



Singapore Online Safety Report 2024

Introduction

This Online Safety Report (**Broadcasting Act 1994, Code Of Practice For Online Safety**), outlines our overall approach to online safety on X as well as the measures and activities to combat harmful and inappropriate content for the period of 1 July 2023 to 30 June 2024 for Singapore.

Our Approach to Online Safety

X's purpose is to serve the public conversation. In line with our mission to promote open conversation, we encourage a variety of perspectives on the platform. This is central to our [Freedom of Speech. Not Reach](#) philosophy, that moves us away from a binary take down/leave up approach to a more reasonable, proportionate and effective moderation process. Violence, harassment, and other similar types of behaviour discourage people from expressing themselves, and ultimately diminish the value of global public conversation. We thus have clear Rules and Policies in place that are designed to ensure all people can participate in the public conversation freely and safely. They apply globally, including to Singapore, and are easily accessible on our Help Center.

Our Rules and Policies are dynamic, and we continually review them to ensure that they are up-to-date, necessary and proportional. Creating a new policy or making a policy change requires in-depth research around trends in online behaviour, developing clear external language that sets expectations around what's allowed, and creating enforcement guidance for reviewers that can be scaled across millions of pieces of content and accounts. We are also committed to providing transparency on our policy development process and rules enforcement philosophy, and provide explanations of these on our Help Center.

While we aim to enable open discussion of differing opinions and viewpoints, we are committed to the objective, timely, and consistent enforcement of our Rules. To enforce our Rules, we use a combination of machine learning and human review.

Our content moderation systems are designed and tailored to mitigate potential harms without unnecessarily restricting the use of our platform and fundamental rights, especially freedom of expression. Content moderation activities are implemented and anchored on principled policies and leverage a diverse set of interventions to ensure that our actions are reasonable, proportionate and effective. Our content moderation systems blend automated and human review paired with a robust appeals system that enables our users to quickly raise potential moderation anomalies or mistakes. This work is led by an international, cross-functional team with 24-hour coverage and the ability to cover multiple languages. These moderation activities are supplemented by scaled human investigations into the tactics, techniques and procedures that bad actors use to circumvent our rules and policies.

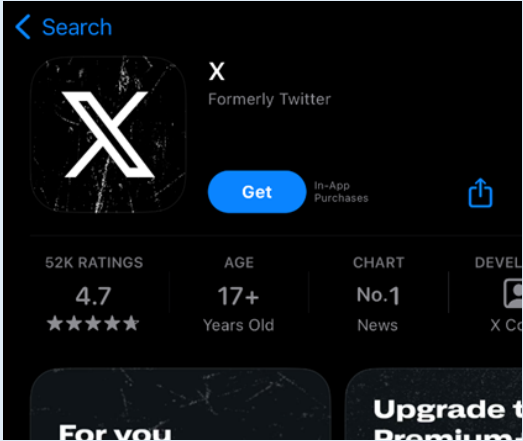
X strives to provide an environment where people can feel free to express themselves. If abusive behaviour happens, we want to make it easy for people to report it to us. When we take [enforcement actions](#), we may do so either on a specific piece of content (e.g., an individual post or Direct Message) or on an account. We may employ a combination of these options. In most cases, this is because the behaviour violates the X Rules.

We always aim to exercise moderation with transparency. Where our systems or teams take action against content or an account as a result of violating our Rules or in response to a valid and properly scoped request from an authorised entity in a given country, we strive to provide context to users.

X is a place for users to share ideas and information, connect with communities, and see the world around them. In order to protect the very best parts of that experience, we provide tools designed to help users control what they see and what others can see about them, so that they can express themselves on X with confidence. Our diverse product-level safety features allow users to modify their experience and engagement on X to ensure each user is able to participate on the platform in a safe and meaningful way.

We recognise that minors are a more vulnerable group by virtue of their age. X, as a service, is not primarily for children, and those below the age of 13 are not permitted to sign up for the service. Users who do not meet our age requirements have their account locked. Parents and guardians are able to access our [Rules and Policies](#) to learn more about how to keep their child's account and experience on X safe, secure and welcoming.

Our Online Safety Measures

Code Obligations:	X's input	Additional Comments and Explanatory Notes on the data provided
Section A: User Safety		
<p>Paragraph 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.</p> <p><i>And</i></p> <p>Paragraph 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.</p>	<p><i>Please provide information on the measures in place for all End-users in Singapore:</i></p> <p>Please refer to our responses to paragraphs 11 to 16 and 23 below on the measures put in place for all End-users in Singapore.</p> <hr/> <p><i>Please provide information on any differentiated or additional measures in place for Children:</i></p> <p>X does not have a line of business which is dedicated to children. Users on X must be aged 13 or above, and those below the age of 13 are not permitted to sign up for the service, as required by our Terms of Service ("TOS"). In addition, the X app is not recommended for those below the age of 17 on Apple's App Store:</p> 	

