

# COVID-19 misleading information policy



1. [Help Center](https://help.twitter.com/) ^ (https://help.twitter.com/)
2. [Platform integrity and authenticity](https://help.twitter.com/en/rules-and-policies#platform-integrity-and-authenticity) ^ (https://help.twitter.com/en/rules-and-policies#platform-integrity-and-authenticity)

## COVID-19 misleading information policy

### Overview December 2021

**You may not use Twitter's services to share false or misleading information about COVID-19 which may lead to harm.**

Even as scientific understanding of the COVID-19 pandemic continues to develop, we've observed the emergence of persistent conspiracy theories, alarmist rhetoric unfounded in research or credible reporting, and a wide range of false narratives and unsubstantiated rumors, which left uncontextualized can prevent the public from making informed decisions regarding their health, and puts individuals, families and communities at risk.

Content that is demonstrably false or misleading and may lead to significant risk of harm (such as increased exposure to the virus, or adverse effects on public health systems) may not be shared on Twitter. This includes sharing content that may mislead people about the nature of the COVID-19 virus; the efficacy and/or safety of preventative measures, treatments, or other precautions to mitigate or treat the disease; official regulations, restrictions, or exemptions pertaining to health advisories; or the prevalence of the virus or risk of infection or death associated with COVID-19.

# What is in violation of this policy?

In order for content related to COVID-19 to be considered violative under this policy, it must:

- advance a claim of fact, expressed in definitive terms;
- be demonstrably false or misleading, based on widely available, authoritative sources; and
- be likely to impact public safety or cause serious harm

## **Tweet Removal**

We may require customers to delete Tweets that are found to violate this policy and are severely harmful. We may also temporarily lock you out of your account before you can Tweet or share information again. These tweets will accrue **2 strikes** in accordance with our strike policy stated below. We will require the deletion of Tweets that contain, for example:

1. False claims about COVID-19 that invoke a deliberate conspiracy by malicious and/or powerful forces, such as:
  - The pandemic is a hoax, or part of a deliberate attempt at population control, or that 5G wireless technology is causing COVID-19.
  - COVID-19 is not a real disease.
  - Immunizations are part of a global surveillance, population control or depopulation effort.
  - Vaccines (in general) are dangerous and the adverse effects that have been covered up by governments/the medical industry.
  - Vulnerable groups (such as pregnant women, the elderly, or children) are being experimented on.
  - That COVID-19 vaccines are causing magnetic reactions in individuals who have been vaccinated.
  - That vaccines approved by health agencies (such as Pfizer's Comirnaty vaccine in the United States) did not actually receive full approval/authorization, and therefore that the vaccines are untested, “experimental” or somehow unsafe.
2. Claims that specific groups or people (or other demographically-identifiable identity) are more or less prone to be infected or to develop adverse symptoms on the basis of their membership in that group;
3. False or misleading claims about potentially harmful and unapproved treatments or preventative measures, for example, that chlorine dioxide or Povidone-iodine can be used as a prophylactic or in the treatment of COVID-19.
4. False or misleading information about official regulations, restrictions, or exemptions pertaining to health advisories.
5. Any efforts to promote, advertise, facilitate the sale of, or provide instructions on how to create fraudulent vaccination cards (or other digital records) or “exemption cards.”
6. False information about widely accepted testing methodologies, such as that PCR tests are unable to detect the virus.
7. False claims that suggest that vaccines contain deadly and severely harmful ingredients.
8. False affiliation - Accounts which misrepresent their affiliation, or share content that falsely represents its affiliation to a medical practitioner, public health official or agency, research institution, or that falsely suggests expertise on COVID-19 issues.

## Tweet Label

When Tweets include misleading information about COVID-19, we may place a label on those Tweets that includes corrective information about that claim. In cases where we determine there is potential for harm associated with the misleading claim, we will disable the ability for others to Retweet, Quote Tweet, or engage in other ways to prevent the spread of the misleading information. These tweets will accrue **1 strike** in accordance with our strike policy stated below.

In some cases we may also add labels to provide context in situations where authoritative (scientific or otherwise) opinion might change or is changing over time, in situations where local context is important, or when the potential for harm is less direct or imminent. We may also apply labels on Tweets linking to content from a third-party website that would otherwise violate our policies if the content were posted directly on Twitter. Tweets with labels that meet this criteria will receive a label that provides credible and authoritative information, but **will not accrue a strike** in accordance with our strike policy stated below.

We may apply labels to Tweets that contain, for example:

1. False or misleading information about preventative measures one can take to avoid infection, such as claims that face masks cause hypoxia or bacteria pneumonia, or do not work to reduce transmission or to protect against COVID-19.
2. False or misleading information suggesting that unapproved treatments can be curative of COVID-19.
3. False or misleading information regarding the safety or science behind approved or authorized COVID-19 vaccines, such as:
  - The vaccines will cause you to be sick, spread the virus, or would be more harmful than getting COVID-19.
  - Tweets that incite fear or misrepresent the ingredients or contents of COVID-19 vaccines.
  - Tweets that mischaracterize the nature and science behind mRNA vaccines, and how they work.
  - Tweets that claim vaccines alter genetic code.
  - Tweets that misrepresent or misuse official reporting tools/statistics.
  - False or misleading claims that people who have received the vaccine can spread or shed the vaccine (or symptoms, or immunity) to unvaccinated people.
4. False or misleading information that misrepresent the protective effect of vaccines, to make claims contrary to health authorities. Claims that misrepresent research or statistical findings pertaining to the severity of the disease, prevalence of the virus, or effectiveness of widely accepted preventative measures, treatments, or vaccines.

