HEARING BEFORE
THE UNITED STATES HOUSE OF REPRESENTATIVES COMMITTEE ON HOMELAND SECURITY

June 26, 2019

Testimony of Monika Bickert
Head of Global Policy Management, Facebook

I. Introduction

Chairman Thompson, Ranking Member Rogers, and distinguished members of the Committee, thank you for the opportunity to appear before you today. My name is Monika Bickert, and I am the Vice President of Global Policy Management at Facebook. In that role, I lead our efforts related to Product Policy and Counterterrorism. Prior to assuming my current role, I served as lead security counsel for Facebook, working on issues ranging from children's safety to cybersecurity. And before that, I was a criminal prosecutor with the Department of Justice for 11 years in Chicago and Washington, DC, where I prosecuted federal crimes including public corruption and gang violence. On behalf of Facebook, I want to thank you for your leadership in combating extremism, terrorism, and other threats to our national security.

I want to start by saying that all of us at Facebook stand with the victims, their families, and everyone affected by recent terrorist attacks, including the horrific violence in Sri Lanka and New Zealand. In the aftermath of such heinous acts, it is more important than ever to stand against hate and violence. We will continue to make that a priority in everything we do at Facebook.

Facebook's mission is to give people the power to build community and bring the world closer together. We are proud that more than two billion people around the world come to Facebook every month to share with friends and family, to learn about new products and services, to volunteer or donate to organizations they care about, or to help in a crisis. But people need to feel safe in order to build this community. And that is why we are committed to fighting any efforts by terrorist groups to use Facebook. That is also why Facebook has rules against inciting violence, bullying, harassing, and threatening others. Our goal is to ensure that Facebook is a place where both expression and personal safety are protected and respected.

II. Facebook's Efforts To Combat Terrorism

On terrorist content, our view is simple: there is absolutely no place on Facebook for terrorism. Our longstanding Dangerous Individuals and Organizations policy bans any organization or individual that proclaims a violent mission or has engaged in acts of violence, including terrorist activity and organized hate. Regardless of whether or not these individuals or groups post content that would violate our policies, we remove their accounts as soon as we find them. They simply are not allowed to use our services under any circumstances. Furthermore, we remove any content that praises or supports terrorists or their actions whenever we become aware of it,

and when we uncover evidence of imminent harm, we promptly inform authorities.

We recognize the challenges associated with fighting online extremism, and we are committed to being part of the solution. We are working to address these threats in three ways: through products that help us stop terrorists at the gate, people who help us implement our policies, and partnerships outside the company which can help us stay ahead of the threat.

A. Products

One of the challenges we face is identifying the small fraction of terrorist content—less than 0.03%—posted to a platform used by more than two billion people every month. Facebook has invested significantly in technology to help meet this challenge and to identify proactively terrorist content, including through the use of artificial intelligence (AI) and other automation. These technologies have become increasingly central to keeping hateful or violent content off of Facebook.

Importantly, we do not wait for ISIS or Al Qaeda to upload content to Facebook before placing it into our internal detection systems. Instead, we proactively go after it. We contract with groups like SITE Intelligence and the University of Alabama at Birmingham to find propaganda released by these groups before it ever hits our site. We put this content, and other content we are able to identify from elsewhere on the Internet, into our matching systems. And once we are aware of a piece of terrorist content, we remove it.

We know that terrorists adapt as technology evolves, and that is why we constantly update our technical solutions. We use these solutions, as well as human expertise, so we can stay ahead of terrorist activity on our platform. We have provided information on our enforcement techniques in the past, and I would like to describe in broad terms some new tactics and methods that are proving effective.

1. Machine Learning Tools

We use machine learning to assess Facebook posts that may signal support for ISIS or Al Qaeda. Our machine learning tools produce a score indicating the likelihood that the post violates our counterterrorism policies, which, in turn, helps our team of reviewers prioritize reviewing posts with the highest scores. The system ensures that our reviewers are able to focus on the most important content first. And when the tool is sufficiently confident that a post contains support for terrorism, we automatically and immediately remove that post.

We have seen real gains as a result of our efforts; for example, prioritization powered by our new machine learning tools has been critical to reducing significantly the amount of time terrorist content reported by our users stays on the platform.

