Helping you feel safer on Twitter
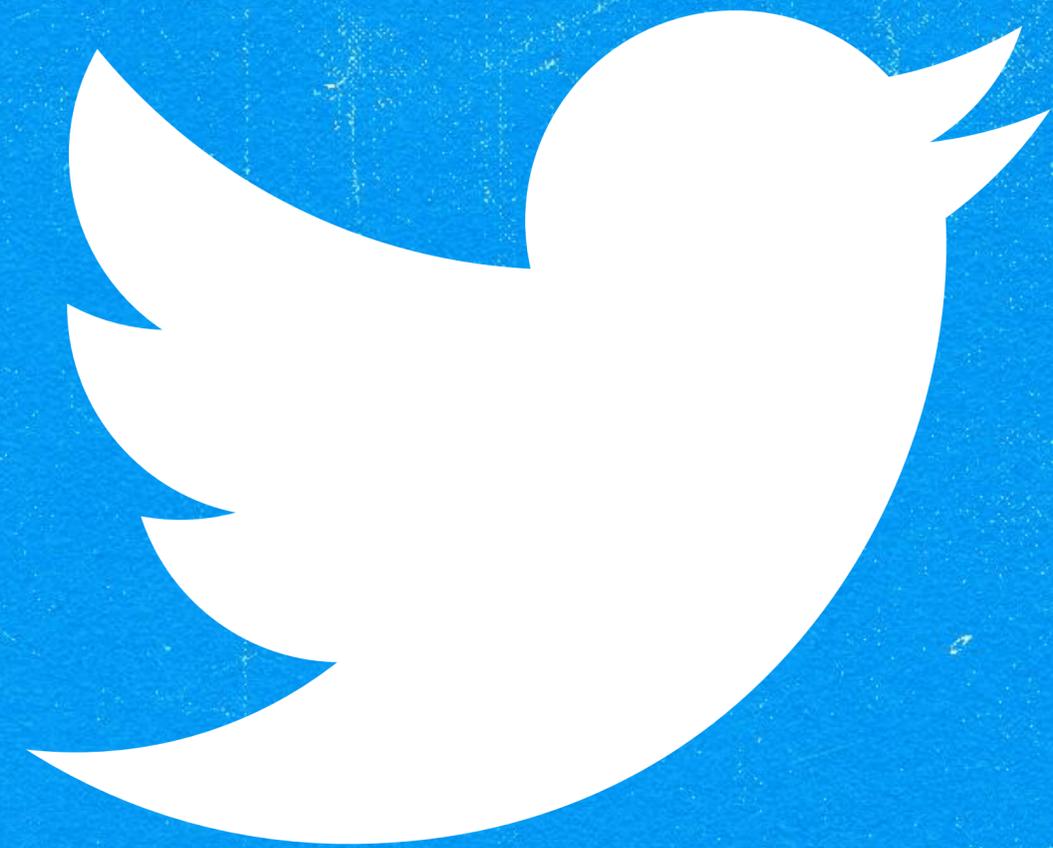
We want people on Twitter to have safe, inclusive, and authentic conversations. We have rules in place and enforce against them, but negative experiences can still happen. So we've created safety tools to help you:

Feel safer

Be in control

Manage your digital footprint

Table of Contents

Feel Safer:

You should feel safer participating in the conversation

Be in Control:

Have control over your experience and what you see on the platform

Manage your Digital Footprint:

Understand and manage the data others have about you

Let's review some of our safety tools

Which of the following are you having trouble with?

An individual account	A conversation with many people	My Twitter feed in general	What others might know about me
Block	Hide Replies	Remove Sensitive Content	Enable 2FA
Mute	Change Who Can Reply	Safety Mode	Protected Tweets
Remove Follower	My DMs		Turn off photo tagging
My DMs			Discoverability
			Data Sharing

Another way to think about these tools

Are you experiencing abuse or harassment on Twitter?

Yes, through Tweet replies

Yes, through messaging (DM)

From one or small number of accounts

From many accounts in a coordinated manner

From one or small number of accounts

From many accounts in a coordinated manner

Hide Replies	**Safety Mode***	**My DMs***	**Safety Mode***
Change who can reply	**Change who can reply**	**Remove Followers**	**My DMs***
Remove Followers	**Block**	**Block**	**Block**
Block	**Report**	**Report**	**Report**
Report			

Only works for accounts you don't follow

Feeling Safe on Twitter

We know at times you have not had the experience you expected. We'll go over options that might help you.

Twitter has rules related to the following:

🚫 **Violence**
Making violent threats or glorifying violence

🚫 **Hateful conduct**
Threatening on the basis of protected categories listed here

🚫 **Terrorism/violent extremism**
Promoting terrorism or violent extremism

🚫 **Suicide or self-harm**
Cannot be promoted or encouraged

🚫 **Child sexual exploitation**
We have zero tolerance for this on Twitter

🚫 **Abusive sharing of sensitive media**
Targeting people with unsolicited images or videos that contain graphic violence, adult content, or hateful imagery

🚫 **Abuse / harassment**
Wishing or hoping someone experiences harm

🚫 **Illegal or specific regulated goods and services**
Promotion of illegal activities including selling, buying, or facilitating transactions of illegal goods or services

You can read more about these Safety policies here.

If you see any violation of our rules, please report it.

We are also working proactively to protect you. We use a combination of human review and technology to help us enforce our rules.

Reporting step by step

How to report on Twitter

An important option for you

Reporting helps us keep conversations on Twitter healthier. You can use the Report function to report individual Tweets, a DM, or an account itself. They will be monitored by our team on a recurring basis.

