

FACEBOOK, INC./\$250 PER 15MIN

**Moderator: Mark Zuckerberg
May 12, 2020
11:00 a.m. PT**

OPERATOR: This is Conference #: 5060848

Operator: Hello and welcome to today's press call. There will be prepared remarks and a Q&A to follow. To ask a question after the prepared remarks have concluded, please press "star," "1."

Now, I'd like to turn the call over to Andrea Saul, who will kick this off.

Andrea Saul: Hi, everyone. Thank you for joining us for the fifth edition of our Community Standards Enforcement Report.

Today on the call, you'll hear from Facebook CEO Mark Zuckerberg, Vice President of Integrity Guy Rosen, Chief Technology Officer Mike Schroepfer, Vice President of Content Policy Monika Bickert, and Vice President of Product at Instagram Vishal Shah. And then, we'll open up the call for questions.

This call is on the record. With that, I will go ahead and kick it over to Mark. Mark?

Mark Zuckerberg: All right. Thanks, Andrea. And hey, everyone, and thanks for joining today. I hope you're all doing OK, and staying healthy and safe as we all navigate this COVID-19 pandemic.

Today, we're publishing our fifth transparency report, outlining how we're doing at finding and removing harmful content. This report covers the period

of October 2019 through March 2020, so that was mostly before we saw the largest impacts from COVID-19.

Guy Rosen, our Vice President of Integrity, who runs all of our products and engineering teams in that area (inaudible) more of the details of this report in just a few minutes.

But before we dive into that I wanted to discuss a few topics up front (inaudible) the content specifically, how this crisis is affecting our overall content moderation work and then some thoughts on the independent oversight board's announcement last weeks of this initial 20 members.

So let's start with COVID-19; since the early days of this outbreak we knew that we were going to have a big role in connecting people with accurate information from trusted sources.

So we so far have directed more than two billion people to resources from health authorities through our COVID information center and other inline educational pop ups across Facebook and Instagram.

We've taken – some of the steps we've taken here we've never taken before to address any kind of crisis, including putting a permanent banner on the top of everyone's Facebook app for more than a month now, directing people to this information center.

We even made the COVID information center a tab in the Facebook app so if you go there frequently, like I do to check on the most recent stats for new cases in different geographic areas, then this COVID-19 information center can fill one of your limited navigation tab slots in the app.

Now beyond connecting people with high quality information, we also have an important responsibility to limit the spread of misinformation to our services and we look at this type of misinformation in two categories.

There is harmful misinformation that could put people in imminent physical danger, so we remove it completely; and then there's broader misinformation

that we don't remove but instead we work with independent fact-checkers to limit its spread across our services.

And in the past couple of months we've removed hundreds of thousands of pieces of harmful misinformation; which includes content claiming that someone has scientifically proven to be a cure for coronavirus or saying that the virus isn't real.

We've partnered with health authorities around the world to identify what type of misinformation risks real danger and having guidance from those independent experts has been very helpful for our policy and enforcement teams in enacting here.

Through other types of misinformation we work with independent fact-checkers to display warning labels and in April alone we put about 50 million labels on pieces of content related to COVID-19 based on 7,500 articles by independent fact-checking partners.

And we have a good sense that these warning labels work, because 95 percent of the time that someone sees content with a label, they don't click through to view that content.

We originally launched this independent fact-checking program about four years ago with 15 partners covering four languages and it's evolved quite a bit since then.

Today we have more than 60 fact-checking partners covering more than 50 different languages around the world.

We've also used machine learning to put in automated systems that can detect and down rank and remove misinformation before it goes viral based on signals from these partners.

And the investments that we've made in recent years have definitely helped us to prepare for this moment and in turn, they've also given us a chance to refine our approach and experiment with new tactics for fighting misinformation as well.

Now beyond misinformation we are seeing an increased prevalence of other types of COVID related content that tries to exploit the crisis for financial gain.

So, since March 1, we've removed more than 2.5 million pieces of content, including – for the sale of masks, hand (sanitizer) (inaudible) and COVID test kits.

And the challenge here is that for everyone trying to exploit the situation, there are also people (trying to) help get masks and other equipment to people who really need it. So, (that) can sometimes to detect what's a fraudulent sale that we should stop and what's a real, well-intentioned offer to help.

And we have started to rely on computer vision technology that we've been building up over the years to find and remove drug sales, to work here – I mean, we're going to keep working hard to take down exploitative sales related to COVID-19 as well.

