The provider should take reasonable steps to secure the removal of a hash from any hash database used by the provider, where the provider determines that the hash relates to ***content*** which is not capable of amounting to ***terrorism content***.

ICU C13.11 The provider should ensure an appropriate policy is put in place, and that measures are taken in accordance with that policy, to secure any hashes of ***terrorism content*** held for the purposes of this Recommendation ICU C13 in accordance with ICU C13.8(g).

Technical configuration

DRAFT FOR CONSULTATION

ICU C13.12 In configuring the technology so that its performance strikes an appropriate balance between **precision** and **recall**, the provider should ensure that the following matters are taken into account:

a) the service's **risk** of harm relating to terrorism, reflecting the **risk assessment** of the service and any information reasonably available to the provider about the prevalence of relevant content that is ***terrorism content*** on the service;
b) the proportion of **detected content** that is a **false positive**;
c) the effectiveness of the ***systems and/or processes*** used to identify **false positives**.

ICU C13.13 The provider should ensure that the performance of the **hash matching technology**, and whether the balance between **precision** and **recall** continues to be appropriate, is reviewed at least every six months.

ICU C13.14 The provider should ensure that a written record is made of how this balance has been struck in configuring the **hash matching technology**, including what information has been considered, and information about reviews and steps taken in response.

Safeguards for freedom of expression and privacy

ICU C13.15 Paragraphs ICU C13.5 to ICU C13.14 of this Recommendation ICU C13 are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***.

ICU C13.16 The following measures are also safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendations ICU C1, ICU C2 and ICU C4A;
b) where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8 (in relation to content moderation);
c) Recommendations ICU D1 and ICU D2, so far as they relate to ***content appeals*** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;
d) Recommendations ICU D8 or ICU D9 (whichever is applicable), ICU D10 and, where they are applicable, ICU D10A (in relation to **content appeals**), and ICU D12; and
e) Recommendation ICU G1 (terms of service: substance (all services)).

ICU C14 Using hash matching to detect intimate image abuse content

Application

ICU C14.1 This measure applies to a ***provider*** in respect of each **service** it provides that enables ***regulated user-generated content*** in the form of photographs, videos or visual images (whether or not combined with written material) to be generated, uploaded or shared, and:

DRAFT FOR CONSULTATION

a) is at high **risk** of intimate image abuse, and:

i) the principal purpose of the **service** is the hosting or dissemination of **regulated pornographic content**;

ii) has more than 700,000 monthly **active United Kingdom users** (see paragraphs 5.7 to 5.10); or

iii) is a **file-storage and file-sharing service;** or

b) is a **large service** and is at medium or high **risk** of intimate image abuse.

Key definitions

ICU C14.2 In this Recommendation ICU C14:

"relevant content" means:

a) any ***regulated user-generated content*** in the form of photographs, videos or visual images (whether or not combined with written material) that:

i) may be ***encountered*** by ***United Kingdom users*** of the service by means of the service, and

ii) is communicated publicly[6] by means of the service; or

b) any material which, if it were present on the service, would be content within sub-paragraph (a);

"unverified hash" means a hash which is not a verified hash;

"verified hash" means a hash which was determined to be of **intimate image abuse content** at the time the hash was uploaded to a database.

Recommendation

ICU C14.3 The provider should ensure that, where technically feasible, **hash matching technology** is used effectively (see ICU C14.8) to analyse relevant content to assess whether it is **intimate image abuse content**. For this purpose, the provider should use:

a) **perceptual hash matching technology**; or

b) where the provider's appropriate set of hashes (as defined in ICU C14.9) does not enable perceptual hash matching for video ***content***, **cryptographic hash matching**.

ICU C14.4 In circumstances where:

[6] Ofcom has published **guidance on content communicated 'publicly' and 'privately' under the Online Safety Act** for this purpose.

DRAFT FOR CONSULTATION

a) relevant content matches with an unverified hash; and

b) either of the following apply:

i) if **perceptual hash matching technology** is used, it is the first time there has been a positive match with that hash at that configuration of the technology (see ICU C14.8(b)); or

ii) if **cryptographic hash matching** is used, it is the first time there has been a positive match with that hash;

the provider should treat this as reason to suspect that the relevant content may be ***illegal content*** and review the relevant content in accordance with Recommendation ICU C1.

ICU C14.5 Where relevant content matches with a hash in circumstances other than those set out in ICU C14.4, the provider:

a) may, depending on the level of assurance the provider has in the detection outcomes achieved by the **hash matching technology**, treat the relevant content as ***illegal content*** and swiftly take it down in accordance with ICU C2.3; or

b) should otherwise treat the match as reason to suspect that the relevant content may be ***illegal content*** and review the relevant content in accordance with Recommendation ICU C1.

ICU C14.6 The provider should ensure human moderators review and assess an appropriate proportion of **detected content**, having regard to:

a) the level of assurance the provider has in the detection outcomes achieved by the **hash matching technology** and any associated ***systems and processes*** (as indicated by the ongoing monitoring, evaluation and quality assurance (including human quality assurance) of the performance of the technology); and

b) to the extent that **detected content** is subject to review for the purpose of Recommendation ICU C1:

i) the potential severity of the harm to those depicted in or ***encountering*** **intimate image abuse content**; and

ii) the overall impact of an incorrect decision that the relevant content is ***illegal content*** on a ***United Kingdom user*** who generated, uploaded or shared the content.

ICU C14.7 For the purposes of ICU C14.3, the provider should ensure that:

a) all relevant content present on the service at the time the **hash matching technology** is implemented is analysed within a reasonable time; and

b) relevant content that is generated on, uploaded to or shared on the service (or that a user seeks to so generate, upload or share) after

DRAFT FOR CONSULTATION

the **hash matching technology** is implemented is analysed before or as soon as practicable after it can be ***encountered*** by ***United Kingdom users*** of the service.

ICU C14.8 For the use of **hash matching technology** to be effective, it should:

a) use a suitable hash function to compare the relevant content to an appropriate set of hashes (see ICU C14.9 to ICU C14.12); and

b) where **perceptual hash matching technology** is used, be configured so that its performance strikes an appropriate balance between **precision** and **recall** (see ICU C14.13 to ICU C14.15).

The set of hashes

ICU C14.9 For the set of hashes to be appropriate, it should:

a) contain hashes of a significant number of original items of ***content***;

b) be proactively updated with reasonable regularity; and

c) be secured from unauthorised access, interference and (to the extent the database is comprised of verified hashes) exploitation (whether by persons who work for that person or are providing a service to that person, or any other person).

ICU C14.10 Where the provider becomes aware of **intimate image abuse content** a hash of which is not included on any database used by the provider, the provider should take reasonable steps to secure that a hash of that ***content*** is added to each such database.

ICU C14.11 The provider should take reasonable steps to secure the removal of a hash from any hash database used by the provider where it has determined that the hash is not of **intimate image abuse content** (see ICU C14.12).

ICU C14.12 For the purposes of ICU C14.11, the provider determines that the hash is not of **intimate image abuse content** in any of the following circumstances:

a) the provider views the ***content*** used to generate the hash and determines it is not **intimate image abuse content**;

b) the provider otherwise reasonably believes that the hash is of ***content*** that is not **intimate image abuse content**;

c) the provider views **detected content** matched with the hash using **cryptographic hash matching technology** and determines that that ***content*** is not **intimate image abuse content**; or

d) the provider views **detected content** matched with the hash using **perceptual hash matching technology** and determines that that ***content*** is not **intimate image abuse content**, provided that the provider reasonably believes that there is an exact match between

DRAFT FOR CONSULTATION

the hash of the **detected content** and the hash included on the database.

Technical configuration

ICU C14.13 In configuring the **hash matching technology** so that its performance strikes an appropriate balance between **precision** and **recall**, the provider should ensure that the following matters are taken into account:

a) the service's **risk** of harm relating to intimate image abuse, reflecting the **risk assessment** of the service and any information reasonably available to the provider about the prevalence of relevant content that is **intimate image abuse content** on the service;

b) the proportion of **detected content** that is a **false positive**; and

c) the effectiveness of the ***systems and/or processes*** used to identify **false positives**.

ICU C14.14 The provider should ensure that the performance of the **hash matching technology**, and whether the balance between **precision** and **recall** continues to be appropriate, is reviewed at least every six months.

ICU C14.15 The provider should ensure that a written record is made of how this balance has been struck in configuring the **hash matching technology**, including what information has been considered, and information about reviews and steps taken in response.

Safeguards for freedom of expression and privacy

ICU C14.16 Paragraphs ICU C14.4 to ICU C14.6 and ICU C14.8 to ICU C14.15 of this Recommendation ICU 14 are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***.

ICU C14.17 The following measures are also safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendations ICU C1, ICU C2 and ICU C4A, and where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8 (in relation to content moderation);

b) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy;

DRAFT FOR CONSULTATION

c) Recommendations ICU D8 or ICU D9 (whichever is applicable), ICU D10 and, where applicable, ICU D10A (in relation to **content appeals**) and ICU D12; and

d) Recommendation ICU G1 (terms of service: substance (all services)).

ICU C15 Crisis response

Application

ICU C15.1 This measure applies to a ***provider*** in respect of each **service** it provides that meets either of the following conditions:

a) the **service** is a **large service** at medium **risk** of one or more of the following **kinds of illegal harm**:

i) terrorism;
ii) hate;
iii) harassment, stalking, threats and abuse;
iv) the foreign interference offence; or

b) the **service** is at high **risk** of one or more of the **kinds of illegal harm** listed in sub-paragraph (a).

Key definitions

ICU C15.2 In this Recommendation ICU C15:

"crisis" means an extraordinary situation in which there is a serious threat to public safety in the United Kingdom which is highly likely to:

a) have resulted (in whole or in part) from a significant increase in relevant illegal content on the service; and/or
b) have caused, or cause, a significant increase in relevant illegal content on the service;

"relevant illegal content" means ***illegal content*** associated with any of the following **kinds of illegal harm**:

a) terrorism;
b) hate;
c) harassment, stalking, threats and abuse;
d) the foreign interference offence.

Recommendation

ICU C15.3 The provider should prepare and apply (but need not publish) a written internal protocol ("crisis response protocol") to identify and respond to a crisis, for the purposes of mitigating and managing:

a) the risks arising from a significant increase in relevant illegal content on the service during a crisis; and
b) where relevant, the increased risk of the service being used for the commission or facilitation of a ***priority offence*** during a crisis.

DRAFT FOR CONSULTATION

ICU C15.4 The provider's crisis response protocol should include, but need not be limited to, the following:

a) indicators, identified by the provider, that the provider will consider in determining whether a crisis is occurring or is likely to occur;
b) how the provider will regularly monitor the indicators referred to in sub-paragraph (a) to determine whether a crisis is occurring or is likely to occur;
c) how the provider will keep the indicators referred to in sub-paragraph (a) under regular review to ensure they remain the relevant indicators to be identified in the crisis response protocol;
d) details of a crisis response team, including representatives of sufficient seniority from relevant internal teams, that the provider will deploy if the provider determines that a crisis is occurring or is likely to occur;
e) how the provider will deploy the crisis response team referred to in sub-paragraph (d) as soon as reasonably practicable if the provider determines that a crisis is occurring or is likely to occur;
f) systems and/or processes, identified by the provider, to mitigate and manage the risks set out in ICU 15.3(a) and (b); and
g) how the provider will deploy the systems and/or processes referred to in sub-paragraph (f) as soon as reasonably practicable if the provider determines that a crisis is occurring or is likely to occur.

ICU C15.5 If Ofcom gives a public statement notice to the provider by virtue of section 175(3) of the **Act**, the provider should consider this, in addition to the indicators referred to in ICU C15.4(a), in determining whether a crisis is occurring or is likely to occur.

ICU C15.6 Where the provider determines that any of the indicators referred to in ICU C15.4(a) no longer remain relevant, or where new indicators are identified or existing indicators require updating, it should update its crisis response protocol accordingly.

ICU C15.7 When the provider determines that the crisis has ended, or 90 days after the crisis began, if earlier, the provider should conduct and record (but need not publish) a post-crisis analysis recording key decisions made during the crisis and assessing whether the crisis response protocol remains appropriate for mitigating and managing the risks set out in ICU 15.3(a) and (b). The provider should then use the post-crisis analysis to make changes to the crisis response protocol, as needed, to address any identified deficiencies.

Safeguards for freedom of expression and privacy

ICU C15.8 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) where they are applicable, Recommendations ICU C3, ICU C4, ICU C5, ICU C6, ICU C7, and ICU C8;
b) Recommendations ICU D1 and ICU D2; and
c) Recommendations ICU D8 or ICU D9 (whichever is applicable), ICU D10, and ICU D10A (where applicable).

DRAFT FOR CONSULTATION

ICU C16 Crisis response (dedicated channel for law enforcement)

Application

ICU C16.1 This measure applies to a ***provider*** in respect of each **service** it provides that is a **large service** at medium or high **risk** of one or more of the following **kinds of illegal harm**:

a) terrorism;
b) hate;
c) harassment, stalking, threats and abuse;
d) the foreign interference offence.

Key definitions

ICU C16.2 In this Recommendation ICU C16:

"crisis" means an extraordinary situation in which there is a serious threat to public safety in the United Kingdom which is highly likely to:

a) have resulted (in whole or in part) from a significant increase in relevant illegal content on the service; and/or
b) have caused, or cause, a significant increase in relevant illegal content on the service;

"relevant illegal content" means ***illegal content*** associated with any of the following **kinds of illegal harm**:

a) terrorism;
b) hate;
c) harassment, stalking, threats and abuse;
d) the foreign interference offence.

Recommendation

ICU C16.3 If the provider determines that a crisis is occurring or is likely to occur, it should ensure there is a dedicated channel for law enforcement to contact the provider on matters related to the crisis.

