AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

S. HRG. 115–316

TERRORISM AND SOCIAL MEDIA: #ISBIGTECHDOINGENOUGH?

HEARING

BEFORE THE

COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION UNITED STATES SENATE

ONE HUNDRED FIFTEENTH CONGRESS

SECOND SESSION

JANUARY 17, 2018

Printed for the use of the Committee on Commerce, Science, and Transportation

Available online: http://www.govinfo.gov

U.S. GOVERNMENT PUBLISHING OFFICE
31–316 PDF WASHINGTON : 2018

SENATE COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION

ONE HUNDRED FIFTEENTH CONGRESS

SECOND SESSION

JOHN THUNE, South Dakota, *Chairman*

ROGER WICKER, Mississippi	BILL NELSON, Florida, *Ranking*
ROY BLUNT, Missouri	MARIA CANTWELL, Washington
TED CRUZ, Texas	AMY KLOBUCHAR, Minnesota
DEB FISCHER, Nebraska	RICHARD BLUMENTHAL, Connecticut
JERRY MORAN, Kansas	BRIAN SCHATZ, Hawaii
DAN SULLIVAN, Alaska	EDWARD MARKEY, Massachusetts
DEAN HELLER, Nevada	TOM UDALL, New Mexico
JAMES INHOFE, Oklahoma	GARY PETERS, Michigan
MIKE LEE, Utah	TAMMY BALDWIN, Wisconsin
RON JOHNSON, Wisconsin	TAMMY DUCKWORTH, Illinois
SHELLEY MOORE CAPITO, West Virginia	MAGGIE HASSAN, New Hampshire
CORY GARDNER, Colorado	CATHERINE CORTEZ MASTO, Nevada
TODD YOUNG, Indiana	JON TESTER, Montana

NICK ROSSI, *Staff Director*
ADRIAN ARNAKIS, *Deputy Staff Director*
JASON VAN BEEK, *General Counsel*
KIM LIPSKY, *Democratic Staff Director*
CHRIS DAY, *Democratic Deputy Staff Director*
RENAE BLACK, *Senior Counsel*

CONTENTS

WITNESSES

APPENDIX

TERRORISM AND SOCIAL MEDIA: #ISBIGTECHDOINGENOUGH?

WEDNESDAY, JANUARY 17, 2018

U.S. SENATE,
COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION,
Washington, DC.

The Committee met, pursuant to notice, at 10 a.m. in room SR–253, Russell Senate Office Building, Hon. John Thune, Chairman of the Committee, presiding.

Present: Senators Thune [presiding], Wicker, Blunt, Cruz, Fischer, Moran, Heller, Inhofe, Lee, Capito, Gardner, Young, Nelson, Cantwell, Klobuchar, Blumenthal, Schatz, Markey, Udall, Peters, Baldwin, Hassan, Cortez Masto, and Tester.

OPENING STATEMENT OF HON. JOHN THUNE, U.S. SENATOR FROM SOUTH DAKOTA

The CHAIRMAN. Good morning. I want to thank everyone for being here to examine what social media companies are doing to combat terrorism, including terrorist propaganda and terrorist recruitment efforts, online.

The positive contributions of social media platforms are well documented. YouTube, Facebook, and Twitter, among others, help to connect people around the world, give voice to those oppressed by totalitarian regimes, and provide a forum for discussions of every political, social, scientific, and cultural stripe. These services have thrived online because of the freedom made possible by the uniquely American guarantee of free speech and by a light touch regulatory policy.

But, as is so often the case, enemies of our way of life have sought to take advantage of our freedoms to advance hateful causes. Violent Islamic terrorist groups like ISIS have been particularly aggressive in seeking to radicalize and recruit over the Internet and various social media platforms.

The companies that our witnesses represent have a very difficult task: preserving the environment of openness upon on which their platforms have thrived, while seeking to responsibly manage and thwart the actions of those who would use their services for evil. We are here today to explore how they are doing that, what works, and what could be improved.

Instances of Islamic terrorists using social media platforms to organize, instigate, and inspire are well documented. For example, the killer responsible for the Orlando nightclub shooting, in which 49 innocent people were murdered and 53 were injured, was report-

edly inspired by digital material that was readily available on social media.

And this issue is not new. Over the course of several years, YouTube hosted hundreds of videos by senior al Qaeda recruiter Anwar al-Awlaki. Although the company promised in 2010 to remove all videos that advocated violence, al-Awlaki's Call to Jihad video, in which he advocates for western Muslims to carry out attacks at home, remained on the site for years. In fact, a *New York Times* report suggested that al-Awlaki videos influenced the Fort Hood terrorist, the Boston Marathon bombers, and the terrorist attacks in San Bernardino and Orlando.

This issue is also international in scope. In response to recent terror attacks in London, British Prime Minister Theresa May has been especially outspoken in calling on social media platforms to do more to combat the kind of radicalization that occurs online. Last fall, for example, she was joined by other European leaders in calling upon social media companies to remove terrorist content from their sites within one to two hours after it appears.

As we'll hear today, the companies before us are increasingly using technology to speed up their efforts to identify and neutralize the spread of terrorist content. In a recent blog post, Facebook said that artificial intelligence now removes 99 percent of ISIS and al Qaeda related terror content even before it can be flagged by a member of the community and sometimes even before it can be seen by any users.

YouTube is also teaming up with Jigsaw, the in-house think tank of Google's parent company Alphabet, to test a new method of counter-radicalization referred to as the Redirect Method. Seeking to redirect or refocus potential terrorists at an earlier stage in the radicalization process, YouTube offers users searching for specific terrorist information additional videos made specifically to deter them from becoming radicalized.

A little over a year ago, Facebook, YouTube, Microsoft, and Twitter committed to sharing a database of unique hashes and digital fingerprints of some of the most extreme terrorist-produced content used for influence or recruitment. By cross-sharing this information, terrorist content on each of the hosts' platforms will be more readily identified, hopefully resulting in faster and more efficient deletion of this material.

Essentially, these companies are claiming they can tag individual videos and photos and, using automation, can kick them off their platforms before they are even seen. We all have a vested interest in their success, and I believe this Committee has a significant role to play in overseeing the effectiveness of their efforts.

I do want to thank Ms. Bickert, Ms. Downs, and Mr. Monje for being here as representatives of their companies.

To Mr. Watts, I look forward to hearing your thoughts about disrupting and defeating terrorism.

With that, I will now recognize the Ranking Member, Senator Nelson, for any opening statement he'd like to make.

Senator Nelson.

STATEMENT OF HON. BILL NELSON,
U.S. SENATOR FROM FLORIDA

Senator NELSON. Mr. Chairman, within a few hours of the Pulse nightclub shooting, I was there on South Orange Avenue in Orlando, and I just want to comment that when a great tragedy occurs such as that, it's encouraging that the community comes together like Orlando did. The same can be said for Boston and so many other places where these tragedies occur, and yet we need to get at the root of the problem, which the Chairman has outlined.

It's the first time that the Commerce Committee has had three of the largest social media companies before us. These social media platforms and those of many other smaller companies have revolutionized the way that Americans communicate, connect, and share information. And, by the way, a comment that the Chairman made about artificial intelligence screening out most of the bad guys' stuff—I wish one of you would explain that. That is encouraging, but it's not quite enough, as the Chairman has outlined.

But at the same time, these platforms have created a new and stunningly effective way for nefarious actors to attack and to harm. It's startling that today, a terrorist can be radicalized and trained to conduct attacks all through social media. And then a terrorist cell can activate that individual to conduct an attack through the internet, creating an effective terrorist drone, in effect, controlled by social media.

So thank you to all of our witnesses for being here and helping explain this and particularly explain what you're doing to rally to the common defense of our people and our country, because using social media to radicalize and influence users is not limited to extremists. Nation states, too, are exploiting social media vulnerabilities to conduct campaigns against this country and to interfere with our democracy.

Now, the Russian hackers, at Vladimir Putin's direction, attempted to influence and did influence the 2016 Presidential election through all of these things that we've been reading about for over a year, and we also know that Putin is likely to do it again. In its January 2017 assessment, the intelligence community said that Putin and his intelligence services see the election influence campaign as a success and will seek to influence future elections. I will be asking Mr. Watts if he would outline what he sees is happening in this 2018 election.

This should be a wake-up call to all of your companies. Indeed, it should be a wake-up call to all Americans, regardless of party. This was an attack on the very foundation of American democracy. We welcome the expertise that each of you bring to the table today. We welcome Mr. Watts and his expertise over many years of how bad actors like Russia use the internet and social media to influence so many things, not just elections.

We even know that Putin is reaching down deep into our government, not just at the top. You remember a few weeks ago, part of the Federal Communications Commission's net neutrality proceeding—half a million comments were traced to Russian IP addresses. That's shocking. That's concerning. We should want to know why these comments were filed. What were they trying to do?

And all of us should be very concerned about what's going to happen next.

In the end, the basic questions that we want to ask are: What have we learned? What are we correcting? What's going to happen in the future, and how can we get ahead of it before it does?

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Nelson.

We do appreciate the great panel of witnesses we have in front of us today. Thank you all for being here.

On my left and your right is Ms. Monika Bickert, who is the Head of Product Policy and Counterterrorism, Facebook; Ms. Juniper Downs, who is Global Head of Public Policy and Government Relations at YouTube; Mr. Carlos Monje, who is Director of Public Policy and Philanthropy, U.S. and Canada at Twitter; and Mr. Clinton Watts, who is a Senior Fellow of the Foreign Policy Research Institute.

So I'm going to start with Ms. Bickert, and we'll just move across the table. If you could confine your oral statements as close to 5 minutes as possible—any additional comments, obviously, will be included as part of the record—that will give us an optimal amount of time for members to ask questions.

So thank you so much to all of you for being here. We look forward to hearing from you.

Ms. Bickert.

STATEMENT OF MONIKA BICKERT, HEAD OF PRODUCT POLICY AND COUNTERTERRORISM, FACEBOOK

Ms. BICKERT. Thank you. Chairman Thune, Ranking Member Nelson, and distinguished members of the Committee, I'm Monika Bickert, and I lead Product Policy and Counterterrorism for Facebook. I'm also a former Federal prosecutor, having spent more than a decade as an Assistant U.S. Attorney for the Department of Justice.

The issues we're discussing here today are of the utmost importance, and on behalf of Facebook, I want to thank you for your leadership in seeking more effective ways to combat extremism, crime, and other threats to our national security.

We share your concerns about terrorists' use of the internet. That's why we remove terrorist content as soon as we become aware of it. It's also why we have a dedicated counterterrorism team of people working across our company. This includes experts like former academics who have spent their career studying terror groups, tracking new trends and tactics. It includes former intelligence and law enforcement officials and prosecutors who have worked in the area of counterterrorism. It also includes engineers who are constantly improving the technology that helps us find and remove terrorist content.

In my written testimony, I describe these efforts in more detail.

I also want to note that we pursue this goal with a mindset that it's important to maximize free expression while keeping people safe online. We work proactively to keep terrorist content off Facebook, primarily through the use of automated systems like image matching and text-based machine learning. Now, more than 99 percent of ISIS and al Qaeda propaganda that we remove from

our service is content that we identify ourselves before anybody has flagged it for us.

Once we're aware of a piece of terrorist content, we remove the vast majority of subsequent uploads within one hour. We do not wait for these global bad actors to upload content to Facebook before placing it into our detection systems. Rather, we work with outside experts who track propaganda released by these groups across the Internet and then send it to us, and we proactively put it in our systems. Often, this means we're able to stop this content from ever being uploaded to Facebook.

However, much of this work cannot be done by machines alone. Accurately removing terrorist content often requires a person to assess it. A photo of an ISIS fighter, for instance, that could be shared by somebody who's a supporter of the group could also be shared by a journalist who's raising awareness or a member of a civil society group who's condemning violence, and we need people to be able to assess that and tell the difference.

We now have more than 7,500 reviewers who assess potentially violating content, including terrorist content, in dozens of different languages. By the end of 2018, we will more than double the 10,000 people who are already working on safety and security issues across our company.

Now, some of these people are responsible for responding to law enforcement requests. We appreciate the critical role that law enforcement plays in keeping people safe, and we do want to do our part. Our global team responds to valid legal requests from law enforcement consistent with applicable laws and our policies, and this includes responding to emergency requests, where we strive to respond within minutes.

We also want to do our part to stop radicalization and disrupt the recruitment process. That's why we've commissioned multiple research efforts over the past three years to understand how online speech can most effectively counter violent ideologies, and we've sponsored efforts to put those learnings into practice. One such example is our peer-to-peer challenging extremism program, which we sponsor with EdVenture Partners, and through that program, we've had more than 6,500 students participate. They've created hundreds of campaigns that have been viewed worldwide more than 200 million times.

No one company can combat the terrorist threat alone. So we partner with others, including companies, civil society, researchers, and governments around the world. Among other things, we work with 11 other companies, including those here, to maintain a shared industry database of hashes, unique digital fingerprints of terrorist content, so that we can all find and remove it faster. We've also recently launched a global internet forum where we can work with smaller companies to help them get better.

In conclusion, let me reiterate that we share your goal of stopping terrorists from using social media, and we're going to keep getting better at it. I'm here today to listen to your ideas and your concerns and to continue this constructive dialog.

Thank you for the opportunity, and I look forward to your questions.

[The prepared statement of Ms. Bickert follows:]

PREPARED STATEMENT OF MONIKA BICKERT, HEAD OF PRODUCT POLICY AND COUNTERTERRORISM, FACEBOOK

Introduction

Chairman Thune, Ranking Member Nelson, 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 head of Product Policy and Counterterrorism at Facebook. Prior to assuming my current role, I served as lead security counsel for Facebook. I am also a former prosecutor, having worked for a decade as an Assistant U.S. Attorney with the Department of Justice. We appreciate the Committee's hard work as it continues to seek more effective ways to combat extremism, crime, and other threats to our national security.

We take all of these threats very seriously. One of our chief commitments is to create and use innovative technology that gives people the power to build community and bring the world closer together. Keeping our community safe on Facebook is critical to this broader mission. We are proud that more than two billion people around the world come to Facebook every month to share with friends and family, to learn about new products and services, to volunteer or donate to organizations they care about, or help in a crisis. The promise of real connection, of extending the benefits of real world connections online, is at the heart of what we do and has helped us grow into a global company.

Being at the forefront of new technology also means being at the forefront of new legal, security, and policy challenges. My team and thousands of other Facebook employees around the world come to work every day to confront these challenges head on. Our goal is to ensure Facebook is a place where both expression and personal safety are protected and respected. We appreciate your commitment to these values as well in your roles as policymakers.

Countering Terrorism on Facebook

I would like to focus my testimony today on the ways Facebook is addressing the challenge of terrorist propaganda and recruitment online.

On terrorist content, our view is simple: There is no place on Facebook for terrorism. Our longstanding policies, which are posted on our site, make clear that we do not allow terrorists to have any presence on Facebook. Even if they are not posting content that would violate our policies, we remove their accounts as soon as we find them. They simply are not allowed to use our services under any circumstances. We also remove any content that praises or supports terrorists or their actions whenever we become aware of it, and when we uncover evidence of imminent harm, we promptly inform authorities.

We recognize the challenges associated with fighting online extremism, some of which I will outline in my comments today. We are committed to being part of the solution, and we are developing strategies built around both technology and human expertise to address these threats.

A. Using Technology to Identify and Remove Terrorist Content

One of the challenges we face is identifying the small fraction of terrorist content posted to a platform used by more than two billion people every month. Our proactive efforts—specifically, the use of artificial intelligence (AI) and other automation—have become increasingly central to keeping this content off of Facebook. We currently focus our most cutting-edge techniques on combating terrorist content about ISIS, Al Qaeda, and their affiliates, and we are working to expand to other terrorist organizations. As we shared recently in a public blog post, 99 percent of the ISIS and Al Qaeda-related terror content that we remove from Facebook is detected and removed before anyone in our community reports it, and in some cases, before it goes live on the site. We do this primarily through the use of automated systems like photo and video matching and text-based machine learning. Once we are aware of a piece of terrorist content, we remove 83 percent of subsequently uploaded copies within one hour of upload.

Importantly, we do not wait for ISIS or Al Qaeda to upload content to Facebook before placing it into our internal detection systems. Rather, we use a variety of techniques, including consulting external experts, to track propaganda released by these groups and proactively insert it into our matching systems. Often, this means we are able to prevent its upload to Facebook entirely.

Because terrorists also adapt as technology evolves, we are constantly updating our technical solutions. I would like to share with you today several specific examples of the ways we are using technology to stay ahead of terrorist activity and combat terrorism online.

1. Image Matching and Language Understanding

When someone tries to upload a terrorist photo or video, our systems look for whether the image matches a known terrorism photo or video. This means that if we previously removed an ISIS propaganda video, for example, we can work to prevent other accounts from uploading the same video to our site.

We also have started experimenting with using AI to understand text that potentially advocates for terrorism. We are working to develop text-based signals to detect praise or support of terrorist organizations. These signals will be incorporated into an algorithm that is in the early stages of learning how to detect similar posts.

2. Removing Terrorist Clusters

We know from the many terrorism academics and experts we work with that terrorists tend to radicalize and operate in clusters. This offline trend is reflected online as well. As such, when we identify Pages, groups, posts, or profiles that support terrorism, we use AI to identify related material that may also support terrorism. As part of that process, we utilize a variety of signals, including whether an account is "friends" with a high number of accounts that have been disabled for terrorism, or whether an account shares the same attributes as a disabled account.

3. Identifying Repeat Offenders

When we disable terrorist accounts, those account owners may try to create new accounts using different identities. We have become faster at using technology to detect new fake accounts created by repeat offenders, or recidivists. Through this work, we have been able to dramatically reduce the time period that terrorist recidivist accounts are on Facebook.

