

October 2023

# Basic Online Safety Expectations

Non-periodic notices issued February 2023  
Key findings

Focus: Child sexual exploitation and abuse

---

Child sexual exploitation and abuse (CSEA) material is class 1 material, as defined under the Online Safety Act by reference to the National Classification Scheme. It includes:

- child sexual exploitation material (CSEM), a broad category of content that encompasses material that sexualises and is exploitative to the child, but that does not necessarily show the child's sexual abuse
- child sexual abuse material (CSAM), a sub-set of child sexual exploitation material that shows a sexual assault against the child.



# Key findings

## Focus: Child sexual exploitation and abuse

On 22 February 2023 eSafety issued legal notices under the Online Safety Act 2021 to **Discord, Google, TikTok, Twitch and Twitter** requiring them to **detail steps they are taking to detect and address online child sexual exploitation and abuse**. The responses to these notices build on the findings from eSafety's first transparency report published in [December 2022](#), which provided transparency and insights on the steps Apple, Microsoft, Skype, Meta, WhatsApp, Snap and Omegle are taking to deal with this material.

In a continuing theme to the first round of notices, the latest responses again showed **significant variation in the steps being taken** by service providers to protect Australian users and the wider public. eSafety closely considered the responses provided by the five recipients of the reporting notices and found that Google and Twitter did not comply with the notices given to them.

This document highlights some of the key findings. The full report contains additional information and context for these key findings, and details the enforcement action taken against Google and Twitter. It is available at [eSafety.gov.au](https://www.esafety.gov.au).



We need the companies to start turning the lights on, so we can get a true sense of the size and scope of this problem.

It's time for all members of the online industry to step up and use their financial, intellectual and technical resources to identify and remove this material from their platforms because even one child sexual exploitation image is one too many?

– eSafety Commissioner



# Detecting livestreamed child sexual exploitation and abuse

Livestreamed CSEA involves the broadcasting of acts of sexual exploitation or abuse of children via webcam or video to people anywhere in the world. The sexual predator controlling the child sometimes charges money for providing access to the livestream.

Despite the availability of technology to help detect child sexual exploitation and abuse in livestreams or video calls, not all companies are using it.

## Use of technology to detect child sexual abuse livestreams\*

### Not using:



**Discord** (public and private servers, direct messages)

### Using:



**Google (YouTube)**



**TikTok**



**Twitch**



**Discord does not monitor or record livestream content or voice chats. Discord has prioritized resources into other forms of CSEA detection, as running models across this type of content at the scale Discord operates would be prohibitively expensive and would operate at the detriment of other Discord safety programs.'**

Discord response to the notice question asking about measures in place to prevent the livestreaming of child sexual exploitation and abuse on Discord.

\*Following publication of this Key findings report X Corp. advised eSafety that its response to questions regarding the detection of livestreaming of CSEA on the Twitter service 'was inaccurate due to an inadvertent error' and provided a revised response to the question. X Corp.'s revised response is [available](#)

## eSafety and third party insights into CSEA harms



This form of child sexual abuse online, along with "self-generated" abuse in livestreams, are all live crime scenes... committed daily on online platforms.'

International Justice Mission (IJM) 2022 [IJM Submission to Public Consultation on the Draft Consolidated Industry Codes of Practice for the Online Industry \(Class 1A and 1B Material\)](#)



Australian children as young as eight are being coerced into performing live-streamed sexual acts by online predators, who often record and share the videos on the dark net and sexually extort victims into producing even more graphic content.'




Australian Federal Police (AFP) 2021 [AFP warn about fast growing online child abuse trend | Australian Federal Police](#)

# Detecting grooming




eSafety is aware that sexual predators use online services to ‘groom’ children. Grooming is predatory conduct to prepare a child or young person for sexual activity, so they can be tricked into sending images or videos or behaving sexually in video livestreams. In addition, the child may be sexually extorted into sending more material or money if the predator threatens to share the image or video.

