

BROADCASTING ACT 1994

CODE OF PRACTICE FOR ONLINE SAFETY – SOCIAL MEDIA SERVICES

1. In exercise of the powers conferred by section 45L of the Broadcasting Act 1994, the Info-communications Media Development Authority (“IMDA”) hereby issues the following online Code of Practice (“Code”).

Title and Commencement

2. This Code is called the Code of Practice for Online Safety – Social Media Services and shall come into effect on 18 July 2023.

Purpose of this Code

3. This Code specifies outcomes that Social Media Services (“Service”) which are designated/will be designated under section 45K(1) of the Broadcasting Act 1994 have to meet to enhance online user safety, particularly for children, and curb the spread of harmful content on their service.
4. The categories of harmful content include:
 - a. Sexual content
 - b. Violent content
 - c. Suicide and self-harm content
 - d. Cyberbullying content
 - e. Content endangering public health
 - f. Content facilitating vice and organised crime

Application

5. This Code applies to Social Media Services which are designated/will be designated under section 45K(1) of the Broadcasting Act 1994.

Definitions

6. For the purpose of this Code, the following definitions shall apply:
 - a. “community guidelines and standards” means guidelines issued by the Service on impermissible content and end-user activity.
 - b. “content moderation” means processes developed and activities taken by the Service to (i) detect, whether through the Service’s systems or in response to user reporting; (ii) assess; and (iii) address harmful content for end-users or content inappropriate for children on the Service in accordance with its

community guidelines and standards such as by removing or restricting access to the content.

- c. “child” means an individual who is below 18 years of age.
- d. “end-user” means Singapore end-user.

Obligations

7. The obligations are categorised into three sections:

Section A - User Safety;
Section B - User Reporting and Resolution; and
Section C - Accountability.

Section A – User Safety

- 8. End-users must be able to use the Service in a safe manner. In this regard, the Service must put in place measures to minimise end-users’ exposure to harmful content, empower end-users to manage their safety on the Service and mitigate the impact on end-users that may arise from the propagation of harmful content.
- 9. Children in particular, may lack the capacity or experience to deal with the information and content available online and will need more protection to ensure a safer online space for them. In this regard, the Service must therefore also have specific measures to protect children from harmful content.
- 10. Measures to comply with the obligations in paragraphs 8 and 9 must include those found in (Ai) and (Aii) below.

(Ai.) Measures for all end-users

Community guidelines and standards and content moderation

- 11. End-users’ exposure to harmful content must be minimised via reasonable and proportionate measures. These measures include, but are not limited to, a set of community guidelines and standards, and content moderation measures that are put in place and effected by the Service. The Service’s community guidelines and standards must address the categories of harmful content in paragraph 4 and must be published.

Empower end-users and improve safety

- 12. End-users must have access to tools that enable them to manage their own safety and effectively minimise their exposure to, and mitigate the impact of, harmful content and unwanted interactions on the Service. Such tools may include:
 - a. Tools to restrict visibility of harmful content and/or unwanted comments.

- b. Tools to limit visibility of the end-user's account, including profile and content, as well as contact and/or interactions with other end-users.
 - c. Tools to limit location sharing.
13. End-users must be able to easily access information related to online safety on the Service. Such information must be easy to understand and must include the availability of tools and local information, including Singapore-based safety resources or support centres, if available. The Service should seek to implement, support and/or maintain programmes and initiatives to educate and raise awareness of such information.
14. End-users who use high-risk search terms such as, but not limited to, terms relating to self-harm and suicide on the Service must be actively offered relevant safety information (stated in paragraph 13) such as, but not limited to, local suicide prevention hotlines, if available.

Proactive detection and removal

15. End-users' exposure to child sexual exploitation and abuse material and terrorism content on the Service must be minimised through the use of technologies and processes. These technologies and processes must proactively detect and swiftly remove child sexual exploitation and abuse material and terrorism content as technically feasible, such that the extent and length of time to which such content is available on the Service is minimised.
16. End-users must be protected from preparatory child sexual exploitation and abuse activity and terrorism activity on the Service through reasonable and proportionate steps taken by the Service to proactively detect and swiftly remove preparatory child sexual exploitation and abuse activity (such as online grooming for child sexual abuse) and terrorism activity (such as glorifying or endorsing terrorist activities and recruitment).

