



@TwitterSafety

Trust & Safety

Lets get started...

We want people to feel safe on Twitter. Last year, we shared that building a Twitter free of abuse, spam and other things that distract from the public conversation is our top priority. Since then, we've made strides in creating a healthier service.

Taking action without reports

People who don't feel safe on Twitter shouldn't be burdened to report abuse to us. Previously, we only reviewed potentially abusive Tweets if they were reported to us. We know that's not acceptable, so earlier this year we made it a priority to take a proactive approach to abuse in addition to relying on people's reports.

Key metrics

>50%

Of abusive content that's enforced is surfaced proactively

16%

Fewer abuse reports after an interaction from an account the reporter doesn't follow

3X

More abusive accounts suspended within 24 hours that are reported to us

45%

Increase in accounts suspended for creating new accounts after a suspension

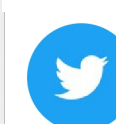
Making our rules easier to understand

Twitter's purpose is to serve the public conversation. Violence, harassment and other similar types of behavior discourage people from expressing themselves, and ultimately diminish the value of global public conversation. We recently updated our Rules to make them simpler and clearer. You can find them at twitter.com/rules



Rules

- [Violence](#) - You may not threaten violence against an individual or a group of people. We also prohibit the glorification of violence.
- [Abuse/harassment](#) - You may not engage in the targeted harassment of someone, or incite other people to do so. This includes wishing or hoping that someone experiences physical harm.
- [Hateful conduct](#) - You may not promote violence against, threaten, or harass other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease
- [Private Information](#) - You may not publish or post other people's private information (such as home phone number and address) without their express authorization and permission. We also prohibit threatening to expose private information or incentivizing others to do so.



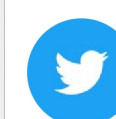
Twitter Safety @TwitterSafety · 5d
Rules should be easy to understand. We heard you, ours weren't. We updated, reordered, and shortened them, so you can know what's not allowed on Twitter. Click through this thread for all our rules, and read our blog to learn more.



Making our rules easier to understand
blog.twitter.com

1,300 1,309 2,729

[Show this thread](#)



Twitter Safety @TwitterSafety · 5d
Note: we may need to change these Rules from time to time in order to support our goal of promoting a healthy public conversation. The most current version is always available here



The Twitter Rules



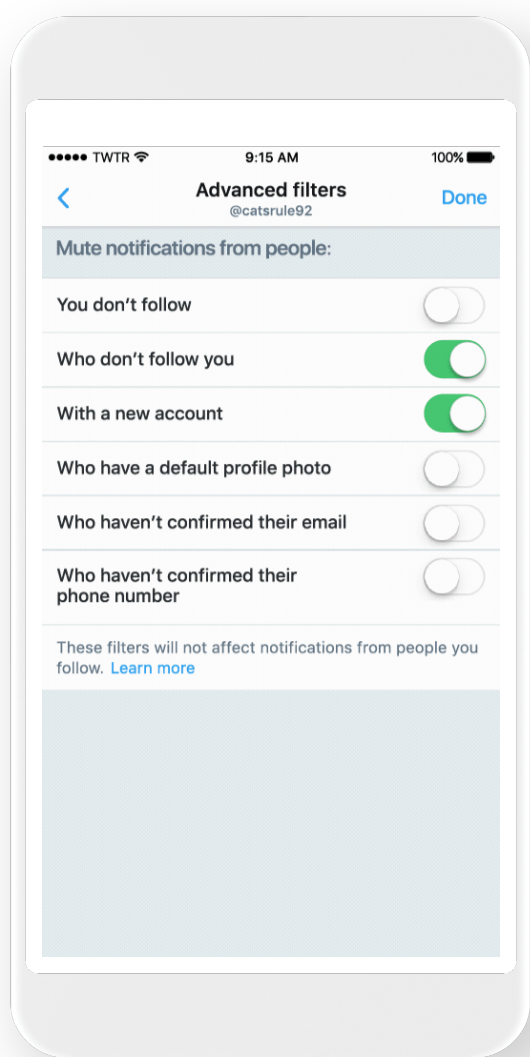


@TwitterSafety

Trust & Safety

Staying safe

To keep you safe and to empower you on Twitter, we build tools so you can control what you see, what others see, who you interact with, and what you can report. Below is a guide to some tools which may help.



