
ONLINE PLATFORMS' RESPONSES TO COVID-19 MIS—AND DISINFORMATION

EU DisinfoLab

1st April 2020



Platforms responses to misinformation on Covid-19

This blogpost is part of a series on Covid19. Please have a look at our dedicated page and subscribe to our mailing list of updates.

In the context of the Covid19 crisis, misinformation on the epidemic have been flooding the internet. The online platforms have been quick to react and demonstrated their ability to take action to limit the spread of such content. Major tech companies even published a common industry statement endorsed by Facebook, Google, LinkedIn, Microsoft, Reddit, Twitter and YouTube, in a move to jointly combat fraud and misinformation on their services.¹

We analysed the blog posts, statements and updates in policies from the platforms to combine a comprehensive overview of the measures undertaken. As the situation is evolving fast, we will update this page regularly. In case you would like to flag any update or inconsistencies, please feel free to reach out to info@disinfo.eu.

From our analysis, we were able to draw and highlight several trends in platforms policies regarding coronavirus disinformation. Overall, it is remarkable how the tech companies have been working closely with public health authorities and governments to face the crisis situation. Yet, this crucial dialogue has just proven the prevalence of social networks on today's information society.

At a time when more than 2 billion people are confined at home, this crisis revealed our reliance on these services to engage and connect with our peers across the world. With such a role lies also an immense responsibility over the public discourse, which should be reflected in any future regulation.

1) Surfacing and prioritizing content from authoritative sources

Most of the platforms we analysed are redirecting users to authoritative sources of content. Searches on Coronavirus surface content from public health authorities, and dedicated information sections and banners are used to highlight information from health agencies. In addition to that, the World Health Organisation partnered with several services to created dedicated channels like on WhatsApp, and Snapchat.

2) Closer cooperation with fact-checkers and health authorities for flagging this informative content to proactively remove disinformation.

In addition to their partnership with fact-checkers, the platforms relied on public health authorities to flag disinformation related to the epidemic. Based on this information, the companies committed to proactively remove content promoting false cures, denial of public health authorities' recommendations and proven scientific facts.

¹

https://twitter.com/fbnewsroom/status/1239703497479614466?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwtterm%5E1239703497479614466&ref_url=https%3A%2F%2Ftechcrunch.com%2F2020%2F03%2F16%2Ffacebook-reddit-google-linkedin-microsoft-twitter-and-youtube-issue-joint-statement-on-misinformation%2F

3) Automated content moderation

Social distancing for preventing the spread of the virus has forced thousands of moderators working from home. To face this, platforms chose to increase the use of AI filters to moderate content, thus leading to a large number of bugs and false positives.

4) Offering free advertising to authorities

The World Health Organisation has been granted free advertising credit by Facebook and was offered help by Google to run ads.

5) Scams

The platforms also had to ban advertising trying to benefit from the crisis. Nonetheless, many scams passed the filters, leading law enforcement and consumer authorities to warn consumers and call on the marketplaces to react.

Information available on 25/03 (last update on 01/04)

Facebook and Instagram²

Prioritizing reliable information

Searches on information and hashtags related to Coronavirus on Facebook and Instagram surfaces educational pop-ups and redirects to information from the World Health Organisation and local health authorities. The WHO has also been granted free advertising credit on the platforms.

Preventing misinformation

In this context, the company is implementing its existing policies. Facebook relies on the work of third-party fact-checkers to debunk false claims related to the coronavirus, limiting the spread of content verified as false. People trying to share an article verified as false are also notified it has been fact-checked.

In addition, false claims flagged by public health authorities are removed. This includes “claims related to false cures or prevention methods — like drinking bleach cures the coronavirus — or claims that create confusion about health resources that are available”³. Also, the company committed to removing hashtags used to spread disinformation on Instagram.

