

## Written evidence submitted by Twitter

**15 December 2020**

Dear Chair,

Thank you for providing the opportunity to provide a written submission regarding our work in response to the rollout of the COVID-19 vaccine; and our wider work on COVID-19 misinformation more broadly.

Please see attached:

- **Our COVID-19 vaccine work**
- **Our wider COVID-19 work in the UK, updated from our submission earlier this year**

Yours sincerely,

Katy Minshall  
Head of UK Public Policy

# COVID-19 vaccine work

## Background

Twitter is a place to freely exchange information and ideas. Our mission is to serve the public conversation and ensure Twitter is a place where people can express themselves freely and safely, and receive authoritative and credible information. Our goal is to ensure people have the context necessary to make informed decisions about what they see on Twitter.

On July 14, [we updated our policy enforcement guidance](#) around misleading information and when we would take action against Tweets. Our primary goal with addressing misleading information about COVID-19 has not changed. We will continue to remove demonstrably false or potentially misleading content that has the highest risk of causing harm.

We also may label Tweets to provide additional context in situations where the risks of harm associated with a Tweet are less severe but where people may still be confused or misled. Our aim is to make it easier to find context and make informed decisions about what people see on Twitter. In line with our existing enforcement approach, Tweets that are labeled under this expanded guidance will have reduced visibility across the service.

Since 2018, we've had in place a prompt that directs individuals to a credible public health resource when someone searches for terms associated with vaccines. In May 2019, we launched an additional search prompt in the UK that serves people with [credible vaccine information from the NHS](#) - the search prompt is available in 37 countries and 15 languages worldwide.

## Policy

As the world continues to fight the COVID-19 pandemic and prepare for the global distribution of vaccines, people continue to turn to Twitter to discuss what's happening and find the latest authoritative health information.

Earlier this year, we [shared our approach](#) around the conversation surrounding COVID-19 on Twitter. We've remained focused on helping people on Twitter find [credible health information](#), verified public health experts, and introduced a [policy](#) (described above) on misleading information specific to COVID-19, just to name a few, and all done in close communication and collaboration with key health agencies and NGOs.

Under our current policy, we [already take enforcement action](#) on Tweets that include false or misleading information about:

- The nature of the virus, such as how it spreads within communities;
- 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; and
- The prevalence or risk of infection or death.

With the public conversation turning towards vaccines and what's next, we're updating our approach to address specifically misleading information around COVID-19 vaccinations.

## Partnerships

Over 2020, we have engaged with governments, civil society organisations and experts from health and research institutions to prepare to counter COVID-19 vaccine misinformation. In the UK, this has included:

- Gathering perspectives and feedback from researchers on COVID-19 and vaccine misinformation, including from King's College London, Oxford University and the Vaccines Confidence Project;
- Meeting with vaccine literacy experts to inform our approach;
- Participating in the Government's Counter-Disinformation Policy Forums.

We're committed to protecting the health of the public conversation on Twitter - ensuring individuals can find information from authoritative sources is a key part of that mission. We welcome opportunities to collaborate on and support public health campaigns with key information about the COVID-19 vaccine.

## Product

To support people in the UK searching for information about the COVID-19 vaccine on our service, it is key to ensure they are met with credible, authoritative content. From 1 March - 20 November we saw over 100k searches, which increased after the Pfizer announcement in early November. We have updated our search prompt to direct people to information to official NHS guidance about the COVID-19 vaccine.



Supporting critical thinking and digital literacy is crucial at this time. The #ThinkBeforeSharing prompt, launched earlier this year, notifies people to open articles they have not opened before sharing them. This has prompted [more reading and more informed Tweeting](#): people open

articles 40% more often after seeing the prompt, and people opening articles before Retweeting has increased by 33%.

## Research

As the only major service to make public conversations available for the purposes of study - including a specific API for conversations around COVID-19 - we welcome research in these key areas and continue to call for future research to be peer-reviewed and undertaken by experts in the field. An excellent example of how our specific COVID conversation API has been used, can be [found here](#), where a leading medical centre used data from the public conversation and the COVID-19 stream endpoint to rapidly develop an in-depth map of COVID-19 attitudes and perceptions in the US. We continue to call for research to be peer reviewed and data to be shared publicly. As we outlined in our [blog](#) in May 2020 "research is rarely peer-reviewed and often does not hold up to scrutiny, further confusing the public's understanding of what's really happening."

## COVID-19 misinformation strategy in the UK

Below is a summary of work undertaken in the UK. It is by no means comprehensive, and so please do let us know if the Committee has any further questions. We share information about our global strategy on an ongoing basis via our [COVID-19 blog](#) and through our partnerships in the UK with government and civil society.

