

Amendments to Illegal Content Codes of Practice for user-touser services and search services

DRAFT FOR CONSULTATION

Published 8 May 2024

Closing date for responses: 17 July 2024

Contents

Section

1.	Understanding this document	3
	Changes to recommended measures for user to user services	
3.	Changes to definitions and interpretation for U2U services	
4.	Changes to recommended measures for search services	16

1. Understanding this document

Background

- 1.1 Ofcom consulted on Protecting people from illegal harms online on 9 November 2023 (our Illegal Harms Consultation). Our consultation included draft illegal content Codes of Practice for user-to-user services (Annex 7) and for search services (Annex 8).
- 1.2 This document sets out proposals to amend those draft Codes of Practice. It is structured as follows:
 - a) We include measures which we intend to change or add. We will aim to amend their numbering in our final decision document so that the approach across the Child Safety Codes of Practice and Illegal Content Codes of Practice is consistent.
 - b) Entirely new measures are labelled as such and are presented in red font.
 - c) Changes are shown by presenting additions in red font and deletions in struck through red font.
 - d) We have reproduced the whole of proposed measures 5A to 5H on complaints in order to show the changes in context.
 - e) We would update the index of recommended measures accordingly.
- 1.3 We are developing the Illegal Harms and Children's Safety Codes in parallel. This means that we are receiving stakeholder feedback and progressing our work on the different Codes at the same time. For the purposes of this consultation, we have not been able to take into account all relevant responses to our Illegal Harms Consultation. The areas where we are in a position to propose changes needed to our draft illegal content Codes of Practice are exceptions. We are still considering most of the responses to our Illegal Harms Consultation and have not yet made any judgments on their merits.
- 1.4 Ofcom is required to develop separate Codes for terrorism (arising from the offences set out in Schedule 5 to the Act), 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 illegal harm. To minimise duplication and simplify the regime for service providers, we have produced one document containing the Codes for terrorism, CSEA and other duties. In our final document, we will identify the relevant Code(s) for each measure in the Index of recommended measures. All of the measures in this document would be a part of all three Codes.

2.Changes to proposed measures for user to user services

a) Content moderation - changes to wording

[4B] Setting internal content policies

Application

- [A4.6] 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

- [A4.7] The provider should set and record (but need not necessarily publish) internal content policies setting out rules, standards and guidelines around:
 - a) what content is allowed on the service and what is not; and
 - b) how policies should be operationalised and enforced.
- [A4.8] The policies should be drafted in such a way that **illegal content** (where it is identifiable as such) is not permitted.
- [A4.9] The provider should:
 - (a) have regard to the **risk assessment** of the service in setting these policies; and
 - (b) have processes in place to update these policies in response to any evidence of new and increasing **illegal harm** on the service (as tracked in accordance with Recommendation A3.13).

In setting such policies, the provider should have regard to:

a) the risk assessment of the service; and

b) information pertaining to the tracking of signals of emerging **illegal** harm.

4F Provision of training and materials to paid moderators

Application

- A4.18 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

- A4.19 The provider should ensure people working in content moderation receive training and materials that enable them to moderate **content** in accordance with Recommendation 4A and the internal content policy set in accordance with recommendation 4B. This measure does not apply in relation to **volunteers**.
- A4.20 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 Recommendation A3.13) information pertaining to the tracking of signals of emerging **illegal harm**; and
 - b) where the provider identifies a gap in moderators' understanding of a specific kind of **illegal harm**, it gives training and materials to remedy this.

b) Content moderation – the following new measure would be added after the last measure in the section

[4G] Provision of materials to volunteer moderators

Application

- [A4.21] 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

- [A4.22] The provider should ensure **volunteers** have access to materials that enable them to moderate **content** in accordance with Recommendation [4A. Having a content moderation function that allows for the swift take down of illegal content] and the internal content policy set in accordance with Recommendation [4B Setting internal content policies].
- [A4.23] 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 Recommendation A3.13); and
 - b) where the provider identifies a gap in **volunteers'** understanding of a specific kind of **illegal harm**, it gives materials to remedy this.
 - c) Reporting and complaints the following changes would be made within measure 5A to 5H, and the section renumbered accordingly.