4. Cross-Platform Collaboration

Because we prohibit terrorists from maintaining a presence anywhere in the family of Facebook applications, we have begun work on systems that enable us to remove terrorist accounts across all of our platforms, including WhatsApp and Instagram. Given the limited data some of our applications collect as part of their service, this ability to share data helps immensely in keeping all of our applications safe.

These are some of our key tools, but there are other tools as well. Our ability to outline them here is, however, constrained by the need to avoid providing a roadmap to bad actors seeking to evade detection.

B. Human Expertise

Identifying terrorist content often requires analyzing the relevant context, and we know we cannot rely on AI alone to identify and remove terrorist content. For example, a photo of an armed man waving an ISIS flag could be propaganda or recruiting material, or it could be an image in a major news story. To understand more nuanced cases, we need human expertise.

Our community of users helps us by reporting accounts or content that may violate our policies—including the small fraction that may be related to terrorism. Our content review teams around the world—which grew by 3,000 people last year—work 24 hours a day and in dozens of languages to review these reports. More broadly, by the end of 2018 we will more than double the number of people working on safety and security, including terrorism issues, from 10,000 to 20,000.

We also have significantly grown our team of counterterrorism specialists. Distinct from our content review teams, we have more than 150 highly trained people who are exclusively or primarily focused on preventing terrorist content from ever appearing on our platform and quickly and identifying and removing it if it does. This group includes former academics who are experts on counterterrorism, former prosecutors and law enforcement agents, investigators and analysts, and engineers. Within this specialist team alone, we speak nearly 30 languages.

C. Partnering with Others

We are proud of the work we have done to make Facebook a hostile place for terrorists. We understand, however, that simply working to keep terrorism off Facebook is an inadequate solution to the problem of online extremism, particularly because terrorists are able to leverage a variety of platforms. We believe our partnerships with others—including other companies, civil society, researchers, and governments—are crucial to combating this threat.

To this end, we have partnered with our industry counterparts to more quickly identify and slow the spread of terrorist content online. For example, in December 2016, we joined with Microsoft, Twitter, and YouTube to announce the development of a shared industry database of "hashes"—unique digital fingerprints for photos and videos—for content produced by or in support of terrorist organizations. The

database now contains more than 40,000 hashes, and the consortium of companies has increased to include twelve companies.

This past summer, we formalized our relationship with industry partners and announced the Global Internet Forum to Counter Terrorism (GIFCT), an endeavor that focuses on knowledge sharing, support for counterterrorism work, and technical cooperation, as represented by the hash consortium. Already, this endeavor has brought together more than 68 technology companies over the course of international working sessions held on three continents. This effort gives structure to our existing and future areas of collaboration and fosters cooperation with smaller tech companies, civil society groups, academics, governments, and international bodies such as the EU and the UN.

We engage with governments and inter-governmental agencies around the world and we recently commissioned a research consortium led by the Brookings Institute and the Royal United Services Institute to examine how governments, tech companies, and civil society can work together to fight online extremism and radicalization. We have learned much through briefings from agencies in different countries about extremist organizations' propaganda mechanisms. We also have participated in and benefited from efforts to support industry collaboration by organizations such as the National Counterterrorism Center (NCTC), the EU Internet Forum, the Global Coalition Against Daesh, and the UK Home Office.

In recent months, we have further expanded our partnerships with several organizations including Flashpoint, the Middle East Media Research Institute (MEMRI), the SITE Intelligence Group, and the University of Alabama at Birmingham's Computer Forensics Research Lab. These organizations report Pages, profiles, and groups on Facebook that are potentially associated with terrorist groups. They also send us photo and video files associated with ISIS and Al Qaeda that they have located elsewhere on the internet. We check this information against our algorithms for file "matches," in order to remove or prevent upload of the files to Facebook in the first instance.

We appreciate the critical role that law enforcement plays in keeping people safe. Our legal and safety teams work hard to respond to legitimate law enforcement requests while fulfilling our responsibility to protect people's privacy and security. We have a global team that strives to respond within minutes to emergency requests from law enforcement. We provide the information that we can in response to law enforcement requests, consistent with applicable law and our policies. For example, in the first half of 2017, we provided information in response to more than 75 percent of the 1,864 requests for emergency disclosures that we received from U.S. law enforcement agencies.

Preventing Recruitment Through Counterspeech

We believe that a key part of combating extremism is preventing recruitment by disrupting the underlying ideologies that drive people to commit acts of violence. That's why we support a variety of counterspeech efforts. Although counterspeech comes in many forms, at its core these are efforts to prevent people from pursuing a hate-filled, violent life or convincing them to abandon such a life.

Over the past three years, we have commissioned research on what types of counterspeech are the most effective at combating hate and violent extremism. Based on that research, we believe the credibility of the speaker is incredibly important. We have therefore partnered with non-governmental organizations and community groups around the world to empower positive and moderate voices. For example, two years ago, we worked with the Institute for Strategic Dialogue to launch the Online Civil Courage Initiative, a project that has engaged with more than 100 anti-hate and anti-extremism organizations across Europe. We also have worked with Affinis Labs to host hackathons in places like Manila, Dhaka, and Jakarta, where community leaders joined forces with tech entrepreneurs to develop innovative solutions to challenge extremism and hate online. Finally, we worked with EdVenture Partners to develop a peer-to-peer student competition called the Facebook Global Digital Challenge (P2P). This is a semester-long university course during which students build a campaign to combat extremism in their area, launch it, track its success, and then submit the results as part of a global competition. The University of Central Oklahoma recently implemented a student-led counterspeech program through P2P that uses social media to encourage people to challenge their beliefs and stereotypes. In less than three years, these P2P projects have reached more than 56 million people worldwide through more than 500 anti-hate and extremism campaigns created by more than 5,500 university students in 68 countries.

Conclusion

In conclusion, let me reiterate our commitment to combating extremism on our platform. We have a responsibility to do all we can to combat these threats, and we are committed to improving our efforts.

Of course, companies like Facebook cannot do this without help. We will continue to partner with appropriate authorities to counteract these threats. By working together, business, government, and civil society can make it much harder for malicious actors to harm us, while simultaneously ensuring that people can express themselves freely and openly. I am here today to listen to your ideas and concerns, and I look forward to continuing this constructive dialogue.

The CHAIRMAN. Thank you, Ms. Bickert.

Ms. Downs.

STATEMENT OF JUNIPER DOWNS, DIRECTOR, PUBLIC POLICY AND GOVERNMENT RELATIONS, YOUTUBE

Ms. DOWNS. Thank you, Senator. 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 Public 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. We see over 400 hours of video uploaded to YouTube every minute. With this comes many benefits to society: unparalleled access to art and culture, news and entertainment, educational materials, a remarkable diversity of viewpoints, and the freedom to exchange ideas. We value this openness. It has democratized how stories and whose stories get told.

We are aware, however, that the very platforms that have enabled these societal benefits may also be used by those who wish to promote hatred or extremism. To that end, I'm pleased to have this opportunity to outline the approach we've taken on these issues.

We've developed rigorous policies and programs to defend against the use of our platform to spread hate or incite violence. YouTube has long had policies that strictly prohibit terrorist content. This includes terrorist recruitment, violent extremism, incitement to violence, glorification of violence, and videos that teach people how to commit terrorist attacks. 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.

We use a mix of technology and humans to remove violent content quickly. Users can alert us to content 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 programs, NGOs who provide highly actionable flags and have expertise on issues like hate speech and terrorism, and, of course, we rely on our technology, which has always been a critical part of our solution. Our image-matching techniques, for example, can prevent the dissemination of violent content by catching re-uploads of known bad content before it becomes public.

Nonetheless, given the evolving nature of the threat, it's necessary for us to continue enhancing our systems. Over the past year, in particular, we've 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 recently deployed classifiers that detect new terrorist content and flag it for review. Machine learning is now helping our human reviewers remove nearly five times as many videos as they were before. Today, 98 percent of the videos we remove for violent extremism were identified by our algorithms.

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 several counterterrorism experts. 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.

Finally, we're creating programs to promote counter-speech on our platforms. Our Creators for Change program supports YouTube creators who are tackling issues like extremism and hate by building empathy and acting as positive role models. Google's Jigsaw group has deployed the redirect method, which uses targeted ads and YouTube videos to disrupt online radicalization.

We also collaborate across the industry. In 2016, we created a hash-sharing database with Facebook, Microsoft, and Twitter, where we share digital fingerprints of terrorist content to stop its spread across platforms. We added seven companies to this coalition in 2017, and our shared database now contains over 50,000 video and image hashes. Last summer, we announced the Global Industry Forum to Counter Terrorism to formalize industry collaboration on research, knowledge sharing, and technology.

No single component or party can solve this problem in isolation. To get it right, we must all work together. We understand the importance of speed and comprehensiveness in our work. Since June, we've removed 160,000 videos and terminated 30,000 channels for violent extremism. We've taken down nearly 70 percent of violent extremist videos within 8 hours of upload and nearly half within two hours. We've reviewed over 2 million videos to make sure we're catching and removing all videos that violate these policies.

We achieved these results through enhanced enforcement by machines and people and collaboration with outside experts. We're deeply committed to working with law enforcement, government, the tech industry, and NGOs to protect our services from being exploited by bad actors. We look forward to continued collaboration with the Committee as it examines these issues.

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

[The prepared statement of Ms. Downs follows:]

PREPARED STATEMENT OF JUNIPER DOWNS, DIRECTOR, PUBLIC POLICY AND GOVERNMENT RELATIONS, YOUTUBE

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 percent 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 percent of the videos we removed for violent extremism were identified by our algorithms. Today, that number is 98 percent. Our advances in machine learning let us now take down nearly 70 percent 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 Inter-

national 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 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.

The CHAIRMAN. Thank you, Ms. Downs.

Mr. Monje.

STATEMENT OF CARLOS MONJE, JR., DIRECTOR, PUBLIC POLICY AND PHILANTHROPY, U.S. AND CANADA, TWITTER

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

I'm here on behalf of Twitter, an open communications platform that allows more than 330 million users to see what's happening in the world and to share viewpoints from every side. Each day, we serve 500 million tweets. We have about 3,700 employees around the world.

Twitter has been at the forefront 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 also giving voice to those who promote a positive message for the future.

Twitter has a zero tolerance policy for terrorist content. This includes not only specific threats of violence, but also promoting terrorism, affiliating with violent extremist groups, and glorifying violence. Our job is to enforce this policy globally, at scale, to evolve to stay one step ahead of the terrorists. We have dramatically improved our ability to implement these rules and have suspended more than 1.1 million terrorist accounts since mid 2015.

Our progress fighting terrorist content is due to our commitment to innovation. While there is no "magic algorithm" for identifying terrorist content, we have increasingly improved the effectiveness of our in-house proprietary technology. Our technology supplements user reports, human review, and it significantly augments our ability to identify and remove bad content from Twitter.

At the beginning of 2015, our in-house technology detected roughly a third of the terrorist accounts that we pulled down at that time. Last year, that number increased dramatically. We identified more than 90 percent of suspensions for terrorism by our internal tools, and 75 percent or three-quarters of those accounts were suspended before they had a chance to tweet even once. Let me repeat that because it's important. We spot more than 90 percent of terrorist accounts before anyone else does, and we stop 75 percent of those accounts before they can spread any of their deplorable ideology.

Of course, like any 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 from being opened to replace those that we suspend while also developing further the tools that prevent the distribution of propaganda in the aftermath of attacks.

Because this is a shared challenge, our industry has established the Global Internet Forum to Counter Terrorism, which is focused on learning and collaboration, on technical cooperation, and research. Twitter sees the forum as a substantial opportunity to en-

sure that smaller companies are not soft targets for terrorists. We have engaged with 68 smaller companies over the past several months to share best practices and learnings, and we plan to grow on that work.

Removing a tweet doesn't eliminate the ideology behind it, so we invest heavily in alternative narratives. Twitter has participated in more than 100 NGO trainings and events around the world since 2015. We work with respected organizations to empower credible, non-governmental voices against violent extremism.

As part of a continuing effort to make Twitter a safe place for open democratic debate, late last year, we broadened our rules to prohibit accounts affiliated with violent extremist groups and to make hateful imagery much harder to find on our platform. We also stepped up our enforcement of abuse reported by witnesses and increased transparency about our enforcement decisions to further educate our users about our terms of service.

Twitter has also devoted significant resources to combat disinformation and election interference by foreign state actors. To prepare for the U.S. midterm elections this year, a cross-functional elections task force is prepared to verify major party candidates as a hedge against impersonation, to maintain open lines of communication with Federal and State election officials, to continually improve and apply our technology to address networks of malicious automation, and to monitor trends and spikes in conversations related to the elections.

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 threats of extremism and those who would harm our democratic process.

Thank you. Thank you for your leadership on these issues. I look forward to this discussion.

[The prepared statement of Mr. Monje follows:]

PREPARED STATEMENT OF CARLOS MONJE, JR., DIRECTOR, PUBLIC POLICY AND PHILANTHROPY, U.S. AND CANADA, TWITTER

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 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.

The CHAIRMAN. Thank you, Mr. Monje.

Mr. Watts.

STATEMENT OF CLINT WATTS, ROBERT A. FOX FELLOW, FOREIGN POLICY RESEARCH INSTITUTE; SENIOR FELLOW, CENTER FOR CYBER AND HOMELAND SECURITY, THE GEORGE WASHINGTON UNIVERSITY; AND NON-RESIDENT FELLOW, ALLIANCE FOR SECURING DEMOCRACY, GERMAN MARSHALL FUND OF THE UNITED STATES

Mr. WATTS. Chairman Thune, members of the Committee, thanks for having me here today.

Ten years ago, it was al Qaeda in Iraq videos on YouTube. A few years later, al Shabaab's deadly rampages played out on Twitter. Shortly after, Facebook groups and Twitter feeds brought the Islamic State to the world's attention and into the homes of new recruits before they scurried off to other social media platforms like Telegram. And 4 years ago, amongst global jihad's social media storm, I stumbled into Russian influence campaigns, their reboot of an old playbook called "Active Measures," which they've deployed across nearly every social media platform with devastating effect.

Social media at its height offered a platform for discussion across diverse audiences and led to uprisings toppling dictators during the Arab Spring. But bad actors with motivation, money, manpower, and know-how will always come to these information gateways to pursue their objectives. Lesser educated populations around the world predominately arriving in cyber space via mobile phones will be particularly vulnerable to the social media manipulation of terrorists and authoritarians.

American focus on the Islamic State social media recruitment or Russian meddling in the Presidential election of 2016 overlooks other indicators of damaging activity. American companies have suffered and remain particularly vulnerable to smear campaigns launched by foreign state actors through malicious false narratives, pushing their stock prices down and decreasing sales through reputational damage.

Beyond just smear campaigns and character assassinations, this committee should take seriously the ability of foreign nations to mobilize violence inside the U.S. through an evolution I would call "Anwar Awlaki Meets PizzaGate." Just a few years ago, Anwar Awlaki, al Qaeda's external operations leader in Yemen, recognized the power of the Internet to recruit and mobilize terrorists in America to conduct violence in the U.S. homeland.

The Islamic State took this to another level with their spokesman, Abu Muhammad al-Adnani, calling on supporters to conduct attacks at home and then further enabling those e-recruits by using a social media battalion to guide those plots remotely. A little over a year ago, America saw an individual consume a false conspiracy on the internet and social media, known as PizzaGate, and then travel to Washington, D.C., to investigate those claims. He arrived at a falsely implicated restaurant and discharged a weapon before being arrested.

Surely, a foreign adversary of the United States sees an opportunity in combining these two scenarios. The greatest concern moving forward might likely be a foreign intelligence service posing as Americans on social media, infiltrating one or both political extremes in the U.S., and then recruiting unwitting Americans to undertake violence against a target of the foreign power's choosing. Social media companies will be better positioned to stop this potential scenario from occurring than U.S. intelligence or Homeland Security that are blind to the technical signatures behind this manipulation.

Social media companies realize the damage of these bad actors far too late. They race to implement policies to prevent the last information attack, but have yet to anticipate the next abuse of their social media platforms by emerging threats. I've offered a range of recommendations for how to counter bad actors using social media in previous testimony. I'll focus on a few issues here today.

The first and most pressing challenge comes in the debate over social media account anonymity versus authenticity. Anonymity of social media accounts has in many cases allowed the oppressed and the downtrodden to speak out about injustice. But over time, anonymity has empowered hackers, extremists, and authoritarians. Under the veil of anonymity, they spread hate, recruit members, and advance divisions in American society.

Social media companies can and should protect the public anonymity of account holders if their user chooses, but they must be able to determine a real person resides behind each persona. Social media companies have better advanced tools recently to certify authenticity. However, the current level of authenticity on the Twitter platform is suboptimal. I'd encourage Twitter to rapidly expand its verification to as many users as possible as quickly as possible.

Closely connected to the issue of account authenticity is the rise of computational propaganda. The negative effects of social bots far outweigh any benefits. The anonymous replication of accounts that routinely broadcast high volumes of misinformation can pose a serious risk to public safety and, when employed by authoritarians, a direct threat to democracy.

Last, social media companies continue to get beat in part because they rely too heavily on technologists and technical detection to catch bad actors. Artificial intelligence and machine learning will greatly assist in cleaning up nefarious activity, but will for the near future, fail to detect that which hasn't been seen before. Those who understand the intentions and actions of criminals, terrorists, and authoritarians must work alongside technologists to sustain the integrity of these social media platforms.

I know it is unreasonable to think that every social media company can and should hire threat analysts for every possible emerging threat. But a variety of rapid outreach approaches with external social media analysts and threat experts positioned outside social media companies could easily be developed or even be collectively sponsored by social media companies. Several models from counterterrorism and cybersecurity could be adopted by Silicon Valley in this regard. I've made other recommendations in the past which I can address during the Q and A.