## Use of language analysis technology to detect likely online grooming

### Not using:

-  **Discord** (public and private servers, direct messages)
-  **Google** (Meet, Chat, Messages, Gmail)
-  **Twitter** (Tweets, direct messages)

### Using:

-  **Google (YouTube)**
-  **TikTok**  
(direct messages, comments on videos/photos/livestreams)
-  **Twitch** (public chat, private messages)



**Twitter is not a service used by large numbers of young people, however we recognise that we need policies to protect against this... We continue to monitor the development of technology... but currently it is not of sufficient capability or accuracy to be deployed on Twitter.’**

Twitter response to the notice question about measures in place to prevent grooming on Twitter.



**11% of young people aged 14-17 report they had been asked on the internet for a photo or video showing their private parts when they didn’t want to.’**

eSafety Commissioner 2022

[Mind the Gap | eSafety Commissioner](#)



**In 2022, we saw a 60% increase in the number of images including children aged 7-10 years old. As ever-younger children become more tech-aware and active online, they become more vulnerable to grooming and abuse by strangers – even in their own bedrooms.’**

Internet Watch Foundation (IWF)

[IWF Annual Report 2022](#)

# Blocking URLs linking to known child sexual exploitation and abuse material

eSafety investigators are aware of platforms being used to distribute thousands of links to child sexual exploitation and abuse sites.

Despite the availability of databases that identify URLs which link to known child sexual exploitation and abuse material and websites that are dedicated to it, some companies are not using them.

## Use of databases to identify URLs linking to known child sexual exploitation and abuse

### Not using:



**Discord**  
(public and private servers, direct messages)



**Google\***  
(YouTube, Drive, Meet, Chat, Photos, Messages, Gmail, Blogger)



**TikTok**  
(direct messages\*\*)

### Using:



**TikTok**  
(on photos/videos, in photo/video descriptions)



**Twitch**



**Twitter**  
(Tweets, direct messages)



**Every URL on the list depicts indecent images of children, advertisements for or links to such content. The list typically contains 5,000 URLs but is subject to fluctuation. The list is updated twice a day to ensure all entries are live. As well as making the internet a safer place for everyone, this initiative can help to diminish the re-victimisation of children by restricting opportunities to view their sexual abuse and may disrupt the accessibility and supply of images to those who seek them out.?**

Internet Watch Foundation (IWF)  
The IWF provides a list of website URLs which they have confirmed contain images and videos of child sexual abuse. [URL list policy](#)

\* If Google detects links to known CSEA it de-indexes them from surfacing on Google search. Google also reports them to NCMEC and hash-matches any CSEA content.

\*\* TikTok stated that it is planning to roll out on direct messages in 2023.

# Community moderation

In user-governed online communities, some service providers use appointed volunteers to actively support content moderation and enforcement of community rules, alongside the tools and resources deployed by the service itself.

These volunteer roles, such as creators, streamers, administrators and moderators, are given administrative power to remove unacceptable material and ban violators. Where there is no standards enforcement policy that outlines the responsibilities and expectations of these volunteers, enforcement of rules can be inconsistent, including with regard to child sexual exploitation and abuse. Some self-appointed creators, streamers, and administrators or moderators can also set up dedicated channels for the exploitation and abuse of children.

In addition, a lack of engagement between volunteer moderators and the Trust and Safety staff of a service increases the risk of sexual predators continuing to abuse and re-victimise children, because they may only be banned from a specific channel or group, rather than the whole service.