Some common reasons people might report include spam, abusive content, impersonation, copyright, or trademark violations. Learn more about rules and reporting here.
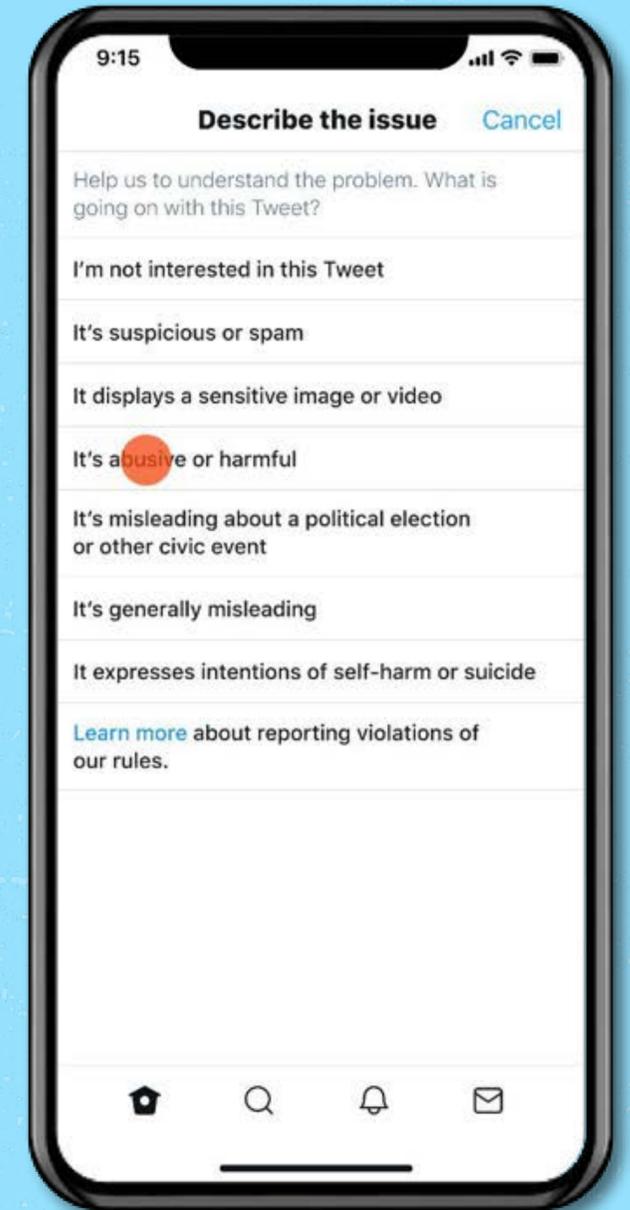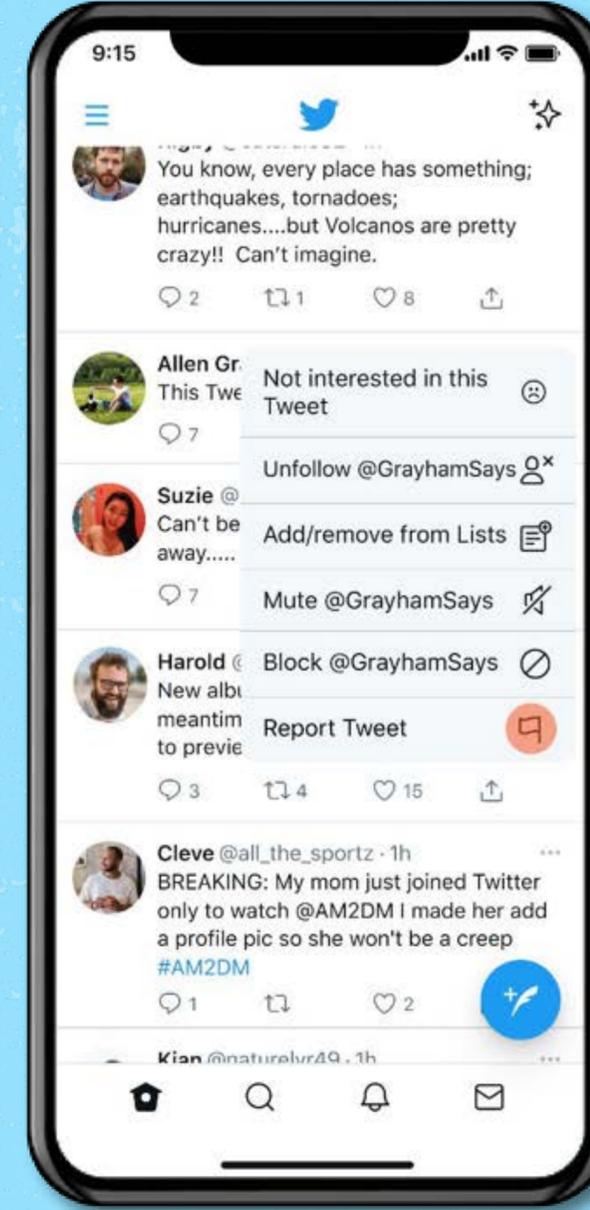
Reporting FAQs

- **How do I submit a report?**
  Take a look at this <u>quick video.</u>

- **Why does my reporting flow look different from the video?**
  We are continually improving our reporting process and some details might not look the same as pictured.

- **Can I report multiple Tweets from one account?**
  Yes.

- **When will we hear back from you?**
  After you submit a report, you will see a confirmation email within 24 hours. If there is violation of our policies, we will take action (<u>ranging from a warning to permanently suspending the account</u>). You may receive a follow up from us when we take action.

- **Can someone report on my behalf?**
  Yes, others can report a Tweet directly or via the <u>report form</u>. Though in some instances, a report by the impacted account allows us to better evaluate the context of the reported content.

After you Report

What if my report does not end up violating the Twitter Rules?

If you file a report and it is not found to be in violation of the Twitter Rules, you may still want to take action. There are a number of tools that can help manage your experience moving forward.