But, in conclusion, some social media companies have done more than others to improve the safety and integrity of their platforms. Others have a lot of work to do to improve their platforms against bad actors. Ultimately, the American consumer will decide whether the benefits of using these services outweigh the risks. Many are walking away from social media applications because they can't trust the information being shared or tolerate the vitriolic user experience.

Social media companies should move aggressively to thwart terrorists and authoritarians exploiting their systems not only because it's what's best for their users and society, but because it's good for business as well.

Thank you for having me.

[The prepared statement of Mr. Watts follows:]

PREPARED STATEMENT OF CLINT WATTS, ROBERT A. FOX FELLOW, FOREIGN POLICY RESEARCH INSTITUTE; SENIOR FELLOW, CENTER FOR CYBER AND HOMELAND SECURITY, THE GEORGE WASHINGTON UNIVERSITY; AND NON-RESIDENT FELLOW, ALLIANCE FOR SECURING DEMOCRACY, GERMAN MARSHALL FUND OF THE UNITED STATES

Ten years ago, it was al Qaeda in Iraq videos on YouTube. A few years later, al Shabaab's deadly rampages played out on Twitter. Shortly after, Facebook groups and Twitter feeds brought the Islamic State to the world's attention and into the homes of new recruits, before they scurried off to other social media applications like Telegram. And four years ago amongst global jihad's social media storm, I stumbled into Russian influence campaigns, their reboot of an old playbook called "Active Measures", which they've deployed across nearly all social media platforms with devastating effect.

Today, disinformation spread on Facebook propels deadly violence in Myanmar against the minority Ronhingya population.[i] The Duterte regime in the Philippines uses social media groups to suppress domestic political opponents.[ii] LTG H.R. McMaster, our National Security Advisor, noted just last week the Kremlin is again using its cyber influence just across our southern border seeking to push their preferred party and politicians to the forefront in Mexico.[iii]

Social media, at its height, offered a platform for discussion across diverse audiences and led to uprisings usurping dictators during the Arab Spring. But bad actors with motivation, money, manpower and know—how will always come to these information gateways to pursue their objectives. Criminals, terrorists and authoritarians see the Internet and social media as a cost effective open doorway into the very heart of their adversaries. Authoritarians worldwide now recognize the power of the Kremlin's social media manipulation, and if left unchecked, will copy and deploy Russia's playbook against their enemies. Lesser—educated populations around the world predominately arriving in cyberspace via mobile phones will be particularly vulnerable to the social media manipulation of terrorists and authoritarians.

American focus on the Islamic State's social media recruitment or Russian meddling in the 2016 Presidential election overlooks other indicators of damaging activity. American companies have suffered and remain particularly vulnerable to smear campaigns launched by foreign state actors through malicious, false narratives pushed by bogus social media personas. These campaigns can cause serious reputational damage sending stock prices plummeting and decreasing sales.

Beyond just smear campaigns and character assassination, this committee should take seriously the ability of foreign nations to mobilize violence inside the U.S. through an evolution I would call "Anwar Awlaki meets PizzaGate". Just a few years ago, Anwar al—Awlaki, al Qaeda in the Arabian Peninsula's leader of exter-

---

[i] Hannah Beech. "Across Myanmar, Denial of Ethnic Cleansing and Loathing of Rohingya." *New York Times.* 24 October 2017. Available at: *https://www.nytimes.com/2017/10/24/world/asia/myanmar-rohingya-ethnic-cleansing.html?_r=0.*

[ii] Lauren Etter. "What happens when the government uses Facebook as a weapon?" *Bloomberg.* 7 December 2017 Available at: *https://www.bloomberg.com/news/features/2017-12-07/how-rodrigo-duterte-turned-facebook-into-a-weapon-with-a-little-help-from-facebook.*

[iii] David Alire Garcia and Noe Torres. "Russia meddling in Mexican election: White House aide McMaster." *Reuters.* 7 January 2018. Available at: *https://www.reuters.com/article/us-mexico-russia-usa/russia-meddling-in-mexican-election-white-house-aide-mcmaster-idUSKBN1EW0UD.*

nal operations, recognized the power of the Internet to recruit and mobilize terrorists in America to conduct violence in the U.S. homeland. The Islamic State took this to another level with their spokesman abu Muhammad al—Adnani calling on supporters to conduct attacks at home[iv] and then further enabling e—recruits by using a social media battalion to guide plots remotely—connecting with, coaching and directing terrorists in the West to specific targets.[v] A little over a year ago, America saw an individual consume a false conspiracy on the Internet and social media, known as PizzaGate, and then travel to Washington DC to investigate these bogus claims. He arrived at a falsely implicated restaurant and discharged a weapon before being arrested.[vi]

Surely a foreign adversary of the United States sees an opportunity in combining these two scenarios. The greatest concern moving forward might likely be a foreign intelligence service, posing as Americans on social media, infiltrating one or both political extremes in the U.S. and then recruiting unwitting Americans to undertake violence against a target of the foreign power's choosing. Social media companies will be better positioned to stop this potential scenario from occurring than U.S. intelligence or homeland security that are blind to the technical signatures behind this manipulation.

The U.S. Government's response to terrorist social media use has been sustained and significant, and their response to state sponsored influence on Americans disjointed and perplexing. In both cases, government officials have pointed to social media companies asking why they would allow their platforms to be used for nefarious purposes.

Social media companies realize the damage of these bad actors far too late. They race to implement policies to prevent the last information attack, but have yet to anticipate the next abuse of their social media platforms by emerging threats seeking to do bad things to good people. In previous testimony to the Senate Homeland Security,[vii] Intelligence,[viii] Armed Services[ix] and Judiciary[x] committees, I've offered a range of recommendations for how to counter bad actors using social media in the pursuit of violence and nefarious influence. Today, I'll focus and reiterate a few of these recommendations.

The first and most pressing challenge comes in the debate over social media account anonymity and authenticity. Anonymity of social media accounts has in many cases allowed the oppressed and the downtrodden to speak out about injustice. It's given the weak a voice against the strong, powerful, and corrupt. But over time, anonymity has empowered hackers, extremists and authoritarians to inflict harm on the public. Under the veil of anonymity, they spread hate, recruit members and advance divisions in American society.

All people, real humans and their virtual personas, have the right to free speech, but this right to free speech does not permit them to endanger society. Account anonymity today allows nefarious social media personas to shout the online equivalent of "fire" in a movie theater. Bad actors and their fictitious and/or anonymous social media accounts can and have created a threat to public safety. This is not protected free speech and many social media companies offer no method to hold these anonymous personas accountable.

[iv] Bulos, Nabih. "Islamic State's taunting speech calls for killing civilians." *LA Times.* 22 September 2014. Available at: *http://beta.latimes.com/world/middleeast/la-fg-islamic-state-taunts-20140922-story.html.*

[v] Rukmini Callimachi, "Not 'Lone Wolves' After All: How ISIS Guides World's Terror Plots From Afar." *New York Times.* 4 Feb 2017. Available at: *https://www.nytimes.com/2017/02/04/world/asia/isis-messaging-app-terror-plot.html?_r=0.*

[vi] Grace Hauck. "Pizzagate shooter sentenced to 4 years in prison." *CNN.* 22 June 2017. Available at: *http://www.cnn.com/2017/06/22/politics/pizzagate-sentencing/index.html.*

[vii] Clint Watts. "Terror in Europe: Safeguarding U.S. Citizens at Home and Abroad." Statement prepared for the Senate Committee on Homeland Security and Government Affairs, 5 April 2016. Available at: *https://www.fpri.org/article/2016/04/terror-europe-safeguarding-u-s-citizens-home-abroad/*

[viii] Clint Watts. "Disinformation: A Primer In Russian Active Measures and Influence Campaigns." Statement prepared for the Senate Select Committee on Intelligence, 30 March 2017. Available at: *https://www.intelligence.senate.gov/sites/default/files/documents/os-cwatts-033017.pdf.*

[ix] Clint Watts. "Cyber-enabled Information Operations." Statement prepared for the Senate Committee on the Armed Services, Subcommittee on Cybersecurity, 27 April 2017. Available at: *https://www.armed-services.senate.gov/download/watts_04-27-17*

[x] Clint Watts. "Extremist Content and Russian Disinformation Online: Working with Tech to Find Solutions." Statement prepared for the Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism. Available at: *https://www.judiciary.senate.gov/download/10-31-17-watts-testimony.*

Social media companies can and should protect the public anonymity of account holders if the user chooses, but they must be able to determine a real, authentic person resides behind each persona accountable for their actions on the platform. Some social media companies have advanced better methods to certify account authenticity. However, the current level of authenticity on the Twitter platform is suboptimal. I'd encourage Twitter to rapidly expand its verification to as many users as possible, as quickly as possible.

Closely connected to the issue of account authenticity is the rise of computational propaganda. The negative effects of social bots far outweigh any benefits. The anonymous, replication of accounts that routinely broadcast high volumes of misinformation can pose a serious risk to public safety and when employed by authoritarians a direct threat to democracy. Social bots should be ceased immediately. For non-automated accounts, reasonable limits on the number of posts any account can make during an hour, day or week should be developed. Even further, human verification systems (CAPTCHA) should be employed by all social media companies to reduce automated broadcasting.

Federal laws governing attribution of political ads and solicitations in television, radio and print should immediately be extended to social media advertising conducted by political campaigns and political action committees. Social media political advertising will continue to grow in every election cycle and U.S. citizens must know the source of the information they consume in any medium—print, radio, television or social media.

Social media companies continue to get beat in part because they rely too heavily on technologists and technical detection to catch bad actors. Artificial intelligence and machine learning will greatly assist in cleaning up nefarious activity, but will for the near future, fail to detect that which hasn't been seen before. Threat intelligence proactively anticipating how bad actors will use social media platforms to advance their cause must be used to generate behavioral indicators that inform technical detection. Those that understand the intentions and actions of criminals, terrorists and authoritarians must work alongside technologists to sustain the integrity of social media platforms. Some social media companies have already moved in this direction.

I'd note it's unreasonable to think that every social media company can and should hire threat analysts for every possible emerging threat. But a variety of rapid outreach approaches with external social media analysts and threat experts positioned outside social media companies could easily be developed or even be collectively sponsored by social media companies. Several models from counterterrorism and cybersecurity could be adopted by Silicon Valley in this regard.

I've made many other recommendations in the past but will close for now and can elaborate further on them during the question and answer session. In conclusion, some social media companies have done more than others to improve the safety and integrity of their platforms. Others have a lot of work to do to improve their platforms against

bad actors. Ultimately, the American consumer will decide whether the benefits of using these services outweigh the risks. Many are walking away from social media applications because they can't trust the information being shared or tolerate the vitriolic user experience. Social media companies should move aggressively to thwart terrorists and authoritarians exploiting their systems not only because its what's best for their users and society, but because it's good for business as well.

The CHAIRMAN. Thank you, Mr. Watts.

Senator NELSON. Mr. Chairman.

The CHAIRMAN. I recognize Senator Nelson for an introduction.

Senator NELSON. We have a new member of the Committee.

The CHAIRMAN. We do, indeed.

Senator NELSON. Senator Jon Tester of Montana. Senator Tester has been wanting to get on this Committee for quite a while, and we are so happy that he finally was able to be appointed to the Committee. He brings a wealth of experience as a senior member of the Appropriations Committee to this committee.

So, welcome, Senator Tester.

The CHAIRMAN. Senator Tester.

Senator TESTER. Mr. Chairman and Ranking Member, I just very quickly—first of all, thanks for the welcome. I look forward to

working on this Committee, although I will tell you I have flashbacks to 2007 right now. Thank you.

The CHAIRMAN. Well, I also want to thank and welcome Senator Tester for joining the Committee. It's always nice to have another rural Senator on this Committee, somebody who is my neighbor and represents a state like mine where there are more cattle than there are people, and, obviously, an area where there are still cell phone and broadband free areas. But we're hoping to change that. Senator Tester can probably remember the days like I can when there were party phone lines. So we've come a long way, but we've got a long way to go, and a lot of the issues that we address on this committee are issues that impact the daily lives of people in his state just like they do with so many members of this Committee.

Senator Tester, welcome. It's good to have you here.

We're going to start with some questions, and I want to direct this to Mr. Monje and Ms. Bickert. As you mentioned, both of you, I think, in your testimony, Google, Facebook, Twitter, and Microsoft announced the formation of a hashing coalition to better identify and ultimately remove offending content. The question is: Is there any shared standard for what constitutes extremist or terrorist content in your coalition?

Mr. MONJE. Thank you, Chairman, for that question. Our companies are constantly working with one another and with civil society and with smaller companies to address the issues that change and evolve and new things that we see around the world. We are constantly adapting how we attack the challenge, and we do rely on the advice and good counsel not only of our peer companies but also of academics and NGOs.

The CHAIRMAN. There's no standard definition, though, that you've agreed upon?

Ms. BICKERT. That's right, Mr. Chairman. I would just add that the companies—we launched the Global Internet Forum in June 2017, but we've actually been working together for a number of years informally. Part of those meetings is discussing what the appropriate standards are, recognizing, of course, that these different products work differently.

But the two types of policies I think you most commonly see are first directed toward the groups having any presence on the platform. So, for instance, at Facebook, if you are Boko Haram, you cannot have a page on Facebook even if you're just talking about the lovely weather. You simply can't be on the platform. And then the other types of policies that you see certainly across the major companies is banning any praise or support of these groups or their actions.

The CHAIRMAN. Ms. Downs, according to the Counter Extremist Project, one single bomb-making video used to instruct the Manchester suicide bomber has been uploaded to YouTube and removed 11 times but continues to resurface as recently as this month. How is it possible for that to happen? Why aren't your hashing efforts working to keep this video off your platform permanently?

Ms. DOWNS. Thank you, Chairman. As I mentioned, we have strict policies against terrorist content, including videos that instruct people on how to commit terrorist attacks. This certainly in-

cludes bomb-making videos, videos instructing people on how to drive vehicles into crowds, and so on. This particular video was caught by our systems. We have used it in our hash-sharing database, and we are catching re-uploads of the video quickly and removing it as soon as those uploads are detected.

The CHAIRMAN. Are your companies—and anybody can answer this. But are your companies, as you start to roll out some of these new counterterrorism programs—how do you have ways of measuring their effectiveness? What is sort of the metric or the standard?

Mr. MONJE. Chairman, at Twitter, we've really doubled down on the technology, on the machine learning, to try to identify and remove content as quickly as possible. So our metric is how many accounts are we taking down, how many accounts are we identifying proactively, and how many are we able to take down before they're able to tweet. And we've seen steady progress in that. We started—we were taking down about a third of our content proactively with our machine learning. Today, that's north of 90 percent, with 75 percent of that coming down before anybody gets a chance to tweet. So that's how we—that's our main metric.

The CHAIRMAN. So it's been reported that ISIS surrogates are using AI bots to carry out recruiting and messaging campaigns, and as you all become more sophisticated in how to prevent and root out, the bad people also become more sophisticated in how to get around, and the threat evolves. So are you seeing that level of sophistication, and, if so, what are you doing to mitigate it—the use of AI against you by these groups?

Does anybody want to take a stab at that?

Ms. DOWNS. In addition to our policies against terrorist content, we have very aggressive and proprietary spam detection systems which would catch massive re-uploads of AI-generated videos. So our long history in fighting spam across our services is an effective technique to get at that behavior.

The CHAIRMAN. Anybody else?

Mr. MONJE. I would just agree with you, Chairman, that it is a cat and mouse game, and we are constantly evolving to meet the challenge. When we, often in the past, would ban an account, suspend an account, they would try to come back and then brag about the fact that they were banned. That became a very strong signal for us which resulted in them being taken down even quicker. So they've stopped doing that.

The CHAIRMAN. My time has expired.

Senator Nelson.

Senator NELSON. Mr. Watts, I'd like you to take my time and inform the Committee with your expertise what the Russians—for that matter, anybody else—can do to us in this coming election?

Mr. WATTS. Thank you, Senator. I think I would start off with—there has been no response from the U.S. Government with regards to Russian influencing campaigns with social media. So, therefore, they have stayed on course with their operations. During non-election years, they tend to focus on social issues and what I would say is audience infiltration. So any organization, entity, social media group that they really want to be able to move or influence later, they begin infiltrating that by just sharing the same content back

with that audience and trying to develop their own content within it.

Beyond just the United States and this Presidential election, I think we should look at all elections worldwide. They've realized that this playbook works very well. It's extremely cost-effective, and there has been almost no downside, at least to this point, of doing it. So you've seen it in Europe, where they continue to seed audience bases.

Anywhere that they can break up a union—so the European Union or NATO—they will continue to seed in those populations. So I would tell you to look at Catalonia or even Scotland, places like that where they see an opportunity to break up an alliance, to create divisions within a democracy—they are moving there. I think Lieutenant General McMaster last week pointed to Mexico as another example of where they've seen some sort of audience infiltration.

The key trigger I always look for is hacking. When they launch widespread hacking against a target, they are making a strategic decision to go after an objective, and that's one thing I would tell everyone to look for on the horizon.

Beyond that, if you want to know where the Russians are going with their influence, you should always look at where they are putting up new state-sponsored outlets. To infiltrate an audience, you have to have a base of content to launch your campaign. So when they add an additional language or for their wire service—let's say Sputnik News—or an RT YouTube channel, that is an audience that they're going to reach for. And I will tell you right now, they are looking very heavily into Latin America. I think they would like to build a capability more in the Middle East moving forward.

Beyond just Russia, they will focus on social issues to win over audiences during non-election years to then be able to pivot them toward whichever candidate or party they want to support moving forward. The goal isn't one party or the other and their victory. The goal is to create divisions inside the United States, and that will be their predominant focus moving forward, further polarizing the information landscape.

