



April 24, 2019

Twitter, Inc.
1355 Market St #900
San Francisco, CA 94103

Dear Chairman Rose, Rep. Lee, Rep. Langevin and Rep. Slotkin,

I write in response to your letter regarding the important issues of fighting violent extremism online and combating terrorist use of the Internet. Terrorist and violent extremist content is fundamentally contrary to our values, our business interest is not served, it breaks the Twitter Rules, and it is a harrowing experience for our users.

All individuals accessing or using Twitter's services must adhere to the policies set forth in the Twitter Rules. Our policies and enforcement options evolve continuously to address emerging behaviors online.

Policies on Terrorism. Individuals are prohibited from making specific threats of violence or wish for the serious physical harm, death, or disease of an individual or group of people. This includes, but is not limited to, threatening or promoting terrorism. A person on Twitter also may not affiliate with organizations that — whether by their own statements or activity both on and off the platform — use or promote violence against civilians to further their causes.

We have now suspended more than 1.4 million accounts for violations related to the promotion of terrorism between August 1, 2015 and June 30, 2018. During the most recent reporting period of January 1, 2018, through June 30, 2018 we suspended a total of 205,156 accounts. Of those suspensions, 91% consisted of accounts that were proactively flagged by internal, proprietary tools.

Policies on Violent Extremist Groups. We do not tolerate groups or individuals associated with them who engage in and promote violence against civilians both on and off the platform. Accounts affiliated with groups in which violence is a component of advancing their cause risk having a chilling effect on opponents and bystanders. The violence that such groups promote could also have dangerous consequences offline, jeopardizing their targets' physical safety.

We prohibit the use of Twitter's services by violent extremist groups. We consider violent extremist groups to be those which identify through their stated purpose, publications, or actions, as an extremist group; have engaged in, or currently engage in, violence (and/or the promotion of violence) as a means to further their cause; and target civilians in their acts (and/or promotion) of violence. An individual on Twitter may not affiliate with organizations that — whether by their own statements or

activity both on and off the platform – use or promote violence against civilians to further their causes.

Policies on Hateful Conduct. An individual on the platform is not permitted to promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories.

We do not allow individuals to use hateful images or symbols in their profile image or profile header. Individuals on the platform are not allowed to use the username, display name, or profile bio to engage in abusive behavior, such as targeted harassment or expressing hate towards a person, group, or protected category.

Under this policy, we take action against behavior that targets individuals or an entire protected category with hateful conduct. Targeting can happen in a number of ways, for example, mentions, including a photo of an individual, or referring to someone by their full name.

When determining the penalty for violating this policy, we consider a number of factors including, but not limited to the severity of the violation and an individual's previous record of rule violations. For example, we may ask someone to remove the violating content and serve a period of time in read-only mode before they can Tweet again. Subsequent violations will lead to longer read-only periods and may eventually result in permanent account suspension. If an account is engaging primarily in abusive behavior, or is deemed to have shared a violent threat, we will permanently suspend the account upon initial review.

* * *

In addition to our policies, we continue to work closely with industry peers, and have made a dramatic investment in corporate resources directed at reviewing and removing any problematic content on our service, including that content associated with terrorism. Additionally, we have participated in more than 100 countering violent extremism trainings with external partners since 2015.

Information Sharing with Industry Peers. We continue to work closely with our industry peers and civil society to address common threats from terrorism globally. In June 2017, we co-launched the Global Internet Forum to Counter Terrorism ("GIFCT"), a partnership among Twitter, YouTube, Facebook, and Microsoft. GIFCT facilitates, among other things, information sharing; technical cooperation; and research collaboration. In the wake of the attack, Twitter and our agency peers shared information, hashes, and unique URLs to more efficiently and effectively remove terrorist content from our platform.

Corporate Investment. We have made the health of Twitter our top priority, and our efforts will be measured by how we help encourage more healthy debate, conversations, and critical thinking on the platform. Conversely, abuse, automation, hateful conduct, terrorism, and manipulation will detract from the health of our platform.

Twitter uses a combination of machine learning and human review to adjudicate abuse reports and whether they violate our rules. Twitter employs extensive content detection technology to identify and police harmful and abusive content embedded in various forms of media on the platform. We use PhotoDNA and hash matching technology, particularly in the context of terrorism or child sexual exploitation. We use these technologies to identify previously identified content in order to surface it for agent review, however, if it is the first time that an image has been seen, it would not necessarily be subject to our technology. It is important to note that we continually expand our databases of known violative content.

Our chief executive officer reorganized the structure of the company to allow our valued employees greater durability, agility, invention, and entrepreneurial drive. The reorganization simplified the way we work, and enabled all of us to focus on health of our platform. At the end of the first quarter 2019, we have a global workforce of 4,100, a substantial portion of whom are directly involved in reviewing and moderating content on Twitter.

Putting a dollar amount on our broader efforts is a complex request. Where we have made acquisitions of companies, where staff work on a series of overlapping issues, where we have invested in technology and tools to support our teams reviewing content — all of these involve expenditure that is not limited to just counterterrorism efforts, but all of this work is leveraged in furtherance of our desire to rid Twitter of terrorist and violent extremist content.

With respect to our teams, we have not structured our teams so that people only work on one certain type of content, be that white supremacist or jihadi content. Tools developed to address one challenge (spam, for instance) have been refined to address other challenges, like violent extremism. Colleagues bring a wide range of expertise and knowledge from a range of disciplines and we continue to believe that is the best way to invest, mindful that the area we hire for today is likely to be different than the challenges that lie ahead.

One of the underlying features of our approach is that we take a behavior-first approach. That is to say, we look at how accounts behave before we look at the content they are posting. This is how we were able to scale our efforts to combat ISIS content far faster than we would have been if we had relied on experts reviewing the content produced by ISIS. As we developed a greater understanding of how ISIS-linked accounts operated, we could further invest in technology to identify accounts,

ultimately allowing us to take action on a large number of accounts before they had even Tweeted.

Our response to the challenges of terrorism and violent extremism on our platform are a company-wide effort and not viewed in isolation from our wider mission to improve the health of the public conversation. We cannot solve the problems posed by violent extremist ideologies by removing content alone, particularly given the clear migration of these bad actors to smaller platforms who do not share their peers commitment to solving this problem.

External Partners. Twitter has consulted with a wide array of experts around the globe to inform our approach. In addition, our team has participated in more than 100 countering violent extremism trainings since 2015, including events in New York, Beirut, Bosnia, Belfast, and Brussels. We have also participated in summits at the White House, the United Nations, London, and Sydney. The trainings work to empower non-governmental organizations and activists across the world to harness the power of our platform to offer positive, alternative narratives to those at risk and their wider communities.

* * *

The issue of terrorism remains one of our highest priorities and we welcome the opportunity to work with you and other Members of Congress to discuss the ways in which we can work together to curb this type of activity that poses a risk to people both on and off the platform.

Sincerely,

A handwritten signature in black ink, appearing to read 'C. Monje, Jr.', with a long horizontal stroke extending to the right.

Carlos Monje, Jr.

Director of Public Policy & Philanthropy, United States & Canada