2. Changes To Facebook Live

Facebook has also made changes to Facebook Live in response to the tragic events in Christchurch. We now restrict users from using Facebook Live if they have violated certain

rules—including our Dangerous Organizations and Individuals policy. We apply a "one strike" policy to Live: anyone who violates our most serious policies will be restricted from using Live for set periods of time—for example, 30 days—starting on their first offense. And we are working on extending these restrictions in the weeks to come, beginning with preventing those same people from creating ads on Facebook.

3. Improvements To Existing Tools And Partnerships

We have improved several of our existing proactive techniques and are now able to detect more effectively terrorist content. For example, our tools to algorithmically identify violating text posts (what we refer to as "language understanding") now work across 19 languages. Similarly, though we have long used image- and video-hashing—which converts a file into a unique string of digits that serves as a "fingerprint" of that file—we now also use audio- and text-hashing techniques for detecting terrorist content.

These improvements in our technical tools and partnerships have allowed for continued and sustained progress in finding and removing terrorist content from Facebook. Since the beginning of 2018, we have taken action on more than 25 million pieces of terrorist content, and we found over 99% of that content before any user reported it.

B. People

We know that we cannot rely on AI alone to identify terrorist content. Context often matters. To understand more nuanced cases, we need human expertise.

One of our greatest human resources is our community of users. Our users help us by reporting accounts or content that may violate our policies—including the small fraction that may be related to terrorism. To review those reports, and to prioritize the safety of our users and our platform more generally, including with respect to counterterrorism, we have more than 30,000 people working on safety and security across the company and around the world. That is three times as many people as we had dedicated to such efforts in 2017. Our safety and security professionals review reported content in more than 50 languages, 24 hours a day.

Within our safety and security team, we have also significantly grown our team of dedicated counterterrorism specialists. Distinct from our content review teams, we have more than 300 highly-trained professionals who are exclusively or primarily focused on preventing terrorist content from ever appearing on our platform and quickly identifying and removing it if it does. This team includes counterterrorism experts, former prosecutors, and law enforcement personnel. Together, they speak over 30 languages and are working 24 hours a day around the world to detect and remove terrorist content.

Because our reviewers are human, our performance is not always perfect. We make mistakes. And sometimes we are slower to act than we want to be. But keeping our platform and our users safe is one of Facebook's highest priorities, and we are always working to improve.

C. Partnerships

We are proud of the work we have done to make Facebook a hostile place for terrorists. We understand, however, that simply working to keep terrorism off Facebook is not an adequate solution to the problem of online extremism, particularly because terrorists are able to leverage a variety of platforms. We believe our partnerships with others—including other companies, civil society, researchers, and governments—are crucial to combating this threat.

In 2017, Facebook co-launched the Global Internet Forum to Counter Terrorism (GIFCT) with YouTube, Microsoft, and Twitter. The GIFCT shares information between the participants and has trained 110 companies from around the globe. Just last week, we held an event in Jordan that brought together more than 100 people from government, industry, and civil society to share best practices.

Through GIFCT we expanded a database—which now contains hashes for more than 200,000 visually distinct images or videos—in which 15 companies share "hashes," or digital fingerprints, to better enable companies to identify noxious terrorist content.

Facebook took over as the Chair of the GIFCT in 2019 and we have worked to expand its capabilities, including increasing hash sharing. In fact, we are freely providing our hashing technology to companies participating in the consortium.

Our efforts to work with others in the industry to tackle the online terrorist threat go further still. On May 15, 2019, in the wake of the tragic New Zealand attacks, Facebook and other tech companies, including Google, Twitter, Microsoft, and Amazon, signed the Christchurch Call to Action. The Christchurch Call expands on the GIFCT and builds on our other initiatives with government and civil society to prevent the dissemination of terrorist and violent extremist content.

Facebook joined with others in the industry to commit to a nine-point plan that sets out concrete steps the industry will take to address the spread of terrorist content. Those steps are:

- (1) **Terms of Use**. We commit to updating our terms of use, community standards, codes of conduct, and acceptable use policies to expressly prohibit the distribution of terrorist and violent extremist content.