Advanced filter settings

You can control what you see and who you interact with on Twitter.

Advanced filter settings allow you to disable notifications from certain types of accounts that you'd like to avoid. In addition, if your account receives a lot of sudden attention, we may insert a notification in your Notifications tab inviting you to adjust these filters to give you more control over what you see.

You can access *advanced filter settings* by navigating to [Settings and privacy > Privacy and safety > Notifications > Advanced filters](#)

What you see

On Twitter we allow you to customise what content you would like to see on your feed.

Block: Restrict specific accounts from contacting you, seeing your Tweets, and following you.

Mute: allows you to remove an account's Tweets from your timeline without unfollowing or blocking that account. You can also mute particular words, conversations, phrases, usernames, emojis, or hashtags.

Safe Search: The Safe Search function removes potentially sensitive content by default as well as accounts people have blocked and muted from search pages.

Sensitive Media: Twitter's default setting is to place potential sensitive material behind a warning. This can be adjusted in settings.

What other people see

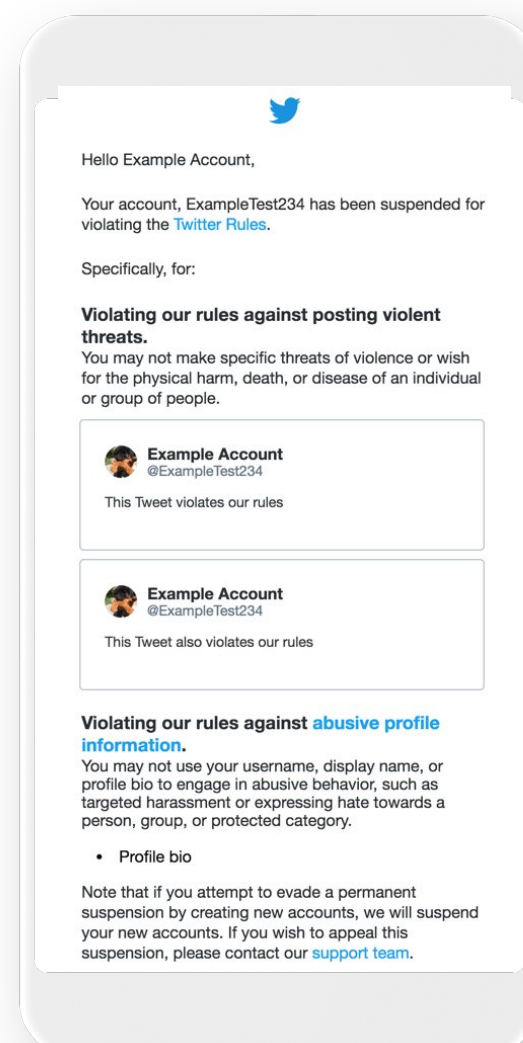
Twitter allows you to control what others can see about you on the platform.

Protect your Tweets: By default all of your Tweets are public but you can choose to protect your Tweets in settings. You will receive a request when people ask to follow you. Protected Tweets will not appear in third-party search engine results.

Photo Tagging: If you don't want to be tagged in other's photographs, we give you the option to disable the ability to do so.

Location Sharing: Twitter lets you select whether to include your location on individual Tweets.

Discoverability: Often the best connections on Twitter are with people you already know. In order to help you make these connections, we use information like your email address and/or phone number to make you more discoverable. You can adjust this setting at any time.



How to report content

Have you seen something you think breaches our policy? You can [report](#) the Tweet, DM or account for us to review.

You can report directly from an individual Tweet or profile for certain violations, including: spam, abusive or harmful content, inappropriate ads, self-harm and impersonation.

You can find more information on Safety Tools at safety.twitter.com

Follow [@TwitterSafety](#) for updates