DRAFT FOR CONSULTATION

D. Reporting and complaints

ICU D1 Enabling complaints, including intimate image content reports

Application

ICU D1.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU D1.2 The provider should have ***systems and processes*** which enable **prospective complainants** to make:

(a) each type of **relevant complaint**; and

(b) ***intimate image content reports***

in a way which will secure that the provider will take appropriate action in relation to them.

ICU D1.3 Appropriate action in relation to ***intimate image content reports*** is taking action in accordance with ICU C2A.

ICU D2 Having easy to find, easy to access and easy to use complaints systems and processes

Application

ICU D2.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU D2.2 The ***systems and processes*** referred to in ICU D1.2 in relation to **relevant complaints** should be operated to ensure that:

a) for **relevant complaints** regarding a specific piece of ***content,*** a reporting function or tool is clearly accessible in relation to that ***content***;
b) processes for making other kinds of **relevant complaints** are easy to find and easily accessible;
c) they are designed so that they only include reasonably necessary steps; and
d) it is possible when making **relevant complaints** to give the provider supporting information.

DRAFT FOR CONSULTATION

ICU D2.3 In designing the ***systems and processes*** referred to in ICU D1.2, including its reporting tool or function, the provider should consider the accessibility needs of its ***United Kingdom user*** base having regard to:

a) the groups of people its **risk assessment** has identified as using the service;
b) in the case of a service that is ***likely to be accessed by children***, the service's **children's risk assessment**;
c) other relevant information the provider holds on its ***United Kingdom user*** base;
d) industry standards and good practice as to the design of the service, to ensure the reporting and complaints process is accessible to disabled people; and
e) comprehensibility, based on the likely reading age of the youngest individual permitted to use the service without the consent of a parent or guardian.

ICU D2.4 For the purposes of ICU D2.3(d), the ***systems and processes*** referred to in ICU D1.2 should be designed for the purposes of ensuring usability for those dependent on assistive technologies including:

a) keyboard navigation; and
b) screen reading technology.

ICU D2.5 The ***systems and processes*** referred to in ICU D1.2 in relation to ***intimate image content reports*** should be operated to ensure that:

a) **prospective complainants** are clearly signposted to the appropriate channel for making an ***intimate image content report***;
b) processes for making an ***intimate image content report*** regarding a specific piece of ***content*** are clearly accessible in relation to that content;
c) they are designed so that they only include reasonably necessary steps; and
d) it is possible when making ***intimate image content reports*** to give the provider:
i) all information necessary for the **report** to be considered an ***intimate image content report***; and
ii) supporting information.

ICU D3 Provision of information prior to the submission of a complaint

Application

ICU D3.1 This measure applies to a ***provider*** in respect of each **service** that is ***likely to be accessed by children*** it provides that is either (or both) of the following:

a) a **large service**; or

DRAFT FOR CONSULTATION

b) at medium or high **risk** of any **kind of illegal harm**.

Recommendation

ICU D3.2 The provider should ensure that the reporting function or tool for **relevant complaints** regarding a specific piece of ***content*** enables **prospective complainants** to easily access information on the following matters prior to the submission of a complaint:

a) whether the provider discloses (either routinely, upon request or otherwise) the fact that a complaint relating to specific ***content*** has been submitted to:

i) the ***user*** that generated, uploaded or shared the ***content*** complained about; or
ii) any ***user*** other than the **complainant**;

and, if so,

iii) the circumstances in which the provider makes the relevant disclosure; and
iv) the information disclosed about the complaint and the person that submitted the complaint; and

b) the information about the complaint, and the person that submitted the complaint, that the provider discloses to a person bringing a **relevant complaint** which is a **content appeal**.

ICU D3A Provision of information prior to the submission of an intimate image content report

Application

ICU D3A.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU D3A.2 The provider should ensure that the ***systems and processes*** referred to in ICU D1.2 in relation to ***intimate image content reports*** enable **prospective complainants** to easily access information on the following matters prior to the submission of an ***intimate image content report***:

a) the information a **prospective complainant** must give the provider in order for the **report** to be considered an ***intimate image content report***, in particular:
i) a declaration that the content present on the service is **intimate image abuse content**;

DRAFT FOR CONSULTATION

ii) a declaration that the **report** is made by the subject of the content or a person acting on that person's behalf;
iii) a declaration that the **report** is made in good faith and is, to the best of the knowledge and belief of the person making the **report**, true;
iv) sufficient information about the content for the provider to identify it; and
v) contact details for the **prospective complainant** making the ***intimate image content report***; and

b) that this is the channel for making ***intimate image content reports,*** and the provider's commitment to ***take down*** **intimate image abuse content** in accordance with ICU C2A.

ICU D4 Appropriate action – sending indicative timeframes

Application

ICU D4.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) at medium or high **risk** of any **kind of illegal harm**.

Recommendation

ICU D4.2 The provider should acknowledge receipt of each **relevant complaint** and provide the **complainant** with an indicative timeframe for deciding the complaint.

ICU D4.3 ICU D4.2 does not apply if:

a) the ***provider's*** acknowledgement is non-ephemeral; and
b) the **complainant** has opted out from receiving non-ephemeral communications in relation to their complaint.

ICU D5 Appropriate action – sending further information about how the complaint will be handled

Application

ICU D5.1 This measure applies to a ***provider*** in respect of each **service** that is ***likely to be accessed by children*** it provides that is either (or both) of the following:

a) a **large service**; or
b) at medium or high **risk** of any **kind of illegal harm**.

DRAFT FOR CONSULTATION

Recommendation

ICU D5.2 In the acknowledgment of receipt of each **relevant complaint**, referred to in Recommendation ICU D4, the provider should set out:

a) the possible outcomes; and
b) confirmation of whether the provider will inform the **complainant** of its decision whether to uphold the complaint and details of any action taken as a result.

ICU D6 Opt-out from communications following a complaint

Application

ICU D6.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) at medium or high **risk** of any **kind of illegal harm**.

Recommendation

ICU D6.2 The provider should enable the **complainant** to opt out of receiving any non-ephemeral communications in relation to a **relevant complaint**.

ICU D7 Appropriate action for relevant complaints about suspected illegal content

Application

ICU D7.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU D7.2 When the provider receives a **relevant complaint** about ***content*** which may be ***illegal content*****:**

a) it should treat the complaint as reason to suspect that the ***content*** may be ***illegal content***, and review the ***content*** in accordance with ICU C1.3; and

b) if Recommendations ICU C4 and ICU C5 are not applicable to the provider, it should consider the complaint promptly.

DRAFT FOR CONSULTATION

ICU D7.3 ICU D7.2 does not apply to a complaint identified as manifestly unfounded in accordance with ICU D13.2.

ICU D8 Appropriate action for relevant complaints which are content appeals – determination (large or multi-risk services)

Application

ICU D8.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**; or
b) a **multi-risk service**.

Recommendation

ICU D8.2 The provider should determine **relevant complaints** which are **content appeals**.

ICU D8.3 The provider should, as a minimum, monitor its performance against performance targets relating to the following:

a) the time it takes to determine the **content appeal;** and
b) the accuracy of decision making,

and should resource itself so as to give effect to those targets.

ICU D8.4 The provider should have regard to the following matters in determining what priority to give to review of a **relevant complaint** which is a **content appeal**:

a) the seriousness of the action taken against the ***user*** or in relation to the ***content*** (or both) as a result of the decision that the ***content*** was ***illegal content*** or **illegal content proxy**;

b) whether the decision that the ***content*** was ***illegal content*** or **illegal content proxy** was made by ***content identification technology*** and, if so:

i) any information that Ofcom has recommended the provider collect about the likelihood of false positives generated by the specific ***content identification technology*** used; and

ii) any other information available about the accuracy of the ***content identification technology*** at identifying similar types of ***illegal content*** or **illegal content proxy**; and

c) the past error rate on the service in relation to **illegal content judgements** or judgements about whether ***content*** is an **illegal content proxy** of the type concerned.

DRAFT FOR CONSULTATION

ICU D8.5 In determining **relevant complaints** which are **content appeals** relating to the ***United Kingdom user*** being **sanctioned**, the provider should consider whether, in the circumstances of the case, the **sanction** that has been imposed is appropriate, having regard to its policy on **sanctions**.

ICU D9 Appropriate action for relevant complaints which are content appeals – determination (services that are neither large nor multi-risk)

Application

ICU D9.1 This measure applies to a ***provider*** in respect of each **service** it provides that is neither a **large service** nor a **multi-risk service**.

Recommendation

ICU D9.2 The provider should determine **relevant complaints** which are **content appeals** promptly.

ICU D10 Appropriate action for relevant complaints which are content appeals – action following determination

Application

ICU D10.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU D10.2 If, in relation to a **relevant complaint** that is a **content appeal**, the provider reverses a decision that ***content*** was ***illegal content*** or **illegal content proxy**, the provider should:

a) so far as appropriate and possible for the purpose of restoring the position of the ***content*** or ***user*** (or both) to what it would have been had the decision not been made, reverse the action taken against the ***user*** or the ***content*** (or both) as a result of that decision;
b) where there is a pattern or significant evidence of ***regulated user-generated content*** being taken down in error, adjust any relevant content moderation guidance if appropriate to ensure it is accurate; and
c) where possible and appropriate, take steps to secure that the use of automated content moderation technology does not cause the same ***content*** to be taken down again.

DRAFT FOR CONSULTATION

ICU D10.3 If, in relation to a **relevant complaint** that is a **content appeal** relating to a ban imposed on a ***user*** under Recommendation ICU H3, the provider does not reverse the determinations made under ICU H3.3(a) and (b), the provider may:

a) decide to reverse the ban against the ***user***, if it has reasonable grounds to believe that the user is a ***child*** and one of the extenuating circumstances referred to in ICU H3.6 applies;

b) decide to reverse the ban against the ***user***, if they were banned for receiving and viewing **CSAM** or **CSAM content proxy** and ICU H3.5 applies;

c) redetermine the duration of the ban in accordance with ICU H3.7 or H3.8 (as applicable), if it determines that the duration originally imposed under ICU H3.7 or H3.8 is not appropriate.

ICU D10A Appropriate action for relevant complaints which are content appeals – action following determination upholding the appeal, except for CSEA content or CSEA content proxy (large or multi-risk services)

Application

ICU D10A.1 This measure applies to a ***provider*** in respect of each **service** it provides that meets either (or both) of the following:

a) a **large service**;

b) a **multi-risk service**.

Recommendation

ICU D10A.2 If the provider upholds, in whole or in part, a **relevant complaint** that is a **content appeal** relating to the ***user*** being **sanctioned** under the provider's policy on **sanctions** (see Recommendation ICU H2), but does not reverse its decision that ***content*** was ***illegal content*** or **illegal content proxy**, the provider should:

a) reverse the action taken against the ***user***, and

b) apply the appropriate **sanction** having regard to its policy on **sanctions** (see Recommendation ICU H2).

This measure does not apply to ***CSEA content*** or **CSEA content proxy** (see ICU D10.3).

ICU D11 Appropriate action for relevant complaints about proactive technology, which are not content appeals

Application

DRAFT FOR CONSULTATION

ICU D11.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU D11.2 This Recommendation ICU D11 applies to **relevant complaints**, which are not **content appeals**, about the use of ***proactive technology*** on the service when:

a) the use of ***proactive technology*** on the service results in ***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
b) the **complainant** 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 blocking ***content*** not of a kind specified in the ***terms of service*** as a kind of ***content*** in relation to which the technology would operate).

ICU D11.3 The provider should inform the **complainant** of:

a) the action the provider may take in response to the complaint; and

b) their right, if they consider the provider to be in breach of contract, to bring proceedings.

ICU D11.4 ICU D11.3 does not apply to a complaint identified as manifestly unfounded in accordance with ICU D13.2.

ICU D12 Appropriate action for all other relevant complaints

Application

ICU D12.1 This measure applies to a ***provider*** in respect of each **service** that it provides.

Recommendation

ICU D12.2 This Recommendation ICU D12 applies to **relevant complaints** that the provider is not complying with:

a) the **illegal content safety duties;**
b) the duty about content reporting set out in section 20 of the ***Act***, so far as it relates to ***illegal content***;
c) the provider's duty in relation to freedom of expression in section 22(2) of the **Act**; or
d) the provider's duty in relation to privacy in section 22(3) of the **Act**.

DRAFT FOR CONSULTATION

ICU D12.3 The provider should nominate a responsible individual or a team to ensure that such complaints are directed to an appropriate individual or team to be processed.

ICU D12.4 **Relevant complaints** should be handled:

a) in a way that protects ***United Kingdom users***; and
b) within timeframes the provider has determined are appropriate.

ICU D12.5 ICU D12.3 and ICU D12.4 do not apply in relation to a complaint identified as manifestly unfounded in accordance with ICU D13.2.

ICU D12A Expedited complaints procedure in relation to intimate image abuse content

Application

ICU D12A.1 This measure applies to a ***provider*** in respect of each **service** that it provides.

Recommendation

ICU D12A.2 In the case of complaints that are made by ***users*** or ***affected persons*** who have made an ***intimate image content report*** and are about the content to which the report relates, the provider should have an expedited complaints procedure in relation to:

a) complaints about content present on a service which **prospective complainants** consider to be ***illegal content***,
b) complaints that the provider is not complying with the **illegal content safety duties**, and
c) complaints that the provider is not complying with the duty about content reporting set out in section 20 of the **Act**.

ICU D12A.3 The provider should set and record performance targets in their ***terms of service*** for resolving the complaints relating to **intimate image abuse content** referred to in ICU D12A.2 that are faster than the performance targets for resolving the **relevant complaints** referred to in ICU D12.2.

ICU D12A.4 The expedited complaints procedure referred to in ICU D12A.2 should be easy to access, easy to use and transparent.