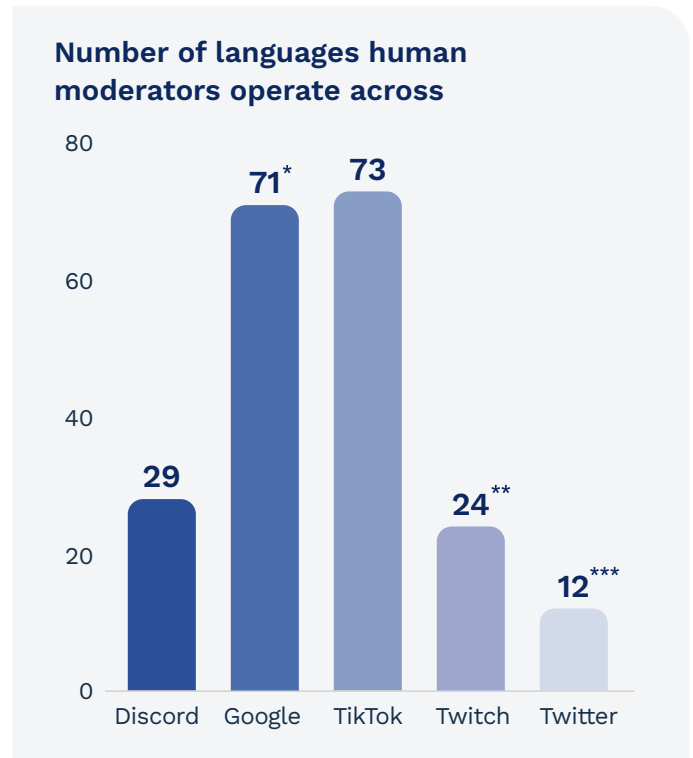


- Discord does not have a standards enforcement policy outlining the responsibilities and expectations of volunteer moderators or administrators
- Twitch does have a standards enforcement policy.
- Professional Trust and Safety staff at Discord and Twitch are not automatically notified when volunteer moderators, administrators or creators take action against child sexual exploitation and abuse material.
- Users on Discord and Twitch are not able to directly report volunteer moderators, administrators or creators for failing to enforce rules.

# Language coverage

Review by human moderators may be required to verify abusive content flagged by technology or reported by users particularly where the material has not been identified previously. This is true for grooming, as well as other harmful content such as hate speech, where effective moderation depends on understanding the language and culture for context.

It's important that service providers have human moderators operating in the languages of the communities that use their services.



\* Google stated that it operates across at least 70 languages and its 'language capability may vary at any time.'

\*\* Twitch noted that additional languages can and will be implemented when there is a need through additional third-party vendors.

\*\*\* Twitter stated that 'The company has the ability to seek and conduct vendor services in a range of additional languages, which include but are not limited to those needed in the event of additional reviews, or for emergencies.'



# User reporting

When illegal content such as child sexual exploitation and abuse material is reported by a user, verifying it and taking action should be done quickly to prevent ongoing or new harm.

The responses to the notices highlighted significant differences in the time service providers take to consider and respond to user reports about child sexual exploitation and abuse material.

Provider	Service or parts of service	Median time for user reported CSEA material to be actioned by the provider
Google	Drive, Meet, Chat, Google Photos, Gmail, Blogger, YouTube, Google Messages	Google did not provide the median time from the point when a user makes a report to the report being actioned.
Twitter	Public posts ('tweets') Direct messages	Twitter did not respond.
TikTok	Photos/videos shared publicly (to everyone or followers only)	5.2 minutes
	TikTok Live	7.7 minutes
	Direct Messages	7.4 hours
Twitch	Twitch	8.22 minutes
Discord	Direct messages	13 hours
	Servers (public)	8 hours
	Servers (private)	6 hours
	Server Livestreams	Discord stated it was unable to calculate the response time as there is no in-service reporting option.



**Survivors surveyed by C3P have generally characterized their own experience reporting CSAM [child sexual abuse material] online as disheartening; exceedingly long delays in responding to their complaints, moderators challenging victims on the veracity of their report or, as is often the case, no response at all.'**

The Canadian Centre for Child Protection (C3P) 2020  
[protectchildren.ca](https://protectchildren.ca)

# Scale of abuse

Although the full size and scope of online child sexual exploitation and abuse is difficult to measure, eSafety knows from the experience of our own investigators and other expert organisations that the scale is significant.



