

HEARING BEFORE THE UNITED STATES SENATE COMMITTEE ON COMMERCE, SCIENCE, & TRANSPORTATION

September 18, 2019

Testimony of Monika Bickert

Vice President for Global Policy Management and Counterterrorism, Facebook

I. Introduction

Chairman Wicker, Ranking Member Cantwell, 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 and Counterterrorism 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 interactions with law enforcement. 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 everyone at Facebook, I would like to express our sympathy and solidarity with the victims, families, communities, and everyone else affected by the recent terrible attacks across the country. In the face of such heinous acts, we remain committed to cooperating with law enforcement and standing with our community against hate and violence. We are thankful to be able to provide a way for those affected by the horrific recent attacks to communicate with loved ones, to organize events for people to gather and grieve, and to raise money to help support these communities as they begin to heal.

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 connect and share with one another. But people need to feel safe in order to build this community. That is why Facebook prohibits harmful conduct on its platform, including hate speech and inciting violence. Our goal is to ensure that Facebook is a place where both expression and personal safety are protected and respected.

We are not aware of any connection between our platform and the recent attacks, but we recognize that we all have a role to play in keeping our communities safe. At Facebook, we have strong policies and invest significant resources to protect our users on and offline.

II. Facebook's Policies Against Hate and Violence

Facebook is committed to protecting our community by removing any content from our services that encourages real-world harm. Because harmful content can take many forms,

we have several policies in place to address these issues, all of which are published in our Community Standards, which define the content that is and is not allowed on our platform.

When we find content that violates our standards, we remove it. We invest in technology, processes, and people to help us identify violations and act quickly to mitigate any impact. There is always room for improvement, but we remove millions of pieces of content every year, much of it before any user reports it. We outline below several of the important steps that we take to prevent violence and keep our users safe.

Prohibition Against Violence and Incitement: We care deeply about our users and we want them to be safe. Therefore, it is critical to our mission to help prevent potential offline harm that may be related to content on Facebook. We remove content, disable accounts, and work with law enforcement when we believe there is a risk of physical harm or direct threats to public safety.

Prohibition of Dangerous Individuals and Organizations: In an effort to prevent and disrupt real-world harm, we do not allow any individuals or organizations that proclaim a violent mission, advocate violence, or are engaged in violence to have a presence on Facebook for any purpose, even if it appears benign. This includes organizations or individuals involved in the following:

- Terrorist activity, both domestic and international;
- Organized hate, including white supremacy and white nationalism;
- Human trafficking; and
- Organized violence or criminal activity.

We do not allow propaganda or symbols that represent any of these organizations or individuals to be shared on our platform unless they are being used to condemn or inform—for example, by media organizations. We do not allow content that praises any of these organizations or individuals or any acts committed by them. And we do not allow coordination of support for any of these organizations or individuals or any acts committed by them.

No Promoting or Publicizing Crime: We prohibit people from promoting or publicizing violent crime, theft, and/or fraud because we do not want to condone this activity and because there is a risk of copycat behavior. We also do not allow people to depict criminal activity or admit to crimes they or their associates have committed.

Policies Against Coordinating Harm: In an effort to prevent and disrupt real-world harm, we prohibit people from facilitating or coordinating future activity, criminal or otherwise, that is intended or likely to cause harm to people, businesses, or animals. People can draw attention to harmful activity that they may witness or experience as long as they do not advocate for or coordinate harm.

Combatting Suicide and Self-Injury: We also use and continue to develop tools and resources to proactively identify and help people who may be at risk of suicide or self-injury. We leverage pattern recognition technology to detect posts or live videos where someone might be expressing an intent to harm themselves. We also use artificial intelligence (AI) to prioritize the order in which our team reviews reported content relating to suicide or self-injury. This ensures we can get the right resources to people in distress and, where appropriate, we can more quickly alert first responders. And we remove content that encourages suicide or self-injury, including certain graphic imagery and real-time depictions that experts tell us might lead others to engage in similar behavior. We also work with organizations around the world to provide assistance and resources to people in distress.