Next, I want to talk for a bit about how our content moderation work has been impacted by COVID-19. Now, as you know, we generally find and remove content that violates our community standards through a combination of automated systems and human review.

And in mid-March we worked with our partners to send our content reviewers home in order to keep them safe for the growing outbreak. And this made the human review part of part our enforcement much more difficult.

So, at first we shifted our review work to a smaller number of full-time employees, who are also working from home. And since then we've worked with some of our partners to allow a majority of contractors doing content review to do their work from home (too).

(Inaudible) protect the privacy of people in our community while reviewers are working from home, and also to make sure that we can put in place the appropriate safeguards to take care of the mental health and wellbeing of these reviewers while they're remote.

And as countries around the world have begun to reopen, we have started allowing some reviewers back into offices on a voluntary basis. And we're going to continue this process in a safe way and we'll continue following all national and local guidance on that front.

Now, since we have a reduced workforce, we're prioritizing reviewing certain types of content, like child safety, suicide and self injury, terrorism and harmful misinformation related to COVID, to make sure that we're handling the most dangerous issues. We're also relying more on our automated system to detect and remove violating content and to disable accounts.

And our machine learning technology has come a long way in recent years and we're able to catch a lot of this harmful content. We're now able to find almost 90 percent of hate speech that we take down on Facebook before it's even reported to us. That's up from just 24 percent when we released our first (of these reports) in 2018, and up from roughly 0 percent a year before that.

On Instagram we are proactively catching about 90 percent of posts that we take down related to suicide and self harm, and that's up from 78 percent last year. And our Chief Technology Officer Mike Schroepfer is joining this call today to talk more about how AI is helping us find the content that breaks our rules.

But I just want to say that I am quite proud of the work that our teams have done to build up these systems over the past few years. It's allowed us to identify more harm proactively and keep our communities safe, including now in the face of a global pandemic.

That said, I know our systems aren't perfect, and they still need to improve. Our effectiveness has certainly been impacted by having less human review during COVID-19, and we do unfortunately expect to make more mistakes until we're able ramp everything back up.

And since today's report only covers through March, it doesn't reflect the full impact of the changes that we've seen and had to make during the pandemic.

So we do expect to see more of an impact in the next report. (And this example, since we have more) people reviewing content, we've prioritized during a first review of more severe reports over doing a second review and appeal of other types of content.

So we expect the number of appeals to be much lower in our next report. As another example, in order to focus on actually removing harmful content, we've temporarily de-prioritized and stopped sampling and labeling content for measurement.

So for this report we're not including some of the (prevalent) metrics for March and we may not be able to calculate the prevalence of violating content until we have more reviewers back online.

Through this pandemic and beyond, we're going to continue reporting our effectiveness and we'll remain transparent with our results. And I want to stress this because in a crisis there can also be a tendency to put your head down and turn inwards.

And we're not going to do that. We're going to keep sharing these reports even if some of our numbers dip in some places because as I've said, I believe these – the transparency in how we're handling the safety of our community is as important as the reports that we make on our quarterly earnings.

Now the last area that I want to discuss today is the independent oversight board's announcement this week of its first 20 members. As you know, the oversight board will rule on some of the most significant and challenging decisions that we face around what can and cannot be shared across our services.

This is going to add an important level of appeals as well as accountability and scrutiny to the content enforcement work that we're talking about today. The first 20 members who will serve on this board bring diverse backgrounds, ideas and experiences; including (as former) judges, a former prime minister, NGO leaders in a number of countries, journalists, and academics.

But one thing they all have in common is a strong belief in the importance of giving people voice and free expression. The oversight board is going to have the power to overturn decisions that we've made on content. As long as they comply with local laws.

Their decisions are going to be final and binding, even if I or anyone else at the company disagree with them. So by establishing this oversight board, we're attempting to engage in the fragile but important process of independent institution building.

I am making a public commitment and putting a stake in the ground that will help keep us true to our values and make it that much harder for me or whoever else runs this company down the line to back away from these principles.

Now it's still early and my hope is that if this is working well, then we're going to be able to expand the oversight board's charter as well. But it's worth noting this structure really is a first of its kind.

I don't think it's something that any other company has done before and I think that it's an important step to make sure that people will always continue having voice and staying safe on our platform. So with that, I'm going to turn this over to Guy to dive a bit deeper into all the numbers that we're sharing today.

