25 Massachusetts Ave., NW
Washington, DC 20001
Phone: 202-346-1100

October 9, 2020

The Honorable Drew Darby
Chairman
House Mass Violence Prevention &
Community Safety Select Committee
Room GW.17
P.O. Box 2910
Austin, TX 78768

Dear Mr. Chairman and Members of the Committee:

Thank you for the opportunity to submit testimony to the committee and to participate in these continuing discussions on the important issues of mass violence prevention and community safety.

At Google, we believe the Internet has been a force for creativity, learning and access to information. Supporting this free flow of ideas is core to our mission to organize and make the world's information universally accessible and useful. Laws strongly protect free expression, yet there are limits such as in the case of violence or extremist speech. This is true both online and off. We take these issues seriously and want to be a part of the solution.

In this testimony to you I will focus on two key areas where we are making progress to help protect people: (i) how we work with governments and law enforcement, and (ii) the enforcement of our policies around terrorism and hate speech.

Working with Government and Law Enforcement

Google appreciates that law enforcement agencies face significant challenges in protecting the public against crime and terrorism. We have long worked with law enforcement to address the policing and judicial needs and requirements related to our products. We have an entire legal team on call around the clock devoted to triaging and responding to law enforcement data production requests. Following the Governor's two round tables last summer, we engaged with the Department of Public Safety staff to answer legal and process questions about user and product information.

Our legal team also has done regular outreach to law enforcement to explain our policies, products and procedures. We always participate and give data disclosure presentations and investigation lab sessions at the annual Dallas Crimes Against Children conference. At this year's Dallas conference, which was held virtually, we presented on Google's data disclosure policies to over 700 law enforcement officers. We also did Google-only outreach in March of last year where we gave multiple presentations to approximately 400-500 Texas law enforcement officers, both state and federal -- specifically, the Northern, Western and Eastern Districts of Texas, as well as presentations to state officers in Dallas and Austin. The state coverage included Austin PD, Dallas PD, Plano PD, Texas AG Criminal Investigations Division, Texas DPS-Criminal Investigations, among others.

In terms of our internal process with regard to threats, when we become aware of statements on our platform that constitute a threat to life or that reflect that someone's life may be in danger, we report this activity to law enforcement agencies. For example, when we have a good faith belief that there is a threat to life or serious bodily harm made on our platform in the United States, the Google CyberCrime Investigation Group (CCIG) will report it to the Northern California Regional Intelligence Center (NCRIC). In turn, NCRIC quickly gets the report into the hands of officers to respond. CCIG is on call 24/7 to make these reports.

Under U.S. law, the Stored Communications Act allows Google and other service providers to voluntarily disclose user data to governmental entities in emergency circumstances where the provider has a good faith belief that disclosing the information will prevent loss of life or serious physical injury to a person. Our team is staffed on a 24/7/365 basis to respond to these emergency disclosure requests (EDRs). We have seen significant growth in the volume of EDRs that we receive from US governmental entities, as illustrated in our transparency report covering government requests for user data[1]. In fact, the number of EDRs submitted from agencies in the US more than doubled from 2017 to 2019. We have expanded our teams to accommodate this growing volume and to ensure we can quickly respond to emergency situations that implicate public safety.

We are also deeply committed to working with government, the tech industry and experts from civil society and academia to protect our services from being exploited by bad actors. An example of this collaboration stems from the tragic events in Christchurch, New Zealand in March, 2019. Those events presented unique challenges, and we had to take unprecedented steps to address the sheer volume of new videos related to the events. In the months following, Google and YouTube signed the Christchurch Call to Action, a series of commitments to quickly and responsibly address terrorist content online. This is an extension of our ongoing commitment to working with our colleagues in the industry to address the challenges of terrorism online. Since 2017, we've done this through the Global Internet Forum to Counter Terrorism (GIFCT), of which Google is a founding company and was its first chair. GIFCT has adopted joint content incident protocols for responding to emerging or active events. GIFCT has also released a transparency report[2] and a counterspeech campaign

[1] https://transparencyreport.google.com/user-data/overview?hl=en&user_requests_report_period=series:requests,accounts;authority:US;time:&lu=legal_process_breakdown&legal_process_breakdown=expanded:0
[2] https://gifct.org/documents/14/Transparency_Report.pdf

toolkit that will help activists and civil society organizations challenge the voices of extremism online.

