

# Protecting children from harms online

Proposed codes at a glance

Welsh version available

## **Proposed Codes at a glance**

This document summarises the proposed Codes measures in this consultation and outlines which services they apply to. Our draft measures together will help create safer experiences online for children in the UK.

We have assessed what measures should be taken by what types of services. We recognise that the size, capacity, functionalities, user base and risks of online services differ widely, and so have not taken a one-size-fits-all approach. Instead, we propose that services who pose the greatest risk to children need to do more to keep them safe.

This document summarises the safety measures that we propose apply to (1) user-to-user services and (2) search services. We set these out in turn, alongside what types of service we are proposing these measures apply to.

Some of these proposed measures would apply to all user-to-user and search services respectively. These measures have been highlighted in pink in the tables below.

Other measures we propose to include in our Illegal Harms Codes and they are highlighted in blue.

Whether some of the measures are recommended for a particular service would depend on the size of the service's user base and on the risk of harm posed to children, based on the service's most recent children's risk assessment. We have proposed the following definitions and terms:

#### Size of UK user base

• Large services: we propose to define a service as 'large' where it has an average user base greater than 7 million per month in the UK, approximately equivalent to 10% of the UK population.

#### Risk level and type

- Multi-risk: this refers to a service that has medium or high risk for two or more kinds of content harmful to children (i.e. at least two across the 4 kinds of PPC, 8 kinds of PC and any kinds of NDC as applicable).
- Other measures are recommended for services with medium or high risk for at least one kind of content harmful to children, from a defined subset of kinds of content as relevant to each measure.

#### **Purpose of services**

• Some measures are also based on the **principal purpose** of the service. This refers to the main activity or objective of the service. Those measures depend on whether a service's principal purpose is to host or disseminate certain kinds of content.

<sup>&</sup>lt;sup>1</sup> Refer to <u>Volume 5, Section 2 'Developing the Children's Safety Codes: Our framework'</u> for more information on our approach to size of service and risk of harm posed to children.

### Measures we have proposed for user-to-user services

The table below sets out our proposed measures for user-to-user services that are likely to be accessed by children and the types of services that should implement each measure. The measures are discussed in more detail in Volume 5, with each grouping of measures representing a section. Governance and Accountability measures are discussed in Volume 4 at Section 11. The draft Children's Safety Codes can be found at Annex 7 (for user-to-user services) and Annex 8 (for search services).

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
	Governance and ac	countability (Volume 4, Section 11)	
GA1	Most senior body to carry out and record an annual review of risk management activities relating to children's safety	All user-to-user services that are large	3A
GA2	Name a person accountable to most senior governance body for compliance with children's safety duties	All user-to-user services	3В
GA3	Written statements of responsibility for senior members who make decisions relating to management of child safety risks	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	3C
GA4	Have an internal monitoring and assurance function to provide independent assurance that measures are effective	<ul> <li>All user-to-user services that are both:</li> <li>Large; and</li> <li>Multi-risk for content harmful to children.</li> </ul>	3D
GA5	Track unusual increases or new kinds of Primary-Priority Content, Priority Content, and Non-designated Content on the service that may be becoming present on the service	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	3E
GA6	Have a Code of Conduct that sets standards for employees around protecting children	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	3F
GA7	Ensure staff involved in the design and operational management of service are sufficiently trained in approach to compliance with children's safety duties	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	3G

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
	Age assura	nce (Volume 5, Section 15)	
AA1	Use highly effective age assurance to prevent children from accessing the entire service	<ul> <li>All user-to-user services:</li> <li>Whose principal purpose is the hosting or the dissemination of one or more kinds of Primary Priority Content.</li> </ul>	None – new measure
		All user-to-user services:	
AA2	Use highly effective age assurance to prevent children from accessing the entire service	<ul> <li>Whose principal purpose is the hosting or the dissemination of one or more kinds of Priority Content, and</li> <li>Who are medium or high risk for one or more of those kinds of Priority Content.</li> </ul>	None – new measure
	Use highly effective age	All user-to-user services:  • Whose principal purpose is not the	
ААЗ	assurance to ensure children are prevented from encountering Primary Priority Content identified on the service	<ul> <li>hosting or the dissemination of one or more kinds of Primary Priority Content, and</li> <li>Which do not prohibit one or more kinds of Primary Priority Content.</li> </ul>	None – new measure
		All user-to-user services:	
<b>AA4</b>	Use highly effective age assurance to ensure children are protected from encountering Priority Content identified on the service	<ul> <li>Whose principal purpose is not the hosting or the dissemination of one or more kinds of Priority Content, and</li> <li>Which do not prohibit one or more kinds of Priority Content, and</li> <li>Are medium or high risk for one or more kinds of Priority Content that they do not prohibit.</li> </ul>	None – new measure
		All user-to-user services that:	
AA5	Use highly effective age assurance to apply relevant recommender system measures in the Code to children	<ul> <li>Are medium or high risk for one or more kinds of Primary Priority Content, and</li> <li>Operate a content recommender system</li> </ul>	None – new measure