Options to consider

Block

Restrict specific accounts from contacting you, seeing your Tweets, and following you. Learn more on the next page.

Filtered Notifications

Get an extra level of control by filtering the notifications you receive. You can filter by using: Quality filter, muted words, and advanced filters.

Sensitive Media

Twitter's default setting is to place potential sensitive material behind a warning. You can opt into viewing this media by adjusting your settings.

Mute

Allows you to remove an account's Tweets from your timeline without unfollowing or blocking it. You can also mute particular words, conversations, phrases, usernames, emojis, or hashtags. Learn more on the next page.

Safe Search

The Safe Search function can help remove potentially sensitive content by default as well as accounts people have blocked and muted, from search results.

Basic Safety Controls

Block

When it becomes too much

Block lets you restrict specific accounts from contacting you, seeing your Tweets, and following your account.

People are not notified when you Block them but if they attempt to visit your profile, they will see a notice that they are Blocked. Note that you'll see labels and warnings if someone you've blocked is speaking in a Space you join.

Learn more about Block here.

Mute

To minimize unwanted noise

Mute is helpful if you want to keep following people but need to turn down their volume or reduce the number of their Tweets in your timeline. They won't know you've muted them and you can unmute them later if you want.

You can mute accounts as well as keywords from appearing in Notifications, on Home timeline, and in conversations. Muting keywords is a proactive way of avoiding certain themes and content.

You can also mute entire conversations, for which you will stop receiving notifications, but those muted conversations can appear in your timeline.

Learn more about Mute here.

How to think about muting and blocking

Muting an account you follow	Muting an account you don't follow	Blocking an account

Alice → follows and mutes → **Bob**

Alice → mutes → **Bob**

Alice → blocks → **Bob**

Alice wants to keep following Bob's account, but she doesn't want to see Bob's Tweets in her timeline.

Alice only wants to see content from Bob that involves her (like mentions and Direct Messages).

Bob won't see that he's been muted by Alice.

Alice doesn't want to follow Bob, and she doesn't want to see Bob's Tweets in her timeline.

Alice isn't interested in Bob's Tweets, but Bob keeps mentioning her and cluttering her notifications timeline.

Bob won't see that he's been muted by Alice.

Alice doesn't want Bob to follow her or see her Tweets. She doesn't want to see Bob's Tweets either.

Alice doesn't want to get any notifications from Bob.

Bob will see that he's been blocked when he visits Alice's profile.

Remove followers

## Remove followers without blocking them

This feature allows people to limit interactions with someone without having to block them. It is intended for everyone but can be specifically helpful for:

- People with public accounts who don't receive "accept" or "reject" follow requests.
- People who are moving from Public to Protected Tweets.
- People who want to curtail interactions with others without the severity of blocking.

It is currently available through web (desktop) only.

FAQs

- **How can I remove a follower?**
  You can remove a follower from your Followers list or on their Profile.

- **Can I remove followers even if my Tweets are public?**
  Yes, followers can be removed regardless of whether their Tweets are Protected or Public.

- **Are people notified if you remove them?**
  No, they will not be notified by Twitter if you remove them as a Follower, but they will be able to Follow you again.

- **How did people with public profiles remove followers before this feature?**
  Before this feature, both people with Public and Protected Tweets could remove a follower by blocking and then unblocking them.
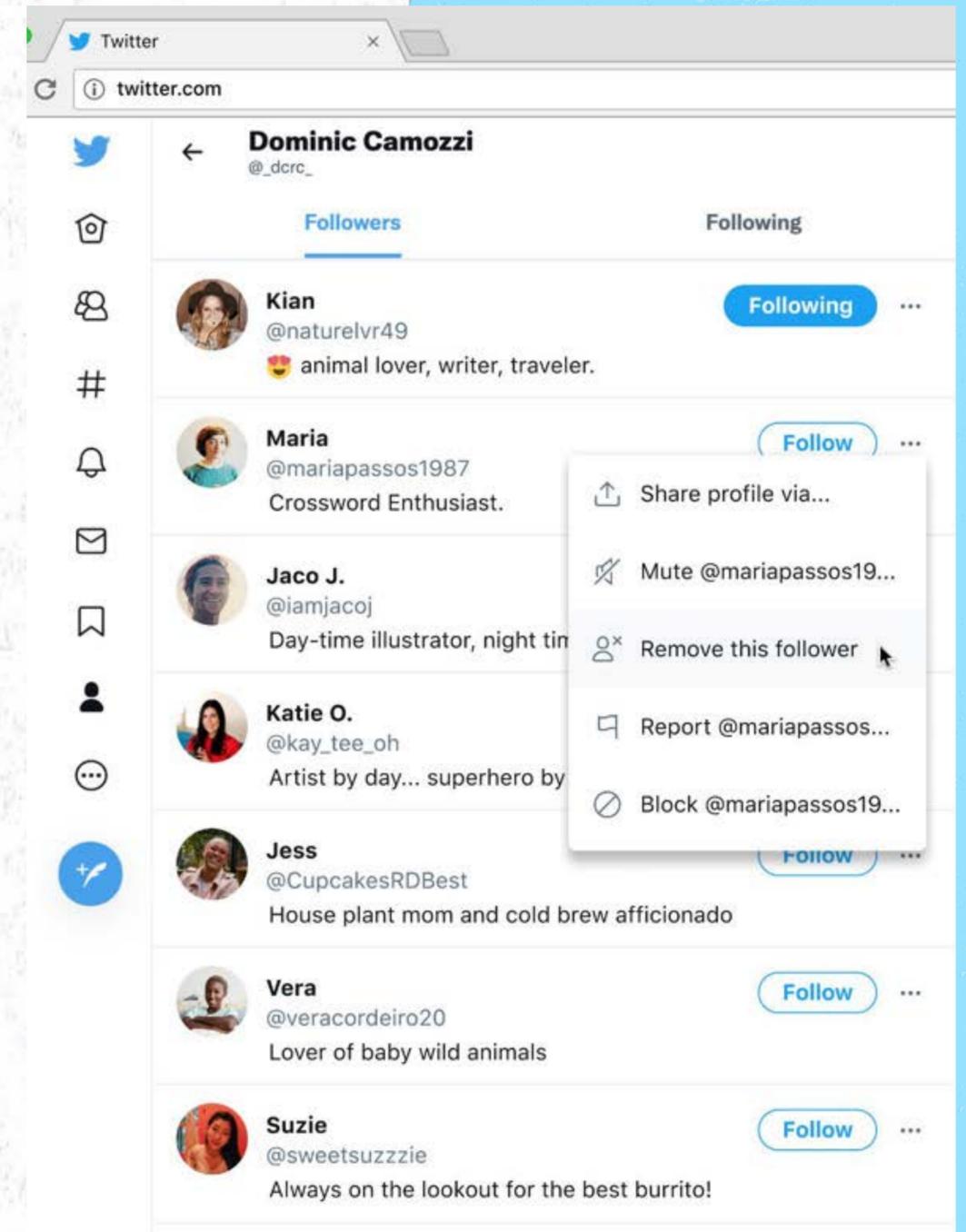
Safety Mode