I would also note that everyone is adopting this technique. You see it in Myanmar. You see it in the Philippines. Any low-level educated population around the world that's on social media, particularly through mobile applications, is highly vulnerable to this. They have not built up the ability to assess information or how to avoid being influenced, and so they're highly vulnerable to this influence technique.

And, last, I would say it's political campaigns and the companies that are going to be hired. If there's not some sort of regulation put around ads in social media, every political campaign, whether it's in the U.S. or around the world, will have to use a dark social media campaign through either Super PACs or candidates to keep up with their competitors, and it will further—it will not only harm the societies in which it's in, but it will actually harm the social media companies and their platforms. They will actually make the environment so terrible and so polarized, as we've seen over the past few years, that it will create just a nasty sense for democracy.

If you want to look at how this has effected in Russia, Russia did this to their own people first before they came, you know, across the ocean. It creates widespread apathy in democracies. It dilutes the line between fact and fiction, and when that happens, you actually cannot keep a democracy moving forward. I think that's what's most dangerous about this entire system, is it's agnostic of party or candidate. Ultimately, it's about breaking our system and turning it against each other.

Senator NELSON. So when you see them dive deep into the instrumentalities of government, such as the example that I gave, that there were half a million comments on the recent FCC rule, and when you see that—you read the public press that they're in 20 states' elections divisions, sketch out what are some of the dastardly things that they could do to undermine America.

Mr. WATTS. The one big thing that they would try and do is an information attack on the integrity of a democratic institution. That's really played out in both of those scenarios. With the FCC, it's "you can't trust the FCC. We need to get rid of these regulatory bodies. You can't trust them. They're trying to mind control you."

The other part is the elections, the integrity of an election. The second campaign they launched in the run-up to 2016 was voter fraud, election rigged. They didn't really care what candidate won. They wanted the American people to think that their vote did not count. The hacking campaign against voter databases—it was to sow doubt such that when you see the narrative of voter fraud, election rigged, you might think, "Oh, maybe my candidate didn't really get elected because my vote didn't count."

So it's about destroying democratic institutions and confidence in the U.S. Government or democratic institutions to govern properly, that the system is always rigged and you can't trust anyone. That's really the focal point of all of those efforts the Russians might run or any authoritarian regime that wants to run a campaign against the U.S. Government.

Senator NELSON. Thank you.

The CHAIRMAN. Thank you, Senator Nelson.

Senator Wicker.

STATEMENT OF HON. ROGER WICKER,
U.S. SENATOR FROM MISSISSIPPI

Senator WICKER. So it wasn't so much an attempt to get one candidate elected over the other. It was knowing there was going to be loser and that relatively half of the population who supported that loser would think their vote hasn't counted.

Mr. WATTS. That was their second campaign. They ran four narratives during the first one which were specific to the candidates. Then in October 2016, they really shifted to the integrity of democratic institutions. So it was twofold. They were running to try and get a candidate they wanted up to election day, and then beyond election day, it was to create mass chaos inside the United States.

Senator WICKER. Thank you for clearing that up. Let me move to Twitter.

Is it Monje? Am I pronouncing it correctly?

Mr. MONJE. Thank you for asking, sir. It's Monje.

Senator WICKER. Monje. OK. Good. Well, let me ask you, then—and is that a Cajun name? You told me you're from New Orleans.

Mr. MONJE. I am from New Orleans, sir. I'm a Saints fan. But my family is from Argentina.

Senator WICKER. Very interesting. Let me ask you about aggregate user data. There are data analytics companies who purchase aggregate user data from all of you, Twitter. Is that correct?

Mr. MONJE. Yes, sir.

Senator WICKER. So, for example, if I am an analytic company, and I want to work for the NFL, for example, I would purchase aggregate user data from Twitter and, using keywords, develop information that might be helpful to the National Football League.

Mr. MONJE. It depends on what it is they plan to use it for. A lot of times, what our data is used most often for is to target advertising.

Senator WICKER. To target advertising. OK. Let me ask you this. Is that same ability to purchase aggregate data available to Federal law enforcement? Is it available to Federal anti-terrorism agencies?

Mr. MONJE. It depends on what the purpose of the use of data is, and we have rules about how any entity, regardless of whether governmental or not, anywhere in the world, can use our data.

Senator WICKER. And what are those rules with regard to terrorism?

Mr. MONJE. With regards to terrorism. I'd have to get back to you on the exact rule—on the exact language of that, sir.

Senator WICKER. OK. Well, because this is pretty important. If a data analytics company wants to purchase data from Twitter, you're willing to sell that to them. What I want to know is if that company is going to supply information to agencies that are seeking information about terrorist activities and that activity is part of this aggregate user data. Will you sell that data to them? Because, frankly, I'm informed that you will not do so.

Mr. MONJE. Well, let me tell you a little bit about what we do with our data, sir, which is we—on our side, on the Twitter side of the equation—are very data-focused and use that data to inform the machines that help us fight the terrorists. We work on a daily basis with law enforcement.

Senator WICKER. That's within the Twitter organization.

Mr. MONJE. Within the Twitter organization. We work on a daily basis with law enforcement, particularly with the FBI, and will respond to any request that they have, as long as they give us the proper legal process, and we are on a first name basis with our counterparts at the FBI.

Senator WICKER. And what would that proper legal process be?

Mr. MONJE. It depends on what they're looking for. It could be a warrant—depending on whether they're looking for private or non-private information on—where they're looking for direct messages.

Senator WICKER. Has Twitter told these data analytic companies that the purchases of this data cannot be used for counterterrorism purposes?

Mr. MONJE. They cannot be used for persistent surveillance of individuals.

Senator WICKER. They can be used to target advertising and products and sales, but they cannot be used to help our anti-terrorism agencies. Is that correct?

Mr. MONJE. We do help our anti-terror agencies, particularly the Federal Bureau of Investigation on a daily basis. And in response——

Senator WICKER. But if a third——

Mr. MONJE. I'm sorry, sir. After you.

Senator WICKER. Go ahead, please.

Mr. MONJE. No and to respond to their request, we have a very fast system that is an input—any time they have information to us, we turn it around as quickly as we can, within hours. We do not allow persistent surveillance of our users. We protect the privacy of our users.

Senator WICKER. You protect the privacy of your users, even if a Federal agency wants to surveil that public information for anti-terrorism purposes?

Mr. MONJE. If an agency comes to us with the right process, and it's according to Federal law, the ECPA——

Senator WICKER. But that's not what I'm talking about. I'm talking about an independent data analytics company.

Mr. MONJE. Yes, sir. So——

Senator WICKER. You will sell that data to them, but you tell that company they can't use it for anti-terrorism purposes. Is that correct?

Mr. MONJE. We're not going to allow any company, whether they're selling cars or cereal or anything, the NFL, to persistently figure out where somebody is in a given time. But we do have news products, data alerts, for law enforcement, for the FBI, that they use.

Senator WICKER. Ms. Bickert, is that the policy of Facebook?

Ms. BICKERT. Thank you, Senator. We don't sell user data outside the context of allowing people to target audiences in their advertisements. That is a capability that is equally available to law enforcement, as it would be to anybody else. Law enforcement can provide us—if they want to find out specifics about an individual user, they can provide us with legal process, and we will respond.

Senator WICKER. What is the privacy concern that supersedes the need to surveil terrorist organizations that participate in Facebook?

Ms. BICKERT. Senator, we absolutely respond to valid law enforcement requests. If it's part of an investigation, and they give us that process, we do respond.

Senator WICKER. Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Wicker.

Senator Klobuchar.

STATEMENT OF HON. AMY KLOBUCHAR,
U.S. SENATOR FROM MINNESOTA

Senator KLOBUCHAR. Thank you very much, Mr. Chairman.

I came in, and three senators told me, Mr. Monje, that you were a Saints fan. Is that correct?

[Laughter.]

Senator KLOBUCHAR. I would just like to note my scarf and who won the game.

Who won the game, Mr. Monje?

Mr. MONJE. It was an excellent game, and it was a spectacular ending for the Vikings.

Senator KLOBUCHAR. Thank you very much. Now, let's get to some serious matters.

So we've had a hearing focused on the election piece of the internet, and I note that, while I want to get to some questions about terrorism as well, there are many ways we can undermine our country and undermine our democracy. One is obvious, with violent attacks. Another is if Americans aren't able to make their own decisions about who they were voting for because they get false information. So that's why Senators Warner, McCain, and I have introduced the Honest Ads Act. We had an entire hearing about this over in Judiciary.

But I would just start with you, Mr. Watts. As you know, right now, there are disclosure rules so radio, TV, print—they all have to keep on file ads of national and political importance, legislative importance, as well as candidate ads, so that opposing campaigns' press can see these ads, as well as disclaimer requirements. Do you think those should apply to social media ads, paid ads as well?

Mr. WATTS. Absolutely. If it does not happen, I mean, both from society and social media's perspective, the conversation will continue to get more polarized and more negative, and people won't be able to trust information on the platform, regardless. So I think it's essential that the ad regulation extend to social media because that's where all advertising is going in the future.

Senator KLOBUCHAR. Exactly. We had $1.4 billion in the last election, and there are projections it's going to go to $3 billion or $4 billion and things like that, and there are literally no rules. We do appreciate that a number of the companies, including the ones here, have stepped up to start putting in place some of their own guidelines and changes. But I do believe that this won't work unless we have guidelines like we have for media. Do you agree with that?

Mr. WATTS. I do. If we don't, it can have a very devastating effect and force all political campaigns essentially to try to do social media manipulation that's maybe not entirely authentic.

Senator KLOBUCHAR. Thank you. Terrorist online recruiting—my state has had its share of that recruiting, especially related—some from ISIS, but in past years, al Shabaab. We've had dozens of prosecutions out of our U.S. Attorney's office, successful ones, where people have actually been recruited to go overseas to fight on behalf of terrorist groups.

What kind of recruiting activity are you able to detect on your platforms, and what can you tell us about the trends? How are they changing their strategies? I remember the FBI showing me the ads targeted at Minnesota with literally airplane tickets from Minnesota to Somalia for terrorists. So tell me what you're doing now and what you see in terms of recruiting and what you can do about it.

Anyone?

[No verbal response.]

Senator KLOBUCHAR. OK. Should I call on people?

Mr. WATTS. Well, I'm sure they don't want to answer as much as I do, so I'll go first.

[Laughter.]

Mr. WATTS. What I would say is that what we should note is that these social media companies here were the forerunners, but they're also the dinosaurs of the social media era, meaning that they're the largest platforms and they have the greatest capability to actually deter this activity. But in the future, if I were a terrorist or an extremist group trying to mobilize, I would go to the smaller social media applications that have the greatest encryption, the largest dissemination capability, and I would focus there and then move to other social media platforms, because there would be less ability for them to deter my activity on the platform.

With that, in terms of the extremists, I think you need to look at what are the social media applications essentially being used by language—language is the key for actually doing recruitment—and where are the populations in each of your states and cities that are refugee populations, immigrant populations, and then how does that sort of play out, and who are they interfacing with overseas.

Senator KLOBUCHAR. One last question here. Throughout the 2016 cycle, Russians worked to influence the U.S. electorate, as I mentioned, and part of it was they did it by searching algorithms to promote misinformation. In the current news era, information is consumed rapidly, and algorithms play a significant role in deciding what content consumers see.

Mr. Monje and Ms. Bickert, what are Twitter and Facebook doing to help ensure the information appearing in search results and on consumers' feeds is free from that kind of exploitation?

Mr. MONJE. Thank you very much, Senator, for that question, and we do quite a bit to protect our search, in particular. More than 95 percent of our users as a default setting have safe search as part of their experience on Twitter. So what we do is when we identify a bot, malicious automation, which is a lot of the ways that this kind of information has promulgated on the internet, is that is severely down ranked so it's very hard to find.

Senator KLOBUCHAR. Ms. Bickert?

Ms. BICKERT. Thank you, Senator. We are increasingly finding new ways to disrupt false news and help people connect with authentic news. We know that's what they want to do. We're also investing in efforts to help people distinguish between the two, which includes basic education and public outreach.

As far as disrupting the false news, oftentimes—because we have a requirement that people have to use Facebook with their authentic name—if we can identify inauthentic accounts—and we're getting much better at that—we can remove those accounts and the false news goes away. The majority of the actors that we see trying to spread disinformation are financially motivated. So that goes a long way.

We're also working with our community to flag false news, send it to third-party fact checkers, and make that content less visible and put it in context. So now, if you come to Facebook and you see a story in your news feed that is an article that has been flagged as potentially false by our community, we will also show you some

related articles underneath it so that you have a sense of where this story sits in the broader spectrum of news.

We're working with responsible publishers to make sure that they know how to most effectively use social media, and then we're also working on user education campaigns.

Senator KLOBUCHAR. Thank you very much.

The CHAIRMAN. Thank you, Senator Klobuchar.

Senator Moran.

STATEMENT OF HON. JERRY MORAN,
U.S. SENATOR FROM KANSAS

Senator MORAN. Mr. Chairman, thank you. Thank you to you and the Ranking Member for conducting this hearing. I think it's one of the most interesting and potentially valuable hearings we will have had, and what a great development it would be if we could reduce the military necessity and the loss of life that comes from military action in fighting terror if we can keep it from occurring in the first place.

So thank you for being here. Thank you for your testimony. Let me ask this. Some of you covered in your testimony collaborative efforts among multiple businesses and groups that involved a shared industry database, which eventually led to the formulation of the Global Internet Forum to Counter Terrorism. I want to know more about that collaboration.

Part of the reason for that question is that my guess is that as larger social media companies become more innovative and effective in what you're attempting to accomplish, preventing terrorism, it would seem to me that other smaller platforms may become the platform of choice in this space. So if you're successful in your efforts, what prevents terrorists from moving to a different platform, and, therefore, what's those smaller platforms' engagement in what you're doing? It's directed at anyone who desires to answer. Or maybe if no one does——

Ms. BICKERT. Thank you, Senator. That is exactly what we were thinking as the large companies, was that we needed to make sure that this movement was industry-wide. With that in mind, we reached out to a number of small companies several years ago. I think we reached out to 18 companies initially. All 18 said yes, they wanted to meet to talk about best practices to counter terrorism.

We then met for more than a year before we ultimately launched the Global Internet Forum. Through that forum, which we launched in June, we've since had five international working group sessions with 68 smaller companies based around the world, and this is an opportunity for us to share expertise and learnings from the larger companies.

Senator MORAN. And let me take that a step further. So what are the smaller companies, smaller platforms, doing? They're a participant in this collaboration? They're doing something similar to what you're telling us that your companies are doing today?

Ms. BICKERT. Yes, Senator. Often, they are learning from what we are experiencing as the larger platforms in terms of the conduct that we see from bad actors, the policies we've put in place, and

how we're thinking about using technology and people to combat those threats.

Senator MORAN. Anyone else?

Mr. MONJE. I would just add that, you know, we've been extremely successful at taking terrorist content off of Twitter. It's a tremendous success for Twitter, but it doesn't eliminate the terrorists and them moving to other platforms like Telegram. It doesn't help everybody. You know, Twitter is a smaller company among the giants, and so we often—because we've had to be creative and innovative in our use of technology—can help be a bridge to the smaller companies and tell them you can make significant progress. You just have to invest in the technology.

Senator MORAN. What evidence do you see that terrorist organizations are changing their behavior as a result of what you're doing?

Mr. WATTS. There is open—you know, in some of their forums right now—they're trying to find a platform where they can go in a secure fashion, communicate, and push their propaganda around the world. So they're actively seeking new platforms.

And I think your question is a great one, which is how do we help these small companies that are developing new social media applications, who don't have the capabilities in terms of security, ward this off, and I don't think there's a good answer for that question. But they are seeking a new home. They just haven't found it yet.

Senator MORAN. Mr. Watts, is the response by terrorist organizations to seek a new home, or are they finding ways to hinder your efforts, or both?

Mr. WATTS. Both. They're looking for a place where they can communicate and organize, but they have to be able to push their propaganda globally in order to recruit and gain resources, so they need some way to do that. They will continue to try and exploit these small applications, but it's tougher for them on these small applications because, globally, there are not as many people on them.

So it's a better problem to have than what we've had in the past, but it really begs the question that, ultimately, one of these social media platforms that's popular overseas will start to gain traction with them, either due to its encryption capabilities or how they can connect with audience or how they can load and share videos. I think this is important across all extremist groups. If you look at some of the platforms that are out there, Reddit, 4Chan, these anonymous platforms, they also can be great tools, and it would be great to see them integrated with the bigger companies that have way more capability to detect that activity.

Senator MORAN. Thank you very much.

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Moran.

Senator Schatz.

STATEMENT OF HON. BRIAN SCHATZ,
U.S. SENATOR FROM HAWAII

Senator SCHATZ. Thank you, Mr. Chairman.

Thank you all for being here. Context—Facebook had $10.3 billion in revenue last quarter, about 23,000 employees, if I'm not mistaken; Twitter, $590 million in revenue last quarter, 3,700 employees. So my question for Facebook and Twitter is: In dollars, as a percentage of revenue, however you want to calculate it, and in terms of employee count, both part-time and full time, how many people, how many dollars are you devoting to this problem?

Ms. Bickert?

Ms. BICKERT. Thank you, Senator. This has been a significant area of investment for us, not just now, but over the past years. But I do want to point to a recent announcement from our CEO, Mark Zuckerberg, after we released earnings last quarter where he specifically pointed to the fact that as we invest in safety and security, people should expect to see Facebook's costs go up. That's reflected in the fact that we now have more than 7,500 people who are working to review terror content and other potential violations.