5A Enabling complaints

Application

A5.1 This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- A5.2 The provider should have complaints processes which enable **United Kingdom users** and **affected persons** to make each type of **relevant complaint** in a way which will secure that the provider will take appropriate action in relation to them.
- Having an easy to find, easy to access and easy to use complaints system

Application

A5.3 This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- A5.4 The provider's systems and processes for making **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 should be easy to find and easily accessible;
 - the number of steps necessary (such as the number of clicks or navigation points) to submit (i) a relevant complaint using the reporting function or tool; and (ii) any other kind of relevant complaint are as few as is reasonably practicable; and
 - d) United Kingdom users and affected persons have the ability when making relevant complaints to provide give the provider with relevant information or supporting material.
- A5.5 In designing its complaints processes for **relevant complaints**, including its reporting tool or function, the provider should have regard to the particular needs of its **United Kingdom user** base as identified in its **risk assessment**. This should include the particular needs of:
 - a) **children** (considering the likely age of the **children** using that service); and
 - b) disabled people.
- A5.6 For the purposes of paragraph 5.5 a), any written information for **users** comprised in the **system or process** should be comprehensible based on the likely reading age of the youngest person permitted to use the service without the consent of a parent or guardian.
- A5.7 For the purposes of paragraph 5.5 b) the **system or process** should be designed for the purposes of ensuring usability for those dependent on assistive technologies including:
 - a) keyboard navigation; and
 - b) screen reading technology.
- X1 Provision of information prior to the submission of a complaint

Application

5X1.1 This measure applies to a **provider** in respect of each **service likely to be accessed by children** it provides.

Recommendation

- The provider should ensure that the reporting function or tool for relevant complaints regarding a specific piece of content enables

 United Kingdom users and affected persons 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 a specific piece of content has been submitted to:
 - the user that generated, uploaded or shared the content complained about; or
 - ii) any other user other than the person that submitted the complaint;

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 user or affected person that submitted the complaint, that the provider discloses to a person bringing a relevant complaint which is an appeal.
- 5C Appropriate action sending indicative timelines

Application

A5.8 This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

A5.9 The provider should acknowledge receipt of each **relevant complaint** and provide the **person that submitted the complaint complainant with an indicative timeframe for deciding the complaint.**

X2 Appropriate action – sending further information about how the complaint will be handled

Application

[5X2.1] This measure applies to a **provider** in respect of each **service likely to be accessed by children** it provides.

Recommendation

- [5X2.2] In its acknowledgment of receipt of each **relevant complaint**, the provider should provide:
 - a) the possible outcomes; and
 - b) confirmation of whether the provider will inform the person that submitted the complaint of its decision whether to uphold the complaint and details of any action taken as a result.
- 5D Appropriate action for relevant complaints about suspected illegal content

Application

A5.10 This measure applies to a **provider** in respect of each **service** it provides.

- A5.11 When the provider receives a **relevant complaint** about suspected **illegal content**:
 - a) if the provider has established a process for content prioritisation and applicable performance targets, it should handle the complaint in accordance with them;
 - b) if the service provider has no process for content prioritisation and applicable performance targets it should consider the complaint promptly; and
 - c) in either case, the provider it should act in accordance with Recommendation 4A (having a content moderation function) in relation to the suspected illegal content.

5E(i) Appropriate action for relevant complaints which are appeals – determination (large or multi-risk services)

Application

- A5.12 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.

- A5.13 For the determination of **relevant complaints** which are **appeals**, the provider should set, and monitor its performance against, performance targets relating to at least the time it takes to determine the **appeal** and the accuracy of decision making, and should resource itself so as to give effect to those targets.
- A5.14 The provider should have regard to the following matters in determining what priority to give to review of a **relevant complaint** which is an **appeal**:
 - a) the severity seriousness of the action taken against the user as
 a result of the decision that the content was illegal content or
 in relation to the content (or both) to which the appeal relates;
 - b) whether the decision that the **content** was **illegal content** was made by **proactive technology** and the likelihood of false positives generated by the specific **proactive technology** used **content identification technology** and, if so:
 - any information that Ofcom has recommended the provider collect about the likelihood of false positives generated by the specific content identification technology used; and
 - any other information available about the accuracy of the content identification technology at identifying similar types of illegal content; and
 - the service's past error rate on the service in relation to making illegal content judgements of the type concerned.

5E(i) Appropriate action for relevant complaints which are appeals – determination (services that are neither large nor multi-risk)

Application

A5.15 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

- A5.16 The provider should determine **relevant complaints** which are **appeals** promptly.
- 5F Appropriate action for relevant complaints which are appeals action following determination

Application

A5.17 This measure applies to a **provider** in respect of each **service** it provides.