Guy Rosen: Thanks, Mark. And good morning, everyone. I'd like to walk you all through some highlights from this fifth edition of our report. I will call out two updates that we've made in this edition.

Firstly, we're including more data on Instagram. We're adding five new policy areas as well as data on appeals on Instagram. Now, Facebook and Instagram use the same reporting tools, detection systems, review systems, measurement infrastructure and we have one central team responsible for developing solutions that span across our services.

(Inaudible) some of (these systems) built for (Facebook) first and later applied to other services. This means that in some of the areas you might see the

systems are further ahead on Facebook than they are on Instagram, and at times that may be reflected in the metrics in this report.

The second (addition) is we're (cutting) data here on our efforts to combat organized hate. This is across both Facebook and Instagram, and so in total, the report now covers 12 policy areas on Facebook and 10 on Instagram.

Now, I'd like to walk through three key metrics that the report covers; prevalence, content action, and proactive rate. The first and perhaps most important metric is prevalence. In other words that is, how much bad content did people actually see? We periodically sample the content that is viewed on our services and we review it to statistically calculate what percent violates our policies. And if something like that were seen by someone, it means we didn't take it down or perhaps we took it down but we got to it too slowly.

And it's important (that we) focus on how much content is seen, not just how much sheer content is out there that violates our rules. For some of the policy areas, especially those that address the most severe safety concerns, the likelihood that people view content that violates these policies is extremely low.

When we sample content that people view to measure prevalence, for these areas we might not find any violating samples at all and so for those cases, we statistically calculate an upper limit. That means in order for us to find a little to no sample, the actual likelihood of seeing this content is between zero and this limit.

In this report, that upper limit is 0.05 percent for all of these violations across both Facebook and Instagram, except for one area on Instagram which (is) higher in Q1 due to how we counted content that we removed related to the death of Iranian Major General Qasem Soleimani. Monica will share more about how that incident manifested on our services.

The second metric is how much content we took actions on. This includes removing content, (applying a) warning screen or disabling accounts, and it reflects how often people violate our policies and how much of it we identified. It also means that it can fluctuate a lot. For instance, if we find and

remove a handful of viral memes (that violate our policy), that might drive the number up one quarter and then the numbers can drop the next (inaudible) will no longer (share) those (memes).

The number can also (inaudible) as we (expand our systems) to proactively detect harmful content and you can see an example in this increase, for example, in drug sale content on Facebook where in Q4 of 2019, we removed 8.8 million pieces of content. That's up from 4.4 million the prior quarter. That's a result of improved detection.

Now the third metric is proactive rate. This typically reflects how effective AI is in a particular policy area as it shows how much of the content that we took action on was detected by our systems before someone reported the content (to us). For hate speech, we continue to expand proactive detection technology, we added more languages and improved the existing detection systems.

And the detection rate – the proactive detection rate for hate speech increased by more than eight points over the past three quarters and almost 20 points in the last year. As a result, we're now able to find more content and we detect almost 90 percent of the hate speech content that we remove on Facebook before people report it to us.

We've made progress on this, but we know we have more work to do, including on (Instagram) for that same (inaudible). To talk more about our progress on AI, I'll now turn it over to Schrep.

Mike Schroepfer: Thanks, Guy. AI is our best tool for keeping users safe at scale. Billions of pieces of content are uploaded to Facebook every day, we use AI to screen each piece in hundreds of milliseconds to make sure content that violates our standards doesn't get posted.

It's an incredibly (different) task done at massive scale against adversaries who are constantly working to (evade) our systems. As Mark mentioned, one of our top priorities during this time is to connect people to accurate information from health experts and keep harmful misinformation about

COVID-19 from spreading on our apps, an area where AI is playing a crucial role.

We're using AI to understand if someone is offering, requesting help in a public post. We can get in front of them (inaudible) the local community via the COVID information help hub. We also know that misinformation labels work, but the challenge is the adversarial nature of this work.

People are uploading slightly different variations of images designed to evade our systems. To address this we shipped a new technology called (Sim Search Net). When a piece of content has been labeled by an independent third-party fact-checking partners as misinformation, this technology can detect near-perfect matches.

For instance, when someone has altered it to try to bypass our systems. It's reviewing billions of photos every day to find copies of false or misleading information that's already been debunked.

