

Written Testimony of Juniper Downs
Director, Public Policy and Government Relations
Senate Commerce Committee
Hearing on “Combating the Spread of Extremist Propaganda”
January 17, 2018

Chairman Thune, Ranking Member Nelson, and distinguished Members of the Committee: thank you for the opportunity to testify at today’s hearing and for your leadership on these difficult issues. My name is Juniper Downs and I serve as the global policy lead for YouTube.

At YouTube, we believe the world is a better place when we listen, share, and build community through our stories. Our mission is to give everyone a voice and show them the world. With this comes many benefits to society — unparalleled access to art and culture, news and entertainment, and educational materials. To put our work in context, it’s important to recognize the scale and goal of our services. More than one and a half billion people come to YouTube every month. We see well over 400 hours of video uploaded every minute. Most of this content is perfectly benign — beauty vlogs, music, comedy. Digital platforms have also become a place for breaking news, exposing injustices, and sharing content from previously inaccessible places.

We value this openness. It has democratized how stories, and whose stories, get told. And has created a platform where anyone can be a creator and can succeed. We are aware, however, that the very platforms that have enabled these societal benefits may also be abused by those who wish to promote hatred or extremism. These challenges are constantly evolving and changing, so our commitment to combat them is similarly sustained and unwavering. To be very clear: using YouTube to incite violence, spread violent extremist propaganda, recruit for terrorism, or celebrate or promote terrorist attacks is strictly and specifically prohibited by our terms of service.

To that end, I am pleased to have this opportunity to outline the approach we have taken on these issues. We have developed rigorous policies and programs to defend the use of our platforms from the spread of hate and incitement to violence. We continue to refine them as we adapt to new and evolving threats. For example, YouTube has long had policies that prohibit terrorist content. This includes: terrorist recruitment, violent extremism, incitement to violence, and instructional content that could be used to facilitate substantial bodily injury or death. Extremism and violence are

not confined to any one community. 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. When we become aware of content that violates these policies, we immediately remove it. Any channel that is dedicated to such content is terminated. We don't allow Foreign Terrorist Organizations (FTOs) to use Google at all — if an account is created by an FTO or its agent, we terminate immediately, regardless of the content it may be sharing.

We also have a strict set of policies for monetizing content on YouTube. We recognize there may be videos that don't break our Community Guidelines, but which advertisers would not want to advertise against. We give advertisers the tools to control where their ads appear.

We use a mix of technology and humans to remove violative content quickly. Users can alert us to content that they think may violate our policies through a flag found below every YouTube video. We have teams charged with reviewing flagged content 24/7 in multiple languages and countries around the world. We also work closely with members of our Trusted Flagger program, which is comprised of NGOs and government agencies with specific expertise who are provided a bulk-flagging tool to alert us to content that may violate our policies. And of course we rely upon our technology, which has always been a critical part of our solution. Our video-matching techniques, for example, can prevent the dissemination of violative content by catching re-uploads of known bad content before it is public.

Nonetheless, given the evolving nature of the threat, it is necessary for us to continue enhancing our systems. We know that no enforcement regime will ever be 100% perfect. Over the past year in particular, we have taken several steps to build on our efforts:

- The first is an investment in machine learning technologies for the detection and removal of violent extremist videos. We have been working on machine learning for years, and recently deployed classifiers that detect terrorist material and flag it for review. Since June, our teams have manually reviewed approximately two million videos to improve this flagging technology by providing large volumes of training examples. Machine learning is now helping our human reviewers remove nearly five times as many videos in violation of our policies than they were previously. Last June, only 40% of the videos we removed for violent extremism were identified by our algorithms. Today, that number is 98%. Our advances in

machine learning let us now take down nearly 70% of violent extremism content within 8 hours of upload and nearly half of it in 2 hours.