ICU D12A.5 When a provider receives a complaint referred to in ICU D12A.2(a) which relates to the continued presence of content on the service which the provider has

DRAFT FOR CONSULTATION

determined to be **intimate image abuse content**, it should ***take down*** that content as soon as reasonably practicable.

ICU D12A.6 The provider should nominate a responsible individual or a team to ensure that complaints referred to in sub-paragraphs (b) or (c) of ICU D12A.2 are directed to an appropriate individual or team to be processed.

ICU D13 Exception: manifestly unfounded complaints

Application

ICU D13.1 This measure applies to a ***provider*** in respect of each **service** that it provides.

Recommendation

ICU D13.2 When the provider receives a **relevant complaint** that is not a **content appeal** or an **age assessment appeal**, it may disregard the complaint only if:

a) the provider has prepared and implemented a policy in accordance with ICU D13.3 to ICU D13.6, setting out the information and attributes that indicate a **relevant complaint** is manifestly unfounded;
b) the provider identifies the complaint as manifestly unfounded in accordance with that policy; and
c) the provider has in place a process to monitor the degree to which the application of the policy incorrectly identifies complaints which are not manifestly unfounded, and to review the policy in accordance with ICU D13.4 to ICU D13.6.

ICU D13.3 In designing a policy for the purposes of ICU D13.2(a), the provider should have regard to:

a) the need to identify manifestly unfounded complaints accurately; and
b) the risks posed to particular groups of vulnerable users if **relevant complaints** are incorrectly identified as manifestly unfounded.

ICU D13.4 The provider should, at minimum, carry out an annual review of the policy to ensure it is not incorrectly identifying **relevant complaints** as manifestly unfounded.

ICU D13.5 If the policy is incorrectly identifying **relevant complaints** as manifestly unfounded, the provider should make changes to it with a view to ensuring its accuracy.

ICU D13.6 The provider should keep a record of its review process and any changes it has made.

DRAFT FOR CONSULTATION

ICU D14 Dedicated reporting channel for trusted flaggers to report fraud

Application

ICU D14.1 This measure applies to a ***provider*** in respect of each **service** it provides that is a **large service** and is at medium or high **risk** of fraud.

Recommendation

ICU D14.2 In this Recommendation ICU D14, a '**recommended trusted flagger**' is each of the following:

a) the City of London police force;
b) the Dedicated Card and Payment Crime Unit (a joint team of the City of London and Metropolitan Police forces);
c) the Department for Work and Pensions;
d) the Financial Conduct Authority;
e) HM Revenue and Customs;
f) the National Crime Agency;
g) the National Cyber Security Centre (a part of the Government Communications Headquarters);
h) the Police Service of Northern Ireland;
i) the Police Service of Scotland (Seirbheis Phoilis na h-Alba).

ICU D14.3 The provider should establish and maintain a dedicated reporting channel for, at minimum, the **recommended trusted flaggers** and relating to, at minimum, fraud, in the circumstances set out in this Recommendation ICU D14.

ICU D14.4 The provider should publish a clear and accessible policy on its processes relating to the establishment of a dedicated reporting channel for, at minimum, the **recommended trusted flaggers**, covering any relevant procedural matters.

ICU D14.5 If a request is made in accordance with the policy by a **recommended trusted flagger**, the provider should ensure a dedicated reporting channel, run in accordance with ICU D14.3 to ICU D14.8, is made available and maintained for, at minimum, **recommended trusted flaggers**. The provider may make an existing dedicated reporting channel available to the **recommended trusted flagger**, if that dedicated reporting channel is run in accordance with ICU D14.3 to ICU D14.8.

ICU D14.6 The provider should engage with the **recommended trusted flagger** at the start of the relationship to understand the **recommended trusted flagger**'s needs with respect to the dedicated reporting channel.

DRAFT FOR CONSULTATION

ICU D14.7 At least every two years, the provider should seek feedback from, at minimum, the **recommended trusted flaggers** with which it has made such arrangements, on whether any reasonable adjustments or improvements might be made to the operation of the dedicated reporting channel.

ICU D14.8 ICU D14.9 applies where the provider receives a complaint from a **trusted flagger** through a dedicated reporting channel established for that **trusted flagger** if the complaint:

a) is about specific ***content*** that is ***regulated user-generated content*** on the service which may be encountered by ***United Kingdom users***; and
b) relates to a matter within the area of expertise of the **trusted flagger**.

ICU D14.9 The provider should treat the complaint as reason to suspect that the ***content*** may be ***illegal content*** and review the ***content*** in accordance with Recommendation ICU C1.

ICU D15 Appropriate action for age assessment appeals (services that are large or multi-risk)

Application

ICU D15.1 This measure applies to a ***provider*** in respect of each **service** that:

a) uses an **age assurance process** to determine which ***United Kingdom users*** of the service are ***children***; and

b) is either (or both) of the following:

i) a **large service**;

ii) a **multi-risk service**.

Recommendation

ICU D15.2 The provider should have regard to the following matters in determining what priority to give to consideration of an **age assessment appeal**:

a) the seriousness of the impact on the ***user*** of any decision made on the basis of the assessment of the ***user***'s age;
b) whether the decision made on the basis of the ***user's*** age was made without human oversight and, if so, information available about the accuracy of the specific technology used in making age assessments;
c) the past error rate on the service in relation to age assessments; and
d) any representations made by the ***user*** as part of the **age assessment appeal** as to the impact on the ***user***'s livelihood from any decision made on the basis of the assessment of the ***user***'s age.

DRAFT FOR CONSULTATION

ICU D15.3 The provider should, as a minimum, monitor its performance against performance targets relating to the following:

a) the time it takes to determine the **age assessment appeal**, and
b) the accuracy of decision making,

and should resource itself so as to give effect to those targets.

ICU D15.4 If the provider determines that the ***user***'s age was incorrectly assessed, it should take any necessary steps to reverse the action taken in relation to the ***user*** as a result of that incorrect assessment.

ICU D15.5 The provider should monitor trends in **age assessment appeals** to help improve any **age assurance process** used on the service.

ICU D16 Appropriate action for age assessment appeals (services that are neither large nor multi-risk)

Application

ICU D16.1 This measure applies to a ***provider*** in respect of each **service** that:

a) uses an **age assurance process** to determine which ***United Kingdom users*** of the service are ***children***; and

b) is neither a **large service** nor a **multi-risk service**.

Recommendation

ICU D16.2 The provider should determine **age assessment appeals** promptly.

ICU D16.3 If the provider determines that the ***user***'s age was incorrectly assessed, it should take any necessary steps to reverse the action taken in relation to the ***user*** as a result of that incorrect assessment.

ICU D16.4 The provider should monitor trends in **age assessment appeals** to help improve any **age assurance process** used on the service.

ICU D17 Livestreaming reporting for imminent physical harm

Application

ICU D17.1 This measure applies to a ***provider*** in respect of each **service** that it provides that meets both of the following conditions:

DRAFT FOR CONSULTATION

a) the service has **one-to-many livestreaming**; and

b) the service is at medium or high **risk** of one or more of the following **kinds of illegal harm**:

i) terrorism;

ii) grooming;

iii) image-based CSAM;

iv) encouraging or assisting suicide (or attempted suicide) and serious self-harm;

v) hate;

vi) harassment, stalking, threats and abuse;

vii) animal cruelty.

Recommendation

ICU D17.2 The provider should ensure that the reporting function or tool for **relevant complaints** regarding a specific piece of ***content*** in a **one-to-many livestream** enables ***United Kingdom users*** to indicate a risk of imminent physical harm to a person (or persons) or an animal (or animals).

DRAFT FOR CONSULTATION

E. Recommender systems

ICU E1 Collection of safety metrics during on-platform testing of content recommender systems

Application

ICU E1.1 This measure applies to a ***provider*** in respect of each **service** it provides that meets both of the following conditions:

a) the provider conducts **on-platform testing** of **content recommender systems** on the service; and
b) the service is at medium or high **risk** of the **kinds of illegal harm** specified in two or more of the following paragraphs:

i) terrorism;
ii) image-based CSAM or CSAM URLs;
iii) encouraging or assisting suicide (or attempted suicide) and serious self-harm;
iv) hate;
v) harassment, stalking, threats and abuse;
vi) drugs and psychoactive substances;
vii) extreme pornography;
viii) intimate image abuse; or
ix) the foreign interference offence.

Recommendation

ICU E1.2 The provider should produce and analyse safety metrics when conducting **on-platform testing** of an actual or proposed **content recommender system design adjustment**.

ICU E1.3 The safety metrics should enable the provider to understand whether a **content recommender system design adjustment** would increase the risk of ***United Kingdom users encountering illegal content***, compared with the existing variant of the **content recommender system**, and include the following (or equivalent):

a) the total number of individual items of ***regulated user-generated content*** that are assessed and identified as ***illegal content*** or as **illegal content proxy** in response to a complaint made during the testing period; and
b) for each such item:

i) the number of times that ***regulated user-generated content*** was displayed to ***users*** (impressions);
ii) the number of unique ***users*** that the ***regulated user-generated content*** was displayed to (reach).

DRAFT FOR CONSULTATION

ICU E1.4 The provider should ensure that:

a) the testing environment is set up in a way that enables the processing of complaints about ***content*** suspected to be ***illegal content*** or **illegal content proxy**;
b) the period during which **on-platform testing** is conducted is sufficient to allow for complaints to be received; and
c) information that is relevant to producing the safety metrics that achieve the outcome described in ICU E1.3 is retained during the testing period.

ICU E1.5 The provider should maintain a **log** of the results of each on-platform test, which should include a record of:

a) the safety metrics produced against each variant of the **recommender system** tested;
b) a description of each variant of the **content recommender system**, including its respective design characteristics; and
c) the design decision taken on which variant of the **recommender system** to deploy following **on-platform testing**.

ICU E1.6 The provider should ensure that the log is:

a) made available and is easily accessible to individuals working for the provider involved directly or indirectly in the development and testing of **content recommender systems**; and
b) referred to by relevant individuals working for the provider in the context of future **content recommender system design adjustments**.

Safeguard for freedom of expression and privacy

ICU E1.7 Recommendation ICU C4A is a safeguard to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***.

ICU E2 Content recommender systems: excluding priority illegal content (terrorism and hate)

Application

ICU E2.1 This measure applies to a ***provider*** in respect of each **service likely to be accessed by children** it provides that meets both of the following conditions:

a) the service is at medium or high **risk** of one or more of the following **kinds of illegal harm**:

i) terrorism;
ii) hate;

DRAFT FOR CONSULTATION

b) a **child-accessible part** of the service has a **content recommender system**.

Key definitions

ICU E2.2 In this Recommendation ICU E2:

"relevant content" means ***regulated user-generated content*** which amounts to one of the **kinds of illegal harm** set out in ICU E2.1(a);

"relevant user" means any ***United Kingdom user*** of the service, other than a ***user*** determined to be an ***adult*** by the use of **highly effective age assurance**.

Recommendation

ICU E2.3 The provider should ensure that any **content recommender system** on a **child-accessible part** of the service is designed and operated so that content indicated potentially to be relevant content on the basis of **relevant available information** is excluded from relevant users' content feeds.

ICU E2.4 For the purposes of ICU E2.3, the provider should:

a) identify what **relevant available information** exists, and
b) ensure that the design of the **content recommender system** takes appropriate account of that **relevant available information**.

ICU E2.5 If the service's **content moderation function** reviews content indicated (for the purposes of this Recommendation) potentially to be relevant content and determines it not to be **illegal content** in accordance with ICU C1.3, this Recommendation no longer applies to that content.

ICU E2.6 This Recommendation does not recommend the use of any specific kind of ***proactive technology***, or the use of ***proactive technology*** to analyse ***user-generated content*** communicated privately or metadata relating to ***user-generated content*** communicated privately.[7]

Safeguards for freedom of expression and privacy

ICU E2.7 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendations ICU B1, ICU C1, ICU C2 and ICU C4A;
b) where they are applicable, ICU C3, ICU C4, ICU C5, C6, ICU C7 and ICU C8;

[7] Ofcom has published **guidance on content communicated 'publicly' and 'privately' under the Online Safety Act**.

DRAFT FOR CONSULTATION

c) Recommendations ICU D1 and ICU D2, so far as they relate to complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy, ICU D10A, and ICU D12.

ICU E3 Content recommender systems: excluding priority illegal content (encouraging or assisting suicide (or attempted suicide) and serious self-harm)

Application

ICU E3.1 This measure applies to a ***provider*** in respect of each **service likely to be accessed by children** it provides that meets both of the following conditions:

a) the service is at medium or high **risk** of encouraging or assisting suicide (or attempted suicide) and serious self-harm;

b) a **child-accessible part** of the service has a **content recommender system**.

Key definitions

ICU E3.2 In this Recommendation ICU E3:

"relevant content" means ***regulated user-generated content*** which amounts to encouraging or assisting suicide (or attempted suicide) and serious self-harm;

"relevant user" means any ***United Kingdom user*** of the service, other than a ***user*** determined to be an ***adult*** by the use of **highly effective age assurance**.

Recommendation

ICU E3.3 The provider should ensure that any **content recommender system** on a **child-accessible part** of the service is designed and operated so that content indicated potentially to be relevant content on the basis of **relevant available information** is excluded from relevant users' content feeds.

ICU E3.4 For the purposes of ICU E3.3, the provider should:

a) identify what **relevant available information** exists, and
b) ensure that the design of the **content recommender system** takes appropriate account of that **relevant available information**.

ICU E3.5 If the service's **content moderation function** reviews content indicated (for the purposes of this Recommendation) potentially to be relevant content and

DRAFT FOR CONSULTATION

determines it not to be **illegal content** in accordance with ICU C1.3, this Recommendation no longer applies to that content.