Using AI to detect these matches enables these partners to focus on new instances of misinformation rather than near identical variations of things they've already seen.

In addition to (Sim Search Net), to help us improve the ability to detect and remove bad ads and commerce listings from medical face masks, hand sanitizer, surface disinfecting wipes, and COVID-19 testing kits – we quickly developed a system using local image features to match similar images to prevent people from avoiding detection by cropping or altering photos and apps.

As Guy mentioned, we're also making rapid progress using AI to detect hate speech in three areas. First, (cell) supervision. We're using the ability to train on datasets that don't have to be manually curated. This allows us to train on massively more data and learn a lot more nuance from it.

(Cross-lingual) understanding. We have the ability now to build classifiers that understand the same concept in multiple languages. And learning in one

language can improve its performance in others. This is particularly useful for languages that are less common on the internet.

And the third area is whole post of multi-modal understanding. This is the ability to train not just on text or images in isolation, but the ability to pull all of them together to understand the post as a whole. This technology is now used at scale to analyze content, (but) to acknowledge how important it is that we work together as an industry to take an open approach on discussing and handling (integrity) issues on our platforms.

We launched technical industry efforts like PyTorch, now one of the most popular deep learning frameworks on the market, along with a Deepfake Detection Challenge that resulted in 2,000 teams from all over the world participating to find a better way to detect deepfakes automatically.

We committed to open sourcing the tools as a result. Today we're open sourcing the first ever hateful memes dataset for an industry-wide challenge to tackle hate. The hateful memes dataset will consist of more than 10,000 newly created examples of this hateful multi-modal content.

By providing these datasets expressly made to help researchers tackle the problem, along with a benchmark, community participation, we are confident that the Hateful Memes Challenge will spur faster progress across the industry in dealing with hateful content and help advance multi-modal machine learning more broadly.

Of course AI isn't our only answer to (harmful content), but (inaudible) to address these challenges as we work to provide the safest communication platform possible.

I will now turn the call over to Monika.

Monika Bickert: Thanks, Schrep, and hello everyone. I'm going to start by expanding upon the COVID-19 updates and protections that Mark mentioned around the policies. We had longstanding policies against content that could contribute to (real world harm), but any time that we're seeing new trends in abusive speech or

behavior, we want to make sure that our teams are ready to apply those existing policies, including our (house) policies, to those new trends as well.

(Now Mark talked a little bit) about how we've applied our policies to exploitative commercial content, like sales of hand sanitizer and face masks. It also means that in the context of COVID-19 we've been applying our policy against the coordination of physical harm to remove posts that advocate for the spread of the virus. We're applying our policies against bullying and harassment to remove claims that a private individual (has the) virus.

(It would also apply in) our policy against misinformation that contribute to the risk of imminent physical harm to take down misinformation that could indeed contribute to the spread of the virus, such as false claims about cures, treatments, the availability of essential services, or the location and the severity of the outbreak.

To date we've removed hundreds of thousands of pieces of misinformation that could have contributed to the risk of imminent physical harm. And as the situation evolves, we'll continue to stay on top of trends and engage with experts like the World Health Organization and other governmental health authorities and we will provide additional guidance where appropriate.

(Inaudible) (this adaptation) of our policies to apply existing policies to new types of abuse. That's not new. That's something that we do regularly, just in the COVID-19 context. What metric that I want to call out for the report is the upper limit on prevalence for terrorism on Instagram, which we've estimated at 0.07 percent in Q1. This was not because of (these spiking terrorist content) on Instagram, rather it was because of the way that we enforced on content related to the death of the Iranian Major General Qasem Soleimani.

Following Soleimani's death in January 2020, we saw a fair amount of content on Instagram that was posted in praise and support of him and of the Islamic Revolutionary Guard Corps. We removed the support of Soleimani and the IRGC in line with our obligation to adhere to U.S. sanctions law, which do apply to the IRGC and its leadership.

However, we left on site posts that referenced Soleimani positively but that stopped short of endorsing his activity or the activity of the IRGC, and even though we left this latter category of content on the site, the content is still (reflected in the prevalence) metric we're sharing today because of the general standards that we use (to find) violations for this category.

Lastly, in addition to the community standards enforcement report, today we released our biannual government request report, which details the numbers of government requests for user data and content restrictions based on local law. And we're also releasing the findings of three human rights impact assessments that we commissioned on the role of our services in Sri Lanka, Indonesia, (and Cambodia).

