



# GLOBAL TRANSPARENCY REPORT

H1 2024





# GLOBAL TRANSPARENCY REPORT H1 2024

---

At X, our purpose is to serve the public conversation, ensuring a safe environment where everyone can participate freely and confidently. With this transparency report, we aim to cement X as a truly safe platform for all. We achieve this through transparency about our content moderation systems, policies, and enforcement philosophy. Our Global Transparency Report covers X Safety Enforcement for the period of January to June 2024.

# X RULES ENFORCEMENT

---

Our mission at X is to promote and protect the public conversation. We believe X users have the right to express their opinions and ideas without fear of censorship. We also believe it is our responsibility to keep users on our platform safe from content that violates our Rules. Violence, harassment, and other similar types of behavior discourage people from expressing themselves, and ultimately diminish the value of global public conversation.

Our policies and enforcement principles are grounded in human rights, and we have been taking an extensive and holistic approach towards freedom of expression by investing in developing a broader range of remediations, with a particular focus on education, rehabilitation, and deterrence. These beliefs are the foundation of “Freedom of Speech, not Freedom of Reach” - our enforcement philosophy, which means we restrict the reach of posts, only where appropriate, to make the content less discoverable as an alternative to removal. The **X Rules and policies** are publicly accessible on our Help Center.

When determining whether to take enforcement action, we may consider a number of factors, including (but not limited to) whether:

- ▶ The behavior is directed at an individual, group, or protected category of people;
- ▶ The report has been filed by the target of the abuse or a bystander;
- ▶ The user has a history of violating our policies;
- ▶ The severity of the violation; and
- ▶ The content may be a topic of legitimate public interest.

To enforce our Rules, we use a combination of machine learning and human review. These systems either take action automatically, or surface content to human moderators based on user reports and/or proactive detection methods.

Our human moderators use important context to make decisions about potential violations. This work is led by an international, cross-functional team with 24-hour coverage and the ability to cover multiple languages. We also have an appeals process for any potential errors that may occur.



# USER REPORTS AND TOTAL ENFORCEMENT ACTIONS

(EXCLUDING PLATFORM MANIPULATION & SPAM)

We strive to create an environment where users feel empowered to express themselves. When abusive behavior occurs, we make it easy for users to report violations of **X Rules** or local laws. Below, you'll find the total number of global user reports from January to June 2024, excluding Platform Manipulation and Spam, which is broken out separately below.

**USER REPORTS**

**224,129,805**

<b>POLICY AREA</b>	<b>TOTAL</b>	<b>% OF TOTAL REPORTS</b>
<b>Abuse &amp; Harassment</b>	<b>81,730,426</b>	<b>36.47%</b>
<b>Child Safety</b>	<b>8,928,019</b>	<b>3.98%</b>
<b>Hateful Conduct</b>	<b>66,898,539</b>	<b>29.85%</b>
<b>Illegal or Certain Regulated Goods and Services</b>	<b>35,695</b>	<b>0.02%</b>
<b>Misleading &amp; Deceptive Identities</b>	<b>5,179,431</b>	<b>2.31%</b>
<b>Non-Consensual Nudity</b>	<b>38,736</b>	<b>0.02%</b>
<b>Private Content</b>	<b>9,902,566</b>	<b>4.42%</b>
<b>Suicide &amp; Self Harm</b>	<b>2,479,097</b>	<b>1.11%</b>
<b>Violent &amp; Hateful Entities</b>	<b>8,932,100</b>	<b>3.99%</b>
<b>Violent Content</b>	<b>40,005,196</b>	<b>17.85%</b>



# ENFORCEMENT ACTIONS

When we take **enforcement actions**, we may do so either on a specific piece of content (e.g., an individual post or Direct Message) or on an account. Below is a summary table of our automated and manual Rule Enforcement Actions and post violation rates that depict the percentage of all posts that were actioned for violating one of our platform rules during the covered time period. This data may include potential false positives and does not attempt to account for false negatives.