The big red emergency button

Currently in Beta, globally available soon

Safety Mode aims to automatically block accounts that appear to violate Twitter Rules or might be using insults, name-calling, strong language, or hateful remarks. It's meant to proactively keep you safe from unwanted interactions.

Accounts are auto-blocked for seven days, but auto-block can be adjusted or undone at any point.

This is an optional and temporary safety control. We're always clear if a block was made by Safety Mode or by you.

Learn more about Safety Mode here and here.

FAQs

- **Who is this feature for?**
  It's for everyone, but especially for marginalized communities and female journalists. Through tools like this, customers can decide where to draw the line in conversations that might impact psychological or emotional safety.

- **Why not just manually block?**
  Safety Mode helps when you experience 'dogpiling' or abuse / spam from many people at once. It can be overwhelming to spend time blocking accounts at this scale so we created this tool to lift the burden for you.

- **When will this feature be available to everyone?**
  We plan to launch globally in early 2022.
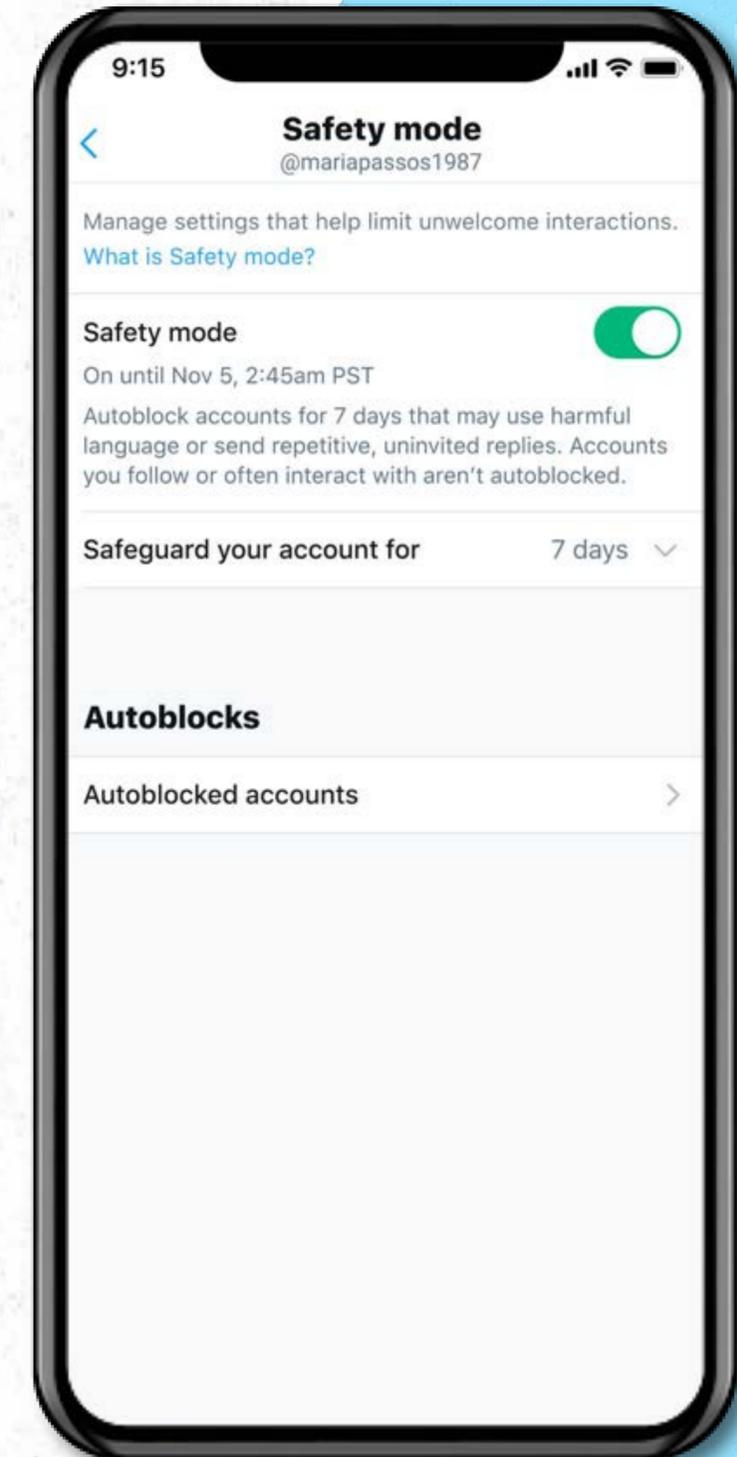
Be in Control

Have control over your messages, your conversations, and what you see on the platform.