² Facebook Coronavirus information center: <https://about.fb.com/news/2020/03/coronavirus/#joint-statement>

³ <https://about.fb.com/news/2020/03/coronavirus/#joint-statement>

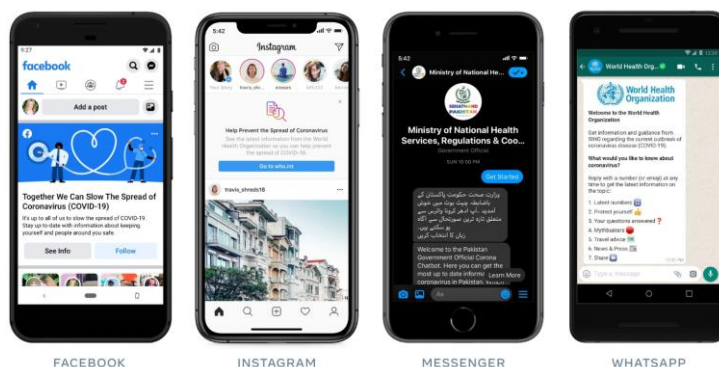
In the meantime, Facebook with the urgency and staff working from home, Facebook increased moderation heavily and relied on automated filters, inevitably leading to more errors, as the company warned in a blogpost.⁴ As anticipated, this situation happened with content from authoritative sources such as BBC being wrongly removed or prevented from upload.⁵

Enforcement

On 30/03, Facebook removed a video⁶ of Brazilian President Jair Bolsonaro endorsing hydroxychloroquine and calling for an end to social-distancing efforts, from its main social network and Instagram. According to the company, the video went against its community standards banning disinformation that could lead to “physical harms”.

WhatsApp

With families, colleagues and relatives increasingly communicating digitally, WhatsApp has been one of the main platforms for conspiracy theories, unverified advice and disinformation going viral on the epidemic. The platform has set up a “WhatsApp Coronavirus Information Hub” in partnership with WHO, UNICEF and UNDP⁷ and suggests usages of the platform for community help during the crisis. In addition, the messaging platform worked with the World Health Organisation to open a dedicated channel with a bot to automatically answer questions and provide reliable information to users.⁸ In the meantime, Facebook’s company has announced a 1Million support grant for fact-checkers (see section on fact-checking responses). According to a WhatsApp’s spokesperson, the company is currently working on AI to remove spam accounts.⁹



⁴ <https://about.fb.com/news/2020/03/coronavirus/#joint-statement>

⁵ <https://twitter.com/guyro/status/1240063821974138881>

⁶ <https://www.bbc.com/portuguese/brasil-52101240>

⁷ <https://www.whatsapp.com/coronavirus>

⁸ <https://thenextweb.com/apps/2020/03/20/world-health-organizations-whatsapp-bot-texts-you-coronavirus-facts/>

⁹ <https://www.irishtimes.com/news/ireland/irish-news/no-plans-to-combat-sharing-of-coronavirus-misinformation-on-whatsapp-1.4204207>

Google and YouTube¹⁰

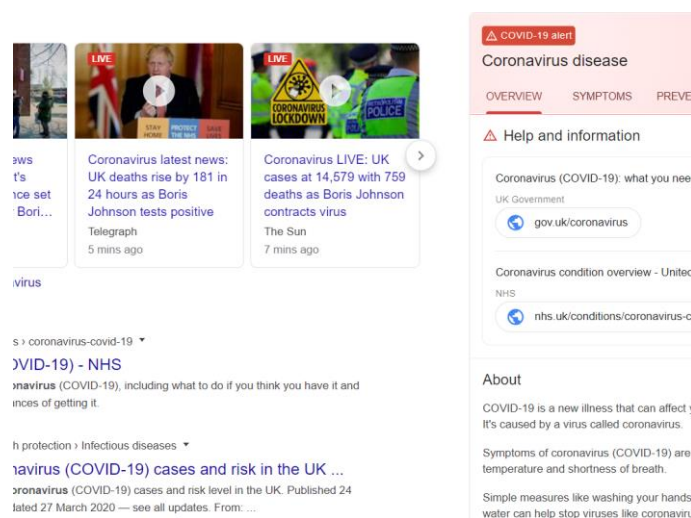
Promoting reliable information

Google partnered with the US government in developing a website dedicated to Covid-19 education and information. Google homepage also promotes measures to limit the spread of the virus such as washing hands. Google search also includes an “SOS Alert” promoting information from official sources on the symptoms and prevention of the disease.

On YouTube, videos from public health agencies appear on the homepage. The video-sharing platform also highlights content from authoritative sources when people search for information on coronavirus, as well as information panels to add additional context.