	<p>Users who do not meet our age requirements have their account locked. In addition, parents and guardians are able to access our Rules and Policies to learn more about how to keep their child’s account and experience on X safe, secure and welcoming.</p> <p>X policies and tools apply uniformly across all of our users and operate to minimise end users’ exposure to harmful content.</p> <p>Nevertheless, although minors represent a minimal fraction of X’s user base, we are fully committed to the protection of this group, who are a more vulnerable audience online, and have a number of specific tools and policies to protect younger audiences on our platform. Further details of these measures can be found in our subsequent responses to paragraphs 17 to 22 below.</p>	
<p>Paragraph 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.</p>	<p><i>Please provide information on the community guidelines and standards in place and how they address the categories of harmful content in paragraph 4.</i></p> <p><i>Please provide information on the content moderation measures in place and how they are enforced or effected. Please include screenshots or evidence where possible.</i></p> <p>Community Guidelines and Standards</p> <p>We have clear Rules and Policies in place that are designed to ensure all users on X can participate in the public conversation freely and safely. They apply globally, including to Singapore, and are easily accessible on our Help Center. All users are required by our Terms of Service to use X in compliance with these Rules and Policies.</p> <p>With reference to the specific categories of harmful content set out under paragraph 4 of the Code, the following X policies operate to minimise end users’ exposure to such harmful content:</p> <p>Sexual Content:</p> <ul style="list-style-type: none"> - Child Safety - Non-consensual nudity - Violent Content (which covers content including that which depicts sexual violence) <p>Violent Content:</p> <ul style="list-style-type: none"> - Violent Content - Violent & Hateful Entities - Perpetrators of Violent Attacks 	

Suicide and Self-Harm Content:

- [Suicide and Self-harm](#)

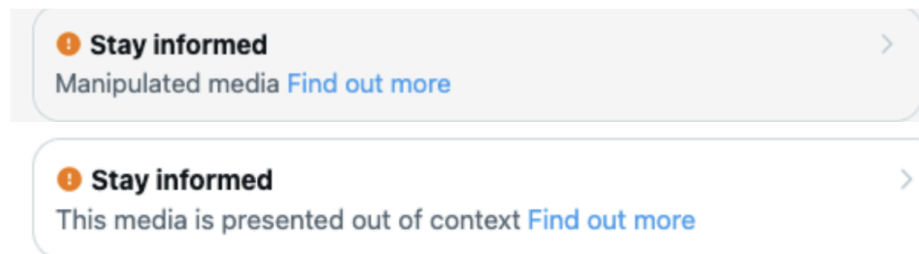
Cyberbullying Content:

- [Abuse and Harassment](#)
- [Hateful conduct](#)
- [Private Content](#)

Content Endangering public health & Content Facilitating Vice and Organised Crime

- [Illegal or certain regulated goods or services](#) (which prohibits the use of X for any unlawful purpose or in furtherance of illegal activities).

In addition, under our [Synthetic and manipulated media policy](#), if we determine a post contains misleading or disputed information that could lead to harm, including widespread confusion on public issues and public safety, we may add a label to the content to provide context. For posts containing media determined to have been significantly and deceptively altered or fabricated, we may add a "Manipulated media" label. For example:

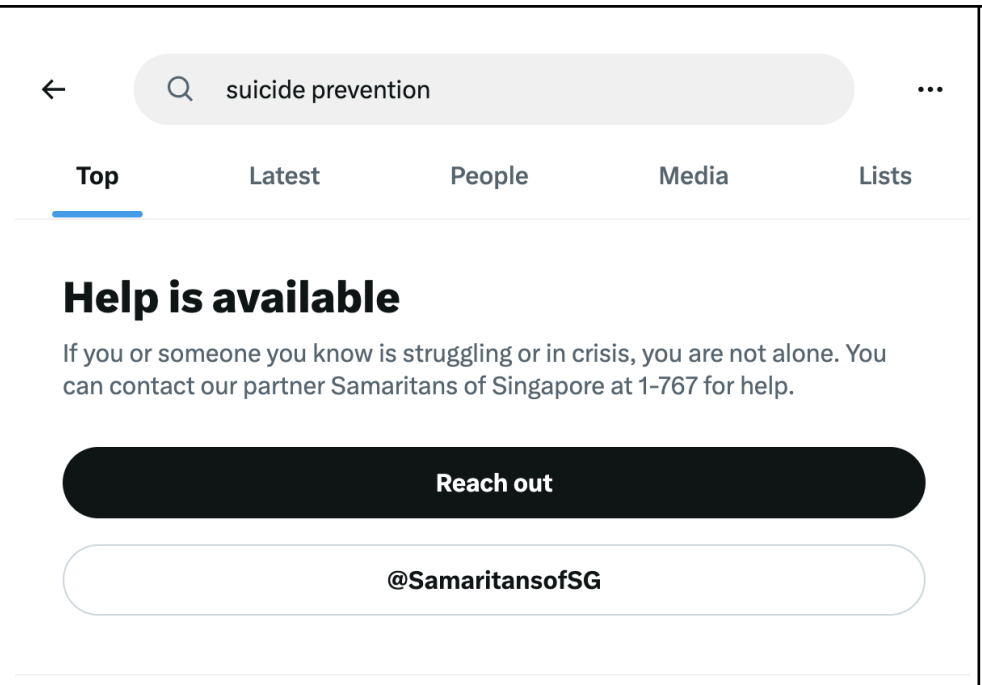


Content Moderation

X employs a combination of heuristics and machine learning algorithms to automatically detect content that violates the X Rules and Policies enforced on our platform. We use combinations of natural language processing models, image processing models and other sophisticated machine learning methods to detect potentially violative content. These models vary in complexity and in the outputs they produce. For example, the model used to detect abuse on the platform is trained on abuse violations detected in the past. Content flagged by these machine learning models are either reviewed by human content reviewers before an action is taken or, in some cases, automatically actioned based on model output. Heuristics are typically utilised to enable X to react quickly to new forms of violations that emerge on the platform. Heuristics are common patterns of text or keywords that may be typical of a certain category of violations. Pieces of content detected by heuristics may also get reviewed by human content

	<p>reviewers before an action is taken on the content. These heuristics are used to flag content for review by human agents and prioritise the order such content is reviewed.</p> <p>Please also see our responses to paragraph 15 below for more details on our other content moderation policies and activities for specific types of policy violations.</p>	
<p>Paragraph 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:</p> <p>(a): Tools to restrict visibility of harmful content and/or unwanted comments</p> <p>(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</p> <p>(c): Tools to limit location sharing</p>	<p>Please provide information on the following, and include screenshots or evidence where possible.</p> <p>All End-users in Singapore: Paragraph 12(a): Tools to restrict visibility of harmful content and/or unwanted comments</p> <p>We provide a suite of tools (here and here) designed to help our users control what they see on X and what others can see about them on X, so that they can express themselves on X with confidence.</p> <p>These include:</p> <ul style="list-style-type: none"> • Unfollow • Filter notifications • Show less often • Mute • Block • Report • Control the media you see in posts • Protect your posts • Photo tagging settings • Discoverability settings • Sharing your location in posts • Sharing precise location through the X app • Media settings (to mark media in users own posts as possibly containing sensitive media) • Tools to authorise and connect to third party application(s) • Home and latest • Autoplay video settings. • Notifications settings • Not interested in Topic suggestions • Safe Search • Controlling how your X information appears in Google search 	

	<p>All End-users in Singapore: Paragraph 12(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</p> <p>Please refer to the suite of tools set out above in our response to paragraph 12(a), which are designed to help our users control what they see on X and what others can see about them on X.</p>	
<p>Paragraph 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.</p> <p><i>And</i></p> <p>Paragraph 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 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.</p>	<p>Please provide information on the measures, and include screenshots or evidence where possible.</p> <p>All End-users in Singapore:</p> <p>The X Rules and Policies, which protect user safety, as well as platform and account integrity, together with all other information related to online safety and security on X, are publicly accessible on our Help Center. We keep our Help Center regularly updated anytime we modify our Rules.</p> <p>For further details on how we’re making X safer please refer to the following: https://help.x.com/en/resources/a-safer-twitter.</p> <p>In addition, we have a dedicated page on our Help Centre which provides dedicated resources for users if they or someone they know is struggling or in crisis. If users input relevant search terms they will see the following pop-up:</p>	



Please provide information on the measures, and include screenshots or evidence where possible.

Children:

As described in our response to paragraph 9 above, X does not permit users below the age of 13 to sign up. The resources described in our response to paragraph 13 above are available to all users of X - including children - and parents and guardians are able to access our [Rules and Policies](#) to learn more about how to keep their child's account and experience on X safe, secure and welcoming.

Paragraph 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.

And

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

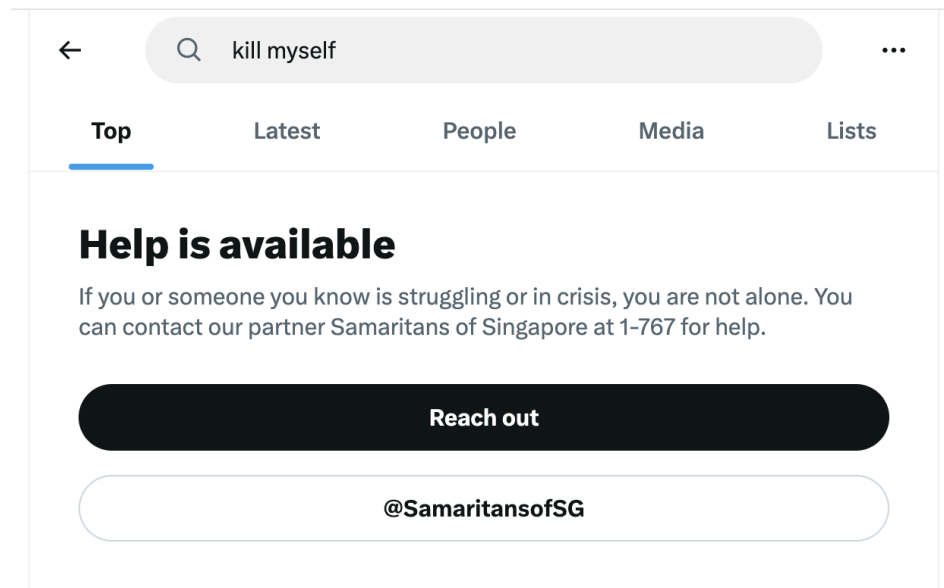
As mentioned in our response to paragraph 13 above, the [X Rules and Policies](#), which protect user safety, as well as platform and account integrity, together with all other information related to online safety and security on X, are publicly

