U.S. Senate Committee on Commerce, Science, and Transportation
Hearing on Terrorism and Social Media
January 17, 2018

Statement of
Carlos Monje, Jr.
Director, Public Policy & Philanthropy, U.S. & Canada
Twitter, Inc.

Thank you Chairman Thune, Ranking Member Nelson, and distinguished members of the Committee and staff.

Twitter has been at the forefront of responding to the evolving challenge of preventing terrorist exploitation of the Internet. Our work in this area will never be complete, as the threats we face constantly evolve. As new challenges emerge, we will continue our efforts to both ensure terrorists don't have a place on Twitter while giving voice to those who promote positive messages for the future.

The Twitter Rules

To be clear, terrorist organizations have no place on Twitter and the promotion of terrorism is against our Rules. The Twitter Rules make clear:

- *You may not make 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. You 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.*

Moreover, our Rules prohibit content that glorifies violence or the perpetrators of a violent act. This includes celebrating any violent act in a manner that may inspire others to replicate it or any violence where people were targeted because of their membership or inclusion in a protected group.

Terrorist Content Removals

Beyond having a clear policy against the promotion of terrorism, we have been tackling the issue of terrorist content on the Twitter platform for many years. As our biennial Twitter Transparency Reports indicate we have made steady progress in this area:

- In 2015, 67,069 accounts suspended
- In 2016: 569,202 accounts suspended
- In 2017: 574,070 accounts suspended

In total, we have suspended more than 1.1 million terrorist accounts since mid-2015.

Technology

Our progress fighting terrorists on Twitter is due to the commitment we have made internally to harness innovation to address the tactics employed by terrorist organizations on our platform. While there is no "magic algorithm" for identifying terrorist content on the Internet, we have increasingly tapped technology in efforts to improve the effectiveness of our in-house proprietary anti-spam technology. This technology supplements reports from our users and dramatically augments our ability to identify and remove violative content from Twitter. Through these efforts we have found success in preventing recently suspended users from coming back onto Twitter.

At the beginning of 2015, this technology was being used to detect roughly one-third of the terrorist accounts we suspended at that time. Last year, that number increased to more than 90 percent of suspensions being flagged by our internal tools. Three-quarters of those suspensions were flagged before the account had a chance to Tweet even once.

As is the case with a determined adversary, as we make it harder for terrorists to use Twitter, their behavior evolves. To stay in front of this, we continue to invest in technology to prevent new accounts being opened to replace those we suspend, while also developing further the tools that prevent the distribution of propaganda in the aftermath of attacks.

Industry Collaboration

Because this is a shared challenge, our industry has established the Global Internet Forum to Counter Terrorism (GIFCT), which has focused on learning and collaboration; technical cooperation; and research. This builds on previous work undertaken by the EU Internet Forum and follows constructive discussions held at the UN General Assembly and the G7 Interior Ministers meeting.

Twitter sees the GIFCT as a substantial opportunity to ensure that smaller companies are not soft targets for terrorists and that the learnings that we have developed are shared and built upon. The GIFCT's initial goal for 2017 was to work with 50 smaller tech companies to to share best practices on how to disrupt the spread of violent extremist material. We have exceeded that goal, engaging with 68 companies over the past several months.

In the coming months, we plan to deepen this collaboration with smaller companies, working directly to educate them about potential technological approaches, sharing expertise from our own operational teams and allowing them to develop a peer network across industry to support their work.

Twitter Countering Violent Extremism Trainings
The GIFCT, through its work with the Tech Against Terrorism and ICT4Peace projects, is a further avenue through which best practices can be shared and our existing company efforts can be further scaled-up. Twitter has participated in more than 100 CVE trainings and events since 2015, including events in Beirut, Bosnia, Belfast and Brussels and summits at the White House, the United Nations and in London and Sydney.