We have 180 people who are focused specifically on countering terrorism. So these are people like the former academics, like Brian Fishman, formerly with the West Point Counterterrorism Research Center, and others.

Senator SCHATZ. So 180 full time; 7,500, it's part of their job?

Ms. BICKERT. Seventy-five hundred are content reviewers. In the area of safety and security, more generally, we have 10,000 people currently. We are looking to be at 20,000 by the end of the year.

Senator SCHATZ. Mr. Monje?

Mr. MONJE. It is fewer than that. But I can tell you that our entire engineering, product, and design team at various stages are all working on this. We're a small team. We have to be supple. We have to be able to shift as the challenges move. The numbers that are really important also to look at are 2 billion users, 400 hours of video every minute, and, for us, 350,000 tweets every minute.

This isn't—in order to make progress on this issue, you do need to have humans, and we have former law enforcement; we have experts; we partner with contractors, consultants, academics——

Senator SCHATZ. I want to give you an opportunity to set the record straight about fake accounts. I've been reading a lot about this. I saw anywhere from 9 percent to 15 percent fake. I saw another USC study that said it's actually 48 million out of your nearly 300 million users. What's the number? How many fake accounts do you have?

Mr. MONJE. We believe that fewer than 5 percent of the accounts on Twitter are fake accounts.

Senator SCHATZ. Now, if you've kind of zeroed in on—let's stipulate that it's 5 percent of almost 300 million, right? If you know they're fake, what's the issue here?

Mr. MONJE. We are—they keep coming back, and they try different methods to get back on the radar screen, actually, and so we are, as a matter of course, consistently fighting malicious automation. We are now challenging 4 million malicious automated accounts a week. That means we are essentially sending them a note saying, "You're acting weird. Can you verify you're a human being?" That's double where we were last year.

Senator SCHATZ. Can I just talk to you about bots a little bit? I know this is a hearing about terrorism, primarily, but I think it's

worth asking what we're doing about active measures. You know, there was public reporting that the Roy Moore campaign went from 27,000 to 47,000 Twitter followers over the weekend, and a substantial portion of those appeared to be located in Russia. We had the take-a-knee thing where, clearly, there was an active measure to try to just sow discord. In other words, you've got bots and bot farms out there that are taking both sides of the argument.

So when we get into a conversation about active measures against our country, I don't think we should think of it as active measures against Democrats, and I don't think we should assume that it's just Russian active measures. We have to think of this as undermining democracy itself and undermining our ability to have our First Amendment rights exercised in any way that's meaningful.

So my question for you is—I mean, this is relatively recent, and it doesn't seem to—you can give us the measure of your activities, you know. Four million accounts are being challenged, and 500,000 accounts have been taken down. But based on results, you're not where you need to be for us to be reassured that you're securing our democracy. To the degree and extent that elected officials and people who vote and our adversaries are participating in your platform, how can we know that you're going to get this right and before the midterms?

Mr. MONJE. Yes, sir. Thank you for that question, and that's exactly the question that we ask ourselves every day. We think we're better prepared for this election than we've ever been. We are continually improving our tools, and we're going to get better, and we're going to report to the American people the results of our efforts.

Senator SCHATZ. Thank you.

The CHAIRMAN. Thank you, Senator Schatz.

Senator Young's not here.

Senator Markey.

STATEMENT OF HON. EDWARD MARKEY,
U.S. SENATOR FROM MASSACHUSETTS

Senator MARKEY. Thank you, Mr. Chairman, very much.

Last month, the FCC gutted the net neutrality rules that protected the Internet as we know it, and as a result, the next Facebook, the next YouTube, the next Twitter will struggle to get off the ground. I strongly oppose that FCC decision, which is why I plan to introduce a Congressional Review Act Resolution of Disapproval, which will undo the FCC's recent actions and restore the 2015 Open Internet Order. My resolution enjoys the support of Democrats. Susan Collins, the Senator from Maine, has indicated that she will vote for it.

My question to each company here is simple. Do you support my CRA resolution which would put net neutrality back on the books?

Mr. Monje.

Mr. MONJE. Yes, sir. Thank you for your leadership on this issue. It's an important issue for our company and for our users.

Senator MARKEY. You would support it. Thank you.

Ms. Downs.

Ms. DOWNS. We support strong enforceable net neutrality protections. We supported the 2015 rules, and we will support any effort to put those rules back in place.

Senator MARKEY. Thank you.

Ms. Bickert.

Ms. BICKERT. Thank you, Senator. Same answer. We will support the CRA, and we also support and will work with anybody who's interested in working to find a way to put those rules back in place.

Senator MARKEY. Thank you. We thank each of you. Thank you so much.

Next question. Bad actors can and do use the internet and social media to acquire weapons, including firearms. That's why in 2016, I wrote a letter to Facebook and Instagram asking why gun sales continue to take place on their sites, even after announcement of self-imposed policy changes aimed at eliminating this type of activity. I was pleased when both Facebook and Instagram announced they would prohibit individual users from buying and selling firearms on their sites.

Yet recent media reports indicate that users are still able to gain access to deadly weapons on social media. Just last month, the Chicago Police Department arrested 50 people in a case involving the sale of illegal guns in Facebook groups.

Ms. Bickert, it appears that gun sales on your platform may have moved into private Facebook groups. How is Facebook working to stop the sale of firearms in that corner of your platform? Notably, the Chicago Police Department said it did not receive cooperation from Facebook during its 10-month investigation. Law enforcement officials reported that Facebook hampered their investigation by shutting down accounts that officers were using to infiltrate the group in question.

Ms. BICKERT. Thank you, Senator. It's certainly an issue that we take seriously, and as a former Federal prosecutor based in Chicago, our relationship with law enforcement authorities is very important to us. We have cooperated with law enforcement and will continue to do so in that case.

We do not allow firearm sales. Enforcement has presented challenges for us, and to get better, one of the things we're doing is working on our technology. Anybody in the community can report gun sales to us, and we will take action, and that's important, and that does happen even in private groups. But we know we need to do more, and that's why we're now using things like image prediction technology to help us recognize when those sales might be taking place.

Senator MARKEY. So since Instagram can turn into Instagun, you know, for someone who intends on using it for nefarious purposes, the answer that you would give to the Chicago Police Department when it said it did not receive cooperation from Facebook during its 10-month investigation is that you did cooperate or that you have now established a policy of cooperation with the Chicago Police Department and every police department across the country?

Ms. BICKERT. Thank you, Senator. I believe they clarified their statement afterwards. We have been cooperative with them from

the beginning, and I would be happy to follow up afterwards with you on that.

Senator MARKEY. That would be helpful. And in terms of the private Facebook groups that this type of activity has migrated to, you are saying as well that you are working to shut that down as well?

Ms. BICKERT. That's right, Senator. This is an area where we recognize enforcement can be challenging, and we have to be proactive in looking for solutions. So we're trying to make it easy for people to report, but also going further to look for this content.

Senator MARKEY. Thank you. And that's why this hearing is so important.

I would thank you, Mr. Chairman, because the internet can be used to spread hate, but it can also be used to spread weapons of war into the hands of those who are the haters and do enormous harm in all of our communities across the country. So thank you.

We thank each of you for your testimony.

The CHAIRMAN. Thank you, Senator Markey.

I would just ask the three, too, that you all, I assume, would support legislation that would put in place rules for an open Internet as well. Would that be true?

Mr. MONJE. Twitter has long been a supporter of net neutrality, and, hopefully, Congress can develop good rules.

Ms. DOWNS. Same answer.

Ms. BICKERT. Same answer, Mr. Chairman.

The CHAIRMAN. Very good.

Next up is Senator Baldwin.

STATEMENT OF HON. TAMMY BALDWIN, U.S. SENATOR FROM WISCONSIN

Senator BALDWIN. Thank you, Mr. Chairman and Ranking Member, for this important hearing.

Much of the conversation today has been focused on addressing foreign terrorist organizations' use of your platforms as tools to recruit and radicalize individuals both here and abroad. I'd like to turn to how you are addressing the use of social media to further domestic extremism. Whether it's the vehicular attack on counter-protestors in Charlottesville this summer or the 2012 shooting at a Sikh temple in my home state of Wisconsin, we've seen numerous individuals subscribing to racist ideologies turning to violence.

Beyond that, there's a disturbing increase in hate crimes in this country, as documented by FBI's limited collection of data from state and local law enforcement. As with other forms of extremism, social media is undoubtedly playing a role in spreading these ideologies and channeling these individuals into violent action. How are your companies working to address the role of social media in furthering domestic extremism, particularly white nationalist or white supremacist violence?

I'd like to start with you, Ms. Bickert.

Ms. BICKERT. Thank you, Senator. I want to be clear that our policies prohibit any group that is either a violent organization—and that's regardless of ideology. So if it is a domestic terror organization, if it's a foreign terror organization, no matter what the ideological underpinning is, they are not allowed on Facebook.

But we also prohibit hate organizations, and these are groups that are propagating hate based on a protected characteristic, like race, religion, gender, gender identity, and so forth. The same consequences under our policies apply. They're not allowed to be on our platform. People cannot praise or support them.

Senator BALDWIN. Ms. Downs.

Ms. DOWNS. Thank you, Senator. Our violent extremism policies apply to violent extremism in all its forms, including white supremacy and other forms of hatred, and we apply our policies against incitement to violence and violent ideology consistently across violent extremism in all its manifestations.

Senator BALDWIN. Mr. Monje.

Mr. MONJE. That's a very similar answer for Twitter as well. We don't allow violent extremist groups. We don't allow glorification of violence. I think it's also when—you know, Charlottesville was a hard day for a lot of folks, and I think what you saw not only online was the very small minority of folks who were saying terrible things, but the vast majority of folks who were coming out to reject it.

Senator BALDWIN. I'm going to turn to a different topic. I'm concerned by President Trump's and Secretary of State Tillerson's reluctance to support, fund, and staff the State Department's Global Engagement Center, which is tasked with coordinating U.S. efforts to counter extremist propaganda and recruitment as well as Russian active measures like disinformation.

I'd like to hear from each of the companies about their experiences working with the Center and how cooperative efforts could be improved.

And, Mr. Watts, what are the national security impacts of this administration's failure to prioritize the Center, especially in the context of Russia?

Why don't we again go right down the line?

Ms. Bickert.

Ms. BICKERT. Thank you, Senator. We are committed to working with governments around the world in promoting and finding counter-speech solutions. We have worked with the Global Engagement Center and others in the U.S. Government. We have found that collaboration to be effective. Often, what we find is that government can be very effective as a convening power for bringing together civil society stakeholders and then industry and researchers to get together and share their knowledge. That's something that we hope to continue in the future.

Ms. DOWNS. Thank you, Senator. Our efforts to combat terrorism on our product obviously start with making sure we're removing the most egregious content. But an equally important part of the strategy is our investment in counter-speech, to do the hearts and minds work to address these issues at their root. So we meet regularly with NGOs and government actors, including the State Department and the Global Engagement Center, to talk about counter-speech and the importance of investing in that work.

Senator BALDWIN. Mr. Monje.

Mr. MONJE. A very similar answer as well, in that, you know, government does have an important role in combating this issue

and not only investing in counter-speech but investing in the groups that are authentic voices in their communities.

Senator BALDWIN. Mr. Watts.

Mr. WATTS. I'm absolutely baffled as to why the Global Engagement Center—they received that mission, from what I understood, in 2016 before the election. Senator Portman, if I recall, was one of the leaders of that, and I had actually communicated with their staff on the Russia issue. At a bare minimum, the U.S. Government needs to have a real-time understanding of what Russia is doing in social media.

The Hamilton 68 platform I've tried to provide to the U.S. Government directly through multiple agencies. I have briefed the U.S. Government since 2014 in different contexts on Russian active measures. I sit here today and I have no answer for you.

I don't understand why we wouldn't, at a minimum, regardless of the outcome of the election in 2016, want to equip our intelligence agencies, our law enforcement agencies, and the Department of Defense with just an understanding—we don't even have to counter—just an understanding of what Russian active measures are doing around the world. There's no excuse for it. I can't understand it.

The CHAIRMAN. Thank you, Senator Baldwin.

Senator Udall.

STATEMENT OF HON. TOM UDALL,
U.S. SENATOR FROM NEW MEXICO

Senator UDALL. Thank you very much, Mr. Chairman. I really appreciate you and the Ranking Member pursuing this very, very important topic.

Terrorism and social media is a challenging and, I think, pressing subject, and I recognize that technology companies cannot solve this alone. But they must do more, and I think that has been highlighted by the questioning you've seen here today. I'm focusing—my first question is similar to Senator Baldwin's. I'm particularly concerned about the explosion of white supremacists online.

In December, after years of posting fantasies about school shootings and hate-filled racist rants over many internet platforms and many other identities, a young man took a gun to a local high school in Aztec, New Mexico, and killed two students before taking his own life. And listening to you, I'm wondering, you know, what can be done in this kind of situation?

Ms. Bickert, in your testimony, you highlighted the efforts that Facebook is taking to counteract ISIS and other foreign terrorists. But can you speak to the efforts Facebook is taking to fight one of the most and biggest—one of the biggest threats to us in the United States, domestic terrorists like white supremacists? I mean, in this kind of situation where you have an individual under various identities taking positions and indicating right on the edge of violence, what can be done in this kind of circumstance, and have you run into situations like this before?

Ms. BICKERT. Thank you, Senator. It's certainly an important issue. We stand against violence in all its forms, and we don't allow any violent organization, regardless of ideology. If we become aware of a threat of violence, credible threats of eminent harm to

somebody, we proactively reach out to law enforcement authorities, and that is something that we have done in cases where we've seen a threat like a shooter. Whatever the ideology is, it doesn't matter. We will proactively provide that to law enforcement.

Senator UDALL. Mr. Watts, do you think more could be done here based on the answers you hear?

Mr. WATTS. In terms of domestic extremism, I side with the social media companies in the sense that it's difficult to understand where to fall, because there's not good leadership from the U.S. Government about what a domestic extremism group is. We have the luxury——

Senator UDALL. Do you think we could do more there, in terms of the government?

Mr. WATTS. Yes. If we delineate more appropriately as a Federal Government, we can then enable the social media companies to effectively draw the line. I don't like the social media companies having to decide what is free speech versus violent speech or extremist versus norm. It puts them in a terrible position. I also don't think it's good for business and their platforms.

At the same time, you know, how do you do that short of a violent threat or an eminent threat? To do that, we would have to have the equivalent of an FTO or a foreign terrorist designation program in the domestic context. I'm not sure how we get there.

Senator UDALL. And that may be something we should consider, is how to urge the government to be more specific here and outline areas where we could work with industry in order to move that along.

Mr. WATTS. I think—so it's difficult, even from the FBI's perspective, that there are two different playbooks. There's the international terrorist playbook and the domestic terrorist playbook, and without that formalization of what an extremist group is or an extremist, individually, is, it's really hard, I think, for any corporation or company to evenly and legitimately enforce any sort of regulation on a user or a group.

Senator UDALL. Mr. Monje, in your testimony, you outlined Twitter rules against terrorism that expressly include that users—and I'm quoting here—"cannot make specific threats of violence or wish for serious physical harm, death, or disease of an individual or group of people," end quote. I'm curious, then, what Twitter's position is on one of the President's video tweets, where he was body slamming a person with the CNN logo superimposed on their face. The video appears to promote serious physical harm to CNN reporters in the context of an alarming increase in violence against reporters in the U.S.

Mr. MONJE. Thank you very much, Senator, for that question. No Twitter user is above the Twitter rules. As we action accounts on any given moment, we are looking whether they are trying to do satire, whether they're trying to do humor, even if it's not successful humor. We also recognize that world leaders do have a special voice, and it is in the public interest for their constituents to hear from them.

Senator UDALL. Well, I don't think this was humor, and I don't think the result—I think if you look at the—what CNN reporters have said since this, there's more violence toward them. There's

more animosity toward them. I think you need to look at it in the whole context, and I would encourage all the companies at this table to take threats to journalists very seriously. I'm extremely concerned when any threats of violence-based reporting that the President finds disagreeable with our President calling U.S. media outlets "the enemy of the people." I think it is up to all of us to safeguard the First Amendment.

Thank you very much, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Udall.

Senator Tester.

STATEMENT OF HON. JON TESTER,
U.S. SENATOR FROM MONTANA

Senator TESTER. Thank you, Mr. Chairman and Ranking Member Nelson. Thanks for having this hearing, and I want to thank the witnesses today.

I think we'll start with you, Mr. Watts. Can you tell me why transparency behind who's paying for political ads and issue-based ads is important?

Mr. WATTS. Yes. I think the number one issue with that is public safety. We saw with Russian active measures one of the primary things they sought to do was actually mobilize the population regardless of the election, mobilize people to protest or counter protest, which can lead to violent confrontations. At the same point, those advertisements, when annotated and noted based on campaign laws, give legitimacy to those advertisements so that the public actually knows what is a real political stance versus a false or manipulated truth or a narrative, they have to stand behind their actions. I think that's important for the public to restore trust and faith in the democratic processes.

Senator TESTER. Could it also have impacts on election results?

Mr. WATTS. Yes. It makes it more difficult for a foreign adversary or even a social media manipulator with a lot of resources and an ax to grind to do character assassination or to tear down social movements.

Senator TESTER. OK. So—and this goes to any one of the other three that wants to answer this. Can you tell me why you don't tell us who's paying for the ads, whether they're political ads or whether they're issue-based ads? Who wants to answer that?

Mr. MONJE. Thank you, Senator. Twitter is very proud that we last year announced industry-leading transparency practices for political advertising.

Senator TESTER. So do you tell people who's paying for the ads?

Mr. MONJE. For electioneering ads, yes, sir.

Senator TESTER. How about issue-based ads?

Mr. MONJE. Issue-based ads are a harder–to crack. It's harder to determine—and we are working with our colleagues, with our peer companies, to try to figure out what the right way to address those issues are.