**In 2022, ESPs submitted 49.4 million images to the CyberTipline of which 18.8 million (38%) were unique. Of the 37.7 million videos reported by ESPs, 8.3 million (22%) were unique.**

National Centre for Missing and Exploited Children (NCMEC) CyberTipline Data 2023  
[missingkids.org](https://missingkids.org)

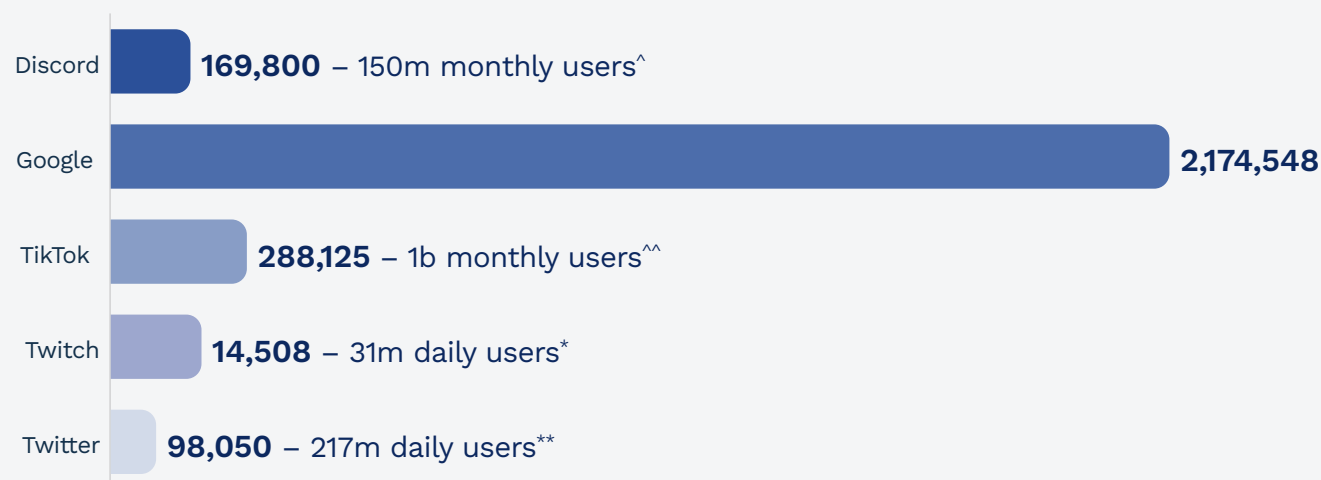


**In “2022...the IWF and partners blocked 8.8 million attempts in one month to access known child sexual abuse material”**

Internet Watch Foundation (IWF)  
[IWF Annual Report 2022](#)

Note: NCMEC’s CyberTipline is the USA’s centralised reporting system for the online exploitation of children. In 2022, the CyberTipline received 31.8 million reports from Electronic Service Providers (ESPs) about apparent child sexual abuse material on their services.

### Child sexual exploitation and abuse reports sent to NCMEC's Cybertipline 2022 (for providers that received notices)



It is important to note that a high reporting figure may indicate the provider is taking the issue seriously and has technology and processes in place to detect and report online child sexual exploitation and abuse material and video livestreaming of abuse on its platforms and services. eSafety is therefore concerned by providers reporting low numbers despite their platforms and services being widely and commonly used.

We can only know the true scale of the global problem if all online services use readily available technologies and human moderation to detect child sexual exploitation and abuse material, video livestreaming of abuse, grooming of children and sexual extortion.

As well as their regulatory requirements in Australia - Twitter, Google, TikTok and Amazon (parent company of Twitch) have endorsed the [Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse](#) and in doing so have committed to taking steps to detect known and new CSEA, including grooming and livestreaming as part of the implementation of those principles. The principles also include steps to share 'meaningful data and insights' on their implementation.

Information obtained in response to eSafety's non-periodic notices can therefore help eSafety understand how providers are implementing the Voluntary Principles.

<sup>^</sup>Discord, 'An update on our business', 2021, accessed 19 January 2023, URL: <https://discord.com/blog/an-update-on-our-business>

<sup>^^</sup>TikTok, 'Thanks a billion!', September 2021, accessed 17 February 2023, URL: <https://newsroom.tiktok.com/en-us/1-billion-people-on-tiktok>

<sup>\*</sup>Twitch, 'Press centre', 2023, accessed 10 February 2023, URL: <https://www.twitch.tv/p/press-center/>

<sup>\*\*</sup> [Twitter Inc Annual Report 2021](#) (annualreports.com)

No average user volumes available for all Google services



[eSafety.gov.au](https://www.esafety.gov.au)