TOTAL ACCOUNT SUSPENSIONS	5,296,870
TOTAL POSTS REMOVED OR LABELED	10,675,980
OVERALL POSTS VIOLATION RATE	0.0123%

POLICY AREA	ACCOUNT SUSPENSIONS	POSTS REMOVED OR LABELED	POST VIOLATION RATE
Abuse & Harassment	1,102,778	2,648,475	0.0030%
Child Safety	2,781,634	14,571	<0.0001%
Hateful Conduct	2,361	4,950,321	0.0057%
Illegal or Regulated Goods & Services	514,095	549,328	0.0006%
Misleading & Deceptive Identities	677,798	62	<0.0001%
Non-Consensual Nudity	52,093	156,498	0.0002%
Private Content	1,719	34,497	<0.0001%
Suicide & Self Harm	2,286	86,184	0.0001%
Violent & Hateful Entities	57,185	153	<0.0001%
Violent Content	104,921	2,235,891	0.0026%



# POLICY SPECIFIC BREAKDOWNS

---

In the following pages, we break down our enforcement actions by policy area over a six-month period from January to June 2024. It is further broken down into the specific actions taken, and if they were performed through automated or human review.

## CHILD SAFETY

At X, we have zero tolerance for child sexual exploitation and are committed to removing media that depicts physical child abuse. We also suspend users who engage with that content to prevent the normalization of violence against children. Along with taking action under our Rules, we also work closely with the National Center for Missing and Exploited Children (NCMEC).

We send a mix of automated and human reviewed reports to NCMEC. We will never stop using the full extent of our capabilities to eliminate these harmful practices.

Read our comprehensive policy that outlines how we make our platform inhospitable for those who seek to exploit minors in any way [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
NCMEC Reports	370,588	35,176	335,412
Accounts Suspended <sup>1</sup>	2,781,634	2,388,683	392,951
Content Removed	14,571	1,645	12,926

<sup>1</sup> Accounts suspended for engaging with (liking, replying, bookmarking, etc.) Child Sexual Abuse Media (CSAM) are not reportable to NCMEC

## ABUSE, HARASSMENT, & HATEFUL CONDUCT

At X, we are committed to a respectful community. We prohibit direct attacks based on race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. We take swift action against any violations to ensure a safe space for everyone. Read more [here](#).

Additionally, we prohibit ongoing behavior and content that targets others with abuse or harassment, or encourages other people to do so. Our policies are designed to ensure everyone can engage safely and respectfully, as defined [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	1,105,139	154	1,104,985
Content Removed	2,247,107	5,679	2,241,428
User Informed Labels	5,351,689	3,966,430	1,385,259



## PLATFORM MANIPULATION & SPAM

We want X to be a place where people can make human connections, find reliable information, and express themselves freely and safely. To make that possible, we do not allow spam or other types of platform manipulation. We define Platform Manipulation as using X to engage in bulk, aggressive, or deceptive activity that misleads others and/or disrupts their experience.

Multiple accounts that exhibit a pattern or sole purpose of the above behaviors, or found to be misusing X product features to disrupt others' experiences, may be suspended by X. Posts may not appear in certain parts of the X product, and/or may not be recommended or amplified by X due to multiple product quality or Rule enforcement reasons. [Learn more](#) about the instances when a Post's visibility may be limited.

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	463,960,085 <sup>2</sup>	461,172,954	2,787,131
User Informed Labels	187,689,809	187,255,600	434,209

<sup>2</sup> Our Platform Manipulation & Spam defenses are primarily proactive or automated. We remove multiple times more accounts than we receive user reports.

## VIOLENT & HATEFUL ENTITIES

At X, we have zero tolerance for threats or promotion of terrorism and violent extremism. We do not allow Violent or Hateful Entities—including terrorist organizations, extremist groups, [perpetrators of violent attacks](#), or

individuals promoting violence—to use our platform. Such content jeopardizes the safety and well-being of targeted individuals. Read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	57,185	35,420	21,765
Content Removed	153	62	91



## VIOLENT CONTENT

X is a place where people can express themselves, show and learn about what's happening, and debate global issues, often sharing images and videos as part of the conversation. However, healthy conversations can't thrive when Violent Speech is used to deliver a message, and not every participant wishes to be exposed to Violent Media.

As a result, we may remove or reduce the visibility of Violent Content to ensure the safety of our users and prevent the normalization or glorification of violent actions. We also do not allow sharing Violent Content in highly visible places such as profile photos, banners or bio. Read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	104,921	4	104,917
Content Removed	1,797,553	1,103,131	694,422
User Informed Labels	438,338	0	438,338

## SUICIDE & SELF HARM

At X, we are dedicated to supporting mental health and preventing harm. We do not permit content that promotes or encourages suicide or self-harm.