Manage Content you See

Stop seeing uncomfortable or irrelevant content

If certain kind of sensitive content (e.g., violence or nudity) makes you uncomfortable, you can navigate to **Privacy and Safety → Content you see** and make sure that the **"Display media that may contain sensitive content"** box is unchecked. *(Can be updated on Web only.)*

If you don't want sensitive media to show up in your search results, you can also enable **Safe Search** mode by clicking into **Search Settings** and hiding sensitive content. *(Can be updated on Web only.)*

Lastly, you can review and modify the **'Interests'** section to manage the themes and topics that Twitter recommends to you on your Home timeline.
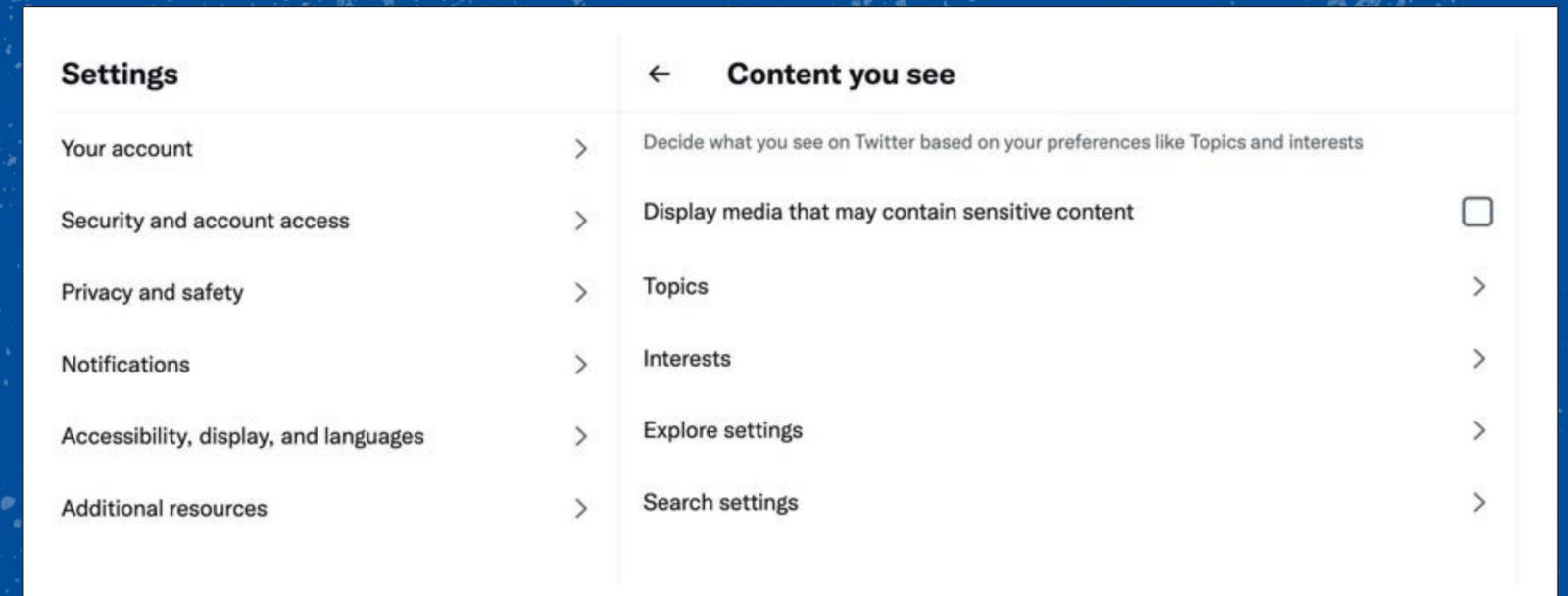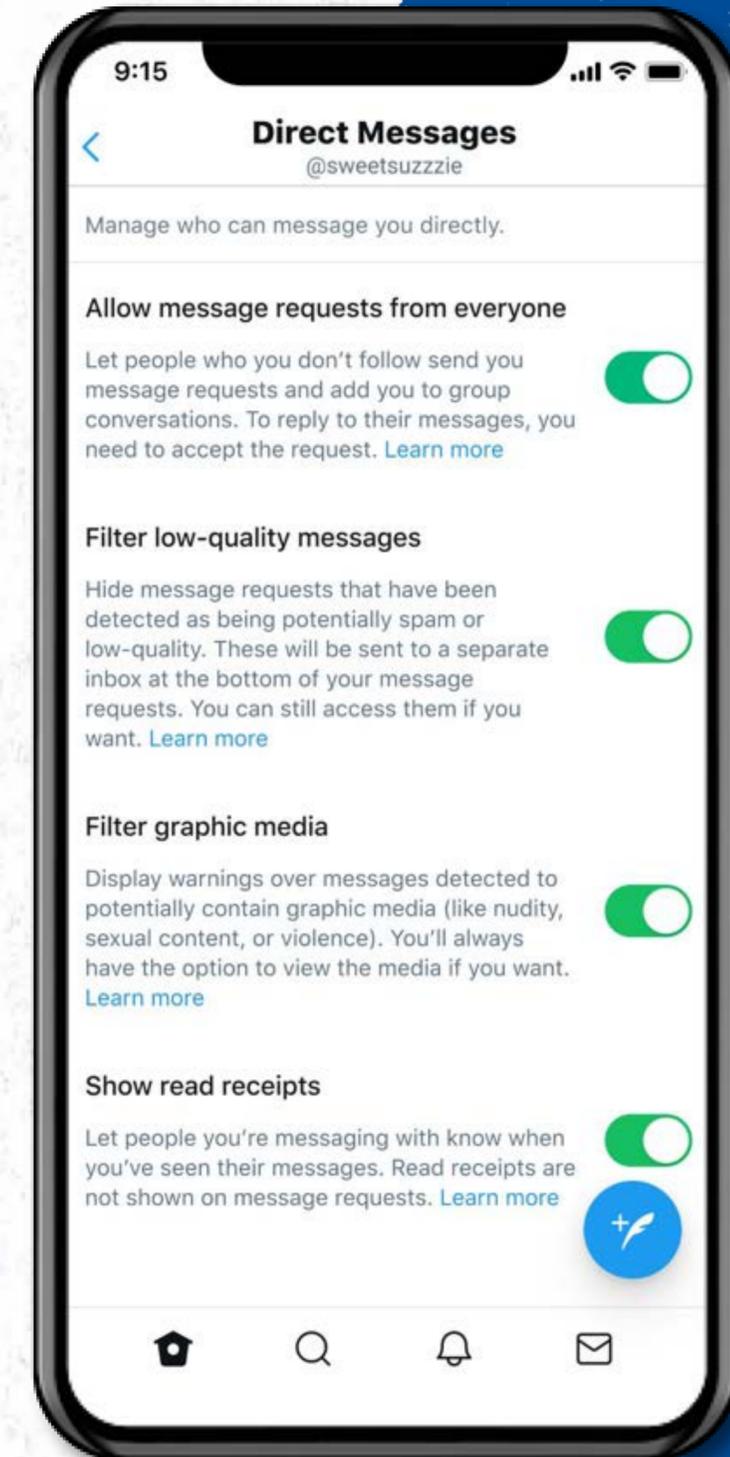
My DMs

