



# California Assembly Bill 587 Transparency Report

Although YouTube does not meet the definition of a "social media platform" as defined per California Assembly Bill 587, YouTube nonetheless is voluntarily producing transparency reports with details on its Terms of Service, specified content categories and corresponding content enforcement as part of YouTube's efforts to provide greater transparency to Californian users. This report is in addition to the various reports we already provide that provide data and shed light on how and why users raise complaints, how we respond to those complaints, how we enforce our Community Guidelines and content policies, and how we comply with local laws to keep users safe. Learn more at [transparencyreport.google.com/youtube-policy](https://transparencyreport.google.com/youtube-policy).

## Information on YouTube's Transparency Report under US-CA 587

Although YouTube does not meet the definition of a "social media platform" as defined per California Assembly Bill 587, YouTube nonetheless is voluntarily producing transparency reports with details on its Terms of Service, specified content categories and corresponding content enforcement as part of YouTube's efforts to provide greater transparency to Californian users. This report is in addition to the various reports we already provide that provide data and shed light on how and why users raise complaints, how we respond to those complaints, how we enforce our Community Guidelines and content policies, and how we comply with local laws to keep users safe.

### Overview

This report outlines information on content actioned by YouTube during the specified reporting period. This report captures information for the period from **July 1, 2023 to September 30, 2023**. We expect to submit reports covering subsequent data periods as specified in California Assembly Bill 587.

### Policies that address harmful content

YouTube's [Terms of Service](#) highlights various policies that outline what is not allowed on YouTube. These policies include YouTube's [Community Guidelines](#), [Advertiser-Friendly Content Guidelines](#), and [related policies](#), which are made accessible in YouTube's [Help Center](#) and apply to all types of content on YouTube.

Our policies cover areas such as [hate speech](#), [harassment](#), and [violent extremism](#) among others. These policies aim to make YouTube a safer community while still giving creators and users the freedom to share a broad range of experiences and perspectives.

### Responding to user reports

YouTube offers YouTube community members an opportunity to report or flag content they believe violates our Terms of Service, which includes our Community Guidelines. Users can report content using YouTube's [flagging feature](#), which is available to all users, globally via computer (desktop or laptop), mobile devices, and other surfaces.

Details on how to report different types of content using YouTube's flagging feature is outlined in YouTube's [Help Center](#). Once we receive a user flag, we review the reported content carefully and take action in accordance with our Terms of Service. Enforcement actions may include removing or restricting content (e.g. age-restrict content not suitable for all audiences), limiting content's monetization eligibility, or applying a [strike](#) to a channel.

---

#### Please note:

\* There are no Medi-Cal languages in which YouTube does offer product features but does not make the Terms of Service available

## Moderating content on YouTube

YouTube relies on machine learning and teams around the world to review flagged videos and remove content that violates our Terms of Service; restrict videos (e.g., [age-restrict](#) content that may not be appropriate for all audiences); or leave the content live when it doesn't violate our guidelines. For example, when content is removed from a user's channel for violating YouTube's Community Guidelines, the user may be issued a [strike](#). If the user receives three strikes in the same 90-day period, their channel may be permanently removed from YouTube. In some cases, YouTube may terminate a channel for a single case of severe abuse, as [explained](#) in the Help Center.

YouTube may also remove content for reasons other than Community Guidelines violations, such as a first-party privacy complaint or a court order. In these cases, users will not be issued with a strike. YouTube also reserves the right to restrict a user's ability to create content at its discretion. A channel may be turned off or restricted from using any YouTube features. If this happens, users are prohibited from using, creating, or acquiring another channel to get around these restrictions. This prohibition applies as long as the restriction remains active on the YouTube channel. A violation of this restriction is considered a circumvention under YouTube's Terms of Service, and may result in termination of all existing YouTube channels of the user, any new channels created or acquired, and channels in which the user is repeatedly or prominently featured.