I hope that you'll take a moment to look at both of these releases. We are very proud of the work that they represent and continued (commitment to transparency) that they demonstrate. And with that, I'll now turn the call over to Vishal.

Vishal Shah: Thanks, Monika. It's been six months since we introduced (Instagram) metrics to this report. At first we shared five policy areas, child nudity and sexual exploitation sales, firearm sales, suicide and self injury and terrorist propaganda. (We're reporting our) (inaudible) in those areas and (as Guy mentioned), we're introducing more areas.

In addition, we're sharing metrics on appeals made to us and decisions we overturn based on those appeals. I'm proud of these new additions and the progress we've made when it comes to keeping people safe on Instagram. In particular, I want call out improvements to our machine learning and media matching technologies. These technologies allow us to find and remove content that is identical or nearly identical to existing violations in our databases.

We focus heavily on this area and (seen a) market improvement in our metrics as a result. On contact related to suicide and self injury, following improvements I mentioned to our media matching capabilities, we increased

our proactive detection rate to almost 90 percent in Q1. And improvement of more than 12 points since the last report.

The volume of content action increased by 40 percent as a result. Like Facebook, we were also able to detect and take down more child nudity and sexually (exploited) content in Q1. With a 53 percent increase in content action to over 1 million pieces of content. We still have work to do here but as always, child safety remains an utmost priority for us.

Elsewhere we remained committed to leading the industry in the fight against online bullying. And in both Q4 of 2019 and Q1 of 2020 we took action on 1.5 million pieces of content. Today we're announcing a number of new features to further our work here. These include giving people the option to remove multiple negative comments or blocking multiple people at once, allowing people to pin positive comments to set the tone for their accounts, and giving people more control over who can tag and mention their account.

Please see the Instagram press center for more information on each of these features. Thanks for listening, everyone. We hope these metrics across our services demonstrate our continued (concerted) effort to keep people safe. I will now open the floor for questions.

Andrea Saul: Operator, can you instruct the participants on the call on how to ask a question?

Operator: We will now open the line for questions. To ask a question press "star" followed by "1". Your first question comes from the line of Danielle Abril from Fortune.

Danielle Abril: Hi, everybody, thank you so much for taking the call today. I had a quick question on hate speech. I know that you guys mentioned that your systems got a lot better at detecting hate speech. But I wondered since the coronavirus outbreak, how much – have you seen any increase around that or perhaps even the political environment around the coronavirus? Basically, have you seen any change in behavior as to what you're actually (seeing) on the platform? Thanks.

Guy Rosen: Hey, this is Guy. Thanks for the question. So this report goes through March so it's probably still too early to assess the full impact of that. We do see an increase in our content action and hate speech for Q1. That's mostly driven by improvements in our proactive detection technology.

We're definitely monitoring and continuing to look closely at how the COVID pandemic is impacting these kind of trends on our platform. And I think in the next report we hope to have more to share.

Operator: Your next question comes from Issie Lapowsky from Protocol. Please go ahead.

Issie Lapowsky: Yes, thanks for taking my question. My question, I think this might be most appropriate for Monika but you guys can tell me. As you have shifted to having full-time staffers moderating some of these more serious violations, which is a job I imagine isn't part of their typical set of responsibilities, can you share what their experience has been like, what sort of feedback you've gotten from them, and how, if at all, it has changed (inaudible) (outlook in yours) on what (inaudible) entails (and what are the resources) (inaudible) (needs)? Thank you.

Monika Bickert: Sure. I should say that some of the – some of the people who are reviewing content who are full-time employees have been doing that job in the past. And some of them have had to be intimately aware of the details of how to do that work because of the nature of their jobs.

So what we've done to make sure they're ready to review content is make sure we're going through with them all the different ways that the content can manifest itself on our platform and why we have the policies that we do. And then, just make sure that we're with them every step of the way, making sure they have the resources that they need.

Like Mark mentioned, we are using of both the full-time employees and some of our contractors, just like we always have. Of course, it's in a – it's different now, but we're still using that mix of full-time employees and contractors in the office work voluntarily and at home work to make sure that we are keeping our community safe.

Operator: Your next question comes from Queenie Wong from CNET. Please go ahead.

Queenie Wong: Thank you so much for taking the time to answer my question. You talked about how AI has helped you proactively detect content that could possibly violate your rules.