- (2) **User Reporting of Terrorist and Violent Extremist Content.** We commit to establishing one or more methods within our online platforms and services for users to report or flag inappropriate content, including terrorist and violent extremist content. We will ensure that the reporting mechanisms are clear, conspicuous, and easy to use, and provide enough categorical granularity to allow the company to prioritize and act promptly upon notification of terrorist or violent extremist content.

- (3) **Enhancing Technology.** We commit to continuing to invest in technology that improves our capability to detect and remove terrorist and violent extremist content

online, including the extension or development of digital fingerprinting and AI-based technology solutions.

- (4) **Livestreaming.** We commit to identifying appropriate checks on livestreaming, aimed at reducing the risk of disseminating terrorist and violent extremist content online. These may include enhanced vetting measures and moderation where appropriate. Checks on livestreaming necessarily will be tailored to the context of specific livestreaming services, including the type of audience, the nature or character of the livestreaming service, and the likelihood of exploitation.

- (5) **Transparency Reports.** We commit to publishing on a regular basis transparency reports regarding detection and removal of terrorist or violent extremist content on our online platforms and services and ensuring that the data is supported by a reasonable and explainable methodology.

- (6) **Shared Technology Development.** We commit to working collaboratively across industry, governments, educational institutions, and NGOs to develop a shared understanding of the contexts in which terrorist and violent extremist content is published and to improve technology to detect and remove terrorist and violent extremist content including by creating robust data sets to improve AI, developing open source or other shared tools to detect and remove content, and by enabling all companies to contribute to the effort.

- (7) **Crisis Protocols.** We commit to working collaboratively to respond to emerging or active events on an urgent basis, so relevant information can be quickly and efficiently shared, processed, and acted upon by all stakeholders with minimal delay. This includes the establishment of incident management teams that coordinate actions and broadly distribute information that is in the public interest.

- (8) **Education.** We commit to working collaboratively to help understand and educate the public about terrorist and extremist violent content online. This includes educating and reminding users about how to report or otherwise not contribute to the spread of this content online.

- (9) **Combating Hate and Bigotry.** We commit to working collaboratively across industry to attack the root causes of extremism and hate online. This includes providing greater support for relevant research—with an emphasis on the impact of online hate on offline discrimination and violence—and supporting the capacity and capability of NGOs working to challenge hate and promote pluralism and respect online.

Our work to combat terrorism is not done. Terrorists come in many ideological stripes—and the most dangerous among them are deeply resilient. At Facebook, we recognize our responsibility to counter this threat and remain committed to it. But we should not view this as a problem that can be “solved” and set aside, even in the most optimistic scenarios. We can reduce the presence of terrorism on mainstream social platforms, but eliminating it completely requires addressing the people and organizations that generate this material in the real world.

III. Fighting Other Harmful Content

Facebook recognizes that terrorist content is not the only threat to our users' safety and well-being. There will always be people who try to use our platforms to spread hate. And we have seen foreign actors trying to interfere with elections by sowing division and spreading false information.

We are also working to address new tools of distortion, including manipulated media. We are developing technologies to identify manipulated content, dramatically reduce its distribution, and provide additional context to inform our community about its falsity. And we have partnered with outside fact-checkers, researchers, and our colleagues across the industry to help with these efforts. We know that people want to see accurate information on Facebook, so we will continue to make fighting misinformation a priority.

We take all of these problems very seriously. Hate of any kind has no place on Facebook. Any organization or individual that espouses violence or hatred violates our standards. A few months ago, we updated our policies to make it clear that all praise, support, and representation of white nationalism or white separatism, in addition to white supremacy, violates our rules. Any such content is removed from our platform under our Dangerous Organizations and Individuals policy. And Facebook does not tolerate attempts to undermine the integrity of an election or suppress voter turnout. These issues are difficult, but we will continue to work to craft policies that protect people; to apply those policies consistently and without bias; and to give voice to a community that transcends regions, cultures, and languages.

IV. Conclusion

Security is an arms race and our adversaries are always evolving their tactics. We constantly have to improve to stay ahead. Though we will never be perfect, we have made progress. And we are committed to tirelessly combating extremism on our platform by regularly reviewing our policies, adopting technical solutions, and strengthening our partnerships with external stakeholders. I appreciate the opportunity to be here today, and I look forward to your questions.