

Facebook response to the European Commission Communication on Covid-19 Disinformation

Report for December 2020

1. Introduction

This report builds off our previous reports to the European Commission, in response to the [Joint Communication](#) for tackling COVID-19 disinformation, and provides an overview of the policies, products, and processes we have deployed to combat COVID-19 misinformation and disinformation across Facebook and Instagram in December 2020.

2. User Engagement with Authoritative Resources and Tools to Raise Awareness

We continue to find new ways to connect people with accurate, reliable and authoritative information. This is a core component of our strategy to combat misinformation because we want to be able to provide our users with the means to decide what to read, trust and share.

We believe that informing people with accurate and authoritative information, as well as more context, is an approach that can be more impactful than the alternative of just removing content. If we simply removed all posts flagged by fact-checkers as false, for example, the content would still be available elsewhere on the internet, other social media platforms, or even discussed around the dinner table. By leaving this content up and surfacing research from fact-checkers or pointing people to reliable information, we're providing people with important information and context.

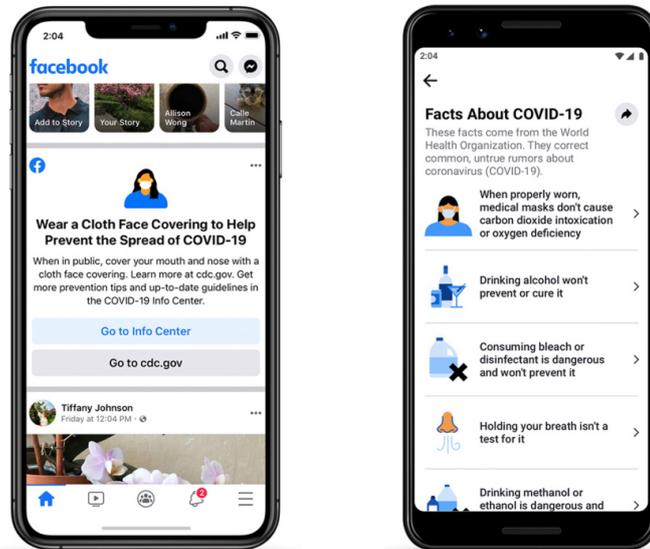
As noted by an international group of [human rights experts](#) (in relation to COVID-19): "it is essential that governments and internet companies address disinformation in the first instance by themselves providing reliable information... Resorting to other measures, such as content take-downs and censorship, may result in limiting access to important information for public health and should only be undertaken where they meet the standards of necessity and proportionality."

During the coronavirus public health crisis, we have been supporting the global public health community's work to keep people safe and informed by connecting them to accurate, reliable, accessible and relevant sources of information about COVID-19. Our COVID-19 Information Center on Facebook provides people with the latest information from health authorities, news, resources, facts, and tips to stay healthy and safe. It is available globally, including all 27 EU member states. More than 130 million people globally, including over 15 million people in the EU, visited the COVID-19 Information Center during the month of December.

Facts About COVID-19

In December, we re-launched a pop-up on Facebook's News Feed to direct users to the [Facts about COVID-19](#) section of our COVID-19 Information Center. The section debunks common myths that have been identified by the World Health Organization, such as drinking bleach will prevent the coronavirus or that holding your breath for 10 seconds without coughing means you don't have

coronavirus. It also features facts about the vaccine testing process, masks, and the vulnerability of young people.



On Instagram, as we do not have a specific Information Centre, we launched a similar pop-up but directed people instead to the websites of the World Health Organization or the Ministry of Health of the different countries for authoritative information.

3. Actions on Misinformation

Our goal is to create a place for expression and give people a voice. Building community and bringing the world closer together depends on people's ability to share diverse views, experiences, ideas and information. Our commitment to expression is paramount, but we recognise that the internet creates new and increased opportunities for abuse. When considering whether to provide more context, allow, reduce distribution, or remove misinformation, we do it in service of one or more of our [Community Standards Values](#): voice, authenticity, safety, privacy, and dignity.