(Aii.) Measures for children

Community guidelines and standards and content moderation

17. Besides harmful content, children's exposure to inappropriate content must also be minimised through reasonable and proportionate measures. These measures include, but are not limited to, a set of community guidelines and standards and content moderation measures put in place and effected by the Service that are appropriate for children. These community guidelines and standards must minimally address the following categories of content, and must be published:
- a. Sexual content
 - b. Violent content
 - c. Suicide and self-harm content
 - d. Cyberbullying content

18. Children must not be targeted to receive content that the Service is reasonably aware to be detrimental to their physical or mental well-being. Such content includes the categories of harmful and/or inappropriate content in paragraphs 4 and 17. In this regard, content targeting refers, but is not limited to, advertisements, promoted content and content recommendations.

Protection for children

19. Children or their parents/guardians must have access to tools that enable them to manage children's safety, and effectively minimise children's exposure to, and mitigate the impact of, harmful and/or inappropriate content and unwanted interactions on the Service. These tools may include the following:

- a. Tools to effectively manage the content that children see and/or their experiences.
- b. Tools to:
 - i. Limit the public visibility of children's accounts, including their profile and content;
 - ii. Limit who can contact and/or interact with children's accounts; and
 - iii. Limit location sharing.

20. Unless the Service restricts access by children, children must be provided differentiated accounts whereby the settings for the tools to minimise exposure and mitigate impact of harmful and/or inappropriate content and unwanted interactions are robust and set to more restrictive levels that are age appropriate by default. Children or their parents/guardians must be provided clear warnings of implications if they opt out of the default settings.

21. Children must be able to easily access information related to online safety on the Service. Such information must be easily understood by children and must include information on tools available to protect children from harmful and/or inappropriate content and unwanted interactions, as well as local information, including Singapore-based safety resources or support centres, if available. The Service should seek to implement, support and/or maintain programmes and initiatives to educate and raise awareness of such information.

22. Children who use high-risk search terms, such as, but not limited to, terms relating to self-harm and suicide, on the Service must be actively offered relevant safety information (stated in paragraph 21) such as, but not limited to, local suicide prevention hotlines, if available.

Section B – User Reporting and Resolution

23. Any individual must be able to report concerning content or unwanted interactions to the Service in relation to the categories of harmful and/or inappropriate content in paragraphs 4 and 17. In this regard, the reporting and resolution mechanism

provided to end-users must be effective, transparent, easy to access, and easy to use.

- a. End-users' reports must be assessed, and appropriate action(s) must be taken by the Service in a timely and diligent manner that is proportionate to the severity or imminence of the potential harm. In particular, timelines must be expedited for content and activity related to terrorism. Appropriate action(s) may include:
 - i. Swiftly removing the reported content or restricting access to the reported content; and
 - ii. Warning, suspending, or banning the account(s) that generated, uploaded, or shared the reported content.
- b. Where the Service receives a report that is not frivolous or vexatious:
 - i. The end-user who submitted the report must be informed of the Service's decision and action taken with respect to that report without undue delay.
 - ii. Should the Service decide to take action against the reported content or account(s), the end-user holding the account(s) that generated, uploaded, or shared the reported content must be informed of the Service's decision and action without undue delay.
- c. The end-users referred to in sub-paragraphs (b)(i) and (b)(ii) must be allowed to submit requests to the Service for a review of the decision and action taken.

Section C – Accountability

24. End-users must have access to clear and easily comprehensible information that enable them to assess the level of safety and related safety measures afforded by the Service and make informed choices.
25. In this regard, the Service must submit to IMDA annual online safety reports on the measures the Service has put in place to combat harmful and inappropriate content, for publishing on IMDA's website. The annual online safety reports must reflect Singapore end-users' experience on the Service, including:
 - a. What steps the Service has taken to mitigate Singapore end-users' exposure to harmful or inappropriate content, including descriptions of specific measures that the Service has in place to enhance online safety for end-users in Singapore in relation to obligations in Sections A and B;
 - b. How much and what types of harmful or inappropriate content end-users in Singapore encounter on the Service; and
 - c. What action(s) the Service has taken on end-user reports.

26. The Service may propose suitable information and metrics to be included in its annual online safety reports. These are subject to agreement by IMDA. These may include but are not limited to:
- a. The number and types of end-user reports received from end-users in Singapore, and the number and types of harmful and inappropriate content removed as a result of end-user reports;
 - b. The time between the Service receiving end-user reports from end-users in Singapore on harmful and inappropriate content and taking action (if any) as an aggregate;
 - c. The number and types of harmful or inappropriate content proactively removed by the Service that are:
 - i. Accessible by end-users in Singapore; and
 - ii. Originated from Singapore.
 - d. The number of accounts suspended or banned in Singapore, and the reasons for suspending or banning accounts in relation to the categories of harmful and inappropriate content in paragraphs 4 and 17.