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> Harms
AA6	Use highly effective age assurance to apply relevant recommender system measures in the Code to children	<ul> <li>All user-to-user services that:</li> <li>Are medium or high risk for one or more kinds of relevant Priority Content (excluding bullying), and</li> <li>Operate a content recommender system<sup>2</sup></li> </ul>	None – new measure
	Content mode	eration (Volume 5, Section 16)	
CM1	Content moderation systems and processes designed to swiftly take action against content harmful to children	All user-to-user services	4A
CM2	Set internal content policies	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	4B
СМЗ	Set performance targets for content moderation function	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	4C
CM4	Have and apply policies on prioritisation of content for review	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	4D
CM5	Ensure content moderation functions are well-resourced	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	4E
CM6	Ensure content moderation teams are appropriately trained	<ul> <li>All user-to-user services that are:</li> <li>Large, or</li> <li>Multi-risk for content harmful to children.</li> </ul>	4F
CM7	Volunteer moderators should be provided with materials for their roles	All user-to-user services that use volunteer moderation and are either:  Large, or  Multi-risk for content harmful to children.	4G³

<sup>&</sup>lt;sup>2</sup> We are also minded to extend this measure for two potential kinds of Non-Designated Content (NDC). Refer to <u>Volume 5, Section 19 'Recommender Systems'</u> for more information.

<sup>&</sup>lt;sup>3</sup> Following a review of evidence, we are proposing the inclusion of Measure 4G in our Illegal Content Codes as well as the equivalent Measure CM7 in our draft Children's Safety Codes.

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
	User reporting and	complaints (Volume 5, Section 17)	
UR1	Have complaints processes which enable users to make relevant complaints for services likely to be accessed by children	All user-to-user services	5A
UR2	Have easy to access and use, and transparent complaints systems	All user-to-user services	5B
UR3	Acknowledge receipt of complaints with indicative timeframe and information on resolution	All user-to-user services	5C
UR4	User-to-user services take appropriate action in response to each complaint	All user-to-user services <sup>4</sup>	5D-5H
	Terms of service/publicly a	vailable statements (Volume 5, Section 18)	
TS1	Terms and statements regarding the protection of children should contain all information mandated by the Act	All user-to-user services	6A
TS2	Terms and statements regarding the protection of children should be clear and accessible	All user-to-user services	6B
TS3	Services should summarise the findings of their most recent children's risk assessment in their terms or statements	All Category 1 services	6AA <sup>5</sup>
	Recommender	systems (Volume 5, Section 19)	
RS1	Ensure that content likely to be Primary Priority Content is not recommended to children	<ul> <li>All user-to-user services that:</li> <li>Operate a content recommender system, and</li> <li>Are medium or high risk for at least one kind of Primary Priority Content.</li> </ul>	None – new measure

-

<sup>&</sup>lt;sup>4</sup> For some types of complaints we make different recommendations for some services, based on their size and risk level and depending on whether they are in scope of separate Age Assurance measures. Refer to <u>Volume 5</u>, <u>Section 19 'User Reporting and Complaints'</u>.

<sup>&</sup>lt;sup>5</sup> Following the publication of our categorisation advice to the Secretary of State in March 2024, we are proposing the inclusion of Measure 6AA in our Illegal Content Codes, as well as the equivalent Measure TS3 in our draft Children's Safety Codes.