We define misinformation as content that is false or misleading. We enforce on misinformation by looking at content or behaviors that violate our Community Standards or content that may be reviewed by our third-party fact-checking partners. We define disinformation as coordinated efforts to manipulate public debate for a strategic goal.

Applying Community Standards to COVID-19 Content

As people around the world confront this unprecedented public health emergency, we want to make sure that our [Community Standards](#) protect people from harmful content and new types of abuse related to COVID-19. We're working to remove content that has the potential to contribute to offline harm, including through our policies prohibiting the coordination of harm, the sale of medical masks and related goods, hate speech, bullying and harassment, and misinformation that contributes to the risk of imminent violence or physical harm. Oftentimes, misinformation can cut across different types of abuse areas; for example, a racial slur could be coupled with a false claim about a group of people and we'd remove it for violating our hate speech policy. The following are updated metrics for the month of December 2020:

- We removed over 401 thousand pieces of content on Facebook and Instagram globally, including over 10 thousand pieces of content in the EU, for containing misinformation that may lead to imminent physical harm, such as content relating to fake preventative measures or exaggerated cures.
- We removed over 240 thousand pieces of content on Facebook and Instagram globally, including over 11 thousand pieces of content in the EU member states, related to COVID-19 and which violated our medical supply sales standards.

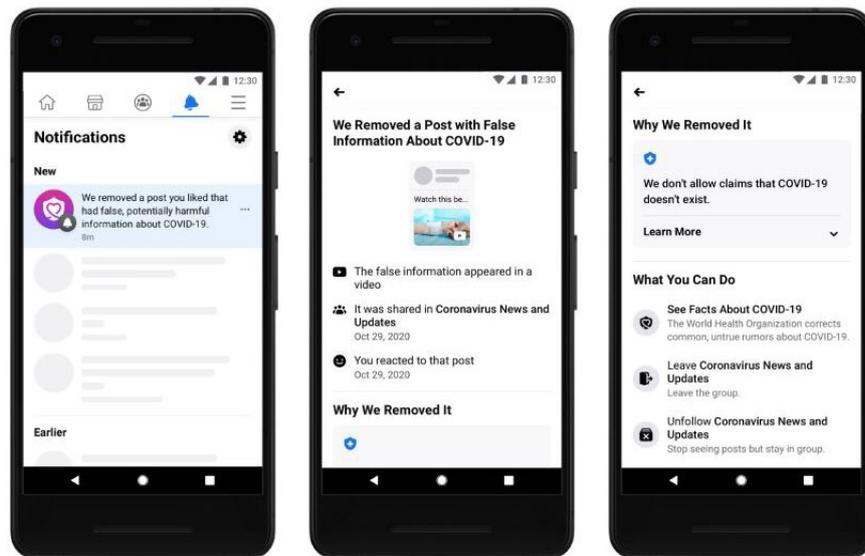
As the situation evolves, we are continuing to look at content on the platform, assess speech trends and engage with experts, and will provide additional policy guidance to our Community Standards when appropriate to keep the members of our community safe during this crisis. These policies, as well as the additional policies listed in our Community Standards apply to content on both Facebook and Instagram, including surfaces such as Groups and Pages.

Keeping People Informed and Limiting Misinformation About COVID-19

In December, we announced that we are updating the messages we send to people who have interacted with misinformation about COVID-19 on Facebook that we've since removed. In April, we started showing these messages in News Feed to people who liked, commented on or reacted to posts with misinformation that we removed for violating our policy. Since then, we've done research to better understand what's most helpful for people, and we've redesigned these as more personalized notifications to more clearly connect people with credible and accurate information about COVID-19.

Now, people will:

- Receive a notification that says we've removed a post they've interacted with for violating our policy against misinformation about COVID-19 that leads to imminent physical harm.
- Once they click on the notification, they will see a thumbnail of the post, and more information about where they saw it and how they engaged with it.
- They will also see why it was false and why we removed it (e.g. the post included the false claim that COVID-19 doesn't exist)
- People will then be able to see more facts about COVID-19 in our [Coronavirus Information Center](#), and take other actions such as unfollowing the Page or Groups that shared this content.