- Second, we are focused on improving and expanding our expertise and resources on these issues. We expanded our Trusted Flagger Program to an additional 50 NGOs in 2017, including to groups like Anti-Defamation League and several counter-terrorism experts such as the Institute of Strategic Dialogue and International Centre for the Study of Radicalization. Working with these organizations helps us to better identify emerging trends and understand how these issues manifest and evolve. In 2018, we will have 10,000 people across Google working to address content that might violate our policies. This includes engineers and reviewers who work around the world, 24/7, and speak many different languages.
- We are taking a tougher stance on videos that may be offensive, but do not violate our policies. Our Community Guidelines prohibit hate speech that either promotes violence or has the primary purpose of inciting hatred against individuals or groups based on certain attributes. Some borderline videos, such as those containing inflammatory religious or supremacist content without a direct call to violence or a primary purpose of inciting hatred, may not cross these lines for removal. But we understand that these videos may be offensive to many and have developed a new treatment for them. Identified borderline content will remain on YouTube behind an interstitial, won't be recommended, won't be monetized, and won't have key features including comments, suggested videos, and likes. Initial uses have been positive and have shown a substantial reduction in watch time of those videos.
- Greater Transparency. We understand that people want a clearer view of how we're tackling problematic content. That's why in 2018, we will be creating a report to provide more aggregate data about the flags we receive and the actions we take to remove videos and comments that violate our content policies.
- Finally, we are creating programs to promote counterspeech on our platforms. We are expanding our counter-extremism work to present counternarratives and elevate the voices that are most credible in speaking out against terrorism, hate, and violence.
 - For example, our Creators for Change program supports creators who are tackling social issues, including extremism and hate, by building empathy

and acting as positive role models. There are 60 million video views of Creators for Change content to date; 731,000 total watch time hours of Creators for Change content; and through ‘Local chapters’ of Creators for Change, creators tackle social challenges specific to different markets.

- Google’s Jigsaw group, an incubator to tackle some of the toughest global security challenges, has deployed the Redirect Method, which uses Adwords targeting tools and curated YouTube videos uploaded to disrupt online radicalization. It focuses on the slice of ISIS’s audience that is most susceptible to its messaging and redirects them towards YouTube playlists of videos debunking ISIS recruiting themes.

We also collaborate across the industry. In 2016, we created a hash-sharing database with Facebook, Twitter and Microsoft, where we share hashes (or “digital fingerprints”) of terrorist content to stop its spread across platforms. Using other companies to give us notice is effective because of the counter-terrorism research showing the pattern of cross-platform abuse and the particularly dangerous nature of this content. We added 7 companies to this coalition in 2017 and our shared database contains over fifty thousand videos and image hashes. Last summer, we announced the Global Industry Forum to Counter Terrorism (GIFCT) to formalize industry collaboration on research, knowledge sharing, and technology. The GIFCT also set a goal of working with 50 smaller tech companies in 2017 to help them better tackle terrorist content on their platforms — and we exceeded that goal. To date, we’ve hosted 68 small companies at workshops through the Tech Against Terrorism Initiative, our partners under the UN Counter Terrorism Executive Directorate. We’ve held workshops for smaller companies in San Francisco and New York, Paris, Jakarta, London, and Brussels.

No single component can solve the problem in isolation. To get this right, we must all work together. Since June, YouTube has removed over 160,000 violent extremist videos and has terminated approximately 30,000 channels for violation of our policies against terrorist content. We achieved these results through tougher policies, enhanced enforcement by machines and people, and collaboration with outside experts. That has become the blueprint for how we tackle this challenge.

While Google’s services can provide real benefits to our users, we recognize that detecting and preventing misuse of those services is critically important. We are deeply committed to working with law enforcement, government, others in the tech industry, and the NGO community to protect our services from being exploited by bad actors. We will only make progress by working together to address these complex

issues at their root. That is why forums like this are so important to underscoring our shared goals and commitments. We look forward to continued collaboration with the Committee as it examines these issues.

Thank you for your time. I look forward to taking your questions.