Paragraph 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.

accessible on our [Help Center](#). We keep our Help Center regularly updated anytime we modify our Rules.

For further details on how we're making X safer please refer to the following: <https://help.x.com/en/resources/a-safer-twitter>

If users input search terms for self-harm or suicide they will see the following pop-up:



Please provide information on the measures, and include screenshots or evidence where possible.

Children:

As described in our response to paragraph 9 above, X does not permit users below the age of 13 to sign up.

The measures described in our response to paragraph 14 above apply to all users of X, and parents and guardians are able to access our [Rules and Policies](#) to learn more about how to keep their child's account and experience on X safe, secure and welcoming.

Paragraph 15: End-users' exposure to child sexual exploitation and abuse material and terrorism content

Please provide information on the measures, and include screenshots or evidence where possible.

on the Service must be minimized 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.

All End-users in Singapore:

Please refer to our response to paragraph 11 for a description of our general content moderation measures which detect and enforce against violations of the X Rules and Policies on our platform.

In addition, we also enforce our [Violent and Hateful Entities policy](#) (which restricts the activities and promotion of terrorist organisations, violent extremist groups, perpetrators of violent attacks or individuals who affiliate with and promote their illicit activities, amongst others) and our [Child Safety](#) policy.

Violent and Hateful Entities enforcement:

X uses various detection tactics to surface and enforce against content and accounts that have a high likelihood of being affiliated with a terrorist organisation. For example, X performs network and behavioural analysis, technical signal analysis, keyword normalisation, and text scraping of media. The highest-precision tactics result in automatic enforcement. Lower-precision automation tactics serve as lead generation to surface content for manual review. Ban evasion detection is used, which refers to automated systems that detect attempts to evade suspension, to automatically suspend a user creating a new account. We also use image labelling, where we can apply rules that either surface or automatically apply enforcements to accounts or posts that share similar images. These rules can also include other decisioning criteria such as when the image is posted by a new or potentially compromised account.

X employs a combination of heuristic-based rules and machine learning models, [specifically](#) natural language processing models, to automatically detect content or accounts that violate our [Violent and Hateful Entities policy](#). These models vary in complexity and in the outputs they produce. We also use heuristic-based rules, which are typically used to enable X to react quickly to new forms of violations that emerge on the platform.

Violent Content:

X employs a combination of heuristics and machine learning algorithms to automatically detect content that violates the [X Rules and policies](#) enforced on our platform. We use combinations of natural language processing models, image processing models and other sophisticated machine learning methods to detect potentially violative content. These models vary in complexity and in the outputs they produce. For example, the model used to detect abuse on the platform is trained on abuse violations detected in the past. Content flagged by these machine learning models are either reviewed by human content reviewers before an action is taken or, in some cases, automatically actioned based on

model output. Heuristics are typically utilised to enable X to react quickly to new forms of violations that emerge on the platform. Heuristics are common patterns of text or keywords that may be typical of a certain category of violations. Pieces of content detected by heuristics may also get reviewed by human content reviewers before an action is taken on the content. These heuristics are used to flag content for review by human agents and prioritise the order such content is reviewed.

Child Sexual Exploitation and Abuse Material (“CSE/CSAM”):

Our [Child Safety policy](#) details our zero tolerance approach for any forms of child sexual exploitation on our platform.

For CSE/CSAM, our proactive efforts are primarily driven by hash matching through which we enforce against known cases of CSE/CSAM. Here is the detailed list of the tools we use to enforce against CSE/CSAM, whether human-generated or synthetic:

- **Hash-sharing:** Our current methods of surfacing potentially violating content for human review include leveraging the hashes provided by the National Center for Missing and Exploited Children (“**NCMEC**”) and industry partners - which makes this the most widely used form of CSE/CSAM detection. We scan all media uploaded to X for matches to hashes of known CSE/CSAM sourced from NGOs, law enforcement and other platforms. Users posting known content are immediately permanently suspended and reported to NCMEC. For videos we use a proprietary hashing algorithm produced by Thorn.
- **New hashes:** Our content moderators also have the ability to add new hashes to our internal database when we find new CSE/CSAM. This allows us to take action against new cases swiftly.
- **Automatic text detection:** We have a variety of tools to assess the likelihood that a post is advertising or promoting the sharing of CSE/CSAM. Some of these defences lead to automatic suspensions while other users are flagged for human review.
- **PhotoDNA and internal proprietary tools:** A combination of technology solutions are used to surface accounts violating our Rules on CSE/CSAM. We continue to report to NCMEC when appropriate.
- **Media Risk Scanning:** We receive a media classifier score through *Safer* and it is used to filter false positive hash matches at the moment. We use a novel classifier model to rate media shared through posts’ likelihood of being CSE/CSAM.

