



TCO Annual Report

2023

The EU [Terrorist Content Online Regulation](#) ("TCOR") is an EU regulation that lays down certain rules to address the misuse of hosting services for the dissemination of terrorist content online. It requires companies offering hosting services ("hosting service providers") in the EU to remove terrorist content within one hour after receiving an official removal order from competent authorities and to prevent any re-upload. To qualify for a removal under the TCOR, content needs to fall under the types of material defined as "terrorist content" under Article 2(7) TCOR. Additionally, TCOR requires hosting service providers to create and publish an annual report about actions taken to address the dissemination of terrorist content online according to Article 7(2) TCOR before March 1. We will update this page before March 1 of each year.

In accordance with Article 7(2) and 7(3) TCOR, we are publishing the following applicable information: 1) information about Roblox's measures in relation to the identification and removal of or disabling access to terrorist content and 2) information about Roblox's measures to address the reappearance online of material that has been previously removed or disabled.

Please note that we've provided information under the relevant categories listed in Article 7(3) TCOR as applicable; omitted categories mean that those cases are not applicable or have not occurred in the previous reporting period. During this reporting period (June 7, 2022 - February 2023), Roblox has not received a TCOR removal order from any competent authorities.

At Roblox, we view transparency as an important part of building a safe and civil platform. Our inaugural TCOR Annual Report reflects our deep commitment towards keeping all Roblox users safe and focuses on our efforts to keep Roblox free from terrorist content as defined by the TCOR.

Policies & Standards

Safety on Roblox is at the core of everything we do. We maintain many systems to keep our users safe and secure; we work tirelessly to maintain a safe, civil, and welcoming environment for all. That includes a robust set of [Community Standards](#) to outline how we expect our users to behave and to be clear up front about what is and isn't allowed on Roblox.

Terrorism and Violent Extremism (TVE) is a unique category and requires a specific enforcement approach. As outlined in our Community Standards, Roblox prohibits content or behavior that supports, glorifies, or promotes any terrorist or extremist organization and their actions. Roblox actively collaborates with global expert organizations like [Tech Against Terrorism](#), and researchers dedicated to these issues, to inform our policies and monitoring practices. In addition to our blanket prohibition on TVE content and behavior, Roblox also has Community Standards that prohibit threats of violence, certain illegal and/or regulated behavior, and actions related to specific harmful behavior off Roblox. This

combination of our Community Standards encompasses “terrorist content” as defined by Article 2(7) of the TCOR. The potential consequences of breaching these policies include warnings, removing content, losing one’s Roblox account, and, in circumstances that present an imminent risk of harm, being reported to the relevant authorities.

01

Information about Roblox’s measures in relation to the identification and removal of or disabling of access to terrorist content.

General Processes & Systems

Roblox takes pride in the processes and systems that it runs to prevent or quickly remove objectionable content and behavior. We have a stringent safety and monitoring system that is continuously active and that we rigorously enforce. Using a combination of 1) monitoring; 2) user reporting; and 3) trusted flaggers, we review the content uploaded to Roblox, and act swiftly to block inappropriate content or behavior. Additionally, per TCOR, Roblox has created a direct channel for competent authorities to submit TCOR removal orders through tco-reports@roblox.com. If Roblox becomes aware of exigent situations that may present an imminent threat or terroristic offense, we take steps to promptly inform relevant authorities.

Once terrorist content is identified, Roblox takes additional steps to remove or disable access to the content. Roblox cooperates with law enforcement and will expedite cases where there may be imminent risk of harm.

Monitoring

Roblox employs a combination of human review and technology review to monitor for terrorist content 24/7. We first review content against industry databases of already-known illegal content (such as terrorist content and child sexual abuse material). In addition to these industry databases of known violative content, we review content against our own previously-removed content. In both instances, we use techniques, including hashes – unique digital signatures of images or videos – to identify when items are exact replicas of or similar to previously-identified violative content.

We also have a team of thousands of trained professionals that review content against our Community Standards. In particular, we have staff that solely focuses on TVE content and behavior, and we hire people with deep expertise in this space who have come from recognized agencies and organizations such as the FBI, Army Intel, and so on. This

team leads the enforcement of TVE-specific policies for our Community Standards and is an in-house expert resource on TVE topics, names, memes, iconography, and more. Our partner organizations help us by providing resources, training, and reporting to us. This expert team has also developed fundamental TVE training for every moderation agent on staff. Agents are trained and tested on how to appropriately recognize and handle TVE content and behavior.

User Reporting

We also actively encourage our users to report any activity they feel concerned about directly to Roblox. They can easily mute or block other community members that they come across and report inappropriate content/behavior using our Report Abuse system located prominently throughout the site and within experiences. See [here](#) for more on our user report functionalities.

Trusted Flagger Program

Roblox has a trusted flagger program, where close and trusted partners can report terrorist content to us. We've made a commitment to them that their reported content will be reviewed and, if necessary, actioned within a short timeframe.

Removing or Disabling Access to Terrorist Content

Once terrorist content is detected on Roblox, we have a specialized team dedicated to handling it. If terrorist content is detected through our content review, it does not appear on Roblox. When we verify that a user report, trusted flagger notice, or government request concerns terrorist content, we will remove that content from our systems and take steps to prevent it from being uploaded again in the future.

02

Information about Roblox's measures to address the reappearance online of material which has previously been removed, or to which access has been disabled because it was considered to be terrorist content, in particular where automated tools have been used.

Roblox has systems in place to prevent the re-upload of content we previously removed. This includes technology that reviews for content that is the same or similar to what we've already taken down.