Controlling who messages you

Direct Messages allow you to have a private conversation or create a group conversation, but sometimes people receive unwanted DMs. DM settings allow you to filter out unwanted messages and limit messages from people you don't follow.

- If the setting is turned off, you will only receive messages from people you follow

- Can toggle the Quality filter to exclude low-quality or potentially spammy DMs

Learn how to utilize these settings <u>here</u>.

FAQs

- **If I delete a DM, is it deleted for other people in the conversation too?**
  No. When you delete a Direct Message or conversation, it is deleted from your account only. Others in the conversation will still be able to see Direct Messages or conversations that you have deleted.

- **If I have blocked an account, can they DM me?**
  No. Blocked accounts are unable to DM you.

9:15

Direct Messages
@sweetsuzzzie

Manage who can message you directly.

Allow message requests from everyone

Let people who you don't follow send you message requests and add you to group conversations. To reply to their messages, you need to accept the request. Learn more

Filter low-quality messages

Hide message requests that have been detected as being potentially spam or low-quality. These will be sent to a separate inbox at the bottom of your message requests. You can still access them if you want. Learn more

Filter graphic media

Display warnings over messages detected to potentially contain graphic media (like nudity, sexual content, or violence). You'll always have the option to view the media if you want. Learn more

Show read receipts

Let people you're messaging with know when you've seen their messages. Read receipts are not shown on message requests. Learn more

Hide Replies

Managing the tone
of your conversation

The Hide Replies features allow you to control who gets visibility in your conversations. If you believe a reply is off-topic or spammy and detracts from the conversation, you can click on the reply and hide it. Replies are not deleted but placed behind an extra click.

Learn how to Hide Replies here.

FAQs

- **Could a feature like this prevent conversation?**
  This feature is meant to promote online safety and civility. Replies will be still accessible for anyone who clicks into the hidden replies section.

- **Can the replier see if I hid their reply?**
  Yes. Anyone who has access to the original Tweet can see which replies were hidden.
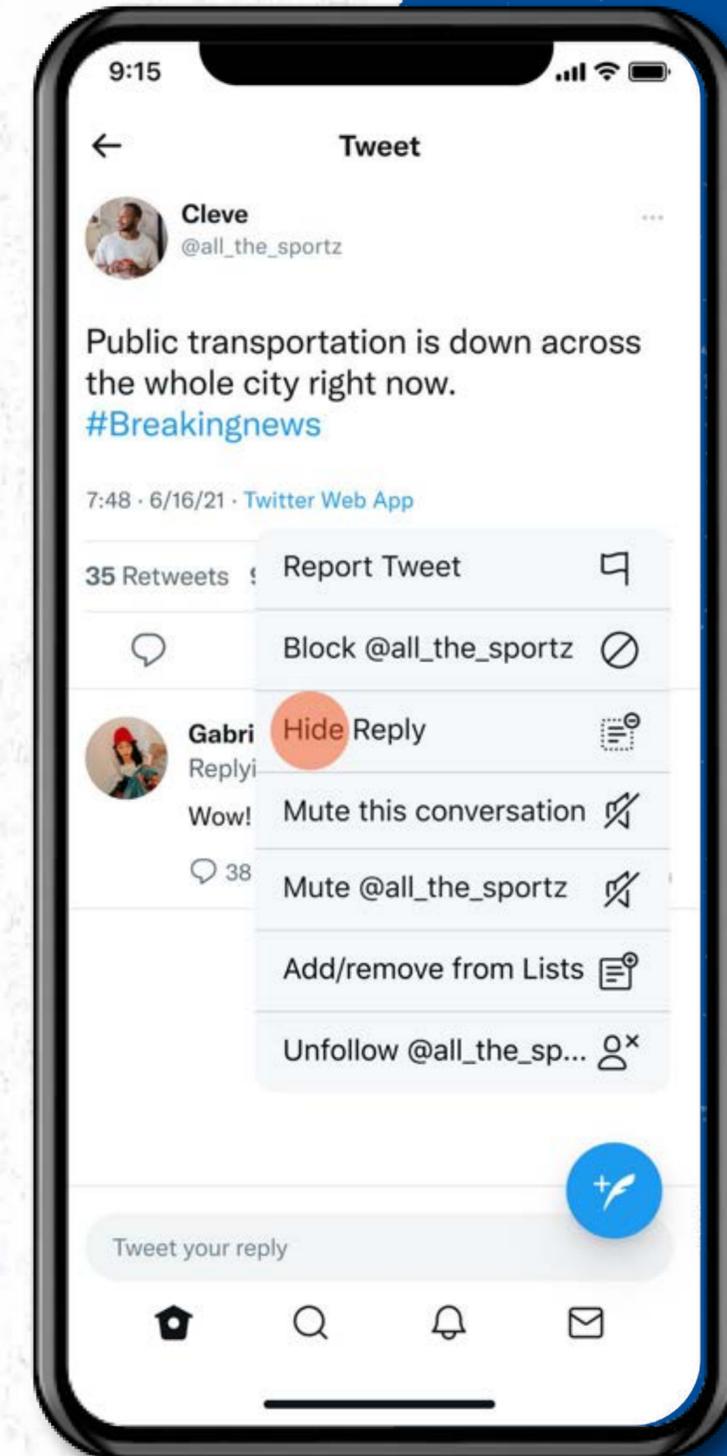
Conversation Settings