ICU E3.6 This Recommendation does not recommend the use of any specific kind of ***proactive technology,*** or the use of ***proactive technology*** to analyse ***user-generated content*** communicated privately or metadata relating to ***user-generated content*** communicated privately.

Safeguards for freedom of expression and privacy

ICU E3.7 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendations ICU B1, ICU C1, ICU C2 and ICU C4A;

b) where they are applicable, ICU C3, ICU C4, ICU C5, C6, ICU C7 and ICU C8;

b) Recommendations ICU D1 and ICU D2, so far as they relate to complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy, ICU D10A, and ICU D12.

DRAFT FOR CONSULTATION

F. Settings, functionalities and user support

ICU F1 Safety defaults for child users

Application

ICU F1.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either of the following:

a) at high **risk** of grooming; or
b) a **large service** that is at medium **risk** of grooming.

Key definitions

ICU F1.1A In this Recommendation ICU F1:

"relevant user" means any ***United Kingdom user*** of the service, other than a ***user*** determined to be an ***adult*** by the use of **highly effective age assurance**;

"relevant user account" means a user account registered to a relevant user.

Recommendation

ICU F1.2 If the service has **network expansion prompt functionality**, the provider should implement **default settings** ensuring that:

a) **network expansion prompts** do not recommend relevant user accounts to **connect** with; and
b) relevant users are not presented with **network expansion prompts**.

ICU F1.3 If the service has **connection lists**, the provider should implement **default settings** ensuring that:

a) **connection lists** do not include relevant user accounts; and
b) **connection lists** associated with relevant user accounts are not displayed to ***users***.

ICU F1.4 If the service has **direct messaging functionality**, the provider should implement **default settings** ensuring that:

a) if the service has **user connection functionality**, relevant user accounts can only receive **direct messages** from user accounts with which they have a **specified connection**;
b) if the service does not have **user connection functionality**, relevant users are provided with a means of actively confirming whether to receive a **direct message** sent from another user account before it is visible to them,

DRAFT FOR CONSULTATION

unless **direct messaging** is a necessary and time critical element of another functionality, in which case before any interaction associated with that functionality begins, relevant users should:

i) be informed that they may receive **direct messages** from user accounts that are not **connected** to that relevant user account when using that functionality; and

ii) having received that information, actively confirm that that they wish to proceed to use that functionality.

ICU F1.5 For the purposes of ICU F1.4, a relevant user account has a **specified connection** with another user account if either:

a) the relevant user operating that relevant user account has taken action to initiate the establishment of a **connection** with the other user account (for example through 'friending', 'following', or 'subscribing'); or

b) the relevant user operating that relevant user account has taken action to confirm the establishment of a **connection** with the other user account (for example by accepting a 'friend request' or a request to 'follow' or 'subscribe' to the other user account).

ICU F1.6 If the service has **automated location information display functionality**, the provider should implement **default settings** ensuring the **location information** associated with a relevant user account is not visible to other ***users*** of the service.

Safeguards for freedom of expression and privacy

ICU F1.7 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendation ICU B1;

b) Recommendations ICU D15 and D16.

ICU F2 Support for child users

Application

ICU F2.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either of the following:

a) at high **risk** of grooming; or
b) a **large service** that is at medium **risk** of grooming.

Key definitions

DRAFT FOR CONSULTATION

ICU F2.1A In this Recommendation ICU F2:

"relevant user" means any ***United Kingdom user*** of the service, other than a ***user*** determined to be an ***adult*** by the use of **highly effective age assurance**;

"relevant user account" means a user account registered to a relevant user.

Recommendation

ICU F2.2 The provider should ensure that, when a ***user*** seeks to disable a **default setting** set out in Recommendation ICU F1 on a relevant user account, that ***user*** is provided with information regarding the potential risk involved before being able to disable the relevant setting. The information provided should assist the ***user*** in understanding the implications of disabling that default setting, including the protections it affords.

ICU F2.3 The provider should ensure that, when a ***user*** operating a relevant user account seeks to respond to a request to establish a **connection**, that ***user*** is provided with the following information before the **connection** is established:

a) the types of interactions that would be enabled through establishing a **connection**; and
b) the options available to take action against a ***user's*** user account, such as **blocking, muting**, **reporting conduct** or equivalent action.

ICU F2.4 The provider should provide the following information when a relevant user sends a **direct message** to or receives a **direct message** from a user account for the first time:

a) a reminder that this is the first direct communication sent to or from that user account (as applicable); and
b) the options available to take action against a ***user*** or user account, such as **blocking, muting, reporting conduct** or equivalent action,

unless **direct messaging** is a necessary and time critical element of another functionality, in which case the relevant user should be provided with this information before any interaction associated with that functionality begins.

ICU F2.5 The provider should provide the following information when a relevant user seeks to **block**, **mute**, **report conduct**, or take equivalent action against a ***user*** or user account:

a) the effect of the action, including the types of interactions that it would restrict and whether the other ***user's*** user account would be notified; and
b) the further options available to limit interaction with another ***user's*** user account or increase the relevant user's safety.

ICU F2.6 The provider should ensure that the information provided in accordance with ICU F2.2 to ICU F2.5 is:

DRAFT FOR CONSULTATION

a) prominently displayed; and
b) clear, comprehensible and easy for a **child user** to understand.

Safeguards for freedom of expression and privacy

ICU F2.7 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendation ICU B1;
b) Recommendations ICU D15 and D16.

DRAFT FOR CONSULTATION

G. Terms of service

ICU G1 Terms of service: substance (all services)

Application

ICU G1.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU G1.2 The provider should include the following in the ***terms of service***:

a) provisions specifying how individuals are to be protected from ***illegal content***, addressing:

i) separately for each of ***terrorism content***, ***CSEA content*** and other ***priority illegal content***, how the provider will minimise the length of time for which any ***priority illegal content*** is present; and

ii) how, where the provider is alerted by a person to the presence of any ***illegal content***, or becomes aware of it in any other way, it will swiftly ***take down*** such content.

b) provisions giving information about any ***proactive technology*** used for the purposes of compliance with an illegal content safety duty set out in section 10(2) or 10(3) of the **Act** (including the kind of technology, when it is used, and how it works);

c) provisions specifying the policies and processes that govern the handling and resolution of **relevant complaints**; and

d) provisions giving information about its policy on banning ***users*** in relation to ***CSEA content*** or **CSEA content proxy**, or both (see Recommendation ICU H3).

ICU G1A Terms of service: substance (large or multi-risk services)

Application

ICU G1A.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**;

b) a **multi-risk service**.

Recommendation

DRAFT FOR CONSULTATION

ICU G1A.2 The provider should include in the ***terms of service*** provisions giving information about its policy on **sanctions** (see Recommendation ICU H2), addressing, at least, the **sanctions** which the provider may impose on ***users*** who generate, upload or share ***content*** that it determines to be ***illegal content*** or **illegal content proxy**, other than ***CSEA content*** or **CSEA content proxy** (see ICU G1.2 for ***terms of service*** relating to ***CSEA content*** or **CSEA content proxy**).

ICU G2 Terms of service: substance (Category 1 services)

Application

ICU G2.1 This measure applies to a ***provider*** in respect of each ***Category 1 service*** it provides.

Recommendation

ICU G2.2 The provider should summarise the findings of its **risk assessment** (including as to levels of risk and as to the nature, and severity, of potential harm) in the ***terms of service***.

ICU G3 Terms of service: clarity and accessibility

Application

ICU G3.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU G3.2 The provider should ensure that the provisions included in the ***terms of service*** in accordance with Recommendation ICU G1 are:

a) easy to find, such that they are:

   i) clearly signposted for the general public, regardless of whether they have registered with or are using the service; and
   ii) locatable within the ***terms of service***;

b) laid out and formatted in a way that helps ***United Kingdom users*** read and understand them;
c) written to a reading age comprehensible for the youngest individual permitted to use the service without the consent of a parent or guardian; and
d) designed for the purposes of ensuring usability for those dependent on assistive technologies, including:

   i) keyboard navigation; and
   ii) screen reading technology.

DRAFT FOR CONSULTATION

H. User access

ICU H1 Removing accounts of proscribed organisations

Application

ICU H1.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Recommendation

ICU H1.2 In this Recommendation ICU H1, "relevant content" means ***content*** that a ***provider*** has determined:

a) is **proscribed organisation content**; or
b) is in breach of ***terms of service*** designed to prohibit **proscribed organisation content** on the service.

ICU H1.3 ICU H1.4 applies where:

a) relevant content has been generated, uploaded or shared using a user account on the service; or
b) the provider has become aware that a user account on the service may be operated by or on behalf of a **proscribed organisation** (including as a result of a **report** or complaint).

ICU H1.4 The provider should consider whether it has reasonable grounds to infer that the user account in question is operated by or on behalf of a **proscribed organisation**.

ICU H1.5 Where the provider has reasonable grounds to infer that a user account is operated by or on behalf of a **proscribed organisation**, it should remove the user account from the service.

ICU H1.6 Reasonable grounds to infer that a user account is operated by or on behalf of a **proscribed organisation** may arise where at least two of the following are true of the **user profile**:

a) the username is the same as: (1) the name of a **proscribed organisation**; or (2) an alias as specified in an order made under section 3(6) of the Terrorism Act 2000;[8]
b) the **user profile** image or any end-user configurable image setting is **proscribed organisation content**;
c) the **user profile** information, such as 'bio' text or other descriptive text, is **proscribed organisation content**.

[8] 2000 c.11.

DRAFT FOR CONSULTATION

ICU H1.7 Reasonable grounds may also arise where one or none of the above is true, but where a significant proportion of a reasonably sized sample of the ***regulated user-generated content*** recently generated, uploaded or shared on the user account is **proscribed organisation content**. What amounts to a reasonable sample size will depend on the amount of ***regulated user-generated content*** generated, uploaded or shared on the account and the nature of the service. "***Regulated user-generated content*** recently generated, uploaded or shared by the account" refers to the newest ***regulated user-generated content*** generated, uploaded or shared by the user account irrespective of date, rather than the ***regulated user-generated content*** generated, uploaded or shared by the user account in a recent date range.

ICU H1.8 References to "***regulated user-generated content***" in ICU H1.7 do not include ***regulated user-generated content*** that has been communicated privately,[9] unless the provider has explicit consent to view the ***content*** in question.

Safeguards for freedom of expression and privacy

ICU H1.9 The following measures are safeguards to protect ***United Kingdom users***' right to freedom of expression and the privacy of ***United Kingdom users***:

za) Recommendation ICU C4A;

a) Recommendations ICU D1 and ICU D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy; and

b) Recommendations ICU D8 or ICU D9 (whichever is applicable) and ICU D10 (in relation to **content appeals**).

ICU H1.10 ICU H1.8 is a safeguard to protect the privacy of ***United Kingdom users***.

ICU H2 User sanctions

Application

ICU H2.1 This measure applies to a ***provider*** in respect of each **service** it provides that is either (or both) of the following:

a) a **large service**;

b) a **multi-risk service**.

Recommendation

ICU H2.2 The provider should prepare and apply a policy in respect of **sanctions** to be imposed on ***United Kingdom users*** who generate, upload or share ***illegal content*** or **illegal content proxy** other than ***CSEA content*** or **CSEA content**

[9] Ofcom has published **guidance on content communicated 'publicly' and 'privately' under the Online Safety Act**.

DRAFT FOR CONSULTATION

proxy, for preventing the future dissemination of ***illegal content*** (for measures in respect of ***CSEA content*** and **CSEA content proxy**, see Recommendation ICU H3).

ICU H2.3 The policy should set out the **sanctions** that the provider may impose on a ***United Kingdom user*** who generates, uploads or shares ***illegal content*** or **illegal content proxy** other than ***CSEA content*** or **CSEA content proxy**.

ICU H2.4 In setting its policy on the circumstances in which it will impose a **sanction** on such a ***user***, and the seriousness of the appropriate **sanction** in those circumstances, the provider should have regard to at least the following factors:

a) the severity of potential harm to ***United Kingdom users*** if they encounter the ***illegal content*** or **illegal content proxy** on the service;
b) whether the ***user*** has previously generated, uploaded or shared ***illegal content*** or **illegal content proxy**, including whether the ***user*** was **sanctioned** for doing so; and
c) the potential impact of the type and duration of the **sanction** on the ***user*** being **sanctioned**.

ICU H2.5 The provider should regularly review and update the policy on **sanctions** to ensure it remains fit for purpose, taking into account at least:

a) the service's **risk assessment**; and

b) any matters arising from **relevant complaints** which have been determined since the policy was last reviewed, which were **content appeals** relating to ***users*** who had been **sanctioned** under the provider's policy on **sanctions**.

Safeguards for freedom of expression and privacy

ICU H2.6 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendation ICU C4A;
b) Recommendations ICU D1 and D2, so far as they relate to **content appeals** or complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the provider is not complying with its duties in relation to freedom of expression or privacy; and
c) Recommendations ICU D8 and D9 (whichever is applicable), ICU D10 and ICU D10A (in relation to **content appeals**), and ICU D12.

ICU H3 Banning users for CSEA

Application

ICU H3.1 This measure applies to a ***provider*** in respect of each **service** it provides.

Key definitions

DRAFT FOR CONSULTATION

ICU H3.2 In this Recommendation ICU H3:

a) "age-appropriate relationship" means a relationship between two ***users*** of the same or similar age;

b) "ban" means to prevent a relevant user from accessing:

i) the **service** on which the relevant user generated, uploaded or shared ***CSEA content*** or **CSEA content proxy**; or on which the relevant user received **CSAM** or **CSAM content proxy**; and
ii) where applicable, other **services** provided by the provider, where the provider has reasonable grounds to believe that the relevant user has access to those services.