When we remove CSE/CSAM content, we immediately report it to the NCMEC. NCMEC makes reports available to the appropriate law enforcement agencies around the world to facilitate investigations and prosecutions.

	<p>In addition, anyone can report potential CSAM content, whether they have an X account or not. Please see our responses to paragraph 23 below for more details on the user reporting tools and processes.</p> <p>In the majority of cases, the consequences for violating our CSE /CSAM policy is immediate and permanent suspension from the platform. In addition, violators will be prohibited from creating any new accounts in the future. Any attempts to circumvent an enforcement action by creating additional accounts or repurposing existing accounts to replace or mimic a suspended account are considered a violation of our ban evasion policy and will result in permanent suspension at first detection.</p>	
<p>Paragraph 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).</p>	<p><i>Please provide information on the measures, and include screenshots or evidence where possible.</i></p> <p><i>All End-users in Singapore:</i></p> <p>Please refer to the above responses to paragraph 15.</p>	
<p>Paragraph 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:</p> <p>(a): Sexual content (b): Violent content (c): Suicide and self-harm content (d): Cyberbullying content</p>	<p><i>Please provide information on the measures, and include screenshots or evidence where possible.</i></p> <p><i>Children:</i> <i>Paragraph 17(a): Sexual content</i></p> <p>We recognise that minors are a more vulnerable group by virtue of their age.</p> <p>As mentioned in our response to paragraph 9, X, as a service, is not targeted at younger users. Under the X Terms of Service, users below the age of 13 are not permitted to sign up for the service.</p> <p>Users who do not meet our age requirements have their account locked.</p> <p>In addition, parents and guardians are able to access our Rules and Policies and other resources in our Help Center to learn more about how to keep their child’s account and experience on X safe, secure and welcoming. This includes a form permitting them to report accounts holders who they suspect as being underage.</p> <p>The X policies detailed in our response to paragraph 11 above apply uniformly across all of our users, and operate to minimise end users exposure to harmful</p>	

content. Please refer to our response to paragraph 11 for further details of these policies.

Our [Child Safety](#) policy is designed to protect minors from sexual and physical abuse, as well as psychological harm that may result from sharing such content.

X strives to keep certain categories of content and accounts out of our recommendations. Neither the Following tab or the For You tab permits sensitive content or inappropriate advertising to be surfaced for accounts under the age of 18.

The following additional measures are also in place:

- **Recommendation Eligibility:** X implements eligibility requirements before it recommends content and accounts (e.g. on the “For you” home timeline). X strives to keep certain categories of content and accounts out of our recommendations. Neither the Following tab or the For You tab permits sensitive content or inappropriate advertising to be surfaced for accounts of known minors.
- **Age Restricted Content - Sensitive Media Settings:** We restrict views of specific forms of sensitive media such as adult content for known minors or viewers who do not include a birth date on their profile with interstitials. For example, instead of seeing the relevant content, a user will be presented with the following:

This Tweet may include sensitive content.

The following media includes potentially sensitive content. [View](#)
[Change settings](#)

- **Age Restricted Content - Content Notices:** X obscures sensitive media behind notices and interstitials. This includes the age restricted content policy which restricts known minors from viewing adult content.

Age-restricted adult content. This content might not be appropriate for people under 18 years old. [Learn more](#)

Age-restricted adult content. This content might not be appropriate for everyone. To view this media, you'll need to [add your birthdate](#) to your profile. Twitter also uses your age to show more relevant content, including ads, as explained in our Privacy Policy. [Learn more](#)

Age-restricted adult content. This content might not be appropriate for people under 18 years old. To view this media, you'll need to [log in](#) to Twitter. [Learn more](#)

- **Age Restricted Content - Safe Search:** X automatically excludes potentially sensitive media content (along with accounts users have muted or blocked) from search results shown to accounts of known minors or without a date of birth.
- **Accounts Defaulted to Protected:** Accounts belonging to known minors will be defaulted to “Protected posts”. This means that known minors will receive a request when new people want to follow them (which they can approve or deny), that their posts will only be visible to their followers, and that their posts will only be searchable by them and their followers (i.e. they will not appear in public searches).
- **DMs Defaulted to Protected:** Accounts belonging to known minors will be restricted to receiving DMs from accounts they follow by default.
- **Age Lock:** Once a new user enters a date of birth that makes them under the age of 18, they will be stopped from re-entering a new date of birth for that account.
- **Suicide and self-harm help:** X has developed a policy prohibiting users from promoting or encouraging suicide or self-harm. When someone searches for terms associated with suicide or self harm, the top search result is a notification encouraging them to reach out for help.