We work with respected organizations such as Parle-moi d'Islam (France), Active Change Foundation (UK), Wahid Foundation (Indonesia), The Sawab Center (UAE), and True Islam (US) to empower credible non-governmental voices against violent extremism. We also continue to play an active role in the task force created by the French Interior Ministry and have attended government-convened summits on CVE hosted by the French Interior Ministry and the Indonesian National Counterterrorism Agency.

We supported the Institute for Strategic Dialogue's "Against Violent Extremism" project, the results of which were published in a report, "The Impact of Counternarratives." The project used *pro bono* Twitter advertising to increase the reach of key NGOs. The campaigns yielded real results. One NGO participant, Average Mohamed, doubled its number of Twitter followers and another, ExitUSA, tripled its Twitter followers.

We also are a member of the Anti-Defamation League's Cyberhate Problem-Solving Lab, which works collaboratively to counter hate speech online.

Extremism
Late last year we broadened our rules to encompass accounts affiliated with violent extremist groups and to cover violent content or hateful imagery displayed in profile information. Our prohibition on the use of Twitter's services by violent extremist groups — i.e., identified groups subscribing to the use of violence as a means to advance their cause — applies whether the purpose or cause of any such group is a political, religious, or social objective.

Accounts affiliated with groups or organizations in which violence is a component of advancing their cause risk having a chilling effect on opponents of that cause who may want to comment on Twitter. In addition, the violence that such groups promote online

could also have dangerous consequences offline, potentially putting the physical safety of Twitter users and others in jeopardy.

The broadening of our policies against violent extremism also includes covering any account that abuses or threatens others through their profile information. In other words, if an account's profile information includes a violent threat or multiple slurs, racist or sexist tropes, or incites fear or otherwise dehumanizes another person, it will be removed. Further, hateful imagery will now be included in the category of "sensitive media" under our rules. This change means that logos, symbols, or images whose purpose is to promote hostility and malice to others based upon their race, religion, disability, sexual orientation, or ethnicity will now be actionable and we will require accounts displaying such imagery to remove such violative media content.

Misinformation
As we have previously described, Twitter has also devoted significant resources to the issue of misinformation and interference in the election context by foreign state actors. We have sought through our Information Quality initiative to enhance our ability going forward to detect and stop such activity and to do our part to protect the democratic process from interference and abuse. We have also undertaken a retrospective review to further the public's understanding of what happened in the 2016 election. As we explained last year, we expect to keep Congress updated on the latest results of that ongoing review as our work progresses. And we made the decision last year not only to offboard both RT and Sputnik as advertisers on our platform, but also to commit to donate the revenue we received from those sources to research into elections and civic engagement on Twitter. We have begun to scope such research needs and are in dialogue with several academic researchers and NGOs in this area. We take these issues seriously and our efforts to address them remain among our highest priorities.

Preparing for the U.S. Midterms
Since 2016 we've had additional elections around the world — such as in France, Germany, and South Korea during 2017 — and we have midterm elections approaching this November in the United States.

To prepare for the U.S. midterm elections, we have organized internally to ensure that our teams are working to address election-related issues as they may arise. Our cross-functional elections task force will be prepared to:

- Verify major party candidates for all statewide and federal elective offices, and major national party accounts, as a hedge against impersonation;

- Maintain open lines of communication to federal and state election officials to quickly escalate issues that arise;
- Address escalations of account issues with respect to violations of Twitter Rules or applicable laws;
- Continually improve and apply our anti-spam technology to address networks of malicious automation targeting election-related matters;
- Monitor trends and spikes in conversations relating to the 2018 elections for potential manipulation activity; and
- Implement our Ads Transparency Center to bring transparency to voters about political ads they see on Twitter.

The companies here today have both shared and unique challenges. And while we are competitors in the marketplace, we are close partners in combating the threat of extremism or those who would do harm to our democratic process. We value the collaboration with our industry peers, and coordinated efforts are driving further progress to degrade the presence of content promoting terrorism.

Thank you, and I look forward to this discussion.

<div align="center">####</div>