Where applicable, to "ban" includes removing the relevant user's user account. Where the provider has reasonable grounds to believe that the relevant user has more than one user account on the service in question, a "ban" also means preventing the relevant user from being able to access the service through any of those user accounts. A "ban" should be understood accordingly;

c) "relevant user" means a ***user*** to which ICU H3.3(b) or ICU H3.4(b) applies.

Recommendation

ICU H3.3 A provider should ban a ***user***, unless ICU H3.6(a) applies, where the provider determines that:

a) ***content*** on the service is ***CSEA*** (following an **illegal content judgement**) or **CSEA content proxy**, and

b) that ***user*** generated, uploaded or shared that content on the service.

ICU H3.4 A provider should also ban a ***user***, unless ICU H3.5 or H3.6(b) applies, where the provider determines that:

a) ***content*** on the service is **CSAM** (following an **illegal content judgement**) or **CSAM content proxy**, and

b) that user received and viewed that content.

ICU H3.5 A provider may decide not to ban a ***user*** it has determined has received and viewed **CSAM** or **CSAM content proxy** where it is satisfied that the ***user*** promptly reported that content to the provider, an appropriate law enforcement agency, or any other dedicated **CSAM** reporting agency[10] and promptly deleted it (if deleting the content is possible).

[10] For example, the Internet Watch Foundation (IWF).

DRAFT FOR CONSULTATION

ICU H3.6 A provider may decide not to ban a relevant user where it has reasonable grounds to believe that such relevant user is a ***child*** and either of the following extenuating circumstances apply:

a) the relevant user generated, uploaded or shared **CSAM** or **CSAM content proxy** of themself:

i) as a result of having been groomed, or
ii) consensually as part of an age-appropriate relationship;

b) the relevant user received **CSAM** or **CSAM content proxy**, uploaded or shared by another ***child*** of themself, consensually as part of an age-appropriate relationship.

ICU H3.7 Where the provider has imposed a ban under ICU H3.3 or ICU H3.4, it should take reasonable steps to ensure that while such a ban is in force, that ***user*** is prevented from regaining access to the service, or **services**, on a continuous basis.

In determining reasonable steps for this purpose, the provider should have regard, in particular, to:

a) the nature of the service in question;
b) what is technically feasible;
c) the resources of the provider; and
d) the provider's risk profile (having regard to its **risk assessment**).

ICU H3.8 Where the provider does not have reasonable grounds to believe a relevant user is a ***child***, a ban imposed on such relevant user under ICU H3.3 or ICU H3.4 should be:

a) permanent, or

b) in exceptional circumstances, of such duration as the provider reasonably considers appropriate in accordance with ICU H3.10.

ICU H3.9 Where the provider has reasonable grounds to believe a relevant user is a ***child,*** a ban imposed on such relevant user under ICU H3.3 or ICU H3.4 should be of such duration as the provider reasonably considers appropriate in accordance with ICU H3.10 and ICU H3.11, which may include the ban being permanent.

ICU H3.10 For the purpose of determining the duration of a ban under ICU H3.8(b) or ICU H3.9, the provider should ensure that the duration of the ban:

a) reflects the nature and severity of harm to ***United Kingdom users*** associated with the sharing, generating or uploading of the ***CSEA content*** or **CSEA content proxy** in question; and

b) protects ***United Kingdom users***, particularly **child users**, from harm by deterring the relevant user, and other ***users***, from sharing, generating or uploading ***CSEA content*** and **CSEA content proxy**.

ICU H3.11 For the purpose of determining the duration of a ban under ICU H3.9, the provider should also ensure that the duration of the ban takes into account:

DRAFT FOR CONSULTATION

a) whether the relevant user has previously:

i) generated, uploaded or shared ***CSEA content*** or **CSEA content proxy**, or

ii) received **CSAM** or **CSAM content proxy** without reporting that content promptly to the provider, an appropriate law enforcement agency, or any other dedicated **CSAM** reporting agency, and promptly deleting it (if deleting the content is possible),

b) whether the relevant user has previously been banned from the service for any of the actions referred to in sub-paragraph (a), and

c) the potential impact of the duration of the ban on the relevant user being banned.

Safeguards for freedom of expression and privacy

ICU H3.12 The following measures are safeguards to protect ***United Kingdom users'*** right to freedom of expression and the privacy of ***United Kingdom users***:

a) Recommendation ICU C4A;
b) where they are applicable, Recommendations ICU C3, ICU C4, ICU C6, ICU C7 and ICU C8;
c) ICU C9.6 to ICU C9.18;
d) Recommendations ICU D1 and D2, so far as they relate to **content appeals** by ***United Kingdom users*** against the decision by the provider to ban a relevant user under this Recommendation ICU H3;
e) Recommendations ICU D8 and D9 (whichever is applicable);
f) ICU D10.2 and D10.3, so far as they relate to **content appeals** by ***United Kingdom users*** against the decision by the provider to ban a relevant user under this Recommendation ICU H3; and
g) ICU H3.4, H3.6(b), H3.7 and H3.8.

DRAFT FOR CONSULTATION

I. [Not used]

[Intentionally left blank]

DRAFT FOR CONSULTATION

J. User controls

ICU J1 User blocking and muting

Application

ICU J1.1 This measure applies to a ***provider*** in respect of each **service** it provides that meets all of the following conditions:

a) the service meets any of Descriptions A to C;
b) ***users*** of the service have **user profiles**; and
c) the service has at least one of the following functionalities:

i) **user connection functionality**;

ii) **posting content functionality**;

iii) **user communication** (including but not limited to: (1) **direct messaging functionality**; and (2) **commenting on content**).

ICU J1.1A Description A is that the service:

a) is a **large service**, and
b) is at medium or high **risk** of one or more of the **kinds of illegal harm** listed in ICU J1.1D.

ICU J1.1B Description B is that the service:

a) has no more than seven million monthly **active UK users**,
b) is ***likely to be accessed by children***, and
c) is at high **risk** of one or more of the **kinds of illegal harm** listed in ICU J1.1D.

ICU J1.1C Description C is that the service:

a) has between 700,000 and seven million monthly **active UK users**,
b) is ***likely to be accessed by children***, and
c) is at medium **risk** of one or more of the **kinds of illegal harm** listed in ICU J1.1D.

ICU J1.1D The **kinds of illegal harm** are:

a) grooming;
b) encouraging or assisting suicide (or attempted suicide) and serious self-harm;
c) hate;
d) harassment, stalking, threats and abuse;
e) controlling or coercive behaviour;
f) cyberflashing.

DRAFT FOR CONSULTATION

ICU J1.1E Where the service meets Descriptions B or C, ICU J1.2 and ICU J1.4 apply in relation to any **child-accessible part** of the service that has any of the functionalities listed in ICU J1.1(c).

Recommendation

ICU J1.2 The provider should offer every applicable user the option to **block** each of:

a) a specific user account, whether or not **connected** to that applicable user's user account; and
b) where the service has **user connection functionality,** all user accounts which are not **connected** to that applicable user's user account.

ICU J1.2A "Applicable user" means:

a) in relation to a service which meets Description A, a registered ***United Kingdom user***;
b) in relation to a service which meets Descriptions B or C, a registered ***United Kingdom user*** other than a ***user*** determined to be an ***adult*** by the use of **highly effective age assurance**.

ICU J1.3 "**Block**" means to take action that will result in:

a) **blocked users** being unable to send **direct messages** from the **blocked account** to the **blocking account**;
b) **blocking users** being unable to send **direct messages** from the **blocking account** to the **blocked account**;
c) the **blocking user** being unable to ***encounter***, by means of the **blocking account**, any **content posted** on the service using the **blocked account** (regardless of where on the service it is posted), including but not limited to: (i) **reactions** to ***content***; and (ii) **content posted** using the **blocked account** which is subsequently posted by another ***user***;
d) the **blocked user** being unable to ***encounter***, by means of the **blocked account**, any **content posted** on the service using the **blocking account** (regardless of where on the service it is posted), including but not limited to: (i) **reactions** to ***content***; and (ii) **content posted** using the **blocking account** which is subsequently posted by another ***user***; and
e) the **blocking account** and **blocked account**, if they were **connected**, no longer being **connected**,

and "**blocking**" is to be read accordingly.

"**Blocking account**" means the user account through which the action resulting in **blocking** has taken place. "**Blocked account**" means the user account that action has been taken against.

"**Blocking user**" means the ***user*** operating the **blocking account**. **"Blocked user**" means the ***user*** operating the **blocked account**.

DRAFT FOR CONSULTATION

ICU J1.4 The provider should offer every applicable user the option to **mute** other user accounts (whether or not **connected** to that applicable user's user account) on the relevant service.

ICU J1.5 "**Mute**" means to take action that will result in the **muting user** being unable to ***encounter*** any **content posted** on the service using the **muted account**, including:

a) **reactions** to **content posted** using the **muted account**; and
b) **content posted** using the **muted account** which is posted by another ***user***,

by means of the **muting account,** unless the **muting user** visits the **user profile** associated with the **muted account**, in which case the **muting user** will experience that **user profile** as if the **muted account** had not been **muted**. "**Muting**" is to be read accordingly.

"**Muting account**" means the user account through which the action resulting in **muting** has taken place. "**Muted account**" means the user account that the action has been taken against.

"Muted user" means the ***user*** operating the **muted account**. "**Muting user**" means the ***user*** operating the **muting account.**

ICU J1.6 For the avoidance of doubt:

a) save for where muting is reciprocal, **muted users** should continue to ***encounter*** the **content posted** using the **muting account**;
b) functionality from the **muted user's** perspective should continue as if the **muting user** had not **muted** the **muted account**; and
c) providers should not at any time notify **muted users,** or otherwise make them aware, that the **muted account** has been **muted** by the **muting user**.

Muting is reciprocal where a ***user*** has through a user account ("A") **muted** a user account ("B"), and a ***user*** has through user account B also muted user account A.

ICU J1.7 The provider should provide information to ***United Kingdom users*** about the availability of the options to **block** and **mute** other ***users*** and the effect of these actions, including the types of interactions or access to ***content*** that it would restrict. That information should be:

a) easy to find; and
b) comprehensible based on the likely reading age of the youngest individual permitted to use the service without the consent of a parent or guardian.

Safeguard for privacy

DRAFT FOR CONSULTATION

ICU J1.8 Recommendation ICU B1 is a safeguard to protect the privacy of ***United Kingdom users***.

ICU J2 Disabling comments

Application

ICU J2.1 This measure applies to a ***provider*** in respect of each **service** it provides that meets both of the following conditions:

a) the service meets any of Descriptions A to C; and
b) the service has **commenting on content functionality**.

ICU J2.1A Description A is that the service:

a) is a **large service**, and
b) is at medium or high **risk** of one or more of the **kinds of illegal harm** listed in ICU J2.1D.

ICU J2.1B Description B is that the service:

a) has no more than seven million monthly **active UK users**,
b) is ***likely to be accessed by children***, and
c) is at high **risk** of one or more of the **kinds of illegal harm** listed in ICU J2.1D.

ICU J2.1C Description C is that the service:

a) has between 700,000 and seven million monthly **active UK users**,
b) is ***likely to be accessed by children***, and
c) is at medium **risk** of one or more of the **kinds of illegal harm** listed in ICU J2.1D.

ICU J2.1D The **kinds of illegal harm** are:

a) grooming;
b) encouraging or assisting suicide (or attempted suicide) and serious self-harm;
c) hate;
d) harassment, stalking, threats and abuse.

ICU J2.1E Where the service meets Descriptions B or C, ICU J2.2 and ICU J2.3 apply in relation to any **child-accessible part** of the service that has **commenting on content functionality**.

Key definition

ICU J2.1F In this Recommendation ICU J2, "applicable user" means:

a) in relation to a service which meets Description A, a registered ***United Kingdom user***;

DRAFT FOR CONSULTATION

b) in relation to a service which meets Descriptions B or C, a registered ***United Kingdom user*** other than a ***user*** determined to be an ***adult*** by the use of **highly effective age assurance**.

Recommendation

ICU J2.2 The provider should offer every applicable user the option of preventing any other ***users*** of the service from **commenting on content** posted on the service using their user account.

ICU J2.3 Applicable users should be able to exercise the option referred to above:

a) when **posting content**; and
b) after having **posted content**.

ICU J2.4 The provider should provide information to ***United Kingdom users*** about the availability of the option to prevent other ***users*** of the service from **commenting on content** posted on the service by the ***United Kingdom user*** concerned and the effect of this action, including the types of interactions or access to ***content*** that it would restrict. That information should be:

a) easy to find; and
b) comprehensible based on the likely reading age of the youngest individual permitted to use the service without the consent of a parent or guardian.

Safeguard for privacy

ICU J2.5 Recommendation ICU B1 is a safeguard to protect the privacy of ***United Kingdom users***.

ICU J3 Notable user and monetised labelling schemes

Application

ICU J3.1 This measure applies to a ***provider*** in respect of each **large service** it provides that meets both of the following conditions:

a) the service is at medium or high **risk** of fraud or the foreign interference offence (or both); and
b) the service labels **user profiles** under one or both of the following: (i) a **notable user scheme**; or (ii) a **monetised scheme**.

Recommendation

ICU J3.2 In this Recommendation ICU J3:

a) **notable user schemes** and **monetised schemes** are referred to together as "**relevant schemes**"; and

DRAFT FOR CONSULTATION

b) a ***user*** whose **user profile** is labelled under a **relevant scheme** is referred to as a "**relevant user**".

ICU J3.3 The provider should have, and consistently apply, internal documented policies regarding the operation of **relevant schemes** on the service, which should, at a minimum:

a) be designed to reduce any **risk** of harm to ***United Kingdom users*** from fraud and/or the foreign interference offence associated with a **relevant scheme**, as identified in the **risk assessment** of the service;
b) set out: (1) the process for considering; and (2) the criteria and thresholds for deciding whether to:
   i) label a **user profile**; and
   ii) remove the label from the **user profile** of a **relevant user**.