Senator TESTER. OK. How about the other two? Do you want to talk about political ads versus issue-based ads and if you're telling us who's paying for them?

Ms. DOWNS. We're working to put more transparency into the election-based advertising system and are taking four steps in ad-

vance of the 2018 midterms. The first is verification. We will require advertisers to identify who they are and where they're from before purchasing advertisements. We'll also launch in-ad disclosures where we notify users of who is running an election-based ad. We'll release a transparency report on election advertising purchased through Google, and we'll also release a creative library to the public where all of those advertisements are made public.

Senator TESTER. And will that release of the transparency report have who's paid for the ads?

Ms. DOWNS. I believe it will, yes, sir, Senator.

Ms. BICKERT. Thank you, Senator. Our answer is substantially similar to my peer companies on the issue of Federal election related ads. And like Mr. Monje, political ads, broadly, is a little bit more complicated, but certainly an area where we think increased transparency is important.

Senator TESTER. Political ads are more complicated than issue-based ads, or the other way around?

Ms. BICKERT. The issue-based—I'm sorry, Senator. The issue-based ads—they're hard to define. But that said, we're very interested in how we can increase transparency, and we look forward to talking to yourself and other policymakers about it.

Senator TESTER. Well, I would just tell you this as an editorial comment. I would agree with Senator Schatz. I don't think this is a Democrat-Republican issue. I think this is a democracy issue, and you guys are smart guys, and just about everybody I read writings of tell me that it's not that difficult, and they're smart people, too. So I would hope that you guys really would put pen to paper, if that's what you do these days, and figure out how you can let people know who's paying for ads. And I think issue-based ads, by the way, are just as important as political-based ads, because those fall into the political category, and I would just say that's important.

Every one of you said that you did not like the FCC decision on net neutrality that came out a month or two ago. During that debate, we had learned that there were bots that dropped comments into the hopper that distorted the whole public comment period. How is that going to be stopped the next time we have a public comment period on a rule that's written by an agency?

Anybody want to answer that? And I'm out of time, so make it quick.

[No verbal response.]

Senator TESTER. I'll tell you what. We'll not occupy the time of the Committee. Give me an answer to that in writing when you go back to your folks.

This is a really important issue. I just want to say this is a really important issue, from a terrorist standpoint, from all the questions that were asked before. But our democracy is at risk here. We've got to figure out how to get this done and get it done right and get it done very quickly, or we may not have a democracy to have you guys up to hear you out.

The CHAIRMAN. Thank you, Senator Tester.

Senator Young.

STATEMENT OF HON. TODD YOUNG,
U.S. SENATOR FROM INDIANA

Senator YOUNG. Well, thank you, Mr. Chairman, for holding this hearing on terrorism and social media.

YouTube went from having 40 percent of its post takedowns last June, being identified by algorithms, AI, and machine learning, to 98 percent today. Twitter went from roughly 33 percent detection of terrorist accounts in 2015 to more than 90 percent of the detections today, again attributable to algorithms, AI, or machine learning. Facebook has stated that nearly 99 percent of ISIS and al Qaeda related content is detected and removed before anyone even reports it.

So what is—Ms. Bickert, Ms. Downs, Ms. Monje, what's responsible for the recent increase in the use of AI and machine learning for this purpose of taking down posts? Is it primarily because of a new commitment to take down posts by your companies, or is it simply that the technology is finally at a place to be effective, or some combination thereof? We'll start with Ms. Bickert.

Ms. BICKERT. Thank you, Senator. It's definitely a timely question. These innovations have been happening over the years. We have seen a lot of improvement, particularly over the past one to two years at Facebook. A lot of these efforts have been in place since I joined the company 6 years ago, such as still image hashing, but it has gotten better.

In the fall of 2016 is when we finally found video hashing to be sufficiently reliable, where we could use it to detect these terror propaganda videos. And for some of them, like a beheading video, that we know violates our policies regardless of how it's shared, we could actually accurately identify it and stop it at the time of upload. That's something we've been trying to do for a while and had not been able to do.

Another area where we've gotten better is in detecting recidivists. So we take down the bad account. They try to come back. That's something that for a variety of reasons has been important to the company for years, but an area where we've made significant progress in the past one to two years.

And then the final advance I'll point to before turning to my colleagues is in the area of natural language understanding. This is hard. We have many different languages that we support on Facebook, and when you train these models, they have to be trained on sufficient data. So this process takes a long time, but we are making progress here, and we're now using it in the area of terrorism where we couldn't before.

Senator YOUNG. Thank you.

Ms. Downs.

Ms. DOWNS. Thank you, Senator. We've always used a mix of technology and humans to enforce our policies, and as technology gets better, we see it doing more of the heavy lifting in detecting the content that violates our policies and needs to be removed. These are a reinforcing loop, where as humans make judgments about what content violates our policies, that feeds back into the training set of data to teach the classifiers and algorithms what they're looking for. So the more content we review over time, the

better and better these classifiers get and the more they're able to detect the content that needs to be removed.

Senator YOUNG. Before I turn to—my apologies—Mr. Monje—I just stepped into the room—I would just note you referenced human judgments and how that feeds into an algorithm to help make more informed decisions moving forward. There won't be time to explore it here, but one of the things I really want to learn more about is what parameters are used to determine, by a human, what is an appropriate or an inappropriate post, and is there transparency, or will there be transparency about that decisionmaking process? But, again, that's for another day since I have 47 seconds left.

Mr. Monje.

Mr. MONJE. Yes, sir. I'd just very briefly—we approach it very similarly to our peer companies and are constantly trying to figure out ways that we can use our technology and feeding it the input so that it can tackle—the AI can tackle increasingly more difficult and more nuance challenges.

Senator YOUNG. OK. I'll just note in the remaining time here that I really enjoyed visiting with Yasmin Green, Director of R and D at Google's Jigsaw group. I'll say that Alphabet is doing some really great work there, and I look forward to working with all of you to improve how we remove this horrible content from the Internet and keep Americans more safe and secure.

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Young.

Next up is Senator Blumenthal.

STATEMENT OF HON. RICHARD BLUMENTHAL, U.S. SENATOR FROM CONNECTICUT

Senator BLUMENTHAL. Thanks, Mr. Chairman. Thank you for holding this hearing and thank you to this really all-star panel for being here today.

Mr. Watts, I find your testimony absolutely chilling. The Internet is a potential monster when it comes to extremists and terrorism, and it requires the kind of inventive and robust investment attitude that, in fact, created the internet. I've been reading a book called *The Innovators* by Walter Isaacson, and it is an inspiring account of how we came to have the Internet and social media, involving heroes whose names have been largely lost to history, including some Nobel Prize winners.

But the point that he makes that I think is so relevant to this discussion is that the Internet itself is the result of a partnership between private industry and inventors, government, and academia, and those partners are as necessary for this effort in combating terrorism and extremism as they were in inventing the platforms themselves.

I want to join in thanking you for your commitment to net neutrality. I also want to thank you for the commitments that your companies have made with varying degrees of enthusiasm to our anti-sex trafficking efforts, most especially SESTA, which, hopefully, will come to a vote. I encourage you to enlist more of your colleagues in that effort.

I want to highlight the importance of the nationalist hate groups and extremist groups that have come to pose a very dire extremist threat. You received a letter signed by 19 civil rights groups, including Muslim advocates, on October 30, 2017. It was co-signed to Facebook, but it's equally applicable to all of your companies. I ask that it be made a part of the record, and I see the Chairman has stepped off, but I'm assuming there will be no objection and it will be made part of the record.

[The information referred to follows:]

October 30, 2017

Mr. Mark Zuckerberg, Chief Executive Officer
Ms. Sheryl Sandberg, Chief Operating Officer
Facebook, Inc.
1 Hacker Way
Menlo Park, CA 94025

Dear Mr. Zuckerberg and Ms. Sandberg,

We, the undersigned civil rights, interfaith, and advocacy organizations write to express our deep concern regarding ads, pages, and hateful content on your platform used to divide our country, and in particular, to promote anti-Muslim, anti-Black, anti immigrant, and anti-LGBTQ animus. We thank you for recent meetings with some of our organizations representing communities that were directly affected by the material on your platform. We appreciate that senior members of your team—including you, Ms. Sandberg—have facilitated these meetings, and we hope that these conversations are the beginning of a serious and ongoing dialogue. Now, it is necessary for Facebook to take critical steps to address the bigotry and discrimination generated on your platform.

As you know, we do not yet have access to all the divisive content targeting communities we represent; therefore, we are only able to cite to the few examples that were leaked to the media.

For example, Russian operatives set up misleading accounts impersonating or posing as American individuals and groups on Facebook to promote Russian propaganda during the American election season. Reports indicate that a Russian Facebook account called "SecuredBorders" posed as a group of U.S. citizens concerned about the increased number of refugees in America. This fake account not only promoted anti-immigrant messaging online, but also managed to organize an in-person anti-refugee rally in Twin Falls, Idaho in August 2016.[1]

In addition, a Facebook page entitled "United Muslims of America" was an imposter account traced back to Russia[2]—the real United Muslims of America is a California-based interfaith organization working at the local level to promote dialogue and political participation.[3] The imposter account smeared political candidates and promoted political rallies aimed at Muslim audiences.[4] In another example, the Internet Research Agency in Russia promoted an anti-Muslim rally thousands of miles away in Houston, Texas where individuals protested outside of a mosque.[5] Additional reports indicate that Facebook offered its expertise to a bigoted advocacy group by creating a case study testing different video formats, and advising on how to enhance the reach of the group's anti-refugee campaign in swing states during the final weeks of the 2016 election.[6] These examples of content on Facebook were not only harmful, but also used to rile up supporters of President Trump.

Furthermore, it has been reported that Russian operatives purchased Facebook ads about Black Lives Matter—some impersonating the group and others describing

---

[1] Geoffrey Smith, "Russia Orchestrated Anti-Immigrant Rallies in the U.S. via Facebook Last Year," Fortune, Sept. 12, 2017, available at *http://fortune.com/2017/09/12/russia-orchestrated-anti-immigrant-rallies-in-the-u-s-via-facebook-last-year/*.

[2] Dean Obeidallah, "How Russian Hackers Used My Face to Sabotage Our Politics and Elect Trump," The Daily Beast, Sept. 27, 2017, available at *https://fwww.thedailybeast.com/how-russian-hackers-used-my-face-to-sabotage-our-politics-and-elect-trump.*

[3] United Muslims of America "About" page, available at *http:/fwww.umanet:orgfabout-us.*

[4] Obeiallah, *supra* note 1.

[5] Tim Lister & Clare Sebastian, "Stoking Islamophobia and secession in Texas—from an office in Russia," CNN Politics, Oct. 6, 2017, available at *http://www.cnn.com/2017/10/05/politics/heart-of-texas-russia-event/index.html.*

[6] Melanie Ehrenkranz, "Facebook Reportedly Used Anti-Muslim Ad as Test Case in Video Formats," Gizmodo, Oct. 18, 2017, available at *https://gizmodo.com/facebook-reportedly-used-anti-muslim-ad-as-test-case-in-1819645900.*

it as a threat.[7] This included ads that were directly targeted to reach audiences in Ferguson, Missouri and Baltimore, Maryland. CNN reports that the Russian Internet Research Agency used these ads in an attempt to amplify political discord and create a general atmosphere of incivility and chaos.[8] This included a fake ad containing an image of an African-American woman dry-firing a rifle, playing on the worst stereotypes regarding African-Americans as threatening or violent.[9]

We were alarmed to see your platform being abused to promote bigotry, and especially disappointed that it has taken media exposure and congressional oversight to give a degree of transparency into your practices. It is important to keep in mind that pervasive bigotry has long existed on your platform, and the Russian operatives simply exploited the hateful content and activity already present. We are concerned about how a platform like Facebook's could operate without appropriate safeguards that take into account how it could be manipulated to further sow divisions in our society.

As a company and social network platform whose mission is "to give people the power to build community and bring the world closer together,"[10] we hope that you understand the gravity of this hateful rhetoric and behavior. During a time when anti Muslim, anti-Black, anti-LGBTQ, and anti-immigrant sentiment has swept the nation, it is more important than ever for companies like yours to take an unequivocal stance against bigotry.

Over the years, many of us have raised concerns about how your platform may have a negative impact on our communities, with disappointing results. For example, we have requested that you address attacks on African Americans and Muslims, organizing by hate groups, and the censorship of Black, Arab, Muslim, and other marginalized voices. As a result of the pervasive presence and organizing by hate groups on your platform-some could not exist as national level entities without it we have repeatedly requested that you convene a gathering with civil rights organizations to discuss appropriate and strategic responses. While you were unable to sufficiently respond to the concerns raised above, Facebook participated in and organized events that stigmatized Muslims and other communities such as a recent convening called "Tech Against Terrorism."

Though in the past you have displayed a willingness to listen to our concerns, we have yet to see meaningful change. It is our hope that recent developments will mark a new chapter in Facebook's commitment to protecting the rights of all who use your platform.

As we continue this important dialogue, we urge you to:

1. Fully disclose to the public all of the ads, pages, events, accounts, and posts you have traced back to Russian operatives targeting African American, LGBTQ, and Muslim communities. In particular, we believe that Facebook has a special responsibility to notify those individuals and organizations who have been impersonated or misrepresented.
2. Bring on an independent third-party team to conduct a thorough and public audit of the civil rights impact of your policies and programs, as well as how the platform has been used by hate groups, political entities, and others to stoke racial or religious resentment or violence. Other leading companies in the industry like Airbnb have made the decision to conduct such an assessment, and we hope you will follow their lead.
3. Regularly convene a new working group of a diverse group of civil rights organizations working to counter bigotry, and solicit input on policies and processes from this group. And, integrate addressing hate into Facebook's corporate structure by:

---

[7] Adam Entous, Craig Tim berg, & Elizabeth Dwoskin, "Russian operatives used Facebook ads to exploit America's racial and religious divisions," The Washington Post, Sept. 25, 2017, available at *https://www.washingtonpost.com/business/technology/russian-operatives-used-facebook-ads-to-exploit-divisions-over-black-political-activism-and-muslims/2017/09/25f4a011242-a21b-11e7-ade1-76d061d56efa_story.html?tid=sm_tw&utm_term=.e49cecc1a834.*

[8] Dylan Byers, "Exclusive: Russian-bought Black Lives Matter ad on Facebook targeted Baltimore and Ferguson," CNN Media, Sept. 28,2017, available at *http://money.cnn.com/2017/09/27fmedia/facebook-black-lives-matter-targeting/index.html.*

[9] Adam Entous, Craig Timberg, & Elizabeth Dwoskin, "Russian Facebook ads showed a black woman firing a rifle, amid efforts to stoke racial strife," The Washington Post, Oct. 2, 2017, available at *https://www.washingtonpost.com/business/technology/russian-facebook-ads-showed-a-black-woman-firing-a-rifle-amid-efforts-to-stoke-racial-strife/2017/10/02/e4e78312-a785-11e7-b3aa-c0e2e1d41e38_story.html?utm_term=.aa2267a2f46c.*

[10] Facebook "About" page, February 4, 2004, available at *https://www.facebook.com/pg/facebookf/about/?ref=page_internal.*

a. Assigning a board committee with responsibility for assessing management efforts to stop hate groups, state actors, and individuals engaged in hate from using your platform and tools;
b. Assigning a senior manager who is a member of Facebook's Executive Team with authority to oversee addressing hate company-wide and name that person publicly and employing staff with expertise in this area to vet advertisements and develop process and procedures the address this issue; and,
c. Creating a committee of outside advisors with expertise in identifying and tracking hate who will be responsible for producing an annual report on the effectiveness of steps taken by Facebook.

4. Develop, with input from diverse civil rights groups and experts, and make public a clear process for how Facebook:
   a. Reviews content constituting hate speech;
   b. Reviews efforts to use Facebook as a platform to stoke identity-based, racial, or religious resentment or violent actions; and,
   c. Responds to complaints about content that reasonably creates fear and chills speech on Facebook.
5. Make public detailed information regarding training and support for anti immigrant, anti-Muslim, anti-black, and anti-LGBTQ organizations, including the monetary value of these services; and establish a fund to provide grants to organizations combating hatred and bigotry.

Thank you in advance for your consideration. Please contact Naheed Qureshi at naheed@muslimadvocates.org with any questions.

We look forward to your reply.

Sincerely,

Arab American Institute (AAI)
Asian Americans Advancing Justice/AAJC
Center for Media Justice
Center for New Community
Color of Change
CREDO
Human Rights Campaign (HRC)
The Leadership Conference on Civil and Human Rights
League of United Latin American Citizens (LULAC)
MoveOn.org
Muslim Advocates
NAACP
NAACP Legal Defense and Educational Fund, Inc. (LDF)
National Center for Lesbian Rights
National Hispanic Media Coalition
National LGBTQ Task Force
National Sikh Campaign
Sikh Coalition
Southern Poverty Law Center

Senator CORTEZ MASTO. No objection.

[Laughter.]

Senator BLUMENTHAL. And I'm not willing to yield a part of my time to address that objection.

The Southern Poverty Law Center has warned that social media has been instrumental to the growth of the alt-right movement, allowing legions of anonymous Twitter users to use the hashtag alt-right to push far right extremism. On YouTube, for example, it's easy to find anti-Semitic content. All of these forms of extremism, often white supremacist extremism, have been allowed to flourish, and they pose a real and present danger.

In the time that I have left, I want to ask about a letter that I wrote to Facebook, Google, and Twitter calling on these companies to individually inform all users who are exposed to false, mis-

leading, and inflammatory posts generated by Russian agents. I'm assuming that none of you have any doubt that the Russians meddled in our 2016 election and attacked our democracy. Any question?