Our Third-Party Fact-Checking Program

For misinformation that does not lead to real world harm, but undermines the authenticity and integrity of our platform, we continue to work with our growing [network of independent third party fact-checking partners](#). We partner with over [80 fact-checking organizations](#) around the world, covering over 60 languages. In the EU and greater Europe, we expanded our program into Romania with AFP. This brings the total number of fact-checkers for the region to 36, covering 26 languages.

Based on the work of our fact-checking partners, we displayed misinformation warning screens associated with COVID-19 related fact-checks on over 19 million pieces of content globally, including over 2.7 million pieces of content in EU member states, in December.

Supporting Media Literacy in Europe

We want to give people the tools to make more informed decisions about the information they choose to consume and share online. In the run up to the Portuguese presidential elections, we launched a media literacy campaign on Facebook and Instagram to help users identify false news. To complement our false news media literacy efforts, we also launched a campaign to help people spot and report hate speech on the platform.

Removing False Claims About COVID-19 Vaccines

Given that COVID-19 vaccines were beginning to roll out around the world in December 2020, we started removing false claims about these vaccines that have been debunked by public health experts on Facebook and Instagram. This is another way that we are applying our policy to remove misinformation about the virus that could lead to imminent physical harm.

This could include false claims about the safety, efficacy, ingredients or side effects of the vaccines. For example, we will remove false claims that COVID-19 vaccines contain microchips, or anything else that isn't on the official vaccine ingredient list. We will also remove conspiracy theories about COVID-19 vaccines that we know today are false: like specific populations are being used without their consent to test the vaccine's safety. Since it's early and facts about COVID-19 vaccines will continue to evolve, we will regularly update the claims we remove based on guidance from public health authorities as they learn more.

We will also continue to help people stay informed about these vaccines by promoting authoritative sources of information through our [COVID-19 Information Center](#).

4. Coordinated Inauthentic Behavior (CIB) and Influence Operations

We know that people looking to mislead others - whether through phishing, scams, or influence operations - try to leverage crises in order to advance their goals, and the COVID-19 pandemic is no different. As the situation evolves, we are actively working to find and stop coordinated campaigns that seek to manipulate public debate across our platforms.

Our approach to Coordinated Inauthentic Behavior (CIB), and Influence Operations (IO) more broadly, is grounded on behavior- and actor-based enforcement. This means that we are looking for specific violating behaviours exhibited by violating actors, rather than violating content (which is predicated on other specific violations of our Community Standards, such as [misinformation](#) and hate speech). Therefore, when CIB networks are taken down, it is based on their behavior, not the content they posted. For a comprehensive overview of our approach, see [here](#).

To date, we have not found evidence of influence operations created to focus specifically on COVID-19. What we have seen is that people behind campaigns opportunistically use coronavirus-related posts among many other topics to build an audience and drive people to their Pages or off-platform sites.

December 2020 Coordinated Inauthentic Behavior Report

Our teams continue to focus on finding and removing deceptive campaigns around the world — both foreign and domestic. In December, we removed 17 networks of accounts, Pages and Groups, the most we've removed in any one month. We took down deceptive campaigns from nearly every continent and shared information about our findings with law enforcement, policymakers and industry partners. Here are a few trends and tactics we saw:

Targeting of domestic audiences: At least 12 out of 17 networks we investigated and took down last month — in Ukraine, Kyrgyzstan, Kazakhstan, Brazil, Argentina, Pakistan, Indonesia and Morocco — targeted domestic audiences in their own countries. At least five networks — from Iran, Ukraine, France, Morocco, and Russia — focused on people outside of their countries. We know these actors will continue to attempt to deceive and mislead people, including making particular viewpoints appear more widely supported or criticized than they are to sway public debate in their respective countries.