The power of a uniquely open service during a public health emergency is clear. The speed and borderless nature of Twitter presents an extraordinary opportunity to get the word out and ensure people have access to the latest information from expert sources around the world. To support that mission, our global Trust & Safety team is continuing its zero-tolerance approach to platform manipulation and any other attempts to abuse our service at this critical juncture.

## Policy

- On March 18, we shared the [1-point criteria](#) our team would be using to assess whether Tweets were considered harmful COVID-19 misinformation and against our rules.
- We broadened our guidance in April on unverified claims that incite people to engage in harmful activity, could lead to the destruction or damage of critical 5G infrastructure, or could lead to widespread panic, social unrest, or large-scale disorder.
- Since introducing our COVID-19 guidance, we have challenged 8.5 million accounts.
- For coordinated harmful activity more broadly, we shared further information in September about how we enforce our rules on groups coordinating to cause harm. For example, impressions on QAnon content dropped more than 50% after we began removing Tweets from trends and recommendations.
- We continue to recommend on our website that people follow @WHO and their local health ministry; and seek out the authoritative health information and ignore the noise.

- We ask people to report to us immediately if they see something suspicious or abusive. Most importantly, we advise people to think before they Tweet.

## Partnerships

- There are now over 40 UK government partners and health authorities onboarded to our Partner Support Portal, which enables trusted partners to expedite a wide range of issues to a dedicated team within Twitter.
- In April, we provided the government with pro bono advertising space and support to promote key #StayHomeSaveLives public health messaging. Not only did this have significant reach, but we also saw people actively participating - committing to #StayHomeSaveLives and the key behaviours required to keep us all safe. Indeed, #StayHome #StaySafe and #StayHomeSaveLives made up 3 of the [top 10 most weeted hashtags in the UK](#) in 2020. People also took to Twitter to show their appreciation for healthcare workers - @NHSuk's Tweet below made the top five most Liked Tweets in 2020:



That was emotional 

8:34 PM · Mar 26, 2020 · Twitter Web App

- We also continue to provide pro bono advertising space and support to a number of charities, including around fact-checking and digital literacy.
- For educators and parents, we recommend they consult our [media literacy guide](#) built in partnership with @UNESCO.

## Product

- Launched six days before the official designation of the virus in January, we partnered with the Department of Health and Social Care to launch a dedicated COVID-19 prompt feature at the top of Search results. This ensures that when you come to the service for information about COVID-19, you are directed to the NHS website where you are met with credible, authoritative content. In May, we developed this further in partnership with the Department for Digital, Culture, Media and Sport - such that when people search for 5G information on Twitter, they are served with a prompt encouraging them to 'Know the facts'; stating that the UK Government has seen no link between 5G and COVID-19 and directing to a new Gov.uk resource.



- Twitter Moments are a curated series of Tweets to tell a story. Through Twitter Moments, we have curated longer-form content that helps tell the full story of what’s happening around COVID-19 globally. We have ensured a live, up-to-date Moment, with the latest Tweets from the key UK government and health accounts about COVID-19, is available at the top of the home timeline for all people on Twitter in the UK. We shared earlier this year that we had seen a 45% increase in usage of these pages globally.



- We continue to highlight credible, timely information about COVID-19 through our Explore tab. Some recent examples (including reach) are:
  - [‘Wearing a face mask is safe, international health authorities and experts say’](#) received 1.14 million impressions.
  - [‘Experts advise that masks are effective in reducing COVID-19 transmission’](#) received 776k impressions.

## Research

- Twitter is the only major service that makes public conversations available for study. Tens of thousands of researchers [have used our public API](#) (technical interface to allow access to our public data) over the past decade.
- To provide transparency on the COVID-19 conversation on Twitter and further support the research community, on April 30 we launched a new program that offers free access to a dedicated [COVID-19 API endpoint](#). In practice, this allows approved developers and researchers to access public conversations about COVID-19 across languages, resulting in a data set that includes tens of millions of Tweets. The data can be used to research a range of topics related to the coronavirus pandemic, including areas like the spread of the disease, the spread of misinformation, crisis management within communities and

more. Statistics for COVID-API endpoint applications:

- 35+ countries are represented by all applicants;
  - 70% are academic researchers;
  - 18% are commercial & non-commercial orgs;
  - 12% are independent developers and/or researchers;
  - 97% of all academic researchers who apply are approved.
- Twitter's public archive of state-backed information operations is the largest of its kind in the industry. We continue to publicly disclose and make datasets [available](#) for research. First launched in October 2018, and continuously updated, the archive has been accessed by thousands of researchers from around the world, who in turn have conducted independent, third-party investigations of their own. In total, there are now over 35 datasets.