Our policy is designed to provide support and resources for those experiencing self-harming behaviors or suicidal thoughts. Read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	2,286	90	2,196
Content Removed	86,184	58	86,126

## MISLEADING & DECEPTIVE IDENTITIES

At X, authenticity is crucial. We do not tolerate the misappropriation of identities or the use of fake identities to deceive others. Our policies are designed to prevent any form of identity fraud or misleading behavior on our platform.

Read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	677,798	470,435	207,363
Content Removed	62	0	62

## NON-CONSENSUAL NUDDITY POLICY

At X, sharing explicit sexual images or videos of someone without their consent is strictly prohibited, whether the content is AI-generated or organically created. This severe violation of privacy, often referred to as revenge porn, poses serious safety and security risks and can lead to significant physical, emotional, and financial hardship for those affected.

To learn more about our non-consensual nudity policy, read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	52,093	520	51,573
Content Removed	156,498	22,653	133,845

## PRIVATE CONTENT

At X, protecting privacy is a priority. You may not threaten, publish, or share other people's private information or media without their express authorization and consent. We enforce strict measures to safeguard personal privacy and prevent unauthorized disclosures. Read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	1,719	4	1,715
Content Removed	34,497	4,047	30,450

## ILLEGAL OR REGULATED GOODS/SERVICES

You may not use our service for any unlawful purpose or in furtherance of illegal activities. This includes selling, buying, or facilitating transactions in Illegal Goods or Services, as well as certain types of regulated goods or services.

Read more [here](#).

	TOTAL ACTIONS	AUTOMATED	HUMAN
Accounts Suspended	514,095	84,482	429,613
Content Removed	549,328	132	549,196



# GOVERNMENT, LEGAL, & LAW ENFORCEMENT REQUESTS

---

At X, government, legal, and law enforcement requests are managed through our **established guidelines**. We provide clear procedures for law enforcement seeking account information and content removal, and may file or serve objections in certain circumstances <sup>3</sup>. For detailed information and to learn about **requests to withhold content**, please **read more**.

<sup>3</sup> For example, X may file or serve objections for requests that are legally defective, overly broad, and/or appear to impermissibly burden free expression.



## INFORMATION AND REMOVAL REQUESTS

This section presents the latest data on government requests for X account information globally, including both emergency and non-emergency inquiries.

	TOTAL NUMBER OF REQUESTS	CASES DISCLOSED	DISCLOSURE RATE
<b>TOTAL INFORMATION REQUESTS</b>	<b>18,737</b>	<b>9,897</b>	<b>52.82%</b>
USA	3,329	2,522	76%
Japan	2,726	1,627	60%
EU	7,872	4,388	56%
UK	635	307	48%
All Others	4,175	1,135	27%

  

	TOTAL NUMBER OF REQUESTS	CASES ACTIONED	ACTION RATE
<b>TOTAL REMOVAL REQUESTS</b>	<b>72,703</b>	<b>51,487</b>	<b>70.82%</b>
Turkey	9,364	6,332	68%
Japan	46,648	36,731	79%
South Korea	5,889	4,296	73%
EU	2,458	1,965	80%
All Others	8,345	2,163	26%

## COPYRIGHT AND TRADEMARK NOTICES

Respecting intellectual property is crucial. At X, we take intellectual property rights seriously, including both copyright and trademark protections. A trademark or copyright can be a word, logo, or phrase that identifies a brand's goods or services.

Trademark law helps prevent unauthorized or confusing use of these trademarks.

To learn more about how we uphold these rights and ensure compliance, read our detailed Trademark policy [here](#) and Copyright policy [here](#).

	TOTAL NOTICES	NOTICES ACTIONED OR WITHHELD	TOTAL URLS ACTIONED OR WITHHELD	WITHHOLDING/ RESTORATION RATE
Copyright Notices	188,661	112,897	836,194	59.84%
Copyright Counter Notices	14,496	1,698	1,698	11.71%
Trademark Notices	70,980	6,296	6,296	8.87%