Election targeting and importance of regulations: At least 12 of 17 influence operations focused on elections, primarily (but not exclusively) on behalf of domestic actors. Deceptive campaigns like these raise a complex challenge by blurring the line between healthy public debate and manipulation. Our teams will continue to find, remove and expose these coordinated manipulation campaigns, but we know these threats extend beyond our platform, and no single organization can tackle them alone. That's why it's critical that we, as a society, have a broader discussion about what is acceptable political advocacy and take steps to deter people from crossing the line. As part of our contribution to this conversation, we outlined recommendations for regulatory and legislative principles against these deceptive campaigns [here](#).

Early detection and limited reach: The vast majority of the networks we removed in December had limited following or were in the early stages of building their audiences when we removed them. The small Iranian network is a good example: many of the recent CIB operations from Iran have continued to become smaller and less effective as we and our peers in the industry find and remove them before they are able to build their audiences, including through improving automated detection systems. To remain active if detected by researchers and platforms, some of the networks we took down last month relied on off-platform websites to host their content while using social media to amplify these domains. Ongoing enforcement against these threat actors across the internet has made these operations less effective in building their following. With each removal, we set back the actors behind these networks, forcing them to re-build their operations and slowing them down.

Fictitious news entities and targeting of the media: At least 7 operations we removed last month ran Pages posing as news entities sharing local updates about current events in the countries they targeted. This continued the trend we've seen over the years. Some of the networks in today's report appeared to also target traditional media to place their stories under fictitious bylines. It's critical that all of us, including journalists and influential public figures, remain vigilant about the messages we amplify and verify information we give credence.

Importance of collaboration with researchers, investigative journalists and local civil society organizations: We continue to see strong collaboration among companies, researchers, law enforcement and investigative journalists looking for these operations. When one of us finds an operation, we share it with others so we can all investigate and take action according to our policies. At least 10 out of 17 networks in this report were found either in collaboration with our external partners or based on open-source reporting that led our teams to uncover the full extent of coordinated inauthentic behavior on our platforms. We worked with disinformation researchers, investigative journalists and civil society organizations to find and remove these operations. We know that our adversaries will keep evolving their tactics. That's why we continue to invest in building partnerships – to find these campaigns earlier in their operation.

We are making progress rooting out this abuse – but as we've said before, it's an ongoing effort. We're committed to continually improving to stay ahead. That means building better technology, hiring more people and working closely with law enforcement, security experts and other companies.

The networks reported below were removed for behaviours that violated our Inauthentic Behaviour Policy. As noted, we have not found evidence of COVID-19 focused influence operations.

- Total number of Facebook accounts removed: 1,957
- Total number of Instagram accounts removed: 707
- Total number of Pages removed: 156
- Total number of Groups removed: 727

Networks removed in December 2020:

1. **Iran:** We removed 4 Facebook accounts, part of a small and largely inactive network from Iran. They targeted primarily Arabic, French and English-speaking audiences globally. This network centered around off-platform typo-squatting domains. The vast majority of this activity was detected and disabled by our automated systems for inauthenticity and spam throughout 2020. We found this operation after reviewing information about some of its activity from FireEye and The Daily Beast.