## Use of automated tools to moderate content

YouTube continues to invest in automated detection systems, and relies on both human evaluators and machine learning to detect and take action on problematic content at scale while simultaneously training their systems on new data. The vast majority of content reviewed and actioned on YouTube is first detected by automated systems. However, after potentially violative content has been detected by automated systems, content moderators may review the content to confirm the decision. This collaborative approach helps improve the accuracy of YouTube's automated systems overtime, as models continuously learn and adapt to content moderator feedback. It also means YouTube's enforcement systems can manage the sheer scale of content that is uploaded to YouTube (over 500 hours of content every minute), while still digging into the nuances that determine whether a piece of content is violative.

We allow users the opportunity to appeal certain enforcement decisions. Users are notified about enforcement decisions and are provided instructions on how to appeal a decision. If a user chooses to submit an appeal, it goes to human review, and the decision is either upheld or overturned.

## Information on actioned content

The following is information on content actioned globally under policies related to YouTube's Terms of Service, between July 1 and September 30, 2023.

### Videos removed

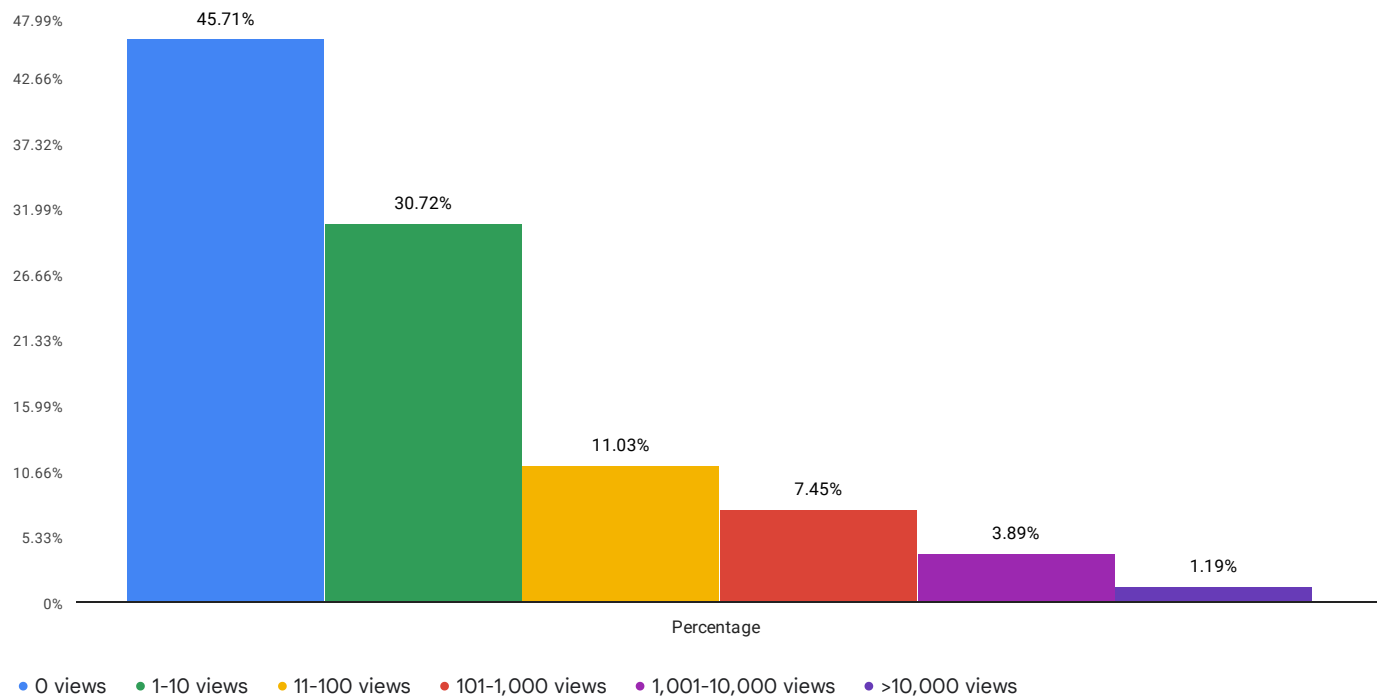
The table at the right shows the number of videos removed by YouTube for violating Community Guidelines policies, for policies relevant to topics specified under California Assembly Bill 587.