   In respect of a **notable user scheme**, the criteria and thresholds should set out how the provider will satisfy itself that:
   i) the user account of a **relevant user** is operated by or on behalf of the person by whom or on whose behalf it is held out as being operated; and
   ii) if that person is held out as holding a particular position or role, that they hold that position or role;
c) set out safeguards to ensure that the **user profile** information (such as username and 'bio' text) of a **relevant user** whose **user profile** is labelled under a **notable user scheme** is not modified so as to suggest the user account is operated by or on behalf of anyone other than the **relevant user**;
d) set out the frequency with which and the circumstances in which the provider will conduct reviews to confirm whether the **user profiles** of **relevant users** continue to qualify to be labelled;
e) set out whether and, if so, how the provider will treat **relevant users** and the **content posted** on the service differently to other users, including in relation to **content recommender systems**, content moderation, and account security;
f) be communicated to relevant individuals working for the provider, including through regular training (in particular when a policy is modified); and
g) be regularly reviewed and updated to ensure the policy remains fit for purpose. As part of regularly reviewing the policy, the provider should, if it considers it appropriate, take into account one or more of the following: user feedback and reporting; user experience testing; and engagement with persons with relevant expertise.

ICU J3.4 The provider should provide the following to ***United Kingdom users***:

a) the following information on the **user profile** of a **relevant user**:
   i) why the **user profile** is labelled; and

DRAFT FOR CONSULTATION

ii) if the provider operates more than one type of **relevant scheme** on the service, the **relevant scheme(s)** under which the **user profile** is labelled; and

b) a user-facing description of the **relevant scheme(s)**, which should:

i) be in writing;

ii) be clear and accessible, and in particular be:

a. easy to find, such that it is clearly signposted for the general public regardless of whether they have signed up to or are using the service;
b. laid out and formatted in a way that helps users read and understand it;
c. comprehensible based on the likely reading age of the youngest individual permitted to use the service without the consent of a parent or guardian; and
d. designed for the purposes of ensuring usability for those dependent on assistive technologies, including keyboard navigation and screen reading technology;

iii) explain how and why **user profiles** are labelled (including different categories of labelling and, in particular, specifying whether a **relevant scheme** is or is not a **notable user scheme**);

iv) explain how and why **relevant users** may have a label removed from their **user profile**; and

v) be consistent with (but need not include) every detail of the provider's internal policies.

DRAFT FOR CONSULTATION

5. Definitions and interpretation

5.1 Terms in **bold** used in these Codes have the meanings set out in table A. The meaning given applies even if the term is also used in the **Act**.

5.2 Terms in ***bold and italics*** used in these Codes have the same meaning as in the **Act**. Table B provides a reference to the provision(s) in the **Act** containing the definition of the relevant term as well as additional notes and references which are intended to assist the reader. In the event of any inconsistency between the **Act** and the information in table B, the **Act** should be regarded as authoritative.

5.3 Terms which are underlined are references to kinds of illegal harms set out in table C.

Table A - Definitions of terms in bold used in these Codes

Term	Meaning
Act	The Online Safety Act 2023 (c.50).
Active United Kingdom user	As defined in paragraph 5.10.
Age appropriate design code	The code of practice issued under section 125 of the Data Protection Act 2018 (c. 12) on 12 August 2020.
Age assessment appeal	A complaint by a ***United Kingdom user*** claiming that their experience on the service is impacted because measures used to comply with a duty set out in section 10(2) or (3) of the **Act** have resulted in an incorrect assessment of the user's age.
Age assurance	***Age verification*** and ***age estimation***.
Age assurance process	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** measures.
Automated location information display functionality	A functionality which displays **location information** including via the following (where relevant): a) shared ***content***; b) **user profile**; and c) functionalities that display the live **location information**.
Block, Blocking	As defined in ICU J1.3.
Blocked user	As defined in ICU J1.3.

Term	Meaning
Blocking user	As defined in ICU J1.3.
Challenge age approach	An **age assurance process** comprised of two stages whereby if, at the end of the first stage of process, the estimated age of a ***user*** is found to be below a specified age, further steps are taken to estimate or verify the exact age of the ***user*** before determining whether the ***user*** is a ***child***.
Child-accessible part (of a service)	A ***user-to-user part*** of a **service** other than any part which **children in the United Kingdom** are not normally able to access as a result of the use on the service of ***age verification*** or ***age estimation***.
Child user	A ***United Kingdom user*** who is under the age of 18.
Children in the United Kingdom	People under the age of 18 in the United Kingdom.
Children's risk assessment	The most recent risk assessment carried out by the provider of a service pursuant to section 11 of the **Act**.
Comment on content	Reply to ***user-generated content***, or generate, upload or share ***content*** in response to another piece of ***user-generated content*** posted on open channels of communication, in such a way that the reply or content (as the case may be) is visually accessible directly from the original user-generated content without navigating away from that user-generated content.
Commenting on content functionality	***User-to-user service*** functionality that allows ***users*** to **comment on content.**
Complainant	The ***United Kingdom user*** or ***affected person*** who made the complaint.
Connect	See **connection**.
Connected (accounts)	Two user accounts with a **connection**.

Term	Meaning
Connection	An established link between two user accounts that one or both of the ***users*** operating those accounts has taken steps to establish. Connections include, but are not limited to: a) established links created when one ***user*** invites another ***user*** to establish a link between the user accounts of the two users that the other user accepts; b) established links created when one ***user*** elects to follow another ***user's*** user account; and c) established links created when one ***user*** elects to subscribe to another ***user's*** user account. The terms "connect", "connected" and "connection" are to be read accordingly.
Connection lists	A list of the user accounts to which a user account is **connected** which is visible to other ***users*** via a **user profile**.
Content appeal	A complaint by a ***United Kingdom user*** about any of the following actions, if the action concerned has been taken by the ***provider*** on the basis that ***content*** generated, uploaded or shared by that user is ***illegal content*** or **illegal content proxy***:* a) the content being ***taken down***; b) the user being given a warning; c) the user being suspended, banned, or in any other way restricted from using the **service**.
Content moderation function	The ***systems and processes*** designed to review, assess and take action in relation to ***content***, including ***content*** a ***provider*** has reason to suspect may be ***illegal content***.
Content posted	***Content*** generated, uploaded and/or shared on open channels of communication by a ***user*** of the service. "Posting content" and "posted content" are to be read accordingly.

Term	Meaning
Content recommender system	An algorithmic system which determines the relative ranking of an identified pool of ***content*** that includes ***regulated user-generated content*** from multiple ***users*** on content feeds. ***Content*** is recommended based on factors that it is programmed to account for, which may include but are not limited to: a) ***user*** feedback, such as interactions with a piece of ***content*** by means of likes, views and shares; b) predicted engagement with ***content*** based on a **user's** consumption history, such as likelihood of liking, sharing, and commenting on a piece of ***content***; c) profile and contextual characteristics, such as age and location; d) ***content*** liked by ***users*** with a similar consumption and engagement history; and e) popularity of a certain piece of ***content***. For the avoidance of doubt, references to **content recommender systems** in these Codes do not include: a) a content recommender system employed exclusively in the operation of a functionality which suggests ***content*** to users in direct response to a search query; or b) a product recommender system; or c) a network recommender system that suggests ***users*** and groups to follow.

CONSULTATION

Term	Meaning
Content recommender system design adjustment	Any iterative and incremental alterations made to the design of an existing variant of a **content recommender system's** underlying model or to the algorithms responsible for content ranking as part of ongoing product management. It does not include design changes that: a) would amount to a significant change and therefore trigger a risk assessment under section 9(4) of the **Act**; or b) are made in connection with a live response to a national security threat or other emergency; c) would not be deployed for ***United Kingdom users*** of the service.
Cryptographic hash matching technology	Technology that detects exact matches to digital ***content*** by comparing cryptographic hash values of the ***content*** against a reference database of cryptographic hash values. A match is identified only where the hash values are identical.
CSAM (child sexual abuse material)	***Content*** that amounts to an offence specified in any of the following paragraphs of Schedule 6 to the **Act**— a) paragraph 1 to 4, 7 or 8; b) paragraph 9 so far as any of the offences it contains are committed in relation to an offence specified in paragraphs 1 to 4, 7 or 8; c) paragraph 10; or d) paragraph 13 so far as any of the offences it contains are committed in relation to an offence specified in paragraph 10. Where the context requires, references to CSAM include material which would be CSAM if it were ***regulated user-generated content*** present on a **service**.

Term	Meaning
CSAM content proxy	***Content*** that a ***provider*** determines to be in breach of its ***terms of service***, where: a) the ***provider*** had reason to suspect that the content may be **CSAM**; and b) the ***provider*** is satisfied that its ***terms of service*** prohibit the type of **CSAM** which it had reason to suspect existed.
CSEA content proxy	***Content*** that a ***provider*** determines to be in breach of its ***terms of service***, where: a) the ***provider*** had reason to suspect that the content may be ***CSEA content***; and b) the ***provider*** is satisfied that its ***terms of service*** prohibit the type of ***CSEA content*** which it had reason to suspect existed. For the avoidance of doubt, **CSEA content proxy** includes **CSAM context proxy**.
Default settings	Automatic settings for functionalities, applicable to a specific user account, which are set by the ***provider*** of a **service** and that can be disabled by a ***user*** operating that user account.
Detected content	***Content*** detected by the use of a **relevant technology** as being (or as likely to be) **target content** (and related expressions are to be read accordingly).
Direct message	A message sent from a user account to a recipient user account that can only be immediately viewed or read on that specific recipient user account.
Direct messaging functionality	***User-to-user service*** functionality that allows ***users*** to send **direct messages**.

Term	Meaning
Existing means to determine the age or age range of a particular user	An existing ***system or process*** designed to determine the age or age range of a particular ***user*** which may be comprised of one or more of the following: a) any measure designed to estimate the age or age range of ***users***; b) any measure designed to verify the exact age of ***users***; and c) a measure which requires a ***user*** to self-declare their age (without more).
False positive	**Detected content** that is not **target content.**
File-storage and file-sharing service	A **service** whose primary functionalities involve enabling ***users*** to: a) store digital content, including images and videos, on the cloud or dedicated server(s); and b) share access to that content through the provision of links (such as unique **URLs** or hyperlinks) that lead directly to the content for the purpose of enabling other ***users*** to ***encounter*** or interact with the content.
Governance body	A body which makes decisions within an organisation, for example a board of directors.
Hash matching technology	Either **perceptual hash matching technology** or **cryptographic hash matching technology.**
Highly effective age assurance	An **age assurance process** that meets each of the criteria listed under: a) ICU B1.3 (technical accuracy), b) ICU B1.4 (robustness), c) ICU B1.5 (reliability), and d) ICU B1.6 (fairness).
Illegal content judgement	A judgement about whether ***content*** is ***illegal content*** or ***illegal content*** of a particular kind, made in accordance with section 192(2) and section 192(5) to (7) of the **Act.**

CONSULTATION

Term	Meaning
Illegal content proxy	***Content*** that a ***provider*** determines to be in breach of its ***terms of service***, where: a) the ***provider*** had reason to suspect that the content may be ***illegal content***; and b) the ***provider*** is satisfied that its ***terms of service*** prohibit the type of ***illegal content*** which it had reason to suspect existed.
Illegal content safety duties	The duties set out in section 10 of the **Act**.
Illegal harm	Harm arising from ***illegal content*** and the commission and facilitation of ***priority offences***.
Intimate image abuse content	***Content*** which amounts to an offence: a) under section 66B of the Sexual Offences Act 2003 (sharing or threatening to share intimate image or film); or b) under section 2 of the Abusive Behaviour and Sexual Harm (Scotland) Act 2016 (asp 22) (disclosing, or threatening to disclose, an intimate photograph or film).
Kind of illegal harm	See the subsection headed 'Risks of illegal harm' below (which begins at paragraph 5.4).
Large service	A **service** which has more than 7 million monthly **active United Kingdom users** (see paragraphs 5.7 to 5.10).
Location information	The geographical location of a device linked to a user account, generated using data from the device, including, but not limited to, GPS data or data about connection with local Wi-Fi equipment.
Log	A record in a form that enables the continuous collection, storage and analysis of information relevant to the operation of algorithmic systems.

Term	Meaning
Monetised scheme	A scheme by which the ***provider*** of a **service** labels the **user profile** of a ***user*** who has made payment to the provider of the service or some other person. Such schemes may be open to all ***users*** and payment may be regular or one-off. ***Users*** participating in the scheme may benefit from access to additional features on the service. The label to indicate that a ***user*** is participating in a **monetised scheme** may appear on that ***user's*** user profile and/or any ***content*** they publish. Services may or may not refer to such schemes as "verification" schemes.
Multi-risk service	See paragraph 5.6.
Mute, Muting	As defined in ICU J1.5.
Muted user	As defined in ICU J1.5.
Muting user	As defined in ICU J1.5.
Network expansion prompt	A recommendation to **connect** with one or more specified user accounts on the relevant service generated by automated means**.**
Network expansion prompt functionality	A functionality that generates **network expansion prompts**. This can include, but is not limited to, recommendations to connect with specific user accounts that have similar interests, that have **location information** which is close geographically, that are associated with the same school or workplace, or that have a user account that is a mutual connection.