2. **Morocco:** We removed 255 Facebook accounts, 93 Pages, 17 Groups and 60 Instagram accounts that originated in Morocco and targeted Morocco and Algeria. We found links between some of this activity and Qualitia Systems, a marketing firm in Morocco, also known as Marketing Digital Maroc. We found this network as part of our investigation into suspected coordinated inauthentic behavior in the region.
3. **Ukraine:** We removed 23 Facebook accounts, 25 Pages, 11 Groups and 19 Instagram accounts that originated in the Luhansk region in Ukraine and targeted Moldova, Kazakhstan, UK, Spain, Kyrgyzstan, Ukraine, Belarus, Germany, and Russia. This network was linked to individuals in the Luhansk region in Ukraine and those associated with Borotba, a political group in Ukraine. It centered around off-platform domains and was early in its audience building when we removed it. We found this activity as a result of reviewing public reporting on a portion of this activity by WELT and netzpolitik.org in Germany.
4. **Ukraine:** We removed 13 Facebook accounts, 31 Pages, six Groups and three Instagram accounts. This domestic-focused network originated in Ukraine and was linked to individuals associated with the non-governmental organization Anti-Corruption Blockpost. The operation had almost no following when we removed it. We found it after reviewing public reporting in Ukraine about some of this activity.
5. **Ukraine:** We removed 27 Facebook accounts, 37 Pages, 21 Groups and 13 Instagram accounts that originated in Ukraine and focused on domestic audiences. Our investigation found links to individuals associated with both the European Solidarity party and Sprava Gromad, an NGO in Ukraine. This network was early in its audience building when we removed it. We found it as part of our investigation into suspected coordinated inauthentic behavior in the region.
6. **Kyrgyzstan:** We removed 193 Facebook accounts, 246 Pages, 50 Groups and 30 Instagram accounts that originated in Kyrgyzstan and targeted domestic audiences. This network was linked to individuals in Kyrgyzstan with backgrounds in media consulting and had a limited following when we removed it. We found it after reviewing public reporting in Kyrgyzstan about a small portion of this activity.
7. **Kyrgyzstan:** We removed 92 Facebook accounts, four Pages, 11 Groups and 30 Instagram accounts. This domestic-focused activity originated in Kyrgyzstan and focused primarily on commenting on their content and also on posts by a political party and popular news Pages. This network had almost no following when we removed it. We found it after reviewing public reporting in Kyrgyzstan about a small portion of this activity.
8. **Kyrgyzstan:** We removed 121 Facebook accounts, 46 Pages, seven Groups and 41 Instagram accounts that originated in Kyrgyzstan and targeted domestic audiences. Our investigation found links to individuals in Kyrgyzstan with a background in media and government, and also a media company called Media Center. This activity centered around the 2020 Parliamentary election and the 2021 snap presidential election while playing on multiple sides of the political debate at once. This network had almost no following when we removed it. We found it after reviewing public reporting about a small portion of this activity.
9. **Kazakhstan:** We removed 31 Facebook accounts, one Group and 28 Instagram accounts that originated in Kazakhstan and targeted domestic audiences. Our investigation found links to individuals associated with the National Security Committee and the Anti-Extremism Unit of the Police Department of the North-Kazakhstan Region. This network appeared to have

ramped up in 2019-2020 following protests in Kazakhstan and had nearly no following when we removed it. We found it after reviewing public reporting in Kazakhstan about a portion of this activity.