- A5.18 If, in relation to a **relevant complaint** that is an **appeal**, the provider reverses a decision that **content** was **illegal content**, the provider should:
 - a) restore the content and/or the user's user account to the position they would have been in had the content not been judged to be illegal content; reverse the action taken against the user or in relation to the content (or both) as a result of that decision (so far as appropriate for the purpose of restoring the position to what it would have been had the decision not been made);
 - b) where necessary to avoid similar errors in future, adjust the relevant content moderation guidance; and
 - c) where applicable, and necessary to avoid similar errors in future, take such steps as are within its power to secure that the use of automated content moderation technology does not cause the same **content** to be taken down again.

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

Application

A5.19 This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- A5.20 This measure relates to **relevant complaints**, which are not **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 **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 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)
- A5.21 The provider should inform the complainant of their right, if they consider the provider to be in breach of contract, to bring proceedings.

5H Appropriate action for all other relevant complaints

Application

A5.22 This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

A5.23 The provider should establish a triage process for **relevant complaints** with a view to protecting **users** from harm, including harm to their rights. A responsible person, team or function should be nominated to lead this triage process and ensure **relevant complaints** reach the most relevant function or team.

A5.24 **Relevant complaints** should be dealt with:

- a) in a way that protects **users** and the provider's compliance with other applicable laws in question;
- b) within timeframes the provider has determined are appropriate; and
- c) in accordance with recommendations 5A to 5H (where applicable).

d) Terms of service - the following new measure would be added after measure 6A

6AA Terms of service: substance (category 1 services)

Application

[6AA.1] This measure applies to a **provider** of a **category 1 service** in respect of each **category 1 service** it provides.

Recommendation

[6AA.2] The provider should summarise the findings of its **risk assessment** (including as to levels of risk and as to nature, and severity, of potential harm) in the **terms of service**.

e) Terms of service – the following change would be made to measure 6B

6B Clarity and accessibility

Application

[A6.4] This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

[A6.5] The provider should ensure that the provisions included in the **terms of service** in accordance with Recommendation 6A are:

- a) easy to find, such that they are:
 - i) clearly signposted for the general public, regardless of whether they have signed up to or are just using the service; and
 - ii) locatable within the terms of service;
- b) laid out and formatted in a way that helps **users** read and understand them;
- c) written to a reading age comprehensible for the youngest person 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.

3. Changes to proposed definitions and interpretation for U2U services

3.1 The following terms would be inserted into the list of defined terms at the appropriate alphabetical point.

Term	Meaning
Category 1 service	Section 95(10)(a) A regulated user-to-user service for the time being included in the part of the register established under section 95(2)(a) of the Act.
Content identification technology	Section 231(2) "Content identification technology" means technology, such as algorithms, keyword matching, image matching or image classification, which analyses content to assess whether it is content of a particular kind (for example, illegal content).
Systems and/or processes	Section 236 any reference to systems and/or processes is to human or automated systems and/or processes, and accordingly includes technologies.

4.Changes to proposed measures for search services

A. Search moderation - changes to wording

[4B] Setting internal content policies

Application

- [A4.7] This measure applies to a **provider** in respect of each **service** it provides that is either (or both) of the following:
 - a) a large general search service; or
 - b) a multi-risk service.

Recommendation

- [A4.8] The provider should set and record (but need not necessarily publish) internal content policies setting out rules, standards and guidelines around:
 - a) what search content is allowed on the service and what is not; and
 - b) how policies should be operationalised and enforced.
- [A4.9] The policies should be drafted in such a way that **illegal content** (where it is identifiable as such) is **deindexed** or **downranked** in accordance with Recommendation 4A (having a **search moderation function**).
- [A4.10] The provider should:
 - have regard to the risk assessment of the service in setting these policies; and
 - have processes in place to update these policies in response to any evidence of new and increasing **illegal harm** on the service (as tracked in accordance with Recommendation A3.13).

In setting such policies, the provider should have regard to:

a) the risk assessment of the service; and

b) information pertaining to the tracking of signals of emerging **illegal** harm.

4F Provision of training and materials to paid moderators

Application

- A4.18 This measure applies to a **provider** in respect of each **service** it provides that is either (or both) of the following:
 - a) a large general search service; or
 - b) a multi-risk service.

Recommendation

A4.19 The provider should ensure people working in content moderation receive training and materials that enable them to moderate **content** in accordance with Recommendation 4A and the internal content policy set in accordance with recommendation 4B. This measure does not apply in relation to **volunteers**.

