



Twitter Report: Staying safe and informed on Twitter during COVID-19

Table Of Contents:

[Executive Summary](#)

[Helping people find reliable information](#)

[Useful links](#)

[COVID-19 Guidance Enforcement](#)

[Advertising on COVID-19](#)

[Political content advertising on Twitter](#)

[Useful links](#)

[Research and data access](#)

[Useful links](#)

Executive Summary

As the global community faces the COVID-19 pandemic together, and with the latest Omicron variant ramping up across the EU after the holiday season, Twitter is helping people to:

- find reliable information and evaluate credible sources,
- connect with others,
- follow what's happening in real time, and
- commit to safe behaviours, from washing hands to wearing a mask.

Twitter has an important role to play as a place for good faith public debate and discussion around critical public health matters. Throughout these unprecedented times, Twitter has continued to adapt and update our policies and enforcement, as well as increase transparency and share more data to ensure experts and the public can better analyse how discussion around COVID-19 continues to evolve. We have kept an updated blog with all relevant information on Twitter's efforts [covid19.twitter.com](https://twitter.com/covid19), and to date, **over 160 million people have visited the COVID-19 curated page, over two billion times.**

Below is an overview of the measures we have taken to protect the health of the public conversation while ensuring we are a collaborative and open partner in endeavours to address the challenging and changing online and offline issues society is facing.



January and February data and updates

- Data on global violations of Twitter's [COVID-19 misleading information policy](#):
 - In January, we **suspended 2,153 accounts, and removed 3,397 pieces of content**
 - In February, we **suspended 336 accounts, and removed 1,828 pieces of content**
- Data on violations of Twitter's [COVID-19 advertising policy](#):

From 1 to 31 January 2022, **34 Promoted Tweets violated** our COVID-19 advertising policy. They were halted and removed from promotion according to our policy.

 - We reviewed all 34 ads and did not find any which contained misinformation.
 - We estimate that about **94% of the violating content during that time was detected by our automated systems**, and approximately 6% was brought into human review and rejected for policy violations.

From 1 to 28 February 2022, **37 Promoted Tweets** violated our COVID-19 advertising policy. They were halted and removed from promotion according to our policy.

 - We reviewed all 37 ads and did not find any which contained misinformation.
 - We estimate that about **89% of the violating content during that time was detected by our automated systems**, and approximately 11% was brought into human review and rejected for policy violations.

Out of altogether 71 promoted Tweets that were halted in January and February, the vast majority targeted several EU countries among other countries - or targeted globally. None of them targeted EU countries exclusively.

January updates

- We've **expanded our misinformation reporting pilot to three new countries, including Spain**.

We also shared key lessons so far, including that [the vast majority of content we take action on for misinformation is identified proactively](#) — either through automation (accounting for more than 50% of all enforcements) or proactive monitoring. On average, only about 10% of misinfo reports were actionable — compared to 20-30% for other policy areas.

To make review of individual reports efficient, we need to build technology to filter and prioritise reports. That's why we're continuing to experiment (instead of rolling out reporting to everyone); we need more research and training data, especially outside of English. We will continue sharing updates.
- We announced investing in **privacy-enhancing technologies (PETs) to pioneer new methods of public accountability** and access to data in a manner that respects and [protects the privacy of the people who use our service](#).
- Continuing to build on our commitment to **sharing meaningful insights of our work**, we've published [the latest update to our Transparency Center](#) with relevant data from 1 January 2021 to 30 June 2021.

One of our key findings was on enforcement: Twitter required account holders to remove 4.7M Tweets that violated the [Twitter Rules](#). Of the Tweets removed, 68% received fewer than 100 impressions prior to removal, with an additional 24% receiving between 100 and 1,000 impressions. In total, impressions on **these violative Tweets accounted for less than 0.1% of all impressions for all Tweets** during that time period.