Are your human reviewers doing anything to proactively detect content that could be slipping through the – through the cracks, like searching through the social network for coronavirus misinformation or looking through private and public groups? Or do these reviewers mainly rely on content being brought to them, either through the AI system or user reports?

Guy Rosen: Hey, this is Guy. Thanks for the question. We use – it's always been a combination of people and the technology.

And the reviewers – there's several sources for the kinds of things that they may be reviewing. One of them is posts that are reported by people in the community. Another one is posts that are flagged by our proactive detection. And in some cases, the AI is at a point where it wants a human to take a look before we make the final decision. It might not be confident enough to actually remove the content independently.

But we also have processes where, through a combination of tools – AI tools and others – we are sort of searching, sweeping the site as it were, trying to find different kinds of content that may violate. If you think about what our proactive detection technology does, it's really trying to locate content that might violate our standards, whether it's posted by individuals, or on a page, or in a group and so forth.

So we use all of those sources and the humans will work in tandem with those tools in order to try to find that content and take down the content that is violating.

Operator: Your next question comes from the line of Casey Newton from The Verge. Please go ahead.

Casey Newton: Hi. So I had a question about industry collaboration. Over the past week, we saw the "Plandemic" video go viral and it appears that it got a lot of attention in Facebook groups that then drove millions of views on YouTube. And I wondered if you all have – could ever envision a time when you informed other platforms about seeing those kind of huge spikes (in traffics) to external sites that might be suggestive of bad actors organizing?

Monika Bickert: Hi, this is Monika. Thanks for the question. We actually do work quite a bit across industry to make sure that we are sharing best (inaudible).

Mark Zuckerberg: Hey Monika, I think you might have muted yourself.

Andrea Saul: While Monika dials back in to answer that question, why don't we go ahead and take the next question and then we can come back to her.

Operator: Your next question comes from the line of Manuel Machado from Observador, please go ahead.

Manuel Machado: Hello, thank you for – thanks again for taking my question. I've been reading these transfer series (reports), but I could not find any information regarding revenues.

Has Facebook revenue grew since the start of the pandemic? If not, by how much has it reduced? If yes, by how much it grew and from which revenues? I think the best person for this question is Mark.

Mark Zuckerberg: So I'm not sure I understand the question; are you asking about our financials? I mean, we typically report in our quarterly earnings. Are you – are you asking about something specific related to how we do content moderation or can you clarify that?

Manuel Machado: Sorry, I was not – if I did not clarify the question. I know you asked – you show that on the reports – for – on the quarterly reports. But specifically (the transfer to reports), we wanted to know if the revenue grew or not since the start of the pandemic, since this is also about transparency and (show how Facebook is going).

Mark Zuckerberg: I think the only data that we're going to share on our financials is what we share on earnings; which it was just a few weeks back and we both shared at the time data from the first quarter in our financials and our CFO Dave Wehner gave some forward-looking statements on what we expect the outlook to be going forward.

Manuel Machado: OK, thank you.

Andrea Saul: And I think we have Monika back if we want to jump back to the question about industry collaboration. Monika, are you there? OK let's...

Monika Bickert: Yes, I'm here, let's try this again. Sorry, I'm not sure what happened there.

But yes, we actually do have regular calls and meetings with industry and we collaborate very broadly across different areas of content policy and we've done that for years and we're doing that now around COVID-19 as well.

That does include sometimes flagging for one another when we see specific pieces of content that might be gaining attention either in the news or on our services. Now of course each platform has different rules and so we each make our decisions independently.

With the video that you mentioned, the "Plandemic" video, that video included a bunch of conspiracy theories including the claim that wearing a mask could activate your own coronavirus and make you sick.

We actually – because that video does violate our policies – we did remove that and we are also seeing people try to post that video again and we're doing our best to identify copies of that video and remove it as well.

Operator: Your next question comes from the line of Karissa Bell from Engadget.

Karissa Bell): Hi, thanks for taking my question. At the beginning of the call, Mark mentioned (prior) (inaudible) (types of) content right now, including content that can cause harm. And I'm curious if content that promotes (inaudible) conspiracy theories or other types of misinformation about vaccines is being prioritized right now, given the links we've seen between accounts that share

this and also sharing misinformation about COVID and conspiracy theories and the “Plandemic” video that was just mentioned?

