X

X Transparency Report for the Regulation (EU) 2021/784 on addressing the Dissemination of Terrorist Content Online January - December 2023

X was founded on a commitment to transparency. This commitment is part of our effort to foster a healthier, more open, and civil global public conversation.

X is reflective of real conversations happening in the world and that sometimes includes perspectives that may be offensive or controversial to others. While we welcome everyone to express themselves in our service, we have clear rules in place to protect the safety of the service and the people using it. The X Rules¹ are in place to help ensure everyone feels safe expressing their beliefs and we strive to enforce them with uniform consistency. When it comes to enforcing these rules, we are committed to being fair, informative, responsive, and accountable. Read more about our approach to policy development and enforcement philosophy in the X Help Center².

In this second iteration of our report that provides information about the EU Regulation on addressing the dissemination of terrorist content online ("TCO"),³ we provide information from January 1, 2023 through December 31, 2023. We will continue to publish this report yearly on our Transparency Report page under "Reports" and "Other".

Identification and removal of terrorist content

We take our responsibility to combat terrorism and violent extremism online seriously, including by working alongside governments, industry, and civil society. In parallel, we commit to continuing to invest in technology that improves our capability to detect and appropriately enforce illegal content or accounts. Our external work with industry partners, such as Global Internet Forum to Counter Terrorism (GIFCT), Tech Against Terrorism, Christchurch Call to Action, and EU Internet Forum (EUIF), helps to identify emerging trends in how terrorists and violent extremists are using online platforms to promote their content.

For a detailed description of X's content moderation activities, including those applied to terrorist content, please see the Description of our Content Moderation Practices in X's DSA Transparency report.⁴

² https://help.x.com/en/rules-and-policies/enforcement-philosophy

¹ https://help.x.com/en/rules-and-policies

³ REGULATION (EU) 2021/784 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 29 April 2021 on addressing the dissemination of terrorist content online.

⁴ https://transparency.twitter.com/dsa-transparency-report.html

Measures to address the reappearance of previously removed content

The vast majority of accounts that are suspended for affiliation with violent and hateful entities are proactively flagged using technology and internal proprietary tools. We employ automated methods to detect and prevent violative content from being re-shared on the platform. This includes labelling violative hashes, as well as using other signals to find recidivist accounts. In addition, we may also carry out manual checks for such content.

X also employs a combination of heuristics and machine learning algorithms to automatically detect content that violates the X Rules and policies⁵ enforced on our platform. The X Rules⁶ prohibit violent threats⁷ and the promotion of violent extremism⁸. Specifically, you may not threaten terrorism and/or violent extremism, nor promote violent and hateful entities. Additionally, there is no place on X for violent and hateful entities, including (but not limited to) terrorist organisations, violent extremist groups, individual perpetrators of violent attacks⁹, or individuals who affiliate with and promote their illicit activities. You may not glorify, praise, or celebrate acts of violence where harm occurred, which includes (but is not limited to) expressing gratitude that someone experienced physical harm or praising Violent entities¹⁰ and Perpetrators of Violent Attacks.¹¹ For further details on the heuristics and machine learning algorithms that X deploys across the platform, please see the Our Own Initiative Content Moderation Activities section in X's DSA Transparency Report.¹²

Our proactive moderation of content violating X's Rules in the context of the Israel/Gaza conflict demonstrates our diligent efforts. Between 7 October 2023 and 7 January 2024, our escalations team has actioned over **778,000 pieces of content** that violate our Terms of Service, including violent speech and hateful conduct and we have taken action under our <u>Violent and Hateful Entities policy</u>¹³ to remove over **5,000** accounts by violent entities in the region, including Hamas, since the start of the conflict.

Removal orders

X has implemented systems and partially automated tooling to expeditiously process removal orders from authorised entities in compliance with the EU Regulation. Content that violates the TCO will be made locally unavailable, or "withheld", in the European Union within one hour of receipt of a valid and properly scoped legal request.¹⁴

⁵ https://help.twitter.com/en/rules-and-policies/twitter-rules

⁶ https://help.twitter.com/en/rules-and-policies/twitter-rules

⁷ https://help.twitter.com/en/rules-and-policies/violent-speech

⁸ https://help.twitter.com/en/rules-and-policies/violent-entities

⁹ https://help.twitter.com/rules-and-policies/perpetrators-of-violent-attacks

¹⁰ https://help.twitter.com/en/rules-and-policies/violent-entities

¹¹ https://help.twitter.com/en/rules-and-policies/perpetrators-of-violent-attacks

¹² https://transparency.twitter.com/dsa-transparency-report.html

¹³ https://help.twitter.com/en/rules-and-policies/violent-entities

¹⁴ https://help.twitter.com/en/rules-and-policies/post-withheld-by-country

Required Data

The data set out below is required under the TCO for the reporting period January 1, 2023 to December 31, 2023.

	Number of instances	
Removals ¹⁵ following specific measures ¹⁶	0	
Removals following removal orders	19 Accounts and 41 Posts based on a total of 20 removal orders	
Removal orders where the content has not been removed	0	
Complaints for reinstatement of removed content granted ¹⁷	2 954	
Reinstatements as a result of administrative or judicial review proceedings	0	
	Number of proceedings	Outcome
Administrative or judicial review proceedings brought by X	0	N/A

	Number of instances
Proactive referrals to law enforcement under Art. 14(5) TCO	17

 ¹⁵ Removals in this section mean both removal and disabling/withholding of access.
 16 Twitter International Unlimited Company (TIUC) has not been identified as having been exposed to terrorist content pursuant to article 5(4) TCO.

To Complaints in this section means user appeals of X's enforcement of its Violent and hateful entities policy.