February updates

- We've launched a **new Help Centre page** that consolidates [Twitter's policies and efforts to address misleading information](#).
- We helped to make the conversation with **government and state-affiliated media accounts** on Twitter [more transparent by expanding state-affiliated labels](#) to more countries, including Germany, Italy, Spain and Serbia and to a majority of countries that Twitter has attributed [state-linked information operations](#) to. We'll also apply labels to the personal accounts of heads of state for these countries.
- With many looking for **credible information due to the conflict in Ukraine**, we're now [adding labels on Tweets linking to state media](#) and reducing the content's visibility.
- We have detailed all the work we've been doing in relation to Ukraine under [this blog post](#).

This report contains information on policies, products, philanthropic activities and actions undertaken from 1 January to 28 February 2022, and complements previous reports submitted in [July](#), [August](#), [September](#), [October](#), [November](#) and [December](#) 2020 as well as in [January](#), [February](#), [March](#), [April](#), [May](#), [June](#), [July and August](#), [September and October](#), [November and December 2021](#).

Helping people find reliable information

We are committed to making sure that when people come to Twitter to see what's happening, they are able to easily find reliable information on both Covid-19 and vaccines.

- As mentioned above, our new [Help Center page](#) was launched in February consolidating Twitter's efforts on misleading information.
- In the EU, COVID-19 search prompts are active in: Austria, Belgium, Cyprus, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Netherlands, Poland, Portugal, Spain, Sweden.
- Vaccine-related prompts have been implemented in Denmark, Spain, Ireland, Italy and Belgium.
- In over 30 countries, we launched '[Twitter Events Pages](#)' that bring together **the latest Tweets from a number of authoritative and trustworthy government, media and civil society sources in local languages**. We regularly update these pages to ensure that people are met with credible information on Twitter.
- We are regularly updating [Twitter topics](#) related to COVID-19 including COVID-19 news, COVID-19 government and public officials, COVID-19 experts.
- **Every week, we serve around 300 million impressions** of content in the [COVID-19 explore tab](#) to over 14 million customers. It now includes three features to further help people find accurate and reliable information around COVID-19: PSA Carousel, Topics and Lists.
- Twitter Lists are being used by the [World Health Organization](#) and the [European Centre for Disease Prevention and Control](#) to organise and prioritise Tweets.

Useful links

- [Our approach around the conversation surrounding COVID-19 on Twitter](#)
- [Our approach to misleading information on Covid-19](#)
- [Updates to our work on COVID-19 vaccine misinformation](#)
- [COVID-19 Misinformation in our Transparency Center](#)



- [Twitter Transparency Report](#)

COVID-19 Guidance Enforcement

Since introducing our [COVID-19 guidance](#) in 2020, we have challenged 11.7 million accounts, suspended 6,599 accounts and removed over 77,287 pieces of content worldwide.

Overview in numbers: Violations of our Covid-19 misleading information policy

2021	January	February	March	April	May	June
Unique accounts suspended (globally)		48	149	260	185	156
Pieces of content removed (globally)		6822	5371	5091	5147	5117
2021	July	August	September	October	November	December
Unique accounts suspended (globally)	215	229	254	819	431	666
Pieces of content removed (globally)	6602	5579	4544	3574	4129	4559

2022	January	February
Unique accounts suspended (globally)	2153	336
Pieces of content removed (globally)	3397	1828



In the month of January, **we suspended 2,153 accounts and removed 3,397 pieces of content globally**, whereas in the month of February, **we suspended 336 accounts, and removed 1,828 pieces of content** globally for violation of our [COVID-19 misleading information policy](#). As the conflict in Ukraine has escalated, it has become a key priority for our internal teams, shifting attention away from COVID-19 to current events in Ukraine, as reflected in the numbers above.

Advertising on COVID-19

Twitter has a strict advertising policy on COVID-19, the details of which can be found [here](#).