Term	Meaning
Notable user scheme	A scheme by which the ***provider*** of a **service** labels the **user profile** of a ***user*** to indicate to other ***users*** that they are notable. "Notable users" include but are not limited to politicians, celebrities, influencers, financial advisors, company executives, journalists, government departments and institutions, non-governmental organisations, financial institutions, media outlets, and companies. The label to indicate that a ***user*** is notable (for example a "tick" symbol) may appear on that ***user's*** user profile and/or any ***content*** they publish. ***Providers*** may or may not refer to such schemes as "verification" schemes.
One-to-many livestream/livestreaming	A ***user-to-user service*** functionality that allows a single ***user*** to simultaneously create and distribute live video depicting real-world events as they occur online, streaming ***content*** (such as live video ***content***) to multiple or potentially unlimited ***users*** in, or very close to, real time. This excludes ***content*** which is communicated privately.

DRAFT FOR CONSULTATION

Term	Meaning
On-platform testing	The process of live testing the operation of different variants of a **content recommender system** on a **service** across a control group and treatment groups comprised of ***users*** of the service. It involves the collection of data to produce metrics relating to certain identified factors, such as commercial or user safety. The methods may include (but are not limited to): a) A/B testing: a randomised control trial in which treatment groups are served content from adjusted variants of the **content recommender system**, and a control group is served content from the existing variant of the **content recommender system**, with a view to comparing their performance against identified metrics. b) Multi Arm Bandit Testing: a continuous experiment that uses machine learning techniques to dynamically allocate ***users*** to the best-performing variant of a content recommender system against a particular metric (e.g., average click-through rate per user) based on real-time data gathered during the test.
Perceptual hash matching technology	Image matching technology which compares the similarity between hashes created from images by means of an algorithm known as a perceptual hash function, to assess whether those images are perceptually similar to each other. This does not include technology which compares similarity through the use of machine learning.
Posting content	See **content posted.**
Posting content functionality	***User-to-user service*** functionality allowing ***users*** to do one or more of generating, uploading or sharing ***content*** on open channels of communication.
Precision	A measure of statistical accuracy, calculated as the proportion of **detected content** that a **relevant technology** has correctly identified as **target content**.

CONSULTATION

Term	Meaning
Produce(s) reproducible results	Produce(s) the same or similar outputs when given the same or similar inputs.
Proscribed organisation	A group or organisation proscribed by the Secretary of State under section 3 of the Terrorism Act 2000.
Proscribed organisation content	***Regulated user-generated content*** which amounts to an offence specified in any of the following paragraphs of Schedule 5 to the **Act**: a) paragraphs 1(a) to (e); b) paragraphs 1(f) to (p) and 3, where the "terrorism" for the purpose of the offence is an action taken for the benefit of a **proscribed organisation**; or c) paragraph 4 so far as any of the inchoate offences relate to an offence falling within points (a) or (b) above.
Prospective complainants	***United Kingdom users*** and ***affected persons*.**
Reaction (to *content*)	Expressing a view on ***content***, including, for example, by: a) applying a "like" or "dislike" button or other button of that nature, b) applying an emoji or symbol of any kind, c) engaging in yes/no voting, or d) rating or scoring ***content*** in any way (including giving star or numerical ratings).
Recall	A measure of statistical accuracy, calculated as the proportion of **target content** analysed by a **relevant technology** that the technology has **detected**.
Recommended trusted flagger	As defined in ICU D14.2.
Regulated pornographic content	***Pornographic content*** other than content of a type described in section 61(6) of the **Act**.
Regulated user-to-user service	A ***user-to-user service*** as defined in section 3 of the **Act**, which is a regulated user-to-user service under section 4 of the **Act** (subject to the disapplication in section 5 of the **Act**).

Term	Meaning
Relevant available information	Information that is reasonably available to the ***provider*** from the operation of the **service**, such as: a) metadata relating to ***content*** (including tags and labels); b) data relating to the **user** uploading, generating or sharing ***content*** (such as whether other ***content*** uploaded, generated or shared by that user has been assessed to be **illegal content**); c) reports about ***content*** which ***United Kingdom users*** or ***affected persons*** consider to be ***illegal content***; d) negative feedback given by ***users*** pursuant to Recommendation PCU E3, in the Protection of Children Code of Practice for user-to-user services (where applicable); e) other indicators generated by technology used on the service.

DRAFT FOR CONSULTATION

Relevant complaints	The following kinds of complaint: a) complaints (including **reports**) by ***United Kingdom users*** and ***affected persons*** about ***content*** present on a service which they consider to be ***illegal content***, excluding ***intimate image content reports***; b) complaints by ***United Kingdom users*** and ***affected persons*** if they consider that the ***provider*** is not complying with a duty set out in the following sections of the **Act**- i) section 10 (illegal content safety duties), ii) section 20 (content reporting) so far as it relates to illegal content; or iii) section 22 (freedom of expression or privacy); c) complaints by a ***United Kingdom 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*** or **illegal content proxy**; d) complaints by a ***United Kingdom user*** of a ***user-to-user 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*** or **illegal content proxy**; e) complaints by a ***United Kingdom 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

Term	Meaning
	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).
Relevant content moderation action	For the purposes of ICU C4.2 and ICU C4.3, as defined in ICU C4.5.
Relevant scheme(s)	For the purposes of Recommendation ICU J3 (notable user and monetised labelling schemes), as defined in ICU J3.2(a).
Relevant technology	The kind of technology specified in the measure in question.
Relevant user	For the purposes of Recommendation ICU J3 (notable user and monetised labelling schemes), as defined in ICU J3.2(b).
Reporting and complaints duties	The duty set out in section 20 of the **Act**, so far as it relates to ***illegal content***, and the duties set out in section 21 of the **Act**, so far as relating to the complaints set out in section 21(4).
Reporting conduct	Making a complaint on the grounds that a ***user*** considers that a ***provider*** is not complying with its duties in relation to ***illegal content*** because the provider is allowing a ***user*** of a **service** it provides to use the features and functionalities of that service to increase the risk of **illegal harm** to individuals.
Reports	Complaints by ***United Kingdom users*** and ***affected persons*** about ***content*** present on a **service** which they consider to be ***illegal content***, made using a reporting function or tool provided by the service.
Risk	See the subsection headed 'Risks of illegal harm' below (which begins at paragraph 5.4).

Term	Meaning
Risk assessment	The most recent risk assessment carried out by the ***provider*** pursuant to section 9 of the **Act**.
Sanction	An action taken against a ***user*****.** This includes: a) giving a warning to the ***user***, b) suspending or banning the ***user*** from using the **service**, or c) in any other way restricting the ***user's*** ability to use the **service**. The term "sanctioned" is to be read accordingly.
Service	A **regulated user-to-user service**.
Specified connection	For the purposes of Recommendation ICU F1, as defined in ICU F1.5.
Target content	***Content*** of the kind the use of a **relevant technology** is designed to identify.
Trusted flagger	An entity which is a **recommended trusted flagger** and any other person: a) whom the ***provider*** has reasonably determined has expertise in a particular **illegal harm**; and b) for whom the ***provider*** has established a dedicated reporting channel.
URL	Uniform Resource Locator, meaning a reference that specifies the location of a resource accessible by means of the internet.
User communication	***User-to-user service*** functionality type that describes functionalities by means of which ***users*** can communicate with one another either synchronously or asynchronously. Includes communication across open and closed channels.
User connection functionality	A ***user-to-user service*** functionality that allows users to create **connections**.

Term	Meaning
User profile	A collection of information that has been shared by a ***user*** and may be viewed by other ***users*** of the service. This can include, but is not limited to, a username, biography or profile picture, as well as ***content*** generated, uploaded or shared by the ***user*** operating the user account associated with the user profile.
Volunteer	An individual who, in relation to the activity in question, is not: a) employed by the ***provider*** or anyone else, b) remunerated, c) acting by way of a business.

Table B - Terms used in these Codes that have the meaning given in the Act

Term	Meaning under the OSA
Adult	Section 236(1)
Affected person	Section 20(5) See also section 21(7).
Age estimation	Section 230(3) and (4) Section 4(4) defines "regulated service". A "regulated user-to-user service" is a "user-to-user service" (see the entry for "user-to-user service" regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5). A "regulated search service" is an "internet service" (see the entry for "internet service" regarding the definition of that term) that is, or includes, a "search engine" (see the entry for "search engine" regarding the definition of that term) (subject to section 3(5) to (7)) that is a regulated search service under section 4 (subject to the disapplication in section 5). Section 79(3) defines "Regulated provider pornographic content". Section 79(2) defines "provider pornographic content".

CONSULTATION

	See the entries for “content”, “internet service”, “pornographic content” and “user” regarding the definitions of those terms.
Age verification	**Section 230(2) and (4)** Section 4(4) defines “regulated service”. A “regulated user-to-user service” is a “user-to-user service” (see the entry for “user-to-user service” regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5). A “regulated search service” is an “internet service” (see the entry for “internet service” regarding the definition of that term) that is, or includes, a “search engine” (see the entry for “search engine” regarding the definition of that term) (subject to section 3(5) to (7)) that is a regulated search service under section 4 (subject to the disapplication in section 5). Section 79(3) defines “Regulated provider pornographic content”. Section 79(2) defines “provider pornographic content”. See the entries for “content”, “internet service”, “pornographic content” and “user” regarding the definitions of those terms.
Category 1 service	**Section 95(10)(a)** Section 95(2)(a) requires Ofcom to establish a register, a part of which sets out the **regulated user-to-user services** which Ofcom considers meet the Category 1 threshold conditions (as specified in regulations made under paragraph 1(1) of Schedule 11 to the Act). Services for the time being included in that part of the register are Category 1 services.
Child	**Section 236(1)**
Combined service	**Section 4(7)** Paragraph 7(2) of Schedule 1 sets out the conditions to be met for a search engine not to be considered a public search engine. See the entry for “search engine” regarding the definition of that term. A “regulated user-to-user service” is a “user-to-user service” (see the entry for “user-to-user service” regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5).

Content	**Section 236(1)** See the entry for "internet service" regarding the definition of that term. See section 232 in relation to "content communicated "publicly" or "privately".
Content identification technology	**Section 231(2)** Under section 231(1), content identification technology is listed as a form of "proactive technology". Section 231(3) describes situations where content identification technology will not be proactive technology.
CSEA content	**Section 59(9)** Schedule 6 lists the relevant offences for determining when content is CSEA content. Sections 59(11) to (14) contain further interpretative provisions. See the entries for "combined service", "content", "regulated user-generated content" and "search content" regarding the definitions of those terms. A "regulated user-to-user service" is a "user-to-user service" (see the entry for "user-to-user service" regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5).
Encounter (in relation to content)	**Section 236(1)**
Illegal content	**Section 59(2)** Section 59(3) sets out when content will amount to a relevant offence. Sections 59(4), (5) and (7) set out what is meant by a relevant offence. Section 59(6) describes offences which are not relevant offences. Sections 59(11) to (15) contain further interpretative provisions. See the entries for "combined service", "content", "regulated user-generated content" and "search content" regarding the definitions of those terms. A "regulated user-to-user service" is a "user-to-user service" (see the entry for "user-to-user service" regarding the definition of that term) which is a

	regulated user-to-user service under section 4 (subject to the disapplication in section 5).
Internet service	**Section 228(1)** and **section 204(1)** Sections 228(2) and (3) describe what is meant by a service that is made available by means of the internet.
Intimate image content report	**Section 20A(2)**
Likely to be accessed by children	**Section 37** Section 35(1) sets out what is meant by a "children's access assessment". Section 35(3) sets out when the "child user condition" is met in relation to a service. Section 36 details the requirement to carry out a children's access assessment. Schedule 3 makes provision about the deadline by which a first "children's access assessment" must be carried out. Section 236 defines a "child" (see also section 35(5)). Section 4(3) defines "Part 3 service". Sections 230(2) and (4) define "age verification". Sections 230(3) and (4) define "age estimation". Section 236(1) defines "user-to-user part" in relation to a "user-to-user service". See the entries for "search engine", "United Kingdom user", "user-generated content" and "user-to-user service" regarding the definitions of those terms.
Pornographic content	**Section 236(1)** See the entry for "content" regarding the definition of that term.
Priority illegal content	**Section 59(10)** Sections 59(8) and (9) define "terrorism content" and "CSEA content". Sections 59(11) to (14) contain further interpretative provisions. See the entries for "combined service", "content", "regulated user-generated content" and "search content" regarding the definitions of those terms. A "regulated user-to-user service" is a "user-to-user service" (see the entry for "user-to-user service" regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5).

Priority offence	Section 59(7)
Proactive technology	Section 231(1) Sections 231 (4) and (6) define "user profiling technology" and "behaviour identification technology". Sections 231(3), (5) and (7) explain when these technologies will not be proactive technology. Sections 231(8) to (13) contain further interpretative provisions. See the entries for "combined service", "content", "content identification technology", "CSEA content", "illegal content", "internet service", "search engine", "terrorism content", "user", "user-generated content" and "user-to-user service" regarding the definition of those terms. Sections 3(4) and 204(1) define "search service". Section 79(2) defines "provider pornographic content". Section 236(1) defines "pornographic content". Section 236(1) defines "personal data". Section 4(4) defines "regulated service". Sections 125(12) and (13) define "accredited technology". A "regulated user-to-user service" is a "user-to-user service" (see the entry for "user-to-user service" regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5). A "regulated search service" is an internet service (see the entry for "internet service" regarding the definition of that term) that is, or includes, a search engine (subject to sections 3(5) to (7) of the Act) (see the entry for "search engine" regarding the definition of that term) that is a regulated search service under section 4 of the Act (subject to the disapplication in section 5 of the Act).
Provider	Section 226 See the entries for "combined service", "internet service", "search engine", and "user" regarding the definition of those terms. Sections 3(4) and 204(1) define "search service".
Regulated user-generated content	Section 55(2) Sections 55(5) to (12) contain interpretative provisions.