Guy Rosen: Hey, this is Guy. We're generally, as Mark said, we're trying to prioritize (a) number of areas that are around sort of real world offline harm. But, it's not that we're not even – not that we're not getting to other areas, this is just a question of sort of the mix and how – what are the things that (our more limited review workforce) is trying to get to.

The – definitely anything in and around and related to COVID related and sort of health related harm is something that's very much top of mind and very much something that we want to prioritize, obviously because of the importance that it has during this time, during the pandemic.

Operator: Your next question comes from the line of Sarah Pryor from Bloomberg. Please go ahead.

Sarah Pryor: Hi, thank you. I know that you spoke earlier in this call about how just the resources of human moderators right now have been restricted because of coronavirus. And as we've seen the virus continue, and it may be more of a long-term thing than a short-term blip, are you actually considering hiring more human moderators to pick up the slack and cover these categories that you can't pay attention to right now?

And I'm also curious, when people – when users report – are you relying less on user reports as you transition more to AI, or are those reports from Facebook and Instagram users still all be reviewed by (humans)? Thank you.

Guy Rosen: Hey, this is Guy. Thanks for the question. So, the – for (priority), we're in the process of bringing the review workforce back online. So, we sent everyone home in order to take care of their health and the health of their communities. But we – we're – as we did that, we worked with our partners to ensure that those reviewers could also get online and be able to work from home.

A majority of our reviewers now are actually able to work from home. So, we have them in addition to the full-time employees who are also working from home. But, there's obviously slight differences in what that sort of work is

like and so we're working hard to make sure that we're prioritizing things in the right way, we're providing the right resources and support to anyone who is doing – who is doing this kind of work.

On the second part of your question, around reports, so there is right now – there is a sort of screen that pops up when people try to report that does note that we are not going to be able to get all reports right now due to the pandemic and due to the reduced workforce.

As we bring everyone back online and get all of the systems that are working again at capacity, we hope to get back to able to looking through all those user reports as well. But, we will be able to – we will be prioritizing the ones that are, for example, the most harmful or the most prevalent, in order to get to the most problematic content that has the most impact on people who use our services.

Operator: Your next question comes from the line of Davey Alba from “The New York Times.” Please go ahead.

Davey Alba: Hi, thanks so much for taking my question. Not to hammer on the “Plandemic” video so much but I just wanted to be – to ask more precisely, I know that you’re (inaudible) wrong information about masks. We’ve also seen that people are re-uploading the video, cutting out that part. And so does that violate your guidelines or is that allowed on Facebook?

Monika Bickert: Hi, this is Monika. Thanks for the question. If there were a number of different parts of that video that could be (fact) by our fact-checkers. So if people are sharing the clip of the video that has the mask claim in, that violates our policy against sharing misinformation that could contribute to the imminent – to the risk of imminent physical harm. So that, we will remove.

If they’re sharing it and it has other parts of the video but not the mask claim, then it’s eligible for our third party fact-checkers to fact check that and if they do rate it – if the fact-checkers do rate it as false, then we label it on our platform and we reduce its distribution.

Andrea Saul: Great. Operator, we have time for one more question.

Operator: Your last question comes from the line of (Zoe Thomas) from (CBC News). Please go ahead.

(Zoe Thomas): Thank you guys for doing this. Just to go back to the AI a bit. You made this big shift into using AI as moderators that you've been kind of doing for some time. But do you think the uptick in using AI now as moderators will mean long term you're limiting how many human moderators you're using to look at content?

Guy Rosen: Hey, this is Guy. So we'll continue to learn from how we're leaning more into automation during this crisis but people are going to remain critical to this work. We can use automated removal when the technology is very confident that a post violates our policies.

But often we need those teams to be working hand in hand. And even in order to build up and train AI we need to have content manually reviewed by reviewers because that's what actually goes into what's called the training data that helps that AI become effective.

So yes, there has been a lot of progress and a lot of content is the (test of our technology) and I think this work over the past year has really helped us respond during this very unprecedented time.

But the humans are going to continue to be a really important part of the equation, both the people who report content to us but also the people on our review teams who are part of keeping people safe on our platform.

Andrea Saul: Great. Thank you everyone for joining our call today and if you have any follow-up questions, you know how to reach our team or you can reach out at press@fb.com. Thanks so much, have a good day.

Operator: This concludes the Facebook press call, thank you for joining; you may now disconnect your lines.

END