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 Recommendation A3.13) information pertaining to the tracking of signals of emerging **illegal harm**; and
- b) where the **provider** identifies a gap in moderators' understanding of a specific kind of **illegal harm**, it gives training and materials to remedy this.

B. Reporting and complaints – the following changes would be made within measures 5A to 5H, and the section renumbered accordingly.

[5A] Enabling complaints

Application

[A5.1] This measure applies to a **provider** in respect of each **service** it provides.

[A5.2] The provider should have complaints processes which enable **United Kingdom users, affected persons** and **interested persons** to make each type of **relevant complaint** in a way which will secure that the **provider** will take appropriate action in relation to them.

[5B] Having an easy to find, easy to access and easy to use complaints system

Application

[A5.3] This measure applies to a **provider** in respect of each **service** it provides.

- [A5.4] The provider's systems and processes for making **relevant complaints** should be operated to ensure that:
 - a) for relevant complaints regarding a specific piece of search content, a reporting function or tool is clearly accessible in relation to that search content;
 - b) processes for making other kinds of **relevant complaints** should be easy to find and easily accessible;
 - the number of steps necessary (such as the number of clicks or navigation points) to submit (i) a relevant complaint using the reporting function or tool; and (ii) any other kind of relevant complaint are as few as is reasonably practicable; and
 - d) United Kingdom users, affected persons and interested persons have the ability when making relevant complaints to give provide the provider with relevant information or supporting material.
- [A5.5] In designing its complaints processes for **relevant complaints**, including its reporting tool or function, the provider should have regard to the particular needs of its **United Kingdom user** base as identified in its **risk assessment**. This should include the particular needs of:
 - a) children (for services likely to be accessed by children and considering the likely age of the children using that service); and
 - b) disabled people.
- [A5.6] For the purposes of paragraph A5.5(a), any written information for users comprised in the system or process should be comprehensible based on the likely reading age of the youngest person permitted to agree to the service's publicly available statement without the consent of a parent or guardian.

- [A5.7] For the purposes of paragraph A5.5(b), the **system or process** should be designed for the purposes of ensuring usability for those dependent on assistive technologies including:
 - a) keyboard navigation; and
 - b) screen reading technology
- [5C] Appropriate action sending indicative timelines

Application

[A5.8] This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- [A5.9] The provider should acknowledge receipt of each **relevant complaint** and provide the **person that submitted the complaint complainant with an indicative timeframe for deciding the complaint.**
- [Y1] Appropriate action sending further information about how the complaint will be handled

Application

[Y1.1] This measure applies to a **provider** in respect of each **service likely to be accessed by children** it provides.

- [Y1.2] In its acknowledgment of receipt of each **relevant complaint**, the provider should provide:
 - a) the possible outcomes; and
 - b) confirmation of whether the provider will inform the person that submitted the complaint of its decision whether to uphold the complaint and details of any action taken as a result.

5D Appropriate action for relevant complaints about suspected illegal content

Application

[A5.10] This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- [A5.11] When the provider receives a **relevant complaint** about suspected **illegal content**:
 - a) if the provider has established a process for search moderation prioritisation and applicable performance targets, it should handle the complaint in accordance with them;
 - b) if the **service** provider has no process for search moderation prioritisation and applicable performance targets it should consider the complaint promptly; and
 - c) in either case, the provider it should act in accordance with Recommendation 4A (having a search moderation function) in relation to the suspected illegal content.
- 5E(i) Appropriate action for relevant complaints which are appeals determination (large or multi-risk services)

Application

- A5.12 This measure applies to a **provider** in respect of each **service** it provides that is either (or both) of the following:
 - a) a large general search service; or
 - b) a multi-risk service.

Recommendation

A5.13 For the determination of **relevant complaints** which are **appeals**, the provider should set, and monitor its performance against, performance targets relating to at least the time it takes to determine the **appeal** and the accuracy of decision making, and should resource itself so as to give effect to those targets

- A5.14 The provider should have regard to the following matters in determining what priority to give to review of a **relevant complaint** which is an **appeal**:
 - a) the severity seriousness of the action taken against the interested person as a result of the decision that the content was illegal content;
 - b) whether the decision that the **content** was **illegal content** was made by **proactive technology** and the likelihood of false positives generated by the specific **proactive technology** used **content identification technology** and, if so,
 - any information that Ofcom has recommended the provider collect about the likelihood of false positives generated by the specific content identification technology used; and
 - any other information available about the accuracy of the content identification technology at identifying similar types of illegal content; and
 - c) the service's past error rate on the service in relation to making illegal content judgements of the type concerned.
- 5E(ii) Appropriate action for relevant complaints which are appeals determination (services that are neither large nor multi-risk)

Application

A5.15 This measure applies to a provider in relation to a service that is neither a large general search service nor a multi-risk service.