Screenshots of notices and interstitials are available [here](#).

**Children:
Paragraph 17(b): Violent content**

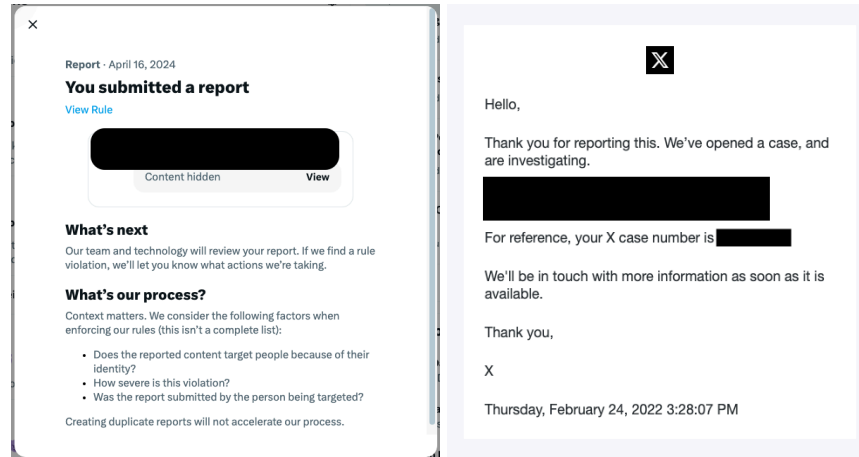
Please refer to our response to paragraph 17(a) above, which details the measure and policies which X has in place, including in respect of violent content. The X policies detailed above additionally operate to minimise children’s exposure to such harmful content.

	<p>Children: Paragraph 17(c): Suicide and self-harm content</p> <p>Please refer to our response to paragraph 17(a) above, which details the measure and policies which X has in place, including in respect of suicide and self-harm content. The X policies detailed above additionally operate to minimise children’s exposure to such harmful content</p>	
<p>Paragraph 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.</p>	<p>Please provide information on the measures, and include screenshots or evidence where possible.</p> <p>Children:</p> <p>X collects each account holder’s date of birth through the neutral presentation of a date of birth prompt. Users must be over the age of 13 to use the site and will be automatically off-boarded if their date of birth indicates that they are under 13.</p> <p>Account holders who enter a date of birth that makes them under the age of 18 will not be permitted to see known adult content on X. Furthermore, once these accounts enter a date of birth that makes them under the age of 18, they will be stopped from re-entering a new date of birth.</p> <p>In addition to the measures described in our response to paragraph 17(a) above, X also prohibits knowingly marketing or advertising certain products and services to minors as detailed in the following policy: Prohibited advertising content for minors. This policy applies to monetization on X and X’s paid advertising products and advertisements containing age-inappropriate content will be tagged as “not family safe” and will be restricted from being shown to users under the age of 21 and signed-out users. If an account does not have a date of birth associated with it, X infers the user’s age based on their interactions with the site.</p> <p>By default, X sets high privacy, safety and security settings for users who access X without logging into an account to help ensure that the experience is appropriate for all users who fulfil X’s minimum age requirement of 13 years old. For example, known sensitive media is not shown and advertising must be tagged as being “family safe” to be shown.</p>	