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
RS2	Ensure that content likely to be Priority Content is reduced in prominence on children's recommender feeds	<ul> <li>All user-to-user services that:</li> <li>Operate a content recommender system, and</li> <li>Are medium or high risk for at least one kind of Priority Content (excluding bullying)<sup>6</sup></li> </ul>	None – new measure
RS3	Enable children to provide negative feedback on content that is recommended to them	<ul> <li>All user-to-user services that:</li> <li>Operate a content recommender system, and</li> <li>Are medium or high risk for at least two kinds of Primary Priority Content and/or Priority Content (excluding bullying)<sup>7</sup>, and</li> <li>Are large</li> </ul>	None – new measure
	User supp	ort (Volume 5, Section 20)	
US1	Provide children with an option to accept or decline an invite to a group chat	<ul> <li>All user-to-user services that:</li> <li>Have group chats, and</li> <li>Are medium or high risk of one or more of: pornographic content, eating disorder content, bullying content, abuse and hate content<sup>8</sup> and violent content.<sup>9</sup></li> </ul>	None – new measure
US2	Provide children with the option to block and mute other users' accounts	<ul> <li>All user-to-user services that:</li> <li>Have user profiles and certain user interaction functionalities, <sup>10</sup> and</li> <li>Are medium or high risk of one of more of: bullying content, abuse and hate content and violent content</li> </ul>	9A

\_

<sup>&</sup>lt;sup>6</sup> We are also minded to extend this measure for two potential kinds of NDC. Refer to <u>Volume 5, Section 19</u> 'Recommender Systems' for more information.

<sup>&</sup>lt;sup>7</sup> We are also minded to extend this measure for two potential kinds of NDC. Refer to <u>Volume 5, Section 19</u> <u>'Recommender Systems'</u> for more information.

<sup>&</sup>lt;sup>8</sup> We use 'abuse and hate' content to refer to the two kinds of content defined in the Act in sections 62(2) and 62(3). A service is considered to have medium or high risk for abuse and hate content if it has medium or high risk for at least one of the two kinds of content defined in the Act in sections 62(2) and 62(3).

<sup>&</sup>lt;sup>9</sup> We use 'violent content' to refer to the three kinds of content defined in the Act in sections 62(4), 62(6) and 62(7). A service is considered to have medium or high risk for violent content if it has medium or high risk for at least one of the three kinds of content defined in the Act in sections 62(4), 62(6) and 62(7).

<sup>&</sup>lt;sup>10</sup> These functionalities include user connections, posting content, and user communication. Refer to <u>Volume 5</u>, <u>Section 20 'User Support Measures'</u> for more information.

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> Harms
US3	Provide children with the option to disable comments on their own posts	<ul> <li>All user-to-user services that:</li> <li>Have comment functionalities, and</li> <li>Are medium or high risk of one or more of: bullying content, abuse and hate content and violent content</li> </ul>	9В
US4	The provision of information to child users when they restrict interactions with other accounts or content	<ul> <li>All user-to-user services that:</li> <li>Have certain functionalities that restrict interactions with other accounts or content, 11 and</li> <li>Are large, and</li> <li>Are multi-risk for content harmful to children</li> </ul>	7B (partly mirrored)
US5	Signpost children to support at key points in the user journey	Intervention point 1 – when children report content  All user-to-user services that  Are medium or high risk of one or more of: suicide content, self-harm content, eating disorder content, or bullying content  Intervention point 2 – when children post or re-post content  Large user-to-user services that  Have posting/re-posting functionalities, and  Are medium or high risk of one or more of: suicide content, self-harm content, eating disorder content, or bullying content, and  Have measures that enable them to identify when a user posts or re-posts suicide, self-harm, eating disorder or bullying content  Intervention point 3 – when children search for harmful content:  All user-to-user services that  Have user-generated content searching, and  Are high or medium risk of one or	None – new measure

<sup>&</sup>lt;sup>11</sup> Refer to <u>Volume 5, Section 20 'User Support Measures'</u> for more information on the functionalities that are applicable to Measure US4.

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
		<ul> <li>Have measures that enable them to become aware of when a user searches using suicide, self-harm or eating disorder related search terms</li> </ul>	
US6	Provide age-appropriate user support materials for children	All User-to-user services that are multi- risk for content harmful to children	None – new measure

## Measures we have proposed for search services

The table below sets out the proposed measures for search services that are likely to be accessed by children.