Directly controlling who can reply

Sometimes people are more comfortable Tweeting when they can choose who can reply.

Using conversation settings frees you up so you can engage in more meaningful conversations and avoid unwelcome replies.

This feature can help you proactively manage a conversation without having to use Block and can prevent spam.

While everyone can see the Tweet and conversation, depending on your setting, not everyone would be able to reply.

**FAQs**

- **When can I change conversation settings?**
  You can modify this before you Tweet but also after you have Tweeted.

- **Can others still engage with the Tweet?**
  Yes. People who can't reply will still be able to view, Retweet, Retweet with Comment, share, and like these Tweets.

- **How is the setting useful for journalists?**
  News outlets and journalists have used the conversation settings to conduct one-on-one interviews, have conversations as a reporting team, and host panel discussions.
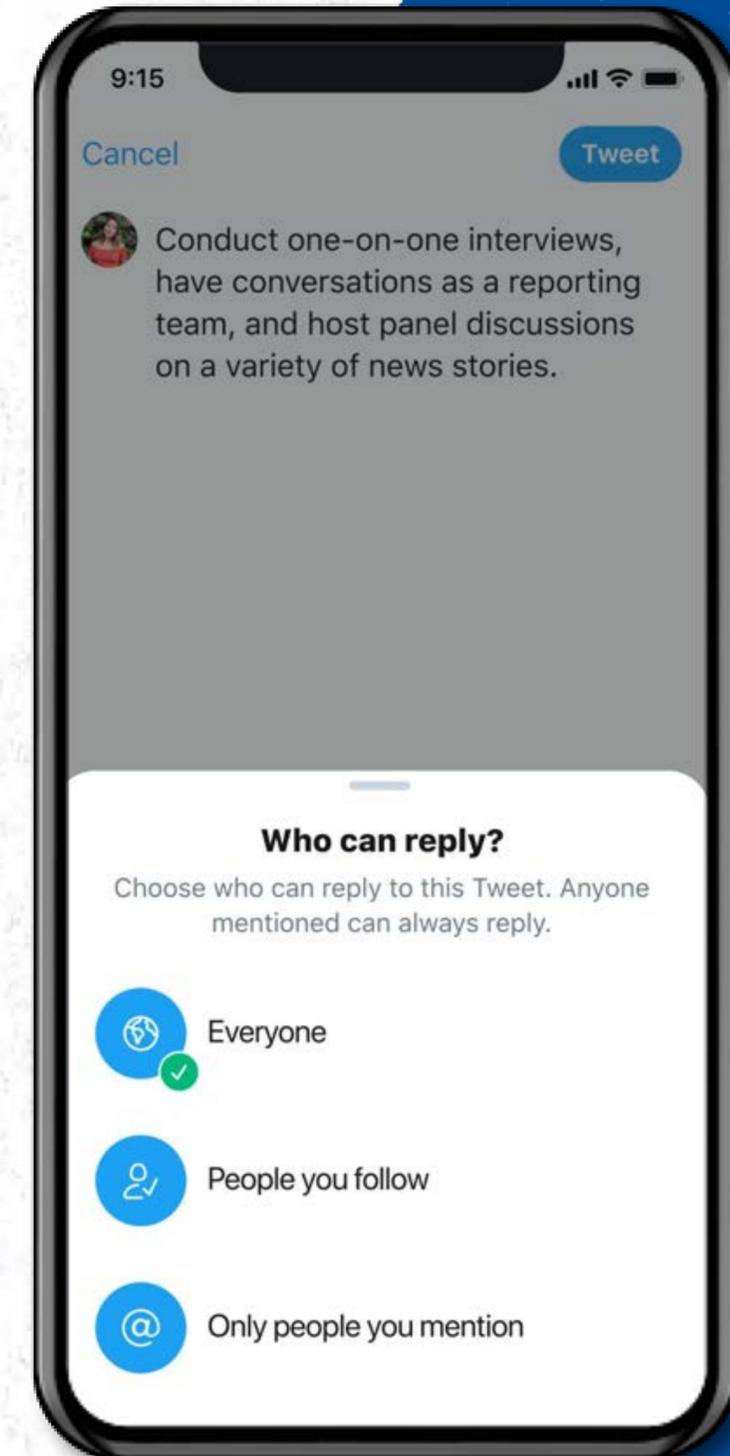
Manage your digital footprint

Keep yourself and your data safe on Twitter. Here's where to start.

Helping secure your account

Use two-factor authentication:

Add an extra layer of security for your account that allows you to add a further security check in addition to your password. Navigate to Settings and privacy → Security and account access → Security.