Preventing misinformation

Google is actively removing disinformation from its services, including YouTube and Google Maps. On YouTube, it removes videos that promote medically unproven cures. In addition, it prevents ads that try to capitalize on the situation.¹¹



Twitter¹²

Promoting reliable information

Twitter has set up a COVID-19 event page with latest from trusted sources to appear on top of the

¹⁰ <https://blog.google/inside-google/company-announcements/covid-19-how-were-continuing-to-help/>

¹¹ <https://support.google.com/adspolicy/answer/6015406?hl=en>

¹² https://blog.twitter.com/en_us/topics/company/2020/An-update-on-our-continuity-strategy-during-COVID-19.html

timeline. The company has also worked on verifying accounts with email addresses from health institutions to provide reliable information on the topic.¹³

Preventing misinformation

Similar to Facebook, Twitter anticipated an increase in automated filtering to prevent the massive spread of coronavirus related information on its platform, assuming it could lead to inevitable errors. In addition, the company broadened the definition of harms on the platform, to include denial of public health authorities recommendations, description of treatment known as ineffective, denial of scientific facts about the transmission of the virus, claims that COVID-19 intends to manipulate people and related conspiracy theories, incitement to actions that could cause widespread panic, or claims that specific group would be more or never susceptible to COVID-19.

Enforcement

Enforcing its policy,¹⁴ Twitter has removed tweets by the presidents of Brazil and Venezuela and former New York City Mayor Rudy Giuliani for violating its ban on content that “goes directly against guidance from authoritative sources of global and local public health information”.

Snapchat¹⁵

Snapchat has used its “Discovery” function to highlight information from partners. This section of the platforms provides vetted content from editors and advertisers. WHO partnered with the platform to create filters and stickers that promote tips and guidance on how to prevent the spread of the virus. In addition, the organisation has a dedicated official account to provide updates and information to the users. For helping teenagers cope with social distancing, Snap has rolled out a dedicated mental health feature.¹⁶

¹³ <https://twitter.com/twittersupport/status/1241155701822476288?s=12>

¹⁴ <https://www.theverge.com/2020/3/30/21199845/twitter-tweets-brazil-venezuela-presidents-covid-19-coronavirus-jair-bolsonaro-maduro>

¹⁵ <https://forbusiness.snapchat.com/blog/coronavirus-how-snapchatters-are-reacting-to-the-news-and-staying-informed>

¹⁶ <https://mashable.com/article/snapchat-snap-coronavirus-here-for-you/?europe=true>



Reddit

The forum has positioned a dedicated subreddit page as a reliable source of information with posts being verified by moderators and contributions from renowned scientists.¹⁷

Telegram¹⁸

The messaging platform which offers secret encrypted conversations and open channels has not announced additional measures regarding Covid19 related disinformation. In its FAQ, the company clarifies that it will process requests to take down illegal public content.

Pinterest

Pinterest has a very strict health misinformation policy, initially adapted to anti-vaxx content.¹⁹ Its community guidelines strictly ban health misinformation from the platform. For this reason, everyone searches for “Coronavirus” on the platform sends to a dedicated page with information from the World Health Organisation. In addition, Pinterest rolled out a new tab called “Today tab” to show users created and trending topics.²⁰ These days, the company presented information related to Coronavirus, in partnership with the World Health Organisation. The company is also preventing ads trying to benefit from the crisis, yet several scams were found on the platform.²¹



¹⁷ <https://www.newsweek.com/coronavirus-reddit-live-updates-news-subreddit-covid-1490536>

¹⁸ <https://telegram.org/faq>

¹⁹ <https://policy.pinterest.com/fr/community-guidelines>

²⁰ <https://www.msn.com/en-us/news/technology/coronavirus-pinterest-launches-new-today-tab-which-will-offer-advice-from-health-agencies/ar-BB11DEju?fullscreen=true#image=1>

²¹ <https://www.buzzfeednews.com/article/craigsilverman/pinterest-is-blocking-coronavirus-searches-and-people-are>

< Coronavirus  

Pins about this topic often violate our Community Guidelines, which prohibit harmful medical misinformation. Because of this, we've limited search results to Pins from internationally-recognized health organizations.

If you're looking for medical advice, please contact a healthcare provider.