[No verbal response.]

Senator BLUMENTHAL. None. And that the investigation of those efforts is not a hoax, or a witch hunt, that this danger is continuing, as Mr. Watts has so dramatically and powerfully outlined, and that they will continue to do it unless they're made to pay a price, and those who colluded and cooperated with them are made to pay a price.

I want to thank Facebook for its substantive response in terms of its commitment to providing consumers with an online tool to inform users if they have interacted with Russian-sponsored pages or accounts. I'm hopeful that Facebook will do even more with more robust steps to further increase transparency in the future, but I am very, very grateful for your beginning.

And I just want to be blunt. I am disappointed by Google's written response. It essentially blew off my concerns by saying the nature of the platform made it difficult to know who has viewed its content. I look forward to responses from Twitter and others. If you want to respond now, I would be eager to hear what your response is to the letter that I wrote.

Mr. MONJE. Thank you, Senator, and we have briefed your staff on our plans, and we'll be rolling out the fulsome response shortly.

Senator BLUMENTHAL. And what will that response be?

Mr. MONJE. We will be working to identify and inform individually the users who may have been exposed to the IRA accounts during the election.

Senator BLUMENTHAL. Thank you. I think it's so tremendously important that we have all hands on deck in dealing with this threat, not only the companies that are represented here, but, again, as Mr. Watts said, some of the smaller actors, some of the newer ones. And there will be others coming that provide, in effect, platforms for hate, extremism, terrorism, division, chaos. In some ways, they are the biggest threat to our democracy today, those groups that want to foster hate.

And, of course, the Russians will continue. They have an asymmetric advantage here. It's an absolutely wondrous investment for Vladimir Putin. He gets more return on the dollar than any other investment he can make in sowing chaos and discord in our democracies, and we must be as inventive as the innovators were, the inventers of the internet, in combating this threat to our democracy.

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Blumenthal.

Senator Cortez Masto.

STATEMENT OF HON. CATHERINE CORTEZ MASTO, U.S. SENATOR FROM NEVADA

Senator CORTEZ MASTO. Thank you.

Welcome and thank you for this conversation. I wanted to start with, unfortunately, a horrific tragedy that occurred in my home town on October 1, 2017. Las Vegas experienced the worst tragedy

that we have ever seen, and it is the worst mass shooting in American history. As we were dealing with the horrific tragedy of the situation and trying to gain information, particularly for law enforcement purposes, unfortunately, a lot of misinformation was being spread after that tragedy on some of your platforms and on the internet, and, particularly, misinformation about the shooter was highlighted on both Google and Facebook.

Obviously, that's incredibly unhelpful for law enforcement, particularly as we move through an unfolding potentially dangerous situation. I know both Facebook and Google cited the need to make algorithm improvements to fight the spread of fake news during a crisis. What do you see as your companies' roles in fighting fake news, especially during a crisis such as a mass shooting or a terrorist attack, and what specific and verifiable metrics can you provide us to ensure our trust in these remedies? I'll start with Ms. Bickert.

Ms. BICKERT. Thank you, Senator. What happened in Las Vegas was horrific, and there were false news stories that we saw that we did address, but not fast enough, and it's an area where we're trying to get faster. We've changed the way that our crisis center operates so that we can make sure that that type of false news story does not appear in the headlines that people are seeing. The crisis center can be incredibly useful during times like this.

In Las Vegas, we saw people using not only our safety check, which allows people to say that they're safe, but also coordinating help, offers of housing and assistance to people throughout the city. So we want to make sure that's working effectively. Things we're doing: removing the bad accounts that are propagating this false news, making algorithmic changes to make news that is likely to be false less visible on the site, providing related articles when people see a news story that has been flagged as something that might be false so that they can see the broad spectrum of information across the internet, and then working with responsible publishers to make sure that they know how to use our tools to get their stories out there.

Senator CORTEZ MASTO. Thank you.

Ms. Downs.

Ms. DOWNS. Thank you, Senator, and my heart goes out to the City of Las Vegas and all the victims of that senseless tragedy.

We take misinformation on our platforms very seriously, and we've made a lot of efforts in our products, from improvements to our ranking algorithms to highlight authoritative sources and to demote low-quality or less reliable sources, particularly when users are seeking news content. We also have strict policies in place against the monetization of news sites that are misrepresenting themselves in order to remove the financial incentive to create and distribute fake news.

Senator MASTO CORTEZ. Thank you.

Mr. MONJE. And a very similar answer for us. I'd only add that one of Twitter's great advantages in the world is that it's fast. It's faster than television news often. We try to arm emergency responders with the knowledge of how to use that as a strength, and so it's one of our key pieces.

During the hurricanes in the Gulf Coast, we were actively working with folks who were responding—they were actually folks in Texas and Houston who were using our platform to identify people to rescue. And so it's one of the strengths of our platform, and, like everyone, it's a continuing challenge to address misinformation.

Senator CORTEZ MASTO. Thank you.

Mr. Watts, would you like to address this, or is there anything else that can be done?

Mr. WATTS. I don't know in terms of the technical things that could be done. But I do think the spread of misinformation so quickly like that—the first thing that you see is what you tend to believe over time. That which you see the most is what you tend to believe as well. It really empowers social media manipulators if you can do amplification through social bots, or if you can generate other systems to push the news quicker than everybody else, and so you see a lot of gaming in terms of trending hashtags and things like that.

I think there has to be some sort of trip that you can put in technically over time—and I'm sure that all these companies are trying to develop—that will tamp that out. When you see an artificial spike in any one of those trends, you should be able to detect it, and I think they're advancing on that. But it's a huge public safety issue, regardless of the threat actor that's employing it.

Senator CORTEZ MASTO. Thank you. I know I'm running out of time, but let me just say this. I had the opportunity to work with Facebook on our Internet Crimes Against Children Task Force in Nevada when I was Attorney General, and I will tell you that for every company that we reached out to, whether it was YouTube, Google, they were willing to work with law enforcement.

So I know there has been a lot of discussion on that interaction that you've had with law enforcement, but I've seen it from one side of it. I know now there is this balance we need to find to figure out how we continue to work together to address these evolving crimes and activity that's happening on the internet, and I'm grateful that you're here, and I look forward to figure out how we can continue to evolve that relationship as well. So thank you.

The CHAIRMAN. Thank you, Senator Cortez Masto.

Senator Lee.

STATEMENT OF HON. MIKE LEE,
U.S. SENATOR FROM UTAH

Senator LEE. Thank you, Mr. Chairman.

Thanks to each of you for being here. We live in an exciting world. We live at a time when the companies represented at this table today 15 years ago were just ideas, and today, they've changed the way we interact with the world around us. Today, these companies have made it possible in ways never imagined just a couple of decades ago for a few people with very little money to have an impact, not only in their community, but across the country and throughout the world. But with that comes a lot of challenges, and those challenges are the reason why we're here today at this hearing.

In some parts of the world, there has been a suggestion that I can summarize only as an effort to make public utility companies

out of social media enterprises that would rather comprehensively attempt to regulate social media, imposing escalating fines and other penalties on companies that fail to report certain types of information to the government. Some of these recommendations for policies like this have been made in the United Kingdom and in the European Union.

To me, this is kind of distressing, in part because I worry about what that would do to private property, what that would do to these thriving businesses that have given so many people so much of an opportunity to be heard. I also worry about what it would do to public safety, the very end sought to be achieved by these proposals. Sometimes when government gets involved and it sets a certain standard in place, that becomes both the floor and the ceiling. Understandably, I would worry about that.

So I'd like to—we'll start with you, Ms. Bickert. Tell me what you think about proposals like that and what some of the risks might be to starting to treat social media companies like public utilities?

Ms. BICKERT. Thank you, Senator. I think whenever we think about regulation, there often are unforeseen consequences, and those can impede our ability to provide services to the people that trust and need our products. I think the big thing for us is that our incentives are often aligned with those of government in terms of creating a safe community.

On this issue, absolutely, there is no question that the companies here do not want terrorists using their platforms. The long-term business interest for Facebook is we need people to have a good experience when they come to Facebook. We need them to like this community and want to be a part of it, and that means keeping them safe and removing bad content. So the incentives are there. These companies are working together to address these challenges, and that's how we think it can work best.

That said, we will continue to have a productive dialog with government. The concerns that you face and what you're hearing from your constituents matter to us very much, and we want to make sure that we're considering that in responding to that.

Senator LEE. In light of the fact that your company and others have—that the progress that your company and others have made in this area does not suggest that some of these proposals are unnecessary, in any event?

Ms. BICKERT. Thank you, Senator. Because our incentives are aligned, the kind of progress that you're going to see is going to happen regardless of what we're seeing from governments, what we're hearing from governments. It's still important to have that dialog. We learn every time that we engage with policymakers. But the incentives exist independently.

Senator LEE. Ms. Downs, would you agree with that?

Ms. DOWNS. Yes. The security and integrity of our products is core to our business model, and that includes the expedient enforcement of all of our content policies. So we are already sufficiently motivated to invest the necessary resources and people in addressing this threat.

Senator LEE. And how might treating you more like a public utility change that dynamic?

Ms. DOWNS. I think the risks that you outlined are important things for policymakers to remain cognizant of. Obviously, the tech industry is incredibly innovative, has created tremendous economic opportunity, and anything that slows down that innovation will cause damage to the ability of the industry to continue to thrive.

Senator LEE. Mr. Monje.

Mr. MONJE. I'd agree with that. We take our responsibility extremely seriously, and it is important to our business to get it right. We measure progress in matters of weeks and months. We move very quickly. So I'd agree with everything that was said.

Senator LEE. I've got one second remaining, if I can just—how do you determine—we'll just go with you, Mr. Monje, since we're already on you. How do you determine what constitutes terrorist or extremist content? For example, do you make this determination internally within your staff? Do you have certain subject matter experts that help you decide that?

Mr. MONJE. Yes, sir. We have former law enforcement officials who are on our team. We also interact with and communicate with governments and NGOs to determine that on an individual basis.

Senator LEE. My time has expired. Thank you very much.

Thank you, Mr. Chairman.

The CHAIRMAN. Thank you, Senator Lee.

Senator Hassan.

STATEMENT OF HON. MAGGIE HASSAN,
U.S. SENATOR FROM NEW HAMPSHIRE

Senator HASSAN. Thank you very much, Mr. Chairman, and thank you to our panelists today.

I want to talk about the See Something, Say Something campaign. The campaign is simple, and many of the thwarted terrorist attacks in the U.S. were stopped because everyday people alerted authorities to suspicious behavior. So I'd like to get a better sense of whether your companies fully embrace this See Something, Say Something campaign. While I understand that most of you shut down accounts that espouse violent extremist propaganda, it's not clear that you proactively report those accounts to law enforcement.

Here's an example of why that makes us less safe. In 2012, Tamerlan Tsarnaev, one of the Boston Marathon bombers, posted on YouTube several videos espousing al Qaeda propaganda under the name Muaz. At the time, the FBI was unaware of this account. However, the FBI had previously investigated Tamerlan, thanks to a tip from the Russians, but found nothing to corroborate the Russians' claims.

In September 2012, Tamerlan applied for U.S. citizenship with DHS. As part of the vetting process, DHS instructed the FBI to run a check on the application, which came up all clear. However, in his application, Tamerlan revealed that he tried to change his legal name to Muaz, the same name as his YouTube account. Eight months later, Tamerlan orchestrated a terrorist attack with his brother that resulted in the death of four people and almost 300 injured.

In hindsight, if YouTube had reported Muaz's troubling social media account to the FBI, then maybe the FBI would have been able to link Tamerlan to Muaz's extremist YouTube account when

Tamerlan was applying for citizenship. That could have prompted the FBI to re-open a closed terrorism investigation just weeks before Tamerlan carried out this awful tragic bombing.

So to Mr. Monje and Ms. Bickert and Ms. Downs, I'd like to understand how and when your companies report extremist accounts to law enforcement, and has it changed since the days of the Boston Marathon bomber?

Mr. MONJE. Thank you, Senator, for that question. When we're aware of an imminent threat, we absolutely do proactively reach out to law enforcement. Whenever they come to us and ask for information, as long as they have the right process, which we are very good with working with them to figure out, we will respond as quickly as we can.

Senator HASSAN. Thank you.

Ms. Downs?

Ms. DOWNS. Thank you, Senator. We also cooperate with law enforcement pursuant to the legal process, including the emergency disclosure provisions, where if we detect any content on our services that poses a threat to life, we proactively report it to law enforcement.

Senator HASSAN. And Ms. Bickert?

Ms. BICKERT. Thank you, Senator. The same answer.

Senator HASSAN. Well, I thank you. I will say that the See Something, Say Something campaign is premised on something a little bit different than what you all just said, because it's premised on "if you think," not "does this meet my definition of eminent danger." But we ask members of the public if they see something suspicious to step up, and what you're all saying is that if it meets certain criteria or if you're asked, and I think that's a little bit different.

So let me follow up with Mr. Watts. As a former Federal law enforcement officer, how would you grade these companies' performance in addressing violent extremist accounts? Do you think they can do more to actively support Federal law enforcement and counterterrorism officials?

Mr. WATTS. Over the last decade or so, they've all done better. Facebook and Google have outpaced Twitter. Twitter, in my opinion, relies too much on technical signatures and doesn't staff on the threat intelligence level to the extent that they should.

Senator HASSAN. Thank you very much.

Thank you. That's all the questions I have, Mr. Chair.

The CHAIRMAN. Thank you, Senator Hassan.

Senator Peters.

STATEMENT OF HON. GARY PETERS,
U.S. SENATOR FROM MICHIGAN

Senator PETERS. Thank you, Mr. Chairman.

And to each of our witnesses, thank you for being here today. It's an important topic, and we appreciate your active involvement in this.

My question concerns the extent that algorithms are used and play a role in the problem and how algorithms can also be used as a solution to this problem that we're dealing with. I was pleased to read Ms. Bickert's post in the Facebook newsroom that Facebook

has started using artificial intelligence to help counter terrorist threat efforts on your platforms.

The speed and breadth of the Internet certainly makes it nearly impossible for humans to keep track of all this. We need to have AI systems to do that, and they need to continually evolve if we're going to be effective in using them. However, it is likely that algorithms may be partly responsible for getting extremist material in front of users, whether it be in search results through Facebook's news feed or YouTube's up next list or elsewhere.

So my question is these algorithms are under your direct control, as all platform providers can control that. What are you specifically doing to learn more about whether and how your algorithms may be promoting extremist content?

I'll start with you, Ms. Bickert.

Ms. BICKERT. Thank you, Senator. The first thing that we need to do is make sure we're removing the terror content, and then it doesn't matter—once you take it out of the equation, then the algorithm has no role in promoting it because the content is simply not available on Facebook. That's something that we do by, as you pointed out, using technology to find the content.

But we don't stop there. After we find an account that is associated with terrorism, if we remove that account, we also fan out from that account. We look at associated content, associated accounts, and we remove those as well. If we can get better in that space, then we can make sure that the content is not appearing before our community.

Senator PETERS. Ms. Downs.

Ms. DOWNS. Thank you, Senator. Absolutely correct that the first priority is making sure that none of this content is on the platform in the first place. At the same time, we also have teams that are protecting our algorithms from being gamed. Obviously, this is a threat to our services and to our users' experience on our services across many issues, and so we have dedicated teams to make sure that people aren't manipulating our systems and that they're working as intended to serve relevant information to users who come to YouTube.

Mr. MONJE. Very similar answer from Twitter. We've been able to use our machine learning, our algorithms, to help identify more than 90 percent of the terrorist content that we've taken down before anybody else brings it to our attention, 75 percent of those before they get to tweet once. And, also, we protect our trends against manipulation. We've done that since 2014, and we continually improve our processes to protect our users' experience.

Senator PETERS. Mr. Watts?

Mr. WATTS. I would just note that any sort of algorithm detection technique is only as good as what's already been seen out in the world, which is part of the reason why the Russians have been more successful in terms of social media manipulation. They understand the terms of service. They have the capabilities to actually beat those systems, and they play within the rules.

The smarter, better-resourced, higher computational people around the world that want to use it will do better. It's kind of like zero day viruses in cybersecurity speak. Cybersecurity protections, anti-virus, is only as good as what has already been seen before in

terms of malware, and so the only way to get in front of that is to combine really smart threat analysts on whatever threat actor it is that's out there with the technologists, and those companies that do that do better in terms of getting in front of these actions.

Senator PETERS. What's your assessment of the companies here and others in the United States?

Mr. WATTS. I think Facebook and Google—I've seen massive increases and much more success in that space. I think Twitter gets beat oftentimes and can continue to get beaten because they rely too heavily on technology, and I don't think they have the partnerships they need to adequately get out in front of it.

Mr. MONJE. If I could respond to that, because——

Senator PETERS. Absolutely.

Mr. MONJE.—because he said it twice, and I disagree. I think there are many external researchers who said that a lot of this terrorist content doesn't—has moved off of our platform. The average ISIS account in 2014 had 177 followers, and now they have 14. They measure their life on Twitter in minutes and hours. We are extremely effective at taking them out. We do have the resources in place and the technology in place to fight the fight.

Senator PETERS. Mr. Watts?

Mr. WATTS. They get beat by a new terrorist group every few years. I mean, al Shabaab—we watched the entire Westgate attack go down on Twitter, monitoring it. We had a key monitoring list that we watch on that. With ISIS and al Nusra in Syria, we were able to build that list of anywhere from 3,000 to 4,000 terrorist accounts at any given time.

They do better after the fact, once they pick up what the signatures are. But the problem is you're always trailing whatever the threat actor is. You're not staying out in front of it, which is why in the cybersecurity space—or even some social media companies are taking this on now—you employ the threat analysts so they can work with the technologists. Otherwise, the technologists are always behind the curve. They have to wait until the group creates enough signatures that they can detect it, and then they can weed it out.