Policies and Enforcement on YouTube for Terrorism and Hate Speech

We have robust policies and programs to defend our platforms to spread hate or incite violence. This includes prohibitions on: terrorist recruitment, violent extremism, incitement to violence, glorification of violence, and instructional videos related to acts of violence. We apply these policies to violent extremism of all kinds, whether inciting violence on the basis of race or religion or as part of an organized terrorist group.

In order to improve the effectiveness of our policy enforcement, we have invested heavily in both technology and people to quickly identify and remove content that violates our policies against incitement to violence and hate speech[3]:

1) YouTube's enforcement system starts from the point at which a user uploads a video. If our technology detects that the video is similar to videos that we know already violate our policies, it is sent for humans to review. If they determine that it violates our policies, they remove it and the system makes a "digital fingerprint" or hash of the video so it can't be uploaded again.

2) Machine learning technology also helps us more effectively identify this content and enforce our policies at scale. However, because hate and violent extremism content is constantly evolving and can sometimes be context-dependent, we also rely on experts to help us identify policy-violating videos. Some of these experts sit at our intel desk, which proactively looks for new trends in content that might violate our policies.

3) This broad cross-sectional work has led to tangible results. Between April and June 2020, YouTube removed over 11.4 million videos for violating our guidelines, of which 95% of these videos were first flagged by machines rather than humans. Of those detected by machines, 53% never received a single view, and just over 81% received fewer than 10. Overall, videos that violate our policies generate a fraction of a percent of the views on YouTube.

Our efforts do not end there, as we are constantly evolving to new challenges and looking for ways to improve our policies. Last year YouTube implemented updates to its hate speech policy to specifically prohibit videos alleging that a group is superior in order to justify discrimination, segregation or exclusion based on qualities like age, gender, race, caste, religion, sexual orientation or veteran status. This would include, for example, videos that promote or glorify Nazi ideology, because it is inherently discriminatory. YouTube also updated its policies to prohibit content denying that well-documented violent events, like the Holocaust or the shooting at Sandy Hook Elementary, took place.

[3] https://www.youtube.com/howyoutubeworks/policies/community-guidelines/

The updated hate speech policy was launched in early June, and as our teams review and remove more content in line with the new policy, our machine learning algorithms will improve in tandem to help us identify and remove such content. Though it can take months for us to ramp up enforcement of a new policy, the profound impact of our hate speech policy update was already evident in the data released the quarter following the update: the number of individual video removals for hate speech saw a 5x spike to over 100,000, the number of channel terminations for hate speech also saw a 5x spike to 17,000, and the total comment removals nearly doubled in Q2 to over 500 million due in part to a large increase in hate speech removals. Additional statistics can be viewed on the YouTube Community Guidelines Enforcement page[4].

Conclusion

We take the safety of our users very seriously and value our close and collaborative relationships with law enforcement and government agencies. We have invested substantial resources to tackle the problem of hate speech. At present, we spend hundreds of millions of dollars annually and have more than 10,000 people working across Google to address content that might violate our policies, which include our policies against promoting violence and terrorism.

We understand these are difficult issues and want to be responsible actors who are a part of the solution. As these issues evolve, Google will continue to invest in the people and technology to meet the challenge.

Sincerely,

Ron Barnes
Head of State Legislative Affairs

[4] https://transparencyreport.google.com/youtube-policy/removals?hl=en&total_removed_videos=period:Y2020Q2;exclude_automated:all&lu=total_removed_videos