Recommendation

- [A5.16] The provider should determine **relevant complaints** which are **appeals** promptly.
- 5F Appropriate action for relevant complaints which are appeals action following determination

Application

[A5.17] This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- [A5.18] If, in relation to a **relevant complaint** that is an **appeal**, the provider reverses a decision that **search content** was **illegal content**, the provider should:
 - a) reverse the action taken against the interested person or in relation to the search content (or both) as a result of that decision (so far as appropriate for the purpose of restoring the position to what it would have been had the decision not been made) restore the search content to the position it would have been in had it not been judged to be illegal content;
 - b) where necessary to avoid similar errors in future, adjust the relevant moderation guidance; and
 - c) where applicable, and necessary to avoid similar errors in future, take such steps as are within its power to secure that the use of automated content moderation technology does not cause the same search content to be deindexed or downranked again to no longer appear in search results or be given a lower priority in search results again.
- Appropriate action for relevant complaints about proactive technology, which are not appeals

Application

[A5.19] This measure applies to a **provider** in respect of each **service** it provides.

- [A5.20] This measure relates to **relevant complaints**, which are not **appeals**, about the use of **proactive technology** on the service when:
 - a) the use of **proactive technology** on the service results in search content being **deindexed**, or **downranked**; and
 - b) the **interested person** considers that the **proactive technology** has been used in a way not contemplated by, or in breach of, the publicly available statement (for example, by **deindexing** or **downranking search content** not of a kind specified in the **publicly available statement** as a kind of **search content** in relation to which the technology would operate).

[A5.21] Where relevant, the **provider** should inform the complainant of their right, if they consider the **provider** to be in breach of contract, to bring proceedings.

5H Appropriate action for all other relevant complaints

Application

[A5.22] This measure applies to a **provider** in respect of each **service** it provides.

Recommendation

- [A5.23] The **provider** should establish a triage process for **relevant complaints** with a view to protecting **users** and **interested persons** from harm, including harm to their rights. A responsible person, team or function should be nominated to lead this triage process and ensure **relevant complaints** reach the most relevant function or team.
- [A5.24] **Relevant complaints** should be dealt with:
 - a) in a way that protects **users** and the provider's compliance with other applicable laws in question;
 - b) within timeframes the **provider** has determined are appropriate; and
 - c) in accordance with Recommendations 5D to 5G (where applicable).

c) Publicly available statements - the following new measure would be added after measure 6A

6AA Publicly available statements: substance (category 2A services)

Application

[X] This measure applies to a **provider** of a **category 2A service** in respect of each **category 2A service** it provides.

[X] The provider should summarise the findings of its **risk assessment** (including as to levels of risk and as to nature, and severity, of potential harm) in the **publicly available statement**.

d) Publicly available statements – the following change would be made to measure 6B

6B Clarity and accessibility

Application

[A6.4] This measure applies to a **provider** in respect of each **service** it provides.

- [A6.5] The provider should ensure that the provisions included in a **publicly** available statement in accordance with Recommendation 6A are:
 - a) easy to find, such that they are:
 - i) clearly signposted for the general public, regardless of whether they have signed up to or are just using the service; and
 - ii) locatable within the publicly available statement;
 - b) laid out and formatted in a way that helps **users** read and understand them;
 - c) written to a reading age comprehensible for the youngest person permitted to use the service without the consent of a parent or guardian;
 - d) designed for the purposes of ensuring usability for those dependent on assistive technologies including:
 - i) keyboard navigation; and
 - ii) screen reading technology.

5. Changes to proposed definitions and interpretation for search services

The following terms would be inserted into the list of defined terms at the appropriate alphabetical point.

Term	Meaning
Category 2A service	Section 95(10)(b) A regulated search service or a combined service for the time being included in the part of the register established under section 95(2)(b) of the Act.
Content identification technology	Section 231(2) "Content identification technology" means technology, such as algorithms, keyword matching, image matching or image classification, which analyses content to assess whether it is content of a particular kind (for example, illegal content).
Systems and/or processes	Section 236 any reference to systems and/or processes is to human or automated systems and/or processes, and accordingly includes technologies.