	A "regulated user-to-user service" is a "user-to-user service" (see the entry for "user-to-user service" regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5). See the entries for "content", "user" and "user-generated content" regarding the definitions of those terms. Section 236(1) defines "identifying content". Section 56 defines "recognised news publisher".
Regulated provider pornographic content	**Section 79(3)** Section 79(2) defines "provider pornographic content". Section 236(1) defines "pornographic content". See the entries for "content", "combined service", "internet service", "search results", "search content" and "user-generated content" regarding the definitions of those terms.
Search content	**Section 57(2)** Sections 57(4) and (5) set out the meaning of "search" and "via search results". Sections 3(4) and 204(1) define "search service". See the entries for "content", "internet service", "search engine", "search results", and "user-to-user service" regarding the definitions of those terms. Section 236(1) defines "paid-for-advertisements". See the entry for "internet service" regarding the definition of that term. Section 56(1) defines "recognised news publisher".
Search engine	**Section 229** Section 57(4) defines "search". See the entries for "internet service" and "user-to-user service" regarding the definitions of those terms.
Search or search request	**Section 57(4)**
Search results	**Section 57(3)** Sections 3(4) and 204(1) define "search service".

	See the entries for “internet service”, “search engine”, “search request” and “user” regarding the definitions of those terms.
Systems and/or processes	**Section 236(1)**
Taking down	**Section 236(1)** See the entries for “content” and “user-to-user service” regarding the definitions of those terms.
Terms of service	**Section 236(1)** See the entries for “United Kingdom user” and “user-to-user service” regarding the definitions of those terms.
Terrorism content	**Section 59(8)** Schedule 5 lists the relevant offences for determining when content is terrorism content. Sections 59(11) to (14) contain further interpretative provisions. See the entries for “combined service”, “content”, “regulated user-generated content” and “search content” regarding the definitions of those terms. A “regulated user-to-user service” is a “user-to-user service” (see the entry for “user-to-user service” regarding the definition of that term) which is a regulated user-to-user service under section 4 (subject to the disapplication in section 5).
United Kingdom user	**Section 227(1)** See the entry for “user” regarding the definition of that term.
User	**Section 227(2)-(3)** Section 227(3) sets out individuals and entities that will not be users for the purposes of the Act. Sections 227(4) to (6) provide further interpretation of terminology used throughout section 227. See the entries for “internet service”, “search engine” and “user-to-user service” regarding the definitions of those terms. Section 57(4) defines “search”. Sections 3(4) and 204(1) define “search service”. Sections 228(2) and (3) describe what is

	meant by a service that is made available by means of the internet.
User-generated content	Section 55(3) and (4) See the entries for "user" and "user-to-user service" regarding the definition of those terms. Section 55(4) provides interpretation of the scope of "content generated, uploaded or shared by a user" and explains when a bot or other automated tool may be regarded as a user of a service.
User-to-user part (of a service)	Section 236(1) See the entries for "user-generated content" and "user-to-user service" regarding the definitions of those terms.
User-to-user service	Section 3(1) and (2) and section 204(1) See the entries for "content", "encounter", "internet service" and "user" regarding the definitions of those terms.

DRAFT FOR CONSULTATION

Risks of illegal harm

Risk of a kind of illegal harm

5.4 A service is at medium or high risk of a kind of illegal harm set out in table C if either:

a) the **risk assessment** of the service identified a medium or high risk (as the case may be) in relation to the offences (taken together) as specified in relation to that kind of harm in table C; or

b) by virtue of a confirmation decision given under section 134 of the Act in relation to a risk of serious harm, the duty set out in section 10(2)(b) or (c) of the **Act** applies in relation to the service as if an illegal content risk assessment carried out by the provider pursuant to section 9 of the **Act** had identified a medium or high risk of serious harm (as the case may be) in relation to that kind of harm.

5.5 In relation to each offence specified in rows 2A to 18 of table C, the offence also, to the extent relevant, includes the offences of encouraging, assisting, conspiring to commit, aiding, abetting, counselling, procuring, attempting, or (in Scotland) inciting or being involved art and part in, the commission of that offence.

Multi-risk services

5.6 A service is a multi-risk service if it is at medium or high risk of two or more kinds of illegal harm set out in table C (excluding the kinds of illegal harm set out in rows 2A, 2B and 2C).

Table C – Kinds of illegal harm

	Kind of illegal harm	Offences
1.	Terrorism	An offence specified in Schedule 5 to the **Act**.
2.	CSEA	An offence specified in Schedule 6 to the **Act**.
2A.	Image-based CSAM	An offence specified in any of paragraphs 1 to 4, 7, 8 or 10 of Schedule 6 to the **Act**, so far as the risk in relation to those offences relates to **CSAM** in the form of photographs, videos or visual images.
2B.	CSAM URLs	An offence specified in any of paragraphs 1 to 4, 7, 8 or 10 of Schedule 6 to the **Act**, so far as the risk in relation to those offences relates to users encountering **CSAM** by means of or facilitated by **URLs** present on the service.
2C.	Grooming	An offence specified in any of paragraphs 5, 6, 11 or 12 of Schedule 6 to the **Act**.
3.	Encouraging or assisting suicide (or attempted suicide) and serious self-harm	An offence under section 2 of the Suicide Act 1961 (assisting suicide etc). An offence under section 13 of the Criminal Justice Act (Northern Ireland) 1966 (c. 20 (N.I.)) (assisting suicide etc). An offence under section 184 of the Online Safety Act 2023 (encouraging or assisting serious self-harm).

DRAFT FOR CONSULTATION

	Kind of illegal harm	Offences
4.	Hate	An offence under any of the following provisions of the Public Order Act 1986— (a) section 18 (use of words or behaviour or display of written material); (b) section 19 (publishing or distributing written material); (c) section 21 (distributing, showing or playing a recording); (d) section 29B (use of words or behaviour or display of written material); (e) section 29C (publishing or distributing written material); (f) section 29E (distributing, showing or playing a recording). An offence under any of the following provisions of the Crime and Disorder Act 1998— (a) section 31 (racially or religiously aggravated public order offences); (b) section 32 (racially or religiously aggravated harassment etc).

DRAFT FOR CONS

	Kind of illegal harm	Offences
5.	Harassment, stalking, threats and abuse	An offence under section 16 of the Offences against the Person Act 1861 (threats to kill). An offence under any of the following provisions of the Public Order Act 1986— (a) section 4 (fear or provocation of violence); (b) section 4A (intentional harassment, alarm or distress); (c) section 5 (harassment, alarm or distress). An offence under any of the following provisions of the Protection from Harassment Act 1997— (a) section 2 (harassment); (b) section 2A (stalking); (c) section 4 (putting people in fear of violence); (d) section 4A (stalking involving fear of violence or serious alarm or distress). An offence under any of the following provisions of the Protection from Harassment (Northern Ireland) Order 1997 (S.I. 1997/1180 (N.I. 9))— (a) Article 4 (harassment); (b) Article 6 (putting people in fear of violence) An offence under any of the following provisions of the Criminal Justice and Licensing (Scotland) Act 2010 (asp 13)— (a) section 38 (threatening or abusive behaviour); (b) section 39 (stalking).
6.	Controlling or coercive behaviour	An offence under section 76 of the Serious Crime Act 2015 (controlling or coercive behaviour in an intimate or family relationship).
7.	Drugs and psychoactive substances	An offence under any of the following provisions of the Misuse of Drugs Act 1971— (a) section 4(3) (unlawful supply, or offer to supply, of controlled drugs); (b) section 9A (prohibition of supply etc of articles for administering or preparing controlled drugs); (c) section 19 (inciting any other offence under that Act). An offence under section 5 of the Psychoactive Substances Act 2016 (supplying, or offering to supply, a psychoactive substance).

DRAFT FOR CONSULTATION

8.	Firearms and other weapons	An offence under section 1(1) or (2) of the Restriction of Offensive Weapons Act 1959 (sale etc of flick knife etc). An offence under any of the following provisions of the Firearms Act 1968— (a) section 1(1) (purchase etc of firearms or ammunition without certificate); (b) section 2(1) (purchase etc of shot gun without certificate); (c) section 3(1) (dealing etc in firearms or ammunition by way of trade or business without being registered); (d) section 3(2) (sale etc of firearms or ammunition to person other than registered dealer); (e) section 5(1), (1A) or (2A) (purchase, sale etc of prohibited weapons); (f) section 21(5) (sale etc of firearms or ammunition to persons previously convicted of crime); (g) section 22(1) (purchase etc of firearms or ammunition by person under 18); (h) section 24 (supplying firearms to minors); (i) section 24A (supplying imitation firearms to minors). An offence under any of the following provisions of the Crossbows Act 1987— (a) section 1 (sale and letting on hire of crossbow); (b) section 2 (purchase and hiring of crossbow). An offence under any of the following provisions of the Criminal Justice Act 1988— (a) section 141(1) or (4) (sale etc of offensive weapons); (b) section 141A (sale of knives etc to persons under 18). An offence under any of the following provisions of the Criminal Justice (Northern Ireland) Order 1996 (S.I. 1996/3160 (N.I. 24))— (a) Article 53 (sale etc of knives); (b) Article 54 (sale of knives etc to minors). An offence under any of the following provisions of the Knives Act 1997— (a) section 1 (unlawful marketing of knives); (b) section 2 (publication of material in connection with marketing of knives).

	Kind of illegal harm	Offences
		An offence under any of the following provisions of the Firearms (Northern Ireland) Order 2004 (S.I. 2004/702 (N.I. 3))— (a) Article 24 (sale etc of firearms or ammunition without certificate); (b) Article 37(1) (sale etc of firearms or ammunition to person without certificate etc); (c) Article 45(1) or (2) (purchase, sale etc of prohibited weapons); (d) Article 63(8) (sale etc of firearms or ammunition to people who have been in prison etc); (e) Article 66A (supplying imitation firearms to minors). An offence under section 36(1)(c) or (d) of the Violent Crime Reduction Act 2006 (sale etc of realistic imitation firearms). An offence under any of the following provisions of the Air Weapons and Licensing (Scotland) Act 2015 (asp 10)— (a) section 2 (requirement for air weapon certificate); (b) section 24 (restrictions on sale etc of air weapons).
9.	Unlawful immigration	An offence under any of the following provisions of the Immigration Act 1971— (a) section 24(A1), (B1), (C1) or (D1) (illegal entry and similar offences); (b) section 25 (assisting unlawful immigration).
10.	Human trafficking	An offence under section 2 of the Modern Slavery Act 2015 (human trafficking). An offence under section 1 of the Human Trafficking and Exploitation (Scotland) Act 2015 (asp 12) (human trafficking). An offence under section 2 of the Human Trafficking and Exploitation (Criminal Justice and Support for Victims) Act (Northern Ireland) 2015 (c. 2 (N.I.)) (human trafficking).

	Kind of illegal harm	Offences
11.	Sexual exploitation of adults	An offence under any of the following provisions of the Sexual Offences Act 2003— (a) section 52 (causing or inciting prostitution for gain); (b) section 53 (controlling prostitution for gain). An offence under any of the following provisions of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2))— (a) Article 62 (causing or inciting prostitution for gain); (b) Article 63 (controlling prostitution for gain).
12.	Extreme pornography	An offence under section 63 of the Criminal Justice and Immigration Act 2008 (possession of extreme pornographic images).
13.	Intimate image abuse	An offence under section 66B of the Sexual Offences Act 2003 (sharing or threatening to share intimate image or film). An offence under section 2 of the Abusive Behaviour and Sexual Harm (Scotland) Act 2016 (asp 22) (disclosing, or threatening to disclose, an intimate photograph or film).
14.	Proceeds of crime	An offence under any of the following provisions of the Proceeds of Crime Act 2002— (a) section 327 (concealing etc criminal property); (b) section 328 (arrangements facilitating acquisition etc of criminal property); (c) section 329 (acquisition, use and possession of criminal property).

DRAFT

	Kind of illegal harm	Offences
15.	Fraud (and financial services)	An offence under any of the following provisions of the Fraud Act 2006— (a) section 2 (fraud by false representation); (b) section 4 (fraud by abuse of position); (c) section 7 (making or supplying articles for use in frauds); (d) section 9 (participating in fraudulent business carried on by sole trader etc). An offence under section 49(3) of the Criminal Justice and Licensing (Scotland) Act 2010 (articles for use in fraud). An offence under any of the following provisions of the Financial Services and Markets Act 2000— (a) section 23 (contravention of prohibition on carrying on regulated activity unless authorised or exempt); (b) section 24 (false claims to be authorised or exempt); (c) section 25 (contravention of restrictions on financial promotion). An offence under any of the following provisions of the Financial Services Act 2012— (a) section 89 (misleading statements); (b) section 90 (misleading impressions).
16.	Foreign interference offence	An offence under section 13 of the National Security Act 2023 (foreign interference).
17.	Animal cruelty	An offence under section 4(1) of the Animal Welfare Act 2006 (unnecessary suffering of an animal).
18.	Cyberflashing	An offence under section 66A of the Sexual Offences Act 2003 (sending etc photograph or film of genitals).

User numbers

5.7 This subsection applies for the purpose of determining whether a **service** has more than a particular number of monthly **active United Kingdom users**.

5.8 A **service** has more than a particular number of monthly **active United Kingdom users**:

a) from such time as the average number of monthly **active United Kingdom users** is more than that number; and
b) until such time as the average number of monthly **active United Kingdom users** has been at or below that number for a continuous period of six months.

5.9 A **service's** average number of monthly **active United Kingdom users** is the mean number of **active United Kingdom users** per month for:

DRAFT FOR CONSULTATION

a) the six-month period ending with the month preceding the time in question; or
b) where the **service** has been in operation for less than six months, the period for which the service has been in operation.

5.10 In this subsection, an **active United Kingdom user** means any ***United Kingdom user*** who has accessed the ***user-to-user part*** of the **service**.

DRAFT FOR CONSULTATION

DRAFT FOR CONSULTATION