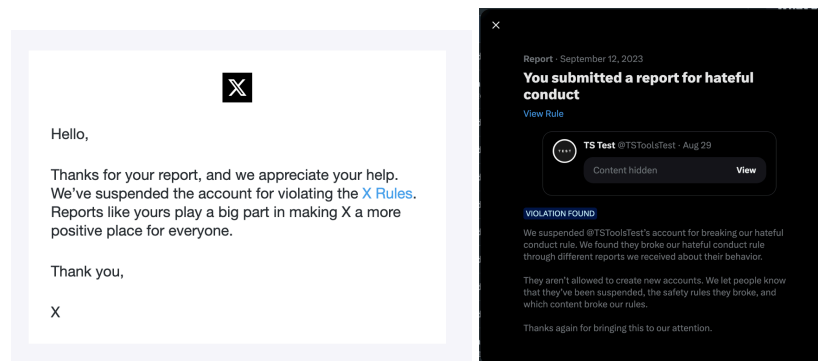
	<p>The X policies detailed in our response to paragraph 11 additionally operate to minimise children’s exposure to such harmful content.</p>	
<p>Paragraph 19: Children or their parents/ guardians must have access to tools that enable them to manage children’s safety, and effectively minimise children’s exposure to, and mitigate the impact of, harmful and/ or inappropriate content and unwanted interactions on the Service. These tools may include the following:</p> <p>(a): Tools to effectively manage the content that children see and/or their experiences.</p> <p>(b): Tools to:</p> <ul style="list-style-type: none"> 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 Limit location sharing. iii. Limit location sharing 	<p>Please provide information on the measures, and include screenshots or evidence where possible.</p> <p>Children: Paragraph 19(a): Tools to effectively manage the content that children see and/or their experiences.</p> <p>X has tools in place which are available to all users to allow them to manage their experience on X. We recognise that minors are a more vulnerable group by virtue of their age. X, as a service, is not targeted at younger users. Under the X Terms of Service, users below the age of 13 are not permitted to sign up for the service.</p> <p>Please refer to our responses to paragraph 12 and paragraph 17 above for details of the tools which X has in place to help our users control what they see on X and what others can see about them on X, and to protect children and to keep children’s accounts and experience on X safe, secure and welcoming. In addition, parents and guardians are able to access our Rules and Policies and other resources in our Help Center to learn more about how to keep their child’s account and experience on X safe, secure and welcoming.</p> <p>Children: Paragraph 19(b): Tools to:</p> <ul style="list-style-type: none"> i. Limit the public visibility of children’s accounts, including their profile and content; <p>Accounts Defaulted to Protected: Accounts belonging to known minors will be defaulted to “Protected posts”. This means that known minors will receive a request when new people want to follow them (which they can approve or deny), that their posts will only be visible to their followers, and that their posts will only be searchable by them and their followers (i.e. they will not appear in public searches). This is in addition to the tools available to all users described in our response to paragraph 12 above.</p> <ul style="list-style-type: none"> ii. Limit who can contact and/or interact with children’s accounts; and <p>DMs Defaulted to Protected: Accounts belonging to known minors will be restricted to receiving DMs from accounts they follow by default. This is in</p>	

	<p>addition to the tools available to all users described in our response to paragraph 12 above</p> <p>iii. Limit location sharing</p> <p>Post location is off by default, and users need to opt in to the service. More details provided here.</p>	
<p>Paragraph 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 parent/ guardians must be provided clear warnings of implications if they opt out of the default settings.</p>	<p>Please provide information on the measures, and include screenshots or evidence where possible.</p> <p>Children:</p> <p>As mentioned in our responses to paragraph 9, X does not have a line of business which is dedicated to children. Users on X must be aged 13 or above, and those below the age of 13 are not permitted to sign up for the service, as required by our Terms of Service. Users who do not meet our age requirements have their account locked. Notwithstanding this, there are certain differentiated measures that apply if we detect that an account belongs to a minor, as described in our response to paragraphs 17, 18 and 19 above.</p> <p>Parents and guardians are able to access the resources in our Help Center to learn more about how to keep their child's account and experience on X safe, secure and welcoming. This includes a form permitting them to report accounts holders who they suspect as being underage.</p>	
Section B: User Reporting and Resolution		
<p>Paragraph 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.</p>	<p>Please provide information on the measures, and include screenshots or evidence where possible.</p> <p>All End-users in Singapore: Paragraph 23(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:</p> <ul style="list-style-type: none"> 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. <p>Users can report content and accounts using our reporting tools. More information available here.</p>	

Once the report is made, the report(er) is sent a notification via in-app notification or email that their report was received.



If the content or account is found to be in violation, we take appropriate action in relation to the reported content/account and user notices are sent with the outcome of the report once the review process is complete and the appropriate remediation is applied. The report(er) is sent a notice via in-app notification or an email that the content they reported was found 'violative' or 'not violative'. If the content/account is found 'violative', the report(ed) user receives notice of violation via in-app notification or email with an invitation to appeal (if eligible). In cases where the user is eligible to appeal, the reported user can submit an appeal from the in-app appeal form or the Help Center webform.



Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

Paragraph 23(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.**

User notices are sent with the outcome of the report once the review process is complete and the appropriate remediation is applied. Report(er) is sent a notice via in-app notification or an email that the content they reported was found 'violative' or 'not violative'.

- ii. Should the Service decide to take action against the report 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.**

If the content/account is found 'violative', the report(ed) user receives notice of violation via in-app notification or email with an invitation to appeal (if eligible) and the reported user can submit an appeal from the in-app appeal form or the Help Center webform. User notices are sent with the outcome of the appeal once the review process is complete and the appropriate remediation is applied.

Your account is suspended

After careful review, we determined your account broke the [X Rules](#). Your account is permanently in read-only mode, which means you can't post, Repost, or Like content. You won't be able to create new accounts. If you think we got this wrong, you can [submit an appeal](#).



Options ▾

We've temporarily limited some of your account features



TestAccount5
@Account5Te93876

What happened?

We have determined that this account violated the [X Rules](#). Specifically, for:

1. Violating our rules against [violent speech](#).

You may not share abusive content, harass someone, or encourage other people to do so.

As a result, we've temporarily limited some of your account features. While in this state, you can still browse X, but you're limited to only sending Direct Messages to your followers — no posts, reposts, follows, or likes. [Learn more](#). Your account will be restored to full functionality in: 12 hours and 0 minutes.