Cooperation with Law Enforcement: Law enforcement plays a critical role in keeping people safe, and we have a long history of working successfully with law enforcement to address a wide variety of threats. As a former federal prosecutor, I know that this cooperation is vital. When we do receive reports or otherwise find content that violates our policies, we remove it. And we proactively reach out to law enforcement if we see a credible threat of imminent harm.

III. Facebook’s Efforts to Combat Violence and Hate

Our efforts to combat violent and hateful content are focused in three areas: developing new technical capabilities for our products, investing in people, and building partnerships.

Product Enhancements: Facebook has invested significantly in technology to help meet the challenge of proactively identifying violent content, including through the use of AI and other automation. These technologies have become increasingly central to keeping hateful and violent content off of Facebook.

We use a wide range of technical tools to identify violent and hateful content. This includes hashes—or digital fingerprints—that allow us to find secondary versions of known bad content; text parsing; digital “fan-outs” to identify profiles, groups, and pages related to those we have identified as problematic; and more holistic machine learning that can assess all aspects of a post and score whether it is likely to violate our Community Standards.

We also know that bad actors adapt as technology evolves, and that is why we constantly update our technical solutions to deal with more types of content in more languages, and to react to the new ways our adversaries try to exploit our products. For example, in response to the tragic events in Christchurch, we made changes to Facebook Live to restrict users if they have violated certain rules—including our Dangerous Organizations and Individuals policy. We now 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. We have also updated our proactive detection systems and reduced the average time it takes for our AI to find a violation on Facebook Live to 12 seconds—a 90% reduction in our average detection time from a few

months ago. Being able to detect violations sooner means that in emergencies where every minute counts, we can assist faster.

Investments in People: We know that we cannot rely on AI alone to identify potentially violent 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 acts of violence. To review those reports, and to prioritize the safety of our users and our platform more generally, 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.

We also have a team of more than 350 people at Facebook whose primary job is dealing with terrorists and other Dangerous Individuals and Organizations. This team includes language and cultural specialists, former law enforcement and intelligence professionals, and academics that have studied these groups and individuals for years. Many of them came to Facebook specifically because they are committed to the mission of keeping people safe.

This team was previously focused on counterterrorism, and we used our most sophisticated tools to predominantly combat ISIS, al-Qaeda, and their affiliates, which were recognized then as posing the greatest threats to our global community. Now, they lead our efforts against all people and organizations that proclaim or are engaged in violence. We are taking the initial progress we made in combatting content affiliated with ISIS, al-Qaeda, and their affiliates, and we are further building out techniques to identify and combat the full breadth of violence and extremism covered under our Dangerous Organizations policy.

Partnerships: We are proud of the work we have done to make Facebook a hostile place for those committed to acts of violence. We understand, however, that simply working to keep violence off Facebook is not an adequate solution to the problem of online extremism and violence, particularly because bad actors can leverage a variety of platforms. We believe our partnerships with other companies, civil society, researchers, and governments are crucial to combatting this threat. For example, our P2P Global Digital Challenge, which engages university students around the world in competitions to create social media campaigns and offline strategies to challenge hateful and extremist narratives, has launched over 600 counterspeech campaigns from students in 75 countries, engaged over 6,500 students, and reached over 200 million people. We're also partnering with Life After Hate, an organization founded by former violent extremists, to connect people who search for terms associated with white supremacy to resources focused on helping people leave behind hate groups.

Our work to combat violence is never complete. Individuals and organizations intent on violent acts come in many ideological stripes—and the most dangerous among them are deeply resilient. We know that bad actors will continue to attempt to skirt our detection with more sophisticated efforts, and we are dedicated to continuing to advance our work and share our progress.

IV. Conclusion

Facebook is committed to helping people build a vibrant community that encourages and fosters free expression. At the same time, we want to do what we can to protect our users from real-world harm and stop terrorists, extremists, hate groups, and any others from using our platform to promote or engage in violence. We recognize that there is always more work to do in combatting the abuse of our site by bad actors, but we are proud of the progress we have made over the last few years. We know that people have questions about what we are doing to continue that progress, and we look forward to working with this Committee, regulators, and others in the tech industry and civil society to continue working on these issues. I appreciate the opportunity to be here today, and I look forward to your questions.