The total number of videos removed for violating Community Guidelines policies during the same period was **8,115,659**.

Harassment & Cyberbullying	513,854
Hateful or Abusive	186,999
Misinformation	51,920
Promotion of Violence or Violent Extremism	163,417

### Videos removed, by views

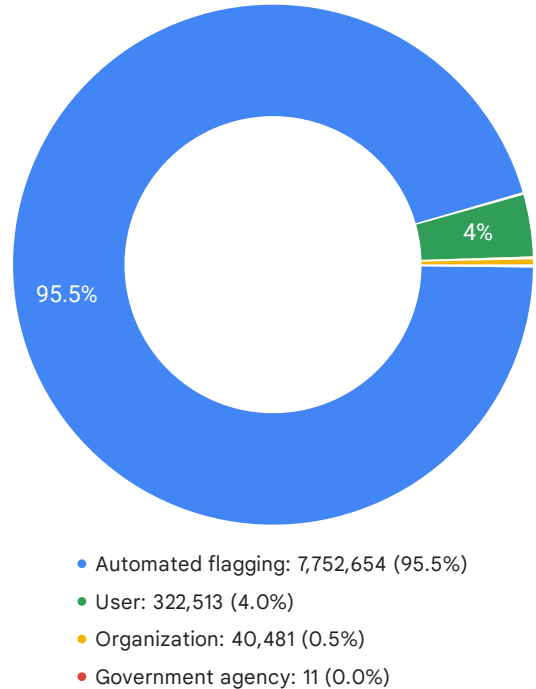
This chart shows the percentage of video removals that occurred before they received any views versus those that occurred after receiving some views.



## Information on actioned content

### Videos removed, by source of first detection

This chart shows the volume of all videos removed by YouTube for violating Community Guidelines policies, by source of first detection (automated flagging or human detection). Flags from human detection can come from a user or a member of YouTube's Priority Flagger program. Priority Flagger program members include NGOs and government agencies that are particularly effective at notifying YouTube of content that violates our Community Guidelines.



### Human flags, by flagging reason

When flagging a video, human flaggers can select a reason they are reporting the video and leave comments or video timestamps for YouTube's reviewers. This chart shows the flagging reasons that people selected when reporting YouTube content, for categories relevant to topics specified under California Assembly Bill 587.

Note that flags related to misinformation are nested under the "Spam or Misleading" category.

Hateful or Abusive	18,419,655
Promotes Terrorism	6,297,704
Spam or Misleading	28,877,696

## Information on actioned content

### Removed videos appealed

The table at right shows the number of appeals YouTube received for videos removed due to a Community Guidelines policy violation, for policies relevant to topics specified under California Assembly Bill 587.

The total number of appeals received for removed videos during the same period was **420,749**.

Creators have 30 days to submit an appeal after the video's removal, so these numbers also include appeals for videos removed during one quarter but appealed in the following quarter.

Policy	Number of appeals
Harassment & Cyberbullying	29,743
Hateful or Abusive	25,498
Misinformation	7,243
Promotion of Violence or Violent Extremism	10,176

### Appealed videos reinstated

The table at right shows the number of appeals YouTube received for videos removed due to a Community Guidelines policy violation, for policies relevant to topics specified under California Assembly Bill 587.

The total number of appeals received for removed videos during the same period was **31,183**.

Policy	Number of reinstates
Harassment & Cyberbullying	1,872
Hateful or Abusive	2,533
Misinformation	821
Promotion of Violence or Violent Extremism	456

### Video demonetizations

The total number of YouTube video demonetizations due to violations of monetization policies was **3,120,510**.

Note: A technical issue affected records on video demonetizations in July-September 2023. Due to this issue, some of the data represented above is estimated.