You can start your countdown and continue to X once you:

- Pass an Arkose challenge
- Delete the content that violates our Rules
 - 1 Post

If you think we've made a mistake, contact our [support team](#). You can learn more about our range of enforcement options [here](#).

[Start](#)

Appeal a locked or suspended account

and

Fill out the form below to appeal your account being locked, suspended, or having certain features limited on X. If you are attempting to submit for other X support options, you'll need to appeal your account restriction first.

Your X username (required) ⓘ

@Account5Te93876

Your email address (required)

This is where we'll contact you.

*****@x****.com

Description of the problem

Tell us if you're having a problem accessing your account, or why you don't believe you violated the X Rules.

testing appeals form

Submit

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

Paragraph 23(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.

Please refer to our responses to paragraph 23(b) above.

Section C: Accountability - Mandatory Information and Metrics

Paragraph 24: End-users must have access to clear and easily comprehensible information that enable them to assess the level of safety and related safety

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

measures afforded by the Service and make informed choices.	Please refer to our responses to paragraphs 11 to 16 and 23 below on the measures put in place for all End-users in Singapore.	
Paragraph 25(b): How much and what types of harmful or inappropriate content end-users in Singapore encounter on the Service	Please provide information and data metrics for: All End-users in Singapore: Please refer to the data supplied in response to paragraph 26 below.	
	Children: N/A	
Paragraph 26	Please see the table below for the metrics which X has agreed to provide to IMDA.	

Mapping of X Corp.'s content categories for types of harmful and inappropriate content, against those in paragraphs 4 and 17 of the Code

<i>Main Policy</i>	Suspended User Global	Suspended User SG
Sexual Content		
Non-Consensual Nudity	106224	849
Child Sexual Exploitation	7,628,750	75,673
Sensitive Media (includes but is not limited to Violent Sexual Conduct, Bestiality & Necrophilia in addition to other content that may fall outside the Sexual Content category, such as Gratuitous Gore)	4,272	12
Violent Content		
Violent & Hateful Entities	190068	314
Violent Speech	308604	816
Perpetrators of Violent Attacks	11,044	26
Sensitive Media (includes but is not limited to Gratuitous Gore. Please see above)	Please see above	Please see above
Suicide and Self-harm Content		
Suicide & Self Harm	4665	8
Cyberbullying Content		
Abuse & Harassment	1656042	1485
Hateful Conduct	7125	23

Private Information & media	2834	16
Content endangering public health & Content Facilitating Vice and Organised Crime		
Illegal or certain regulated goods and services	936384	1885
Financial Scam	66,538	110
Synthetic & Manipulated Media	4	0
Misc.		
Ban Evasion	8,409	16

Report Main Policy	Report Count
Sexual Content	
Private Information & media (includes non-consensual nudity)	84260
Sensitive Media (includes but is not limited to Violent Sexual Conduct, Bestiality & Necrophilia in addition to other content that may fall outside the Sexual Content category, such as Gratuitous Gore)	78908
Child Sexual Exploitation	25108
Violent Content	
Violent & Hateful Entities	25994
Violent Speech	65866
Suicide and Self-Harm Content	
Suicide & Self Harm	8098
Cyberbullying Content	
Abuse & harassment	228148
Hateful Conduct	246721

Private Information & media	84260
Content Endangering Public Health & Content Facilitating Vice and Organised Crime	
N/A	

User Reporting: Turnaround Time on Singapore End-User Reports	
Median time needed for taking action on end-user reports in Singapore for all Terms of Service violations.	15.06 hrs

	Auto Enforced		Manual Enforced		Manual Enforced - Proactive Detection	
<i>Main Policy</i>	Posts + Profile Elements Deletions	Posts + Profile Elements Deletions (SG)	Posts + Profile Elements Deletions	Posts+ Profile Elements Deletions (SG)	Posts + Profile Elements Deletions	Posts + Profile Elements Deletions (SG)
Sexual Content						
Non-Consensual Nudity	51,632	109	226,439	1,699	89	0
Child Sexual Exploitation	2,083	6	16,901	29	1	0
Sensitive Media (includes Violent Sexual Conduct, Bestiality & Necrophilia in addition to other content that may fall outside the Sexual Content category)	1,199,517	9940	698,488	3723	330,302	658
Violent Content						
Violent & Hateful Entities	62	0	171	0	1	0
Violent Speech	1,875,060	3,810	1,008,441	2749	10,277	34
Perpetrators of Violent Attacks	158	0	298	0		
Suicide and Self-Harm Content						

Suicide & Self Harm	92	15	128,891	276	5,159	5
Cyberbullying Content						
Abuse & Harassment	9,166	21	3,332,775	4409	16,688	38
Hateful Conduct	3,128	9	26,947	70	344	0
Private Information & media	5,036	11	39,244	166	8,435	7
Content Endangering Public Health & Content Facilitating Vice and Organised Crime						
Illegal or certain regulated goods and services	334	2	896,168	1266	3,572	2
Synthetic & Manipulated Media			9	0		