When a label is applied to a Tweet, this typically entails:

- Presenting a warning message on the Tweet
- Showing an additional prompt to warn people before sharing or liking the Tweet;
- Reducing the visibility of the Tweet on Twitter and/or preventing it from being recommended;
- Turning off likes, replies, and Retweets; and/or
- Providing a link to additional explanations or clarifications, such as in a curated landing page or relevant Twitter policies.

# What is not a violation of this policy?

We seek to protect robust, public debate about the response to COVID-19, and recognize that the state of scientific knowledge about certain aspects of the pandemic and public response to it (including the development of vaccines) is still evolving. In the absence of other policy violations, the following are generally not in violation of this policy:

- **Strong commentary, opinions, and/or satire**, provided these do not contain false or misleading assertions of fact.
- **Campaigns against official advisories or recommendations**. People are entitled to organize and campaign around matters that are important to them, so long as they're not advancing false and harmful misinformation in the process.
- **Counterspeech**. We allow for direct responses to misleading information which seek to undermine its impact by correcting the record, amplify credible information, and educate the wider community about the prevalence and dynamics of misleading information.
- **Personal anecdotes or first-person accounts**. We do not enforce the COVID-19 misinformation policy on reports or first-person accounts of side-effects, injury, or harm alleged to have been caused by COVID-19 vaccines, or other matters pertaining to COVID-19.
- **Public debate about the advancement of COVID-19 science and research**, including debate about research related to COVID-19, such as the effectiveness of treatments and mitigation measures, so long as the claims don't intentionally misrepresent research findings.

# Who can report violations of this policy?

While we have recently launched a customer report feature (<https://twitter.com/TwitterSafety/status/1427706890113495046>) in some countries, we primarily enforce this policy in close coordination with trusted partners, including public health authorities, NGOs and governments, and continue to use and consult with information from those sources when reviewing content.

We also leverage proactive detection using a combination of keyword heuristics and machine learning models to identify harmful forms of COVID-19 related misleading information.

# What happens if you violate this policy?

## Strike system

The consequences for violating our COVID-19 misleading information policy depend on the severity and type of the violation and the account's history of previous violations. In instances where accounts repeatedly violate this policy, we will use a strike system to determine if further enforcement actions should be applied. We believe this system further helps to reduce the spread of potentially harmful and misleading information on Twitter, particularly for high-severity violations of our rules.

Repeated violations of this policy are enforced against on the basis of the number of strikes an account has accrued for violations of this policy:

- 1 strike: No account-level action
- 2 strikes: 12-hour account lock
- 3 strikes: 12-hour account lock
- 4 strikes: 7-day account lock
- 5 or more strikes: Permanent suspension

If you believe that your account was locked or suspended in error, you can [submit an appeal](https://help.twitter.com/forms/general?subtopic=suspended). (<https://help.twitter.com/forms/general?subtopic=suspended>)

## Reducing Visibility of Content Related to COVID-19 Misinformation

We may reduce the visibility of tweets or accounts that we believe with high confidence to be in violation of the COVID-19 misinformation policy. Limitations on the visibility of Tweets and/or accounts resulting from this framework expire automatically after a limited period of time, but may be re-applied (either manually or automatically) if we determine that the accounts have continued to violate the policy. We may reduce the visibility by:

- Making Tweets and Retweets from those accounts ineligible to appear in certain parts of the Twitter product (such as top Search results)
- Displaying Replies from the account in a lower position in conversations
- Excluding Tweets by the account and/or the actual account in email or in-product recommendations

## Permanent Suspension

We may immediately permanently suspend accounts that represent the following violative behaviors:

**False affiliation:** If the account is determined to misrepresent their affiliation, or share content that falsely represents its affiliation as a medical practitioner, public health official or agency, research institution, or that falsely suggests expertise on COVID-19 issues.

**Repeated Violations:** If we determine that the account repeatedly violates the COVID-19 misinformation policy over a 30-day time period, or if we have determined that the account has been set up for the expressed purpose of Tweeting false or misleading information about COVID-19.

## Additional resources

Learn more about our work to fight misleading information about COVID-19 [here](https://blog.twitter.com/en_us/topics/company/2020/covid-19.html) ([https://blog.twitter.com/en\\_us/topics/company/2020/covid-19.html](https://blog.twitter.com/en_us/topics/company/2020/covid-19.html)), and our expanded approach to COVID-19 vaccine misleading information [here](https://blog.twitter.com/en_us/topics/company/2020/covid19-vaccine.html) ([https://blog.twitter.com/en\\_us/topics/company/2020/covid19-vaccine.html](https://blog.twitter.com/en_us/topics/company/2020/covid19-vaccine.html)).

Learn more about [our range of enforcement options](#) and [our approach to policy development and enforcement](#).

## Share this article