Overview in numbers: Violations of our Covid-19 advertising policy

2021	January	February	March	April	May	June
Number of promoted Tweets that violated Twitter's COVID-19 policy	864	977	945	869	745	729
2021	July	August	September	October	November	December
Number of promoted Tweets that violated Twitter's COVID-19 policy	338		84		64	

2022	January	February
Number of promoted tweets that violated Twitter's COVID-19 policy	34	37

From 1 January to 28 February 2022, **altogether 71 Promoted Tweets violated our COVID-19 advertising policy**. They were halted and removed from promotion according to [our policy](#).

Political content advertising on Twitter

In addition to the above highlighted restrictions, consistent with our belief that political reach should be earned and not bought, [Twitter prohibits the promotion of political content](#).

Useful links

- [Advertising during Covid-19](#)
- [Twitter Ads Policies](#)
- [Twitter Ads Policy update log](#)
- [Political content FAQs](#)
- [Brand communication in times of crisis](#)



Research and data access

Twitter is the largest source of real-time social media data, and we make this data available to the public for free through our [public API](#). No other major service does this. You can [find out more here](#).

To further support our ongoing efforts to protect the public conversation, and help people find authoritative health information around COVID-19, Twitter released a [specific COVID-19 API endpoint](#) into Twitter Developer Labs to enable approved developers and researchers to study the public conversation about COVID-19 in real-time.

January and February data and updates

Twitter [Machine Learning Ethics, Transparency and Accountability \(META\)](#) team is a dedicated group of engineers, researchers, and data scientists collaborating across the company to assess downstream or current unintentional harms in the algorithms we use and to help Twitter prioritize which issues to tackle first. Its recent **update focuses on algorithmic inequalities**.

- [Through research into algorithmic inequalities](#) that may exist on Twitter, [we have published an update](#) that our algorithms prefer people with high follower counts - in fact, that **the top 1% generate 80% of all impressions**. Inequality metrics as this give us a different tool to understand big picture disparities. The teams utilise this data to identify ways to show a diversity of content and people get the same opportunities to be heard, including developing alternatives on how we select content so that we aren't repeatedly preferring the 1% of users, or creating feedback loops that exclude others.
- Previous highlighted work from the teams include [algorithmic amplification of political content](#), [choice and transparency on image cropping](#).
- Our Product and Engineering teams are in the early stages of **exploring the addition of Tweet & Profile labels** (such as misinformation labels, government account labels, etc) to the Twitter API v2. To make sure we're incorporating the needs and use cases, [we are collecting inputs](#) on how to build with these labels in the v2 API. This is an early exploratory stage for us to understand what Tweet / Profile labels would be valuable to developers and researchers.
- In February, **Yale scientist published how social media led to long COVID discovery**. [Akiko Iwasaki's active Twitter use might have led her to the treatment for long COVID](#) after finding a Facebook poll made by individuals with the condition. Iwasaki highlighted the value of Twitter's role in science, allowing a direct line of communication between scientists and patients.
- With the **Academic Research product track**, **qualified researchers have access to all v2 endpoints released to date**, as well as new features and [guides](#). The release of the Academic Research product track is just a starting point. Further information in this [blogpost](#) and in this [thread](#).
- Over 100 researchers and developer teams, representing 92 different academic institutions and universities around the world were granted access to the COVID-19 stream.



- More than half of those approved for this stream are **focused on studying disinformation and misinformation around the facts of coronavirus**, see examples in [previous reports](#) under “How researchers studied COVID-19 on Twitter”.
- Since the Twitter API was introduced, academic researchers have used data from the public conversation to study topics as diverse as the conversation on Twitter itself, including [attitudes and perceptions about COVID-19](#) and [efforts to promote healthy conversation online](#).
- We've been working closely with developers to ensure their services are able to have the widest impact, and reach the most people while operating in compliance with Twitter's developer policies - see COVID-19 related examples in [previous reports](#).

Useful links

- [Academic Research Access](#)
- [Updated our Twitter Developer Policy](#)
- [COVID-19 stream endpoint](#)
- [Apply for access to the Twitter API](#)