

X's India Transparency Report: User Grievances & Proactive Monitoring September 2023

Overview

X was founded on a commitment to transparency. This commitment is part of our effort to serve the public conversation and to increase its collective health, openness, and civility around the world.

X is reflective of real conversations happening in the world and that sometimes includes perspectives that may be offensive, controversial, and/or bigoted to others. While we welcome everyone to express themselves on our platform, we do not tolerate behavior that harasses, threatens, dehumanizes, or uses fear to silence the voices of others. We have the X Rules¹ in place to help ensure everyone feels safe expressing their beliefs and we strive to enforce them with uniform consistency. We are continually working to update, refine, and improve both our enforcement and our policies, informed by in-depth research around trends in online behavior both on and off X, feedback from the people who use X, and input from a number of external entities.

When it comes to enforcing these rules, we have a global team that manages this with 24/7 coverage in every language that X supports and endeavors to take into account regional and cultural context and complexities of the various markets. X is committed to being fair, informative, responsive, and accountable and our goal is to apply the <u>X Rules</u> objectively and consistently. Enforcement actions are taken on content that is determined to violate the <u>X Rules</u>. Read more about our approach to policy development and enforcement philosophy in the <u>X Help Center</u>.²

In compliance with Rule 4(1)(d) of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, X is required to publish a monthly compliance report which includes the details of complaints from users via the India grievance mechanism and action taken on them, as well as information related to X's proactive monitoring efforts under the IT Rules. This report covers the reporting period from July 26, 2023 to August 25, 2023. We will continue to publish this report on a monthly basis, and will make improvements over time, based on feedback received from the government, or in accordance with internal changes that allow for us to provide more granular data.

¹ <u>https://help.twitter.com/en/rules-and-policies; https://help.twitter.com/en/rules-and-policies/x-rules</u>

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

How X Processes Grievances from Users in India

In India, X users can report grievances via the Grievance mechanism by using the contact details available on the <u>Grievance Officer - India</u> page. X does not require such reporters to be registered with X in order to file a grievance. Separately, users can report directly from the post or account in question while logged into X, or reports can be made via <u>X's Help Center</u>.

Each user complaint we receive via the India Grievance channel is assessed under the <u>X Terms of</u> <u>Service</u> and the <u>X Rules</u> and any content that is determined to be in violation is actioned in line with our <u>range of enforcement options</u>. Where appropriate, we also assess whether the grievance meets local legal requirements for X to action the reported content, based on a court order. Users generally have the option to appeal an enforcement action via our <u>appeals process</u>. NOTE: Law enforcement requests should, ideally, be submitted at X's Legal Request Submission site, available at <u>t.co/lr</u> or <u>legalrequests.twitter.com</u>, to facilitate efficient processing.

To process grievances, we have a dedicated team of trained individuals. These individuals take appropriate action after carefully reviewing the grievance, and any available context, in close detail. All of these team members are fluent in one or more of the languages spoken in India and participate in regular training on our tools and policies. Updates about significant current events or internal rules and policy changes are shared with all agents in real time, to give guidance and facilitate balanced and informed decision-making.

Grievance Officer - India Data

Below is the data we received via our Grievance Officer - India channel between **July 26, 2023 through to August 25, 2023** that included content on X. This includes complaints received from individual users with accompanying court orders.³

NOTE: We receive complaints in our Grievance Officer - India channel that relate to account verification, account access, or seeking assistance or information regarding an account or X's enforcement actions that are not included in the data below. The majority of complaints received in this channel during this reporting period fell into these categories.

NOTE: The data in these reports is as accurate and comprehensive as possible at the time of reporting. It includes data on requests received and processed during the reporting period.⁴

| Issue TypeTotal Number of Grievances⁵Total Number of URLs Actioned6 |
|---|
|---|

³ Data on government requests can be found in our existing <u>X Transparency Center</u>.

⁴ We have removed all duplicate requests from the data provided in this report, so that only unique requests are counted.

⁵ Please note that each grievance may identify multiple items. For example, a single grievance may ask us to remove individual posts, an entire account, or both.

⁶ We may not remove content or all aspects of a grievance for a variety of reasons. For example, the grievance may fail to identify content on X, the content may have been removed by the user that posted it, or we find no violation of our Terms of Service (TOS) or the X Rules. Our analysis, and any reasons for not taking action, are specific to each grievance.

| | | - |
|---|------|------|
| Abuse / Harassment | 1267 | 2734 |
| Child Sexual Exploitation | 43 | 11 |
| Defamation | 6 | 0 |
| Hateful Conduct | 62 | 33 |
| Illegal Activities ⁷ | 7 | 4 |
| Impersonation | 24 | 18 |
| IP-related Infringement ⁸ | 3 | 0 |
| Misinformation / Synthetic and Manipulated Media | 5 | 4 |
| Promoting Suicide or Self-Harm | 23 | 0 |
| Terrorism / Violent Extremism | 0 | 0 |
| Sensitive Adult Content | 0 | 0 |
| Privacy Infringement | 27 | 23 |
| | | |

In addition to the above, we processed **78** grievances which were appealing account suspensions. These were all resolved and the appropriate responses were sent. We overturned **10** of these account suspensions after reviewing the specifics of the situation. The remaining reported accounts remain suspended. We received **10** requests related to general questions about accounts during this reporting period.

X's Proactive Monitoring Efforts

There are many measures that have been in place for a long time on our platform that relate to the mitigation of child sexual exploitation and terrorism activity. X does not tolerate any material throughout the service that features or promotes <u>child sexual exploitation</u>. This includes media, text, illustrations, or computer-generated images. When we remove content, we immediately report it to the National Center for Missing and Exploited Children (NCMEC). NCMEC makes reports available to the appropriate law enforcement agencies around the world to facilitate investigations and prosecutions.

X (when it was referred to as "Twitter") was one of the <u>founding members</u> of, and continues to participate in, the <u>Global Internet Forum to Counter Terrorism</u>.

A vast majority of all accounts that are suspended for the <u>promotion of terrorism</u> and <u>child sexual</u> <u>exploitation</u> are proactively flagged by a combination of technology and other purpose-built internal proprietary tools.

⁷ This includes requests related to money laundering, gambling, hacking, doxxing, or similar issues.

⁸ This includes Copyright, Trademark, or other IP-related issues.

Proactive Monitoring Data⁹

| Issue | Total Accounts Suspended | |
|---|--------------------------|--|
| Child Sexual Exploitation and similar content | 1,280,107 | |
| Promotion of Terrorism | 2,307 | |

⁹ This data represents global actions taken, and not just actions related to content from India. "Proactive Monitoring" refers to content proactively identified by employing internal proprietary tools and industry hash sharing initiatives.