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
	Governance and account	ability (Volume 4, Section 11)	
GA1	Most senior body to carry out and record an annual review of risk management activities relating to children's safety	All large general search services	3A
GA2	Name a person accountable to most senior governance body for compliance with children's safety duties	All search services	3В
GA3	Written statements of responsibility for senior members who make decisions relating to management of child safety risks	<ul> <li>All search services that are:</li> <li>Large general search services, or</li> <li>Multi-risk for content harmful to children</li> </ul>	3C
GA4	Have an internal monitoring and assurance function to provide independent assurance that measures are effective	<ul> <li>All search services that are both:</li> <li>Large, and</li> <li>Multi-risk for content harmful to children</li> </ul>	3D
GA5	Track unusual increases or new kinds of PPC, PC and NDC on service	<ul> <li>All search services that are:</li> <li>Large general search, or</li> <li>Multi-risk for content harmful to children</li> </ul>	3E

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
GA6	Have a Code of Conduct that sets standards for employees around protecting children	<ul> <li>All search services that are:</li> <li>Large general search services, or</li> <li>Multi-risk for content harmful to children</li> </ul>	3F
GA7	Ensure staff involved in the design and operational management of service are sufficiently trained in approach to compliance with children's safety duties	<ul> <li>All search services that are:</li> <li>Large general search services, or</li> <li>Multi-risk for content harmful to children</li> </ul>	3G
	User reporting and comp	plaints (Volume 5, Section 17)	
UR1	Have complaints processes which enable users to make relevant complaints for services likely to be accessed by children	All search services	5A
UR2	Have easy to access and use, and transparent complaints systems	All search services <sup>12</sup>	5B
UR3	Acknowledge receipt of complaints with indicative timeframe and information on resolution	All search services	5C
UR5	Search services take appropriate action in response to each complaint	All search services <sup>13</sup>	5D-5H
	Terms of service/publicly availab	ole statements (Volume 5, Section 18)	
TS1	Terms and statements regarding the protection of children should contain all information mandated by the Act	All search services	6A
TS2	Terms and statements regarding the protection of children should be clear and accessible	All search services	6B
TS3	Services should summarise the findings of their most recent children's risk assessment in their terms or statements	All Category 2A services	6AA

 $<sup>^{\</sup>rm 12}$  With the exception of measure UR2(e) that does not apply to search services

<sup>&</sup>lt;sup>13</sup> For some types of complaints we make different recommendations for some services, based on their size and risk level. Refer to <u>1.250 and 1.251 in Volume 5, Section 5 'User Reporting and Complaints'</u>.

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
	Search content modera	ation (Volume 5, Section 16)	
SM1	Have moderation systems and processes in place to take appropriate action on Primary Priority Content, Priority Content and Non-Designated Content  A) When Primary Priority Content has been identified, downrank and/or blur the search content  B) When Priority Content and Non-Designated Content has been identified, consider if it is appropriate to downrank and/or blur the search content	All search services	4A
SM2	When a user is believed to be a child, filter identified Primary Priority Content out of their search results through a safe search setting. Users believed to be a child should not be able to turn this setting off	All large general search services	4A
SM3	Set and record internal content policies	<ul><li>All services that are:</li><li>Large general search services, or</li><li>Multi-risk search services</li></ul>	4B
SM4	Set performance targets for search moderation functions	All services that are:  Large general search services, or  Multi-risk search services	4C
SM5	Develop and apply policies on prioritisation of content for review	<ul><li>All services that are:</li><li>Large general search services, or</li><li>Multi-risk search services</li></ul>	4D
SM6	Ensure search moderation functions are sufficiently resourced	<ul><li>All services that are:</li><li>Large general search services, or</li><li>Multi-risk search services</li></ul>	4E
SM7	Ensure people working in search moderation receive training and materials	<ul><li>All services that are:</li><li>Large general search services, or</li><li>Multi-risk search services</li></ul>	4F

No.	Description of proposed measure	Services we propose this will apply to	Equivalent measure for <u>Illegal</u> <u>Harms</u>
	Search features, functionalities ar	nd user support (Volume 5, Section 21)	
SD1	Offer users a means to easily report predictive search suggestions relating to Primary Priority Content and Priority Content	All large general search services	7A
SD2	Provide crisis prevention information in response to known Primary Priority Content-search requests regarding suicide, self-harm and eating disorders	All large general search services	7C
	User support (V	olume 5, Section 20)	
US6	Provide age-appropriate user support materials for children	All search services that are multi-risk for content harmful to children	None – new measure