10. **Argentina:** We removed 663 Facebook accounts and 388 Instagram accounts that originated in Argentina and focused on domestic audiences. This network focused primarily on inauthentically amplifying posts and news articles related to Sergio Berni, Buenos Aires' Minister of Security and had almost no following when we removed it. We found it after reviewing public information about a small portion of this activity shared by an open-source [researcher](#). Our assessment benefited from additional findings shared with us by FireEye, a cybersecurity company.
11. **Brazil:** We removed 25 Facebook accounts, three Pages, and 10 Instagram accounts that originated in Paraná and targeted two municipalities in Paraná (Almirante Tamandaré and Colombo) focusing on the 2020 regional elections. Our investigation found links to Continental, a PR Agency based in Curitiba, and other individuals in the state of Paraná in Brazil. This network had almost no following when we removed it. We found this activity as part of our investigation into suspected coordinated inauthentic behavior in the region.
12. **Brazil:** We removed 34 Facebook accounts and 18 Instagram accounts that originated in Brazil and focused on three municipalities in the state of Espírito Santo (Serra, Vitória and Cariacica) to amplify the Pages and posts related to mayoral candidates in each town. Our investigation found links to AP Exata Intelligence in Digital Communications, a public relations firm with offices in Brasília, Vitória and Braga in Portugal. This network had nearly no following when we removed it. We found it as part of our investigation into suspected coordinated inauthentic behavior in the region.
13. **Pakistan:** We removed 27 Facebook accounts, seven Pages, and 23 Instagram accounts that originated in Pakistan and targeted domestic audiences. This network created Pages posing as news entities and was early in building its audience when we removed it. We found it as a result of our investigation into the suspected coordinated inauthentic behavior with some limited links to the network we [removed](#) in August 2019.
14. **Indonesia:** We removed 107 Facebook accounts, 58 Pages, and 34 Instagram accounts that originated in Indonesia and targeted domestic audiences. This network focused primarily on the situation in West Papua and had limited following when we removed it. We found this activity as part of our investigation into suspected coordinated inauthentic behavior in the region.
15. **France:** We removed 84 Facebook accounts, 6 Pages, 9 Groups, and 14 Instagram accounts that originated in France and targeted primarily the Central African Republic and Mali, and to a lesser extent Niger, Burkina Faso, Algeria, Cote d'Ivoire and Chad. Our investigation found links to individuals associated with the French military. This network had nearly no following when we removed it. We found this activity as part of our investigation into suspected coordinated inauthentic behavior in the region. (Originally [announced](#) on December 15, 2020)
16. **Russia:** We removed 61 Facebook accounts, 29 Pages, 7 Groups and 1 Instagram account that originated in Russia. They targeted primarily the Central African Republic (CAR), and to a lesser extent Madagascar, Cameroon, Equatorial Guinea, Mozambique, South Africa, and the CAR diaspora in France. We found links to individuals associated with past activity by the Internet Research Agency (IRA) and previous operations we attributed to entities associated

with Russian financier Yevgeniy Prigozhin, who was [indicted](#) by the US Justice Department. This network was early in its audience building when we removed it. We found this activity as a result of our investigation into the suspected coordinated inauthentic behavior with links to the network we [removed](#) in October 2019. Our assessment benefited from information shared with us by researchers at Graphika. (Originally [announced](#) on December 15, 2020)

17. **Russia:** We removed 197 Facebook accounts, 122 Page, 16 Group and 15 Instagram accounts that originated in Russia and focused primarily on Libya, Sudan, and Syria. We found links to individuals associated with past activity by the Russian Internet Research Agency (IRA). We found this activity as a result of our proactive internal investigation into the suspected coordinated inauthentic behavior in the region — with some limited links to the networks we [removed](#) in October 2019. (Originally [announced](#) on December 15, 2020)

A detailed report on the networks taken down and examples of content they posted can be found [here](#). Previous reports can be found [here](#).

5. Advertising

As the COVID-19 situation develops, we have implemented a variety of measures to prevent ads from being used to spread misinformation; to prevent ads from promoting content that could contribute to physical harm; to prohibit exploitative or deceptive ads; and provide transparency on ads about health issues. We have applied our [Advertising Policies](#) to new types of abuse that we're seeing on the platform. We have made adjustments to our enforcement protocols to prevent people from exploiting the COVID-19 pandemic, and continue adapting or removing temporary bans on specific products as the situation stabilizes.

Updating Our Ad Policy for COVID-19 Vaccines

Given the recent approval of COVID-19 vaccines, we want people to be able to safely promote information about these vaccines on Facebook. We will now allow ads that highlight the ability of a COVID-19 vaccine to prevent someone from contracting the virus, as well as ads promoting ways to safely access a COVID-19 vaccine. We'll continue to prohibit content that tries to exploit the pandemic for commercial gain. And ads or organic posts that promote the sale of a COVID-19 vaccine, such as attempts to sell COVID-19 vaccine kits or expedited access to the vaccine, will be rejected. We will also reject ads that claim the vaccine is a cure for the virus.

It will take some time to train our systems and teams on these policies, and we expect enforcement to ramp up over the coming weeks and months.

For a full list of our Advertising Policies about COVID-19, see [here](#).