Enable password reset protection:

Select this option so that you're asked for the associated email address and phone number for the account. This will enable a password reset email or text. Use a strong password that you won't use on other websites. Navigate to Settings and privacy → Security and account access → Security.

Consider protecting your Tweets:

When you sign up for Twitter, your Tweets are public and able to be seen by all. But, you can also choose to make your Tweets visible only to your followers. You'll receive a request whenever someone wants to follow you. Navigate to Settings and privacy → Privacy and safety →Audience and tagging.

Beware of suspicious messages:

Twitter will only reach out to you in-app or via an email from the following domain https://twitter.com. If you're in any doubt, go directly to Twitter.com and log in there.

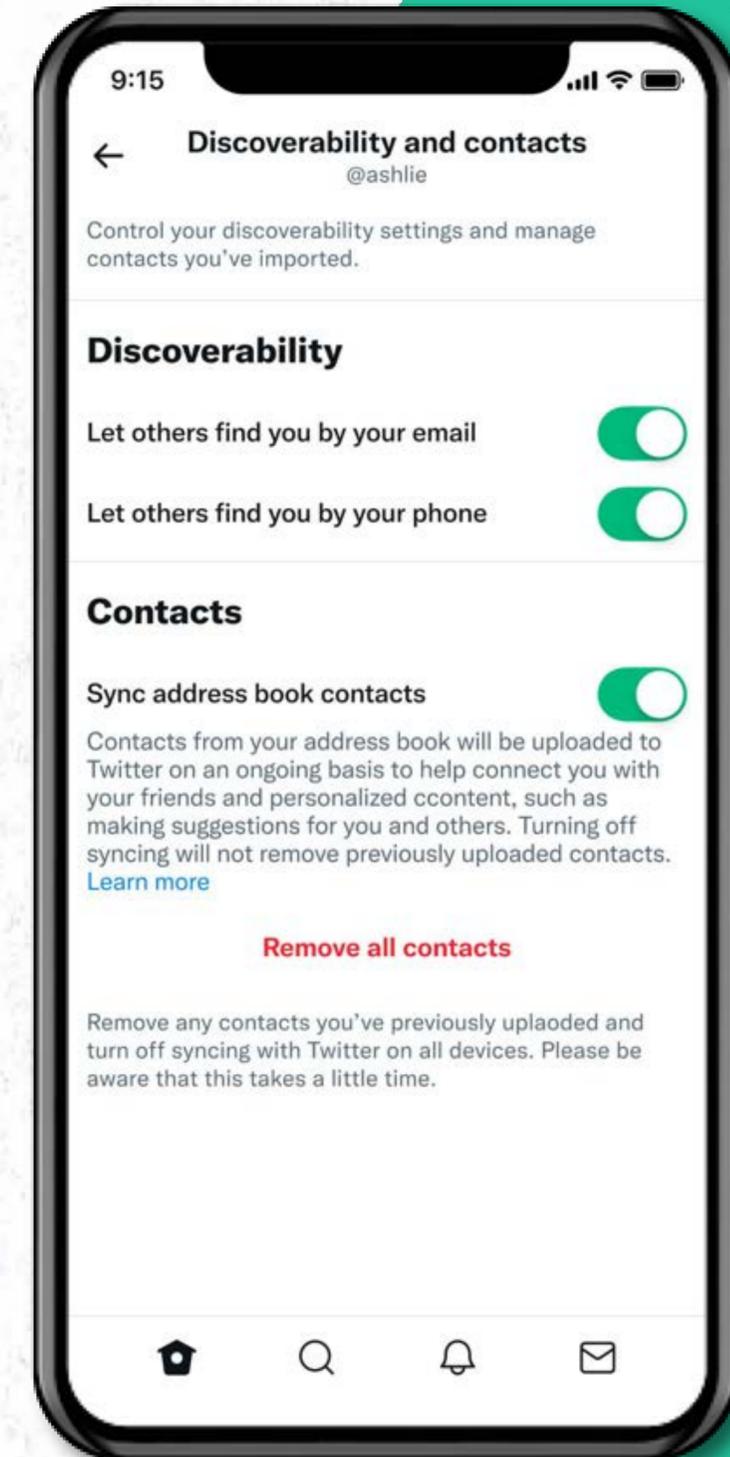# Controlling your public visibility

Turn off photo tagging:

To minimize sharing your name to the public, you can disable a feature that allows anyone to tag you in a photo or limit tagging by only those you follow. This can be found in Settings and privacy → Privacy and safety →Audience and tagging.

Discoverability by others:

When someone new joins Twitter, your account may be suggested for them to follow if they, and Twitter, have access to your email and/or phone number. If you don't want to be suggested to new accounts based on this data, you can disable your discoverability by email and/or discoverability by phone number.

9:15

Discoverability and contacts
@ashlie

Control your discoverability settings and manage contacts you've imported.

Discoverability

Let others find you by your email

Let others find you by your phone

Contacts

Sync address book contacts

Contacts from your address book will be uploaded to Twitter on an ongoing basis to help connect you with your friends and personalized content, such as making suggestions for you and others. Turning off syncing will not remove previously uploaded contacts.
Learn more

Remove all contacts

Remove any contacts you've previously uplaoded and turn off syncing with Twitter on all devices. Please be aware that this takes a little time.

It's important for you to have access to all your data

Twitter believes that everyone who uses our services should understand and have meaningful controls over their data.

Navigate to **Settings and privacy → Privacy and safety** to learn and adjust your settings. You can learn more about privacy in our Privacy Center as well.

To see all the data that Twitter has on file about you, navigate to:

Settings and privacy → Your Account. You will be asked to provide your account information and will be able to **download an archive of your data**.

To learn more about how to create a safe space for yourself on Twitter, view these Safety videos or visit the Safety & Security hub.