They're getting better all the time, but AI and machine learning, even with its advancements, can only detect what's already been seen before. And what humans are very good at, at least up until now and until they become autonomous machines out there—they're good at gaming systems and figuring out ways around it.

So I think in the case of the Russians, for example—and I've seen the takedowns of their accounts by Twitter, and they are hardly making a dent in what I'm seeing in terms of flows. I can't confirm all those accounts that are out there, but, you know, I hear about troll farms. Why do we think there's only one?

So I think in terms of moving forward, there has got to be a much bigger focus for those social media companies on putting threat analysts and pairing them together—and I know both analysts that have gone to Facebook and Google in that space, and they haven't—you know, some have been there longer than others, but I think that's the right approach moving forward.

Senator PETERS. Thank you. I'm out of time. Thank you so much.

The CHAIRMAN. Thank you, Senator Peters.

Senator Cruz.

STATEMENT OF HON. TED CRUZ,
U.S. SENATOR FROM TEXAS

Senator CRUZ. Thank you, Mr. Chairman.

Welcome to each of the witnesses. I'd like to start by asking each of the company representatives a simple question, which is do you consider your companies to be neutral public fora?

Ms. Bickert.

Ms. BICKERT. Thank you, Senator. The mission of our company is to connect people. We do not look at ideology or politics. We want people to be able to connect and share who they are.

Senator CRUZ. I'm just looking for a yes or no, whether you consider yourself to be a neutral public forum.

Ms. BICKERT. We do not have any policies about political ideology that affect our platform.

Senator CRUZ. Ms. Downs.

Ms. DOWNS. Yes, our goal is to design products for everyone, subject to our policies, and on occasions they impose on the types of content that people may share on our product.

Senator CRUZ. So you're saying you do consider YouTube to be a neutral public forum.

Ms. DOWNS. Correct. We enforce our policies in a politically neutral way. Certain things are prohibited by our community guidelines, which are spelled out and provided publicly to all of our users.

Senator CRUZ. Mr. Monje.

Mr. MONJE. Yes, sir.

Senator CRUZ. Well, let me focus for a minute, Mr. Monje. As you know, there have been several videos that were released in recent weeks that I and a lot of other people thought were highly troubling, and so I want to give you an opportunity to respond to that.

One individual, Abhinav Vadrevu, described as a former Twitter software engineer, was captured on video saying the following, quote, "One strategy is to shadow ban so you have ultimate control. The idea of a shadow ban is that you ban someone but they don't know they've been banned, because they keep posting and no one sees their content. So they just think that no one is engaging with their content, when, in reality, no one is seeing it." Is that a practice that occurs at Twitter?

Mr. MONJE. No, sir. We do not shadow ban users.

Senator CRUZ. Why would this individual described as a former Twitter software engineer say that?

Mr. MONJE. Thank you for the opportunity to respond, Senator, about this. These folks were caught on video. They weren't speaking on behalf of the company. They were speaking in their personal capacity. We do not shadow ban folks. What we do do is if an account is spamming, meaning engaging in malicious automation, we will hide—make it harder for them to find—to be found on our platform.

If I could continue, sir, that was one of the reasons why the efforts that we saw with the Russian misinformation didn't hit as big a mark as they were hoping for. We were able to stop that in real time. The other thing, sir—I'm sorry sir.

Senator CRUZ. Another individual named Mo Norai, a Twitter content review agent, was quoted on a video as saying "On stuff like that, it was more discretion on your viewpoint, I guess how you felt about a particular matter. Yes, if they said this is, quote, 'pro-Trump,' I don't want it because it offends me, this, that, and I say I ban the whole thing, and it goes over here and they're like, 'Oh, you know what? I don't like it, too. You know what? Mo's right. Let's go. Let's carry on. What's next?'"

Is that individual describing a practice that occurs at Twitter?

Mr. MONJE. No, sir. We use algorithms as a way to—if we see an account that is being abusive, that also will be down ranked. If they're engaging in targeted abuse against minorities, if they're being—if they're consistently violating our terms of service but they haven't crossed the line into being suspended, we'll make it less visible. But what we won't do is make—your followers will always be able to see you, and we are not—we ensure that—if you go on Twitter at any moment, you can see—you can see arguments on all sides of the issue.

Senator CRUZ. Wait. I want to make sure I'm understanding you right. You're saying for some people who are posting, you will restrict viewership only to those who are actively following them?

Mr. MONJE. If we believe that they're engaged in malicious automation, if we believe that they're violating our terms of service when it comes to abuse.

Senator CRUZ. So is it your position that the individuals that are subject to this form of censorship are extremists or fringe? Is that what you're telling us?

Mr. MONJE. It depends on the user. I can tell you that this is not something that we hide from the public. This is out in the open, the fact that we will reduce the visibility of tweets that are abusive or that are engaged in malicious automation.

Senator CRUZ. Well, let me ask, what about Congresswoman Marsha Blackburn? Is she someone you would consider somehow abusive or fringe or otherwise?

Mr. MONJE. No, sir.

Senator CRUZ. Well, then, why did Twitter restrict and censor her announcement video announcing as a candidate for the U.S. Senate?

Mr. MONJE. I want to be very clear about that, sir, and thank you for the question. We never removed her tweet, and what she did do is advertise on our platform. We do, like many platforms, have a higher standard when it comes to advertising, because we are putting in front of people things they didn't ask to see. Her video was reported to us. There was a decision that was made that was later reversed because of some of the language that was used in her account. It was a mistake, and we acknowledged it.

Senator CRUZ. So her announcement was censored because it was pro-life content. Has Twitter ever censored anyone for pro-choice content?

Mr. MONJE. She was never censored.

Senator CRUZ. So you're saying nothing happened to her tweet?

Mr. MONJE. Her tweet got a lot of attention on the organic side. We action our accounts, and we take our terms of service very seri-

ously. Sometimes we make the wrong decision. We have action on all sides of issues, and we strive to be better every day.

Senator CRUZ. Let me ask a final question, because my time has expired.

Ms. Downs, I'd like to know—what is YouTube's policy with respect to Prager University and the allegations that the content Prager University is putting out are being restricted and censored by YouTube?

Ms. DOWNS. As I mentioned, we enforce our policies in a politically neutral way. In terms of the specifics of Prager University, it's a subject of ongoing litigation, so I'm not free to comment on the specifics of that case.

Senator CRUZ. Well, I will say the pattern of political censorship that we are seeing across the technology companies is highly concerning, and the opening question I asked you, whether you're a neutral public forum—if you are a neutral public forum, that does not allow for political editorializing and censorship, and if you're not a neutral public forum, the entire predicate for liability immunity under the CDA is claiming to be a neutral public forum. So you can't have it both ways.

Thank you.

The CHAIRMAN. Thank you, Senator Cruz.

I think we've exhausted all the questions. Thank you all for being here. It has been a very informative session. We all know that the Internet is an incredibly powerful tool which offers enormous benefits to people globally. But we also realize we live in a dangerous world, and that there are people out there who want to do harm and do bad things and are looking for any means in order to accomplish those. Of course, we know that in the modern world, cyber has become increasingly the tool of choice for a lot of bad actors.

So we appreciate your informing us about steps that you're taking to try and police some of that bad behavior. As I said earlier, you know, we have constitutional protections and a Bill of Rights, and we also have—I think we want to make sure that we have a light touch when it comes to regulating the internet, and that's certainly something that I hope that this Committee will continue to support, and that those at regulatory agencies will adopt as well. But we also want to make sure that we are doing what we can to keep our country safe.

So we appreciate the efforts that you have undertaken already and those that you—as you continue to develop and look at ways to combat some of these threats that we face, and we hope that, working together as partners, that we can do a better job, and there's always room for improvement. So thank you for what you've done and for what you continue to do, and we'll look forward to discussing, I'm sure, in the future, as the threats continue to evolve, things that we can do better.

Thank you all for being here.

I'm going to just say that before we close, I've got a letter from the Consumer Extremism Project highlighting its work on combating radicalization online, and I'm going to enter that into the record, and also enter a piece by the *Wall Street Journal* authored by the Counter Extremism Project's Senior Advisor, Dr.—let me see

if I can say this right here—Hany Farid, underscoring his work on this important issue.

[The information referred to can be found in the Appendix.]

The CHAIRMAN. We'll keep the hearing record open for a couple of weeks. Senators are encouraged to submit any questions that they have for the record, and upon receipt of those questions, we ask our witnesses to submit their written responses to the Committee as quickly as possible.

Thank you all for being here. This hearing is adjourned.

[Whereupon, at 12:25 p.m., the hearing was adjourned.]

A P P E N D I X

PREPARED STATEMENT OF THE COUNTER EXTREMISM PROJECT

Thank you Chairman Thune, Ranking Member Nelson and distinguished members of the Committee for holding this hearing to examine the commitment and performance of Google/YouTube, Facebook and Twitter in permanently removing persistent and dangerous extremist and terrorist content on their platforms.

The Counter Extremism Project (CEP) is a not-for-profit, non-partisan, international policy organization formed in 2014 to combat the growing threat from extremist ideologies. Since its inception, CEP has pioneered efforts to combat extremists' radicalization and recruitment tactics online, and has persistently called upon Internet and social media companies, to take definitive action and adopt policies to stop the misuse of their platforms that has cost many lives around the world.

The Internet promised to democratize access to knowledge, spread great ideas, and promote tolerance and understanding around the globe. This promise, however, is being poisoned by the rise of trolling, cyber-bullying, revenge porn, fake news, child exploitation, hate, intolerance, and extremist and terrorist propaganda.

The horrific aftermath of extremists' weaponization of the Internet and social media platforms stretches from Paris, to Brussels, to London, Orlando, San Bernardino, Istanbul, Beirut, Cairo, and New York Only a few years ago, big technology companies flatly denied the existence of this problem. And while their tone has undoubtedly changed, CEP remains concerned about the level of commitment, consistency and transparency that will be required to overcome the systematic misuse by these platforms. While big social media platforms acknowledge the existence of radicalizing content that violates their stated terms of service, their response to date has followed a familiar pattern utilized in response to other discoveries of abuse: denial, followed by half-measures and attempts to spin the issue in the media, and finally, reluctant action when faced with threats to their bottom-line or possible regulatory action.

Make no mistake. There is no question that reigning in online abuses is challenging. There is also no question, however, that we can and must do more than we are to mitigate the harm that is being seeded and fueled online, while maintaining an open and free Internet where ideas can be shared and debated.

To cite but a few examples of ongoing problems of moderation. After it was determined that Manchester suicide bomber Salman Abedi, who killed 22 people on May 22, 2017 relied in part on ISIS bomb-making instructional videos on YouTube to build his explosive device, Google declared that bomb-making videos had no place on the platform. However, that same video was still on YouTube almost two months after the suicide bomb attack CEP has determined that the bomb making video has been uploaded (and removed) from Google platforms at least 11times since, most recently on January 9, although the actual number is most likely much higher.

The ISIS video "The Religion of Kufr is One," which shows multiple executions by firearms and a hanging-clear violations of YouTube's terms of service-has been uploaded and removed from YouTube at least six times since May 30, 2016.

Google/YouTube's process for detecting and removing terror content is still allowing prohibited content to be repeatedly uploaded and stay live for a sufficient period of time for hundreds people to view and download. Experience has shown that if most of the sharing of a video can take place in the first few hours it is available, meaning if the content is not removed quickly and uploads prevented, the moderating process has failed.

The Committee will no doubt hear today from Google, Facebook and Twitter about improvement they have made and pledges of more action. CEP notes that in November, Google/YouTube removed the lectures and sermons from al-Qaeda operative Anwar ai Awlaki from its platform. That decision, however, followed a multi-year CEP campaign that included direct outreach to the leadership of Google, a sustained effort to highlight issues in the media, including via op-eds in *USA Today* and Fox News, and a series of original reports detailing Awlaki's ubiquitous presence on YouTube and other Internet platforms, as well as his radicalizing influence on U.S.

and European terrorists. For example, CEP researchers identified 90 extremists in the U.S. and Europe with ties to al Awlaki, including Said and Cherif Kouachi, who carried out the Charlie Hebdo attacks; Omar Mateen, who killed 49 people in Orlando; Ohio State car attacker Abdul Razak Ali Artan; Boston Marathon bombers Dzhokhar and Tamerlan Tsarnaev; and many others.

One action does not constitute a lasting solution. Industrywide standards are needed to ensure the timely and permanent removal of dangerous content, especially when produced by groups and individuals on the State Department's Foreign Terrorist Organizations list, the Treasury Department's Specially Designated Nationals and Blocked Persons list, and the United Nations Security Council Sanctions list, and individuals with demonstrable links to violence. There is no shortage of extremists online—Turki ai-Binali, Abdullah Faisal, Yusuf al-Qaradawi and Ahmad Musa Jibril are notable examples. They must be subject to the same treatment and their content should be swiftly and permanently removed.

Existing technology can also assist with the enforcement of new polices and prevent the re-upload of material from known extremists such as Mr. Awlaki. Dr. Hany Farid, a professor of computer science at Dartmouth College who advises our organization, developed an algorithm called eGLYPH that quickly and accurately identifies for removal known extremist material on the Internet and social media platforms. This technology is based on software developed by Dr. Farid and Microsoft almost a decade ago called PhotoDNA. In 2016 alone PhotoDNA was responsible for the take-down of over 10 million child pornography images around the world, based on known images as determined by the National Center for Missing and Exploited Children (NCMEC).

It has already been proven that technology exists to effectively, aggressively, and consistently filter content that is either illegal or an explicit violation of a company's terms of service. While there is no question that reasonable people can disagree about the extremist-nature of some content, we can all agree that videos of a murder, videos of beheadings, videos with explicit calls to violence, or videos on how to build a suicide vest are extremist in nature—the worst-of-the-worst—and violate terms of service of all major tech companies. It is important to understand that technologies like eGLYPH, which was offered to all three of the companies before you today for free, simply allow companies to effectively, consistently, and transparently enforce their own terms of service, their standards for what is and what is not allowed on their networks.

There is no technological, economic, or legal reason why we cannot purge major online platforms of the worst-of-the-worst extremist content that grows more pernicious each year. There is no reason why we can't significantly disrupt global online radicalization and recruitment by hate and extremist groups. And, there is no doubt that this can be accomplished in a thoughtful, effective, and transparent manner, while respecting the privacy and rights of every user.

Lawmakers and the public should demand that tech finally implement industrywide standards and policies that ensure the timely and permanent removal of dangerous extremist and terrorist material, establish measurable best practices and *transparently* deploy proven technologies to prevent there-upload of materials already determined to violate company policies. If tech fails to act, then it is time for regulators to promulgate measures to force the industry to take necessary action to protect the public.

How Algorithms Can Help Beat Islamic State

Hany Farid 'changed the world' by combating child porn. Now his software could suppress terrorists online.

By Joseph Rago, Hanover, N.H.

You can't blame the message on the medium, not exactly. But maybe, all things considered, arming everyone with pocket supercomputers, and then filtering most of human experience through social-media feedback loops, wasn't the greatest idea.

America recently endured the most electronic and media-saturated presidential campaign in memory, with its hacks, private servers, secret videotapes, fake news, troll armies and hour-by-hour Internet outrage across all platforms. And however glorious modern communications may be, they've also empowered a cast of goons, crooks and jihadists to build audiences and influence worldwide.

A technological solution, at least to that last problem, may lie 2,600 miles east of Silicon Valley, in a computer-science laboratory at Dartmouth College. Prof. Hany Farid, chairman of the department, creates algorithms that can sweep digital net-

works and automatically purge extremist content—if only the tech companies will adopt them.

"If you look at recent attacks, from Orlando to San Bernardino to Nice to Paris to Brussels," Mr. Farid says, "all of those attackers had been radicalized online. They weren't going to Syria. They watched YouTube videos."

He continues: "The dark side of the open Internet is that truly fringe and harmful ideas now are mainstream, or at least accessible to 7½ billion people." Yet "whenever we have one of these attacks, we just wring our hands for a few weeks and then wait for the next one to happen."

Social networks have created "a new environment for radicalization and recruitment," says David Ibsen, executive director of the Counter Extremism Project, a nonprofit research and advocacy organization to which Mr. Farid is a senior adviser. Terror groups weaponized Twitter, Google, Facebook and other forums to plan or encourage violence; to discover the vulnerable or disaffected; and to publish professional, sophisticated and carefully presented propaganda.

Islamic State is basically a digital-first media startup. (By comparison, al Qaeda was MySpace.) ISIS content is beamed out globally and becomes refractory across the viral web. Some videos show vignettes of ISIS bureaucrats delivering social services or its fighters talking about the battle between belief and unbelief. Others are more savage—beheadings, stonings, drownings, other torture and combat operations.

Mr. Farid slipped into this world slant-wise. He's a founder of the computer-science field known as digital forensics. In the late 1990s as a postdoctoral researcher, he was among the first to recognize that mathematical and computational techniques to authenticate digital images and other media would be useful to society.

Because images so powerfully change what we are willing to believe, the modern era requires a scientific method to ensure we can trust them. How can we prove, for example, that digital photographs aren't forgeries so they are admissible as evidence in court? Images are increasingly important in cellular, molecular and neurological medicine, Mr. Farid notes, and tampering has led to more than one research-and-retraction scandal. Unscrupulous stringers sometimes file doctored photos with news organizations, and unscrupulous motorists sometimes photoshop pictures to exaggerate fender-benders for insurance claims.

Mr. Farid explains how image authentication works: "We think about how light interacts in the physical world; what happens when that light hits the front of the lens and gets focused and goes through an optical train; what happens when that light hits an electronic sensor and gets converted from an analog to a digital and then goes through a postprocessing and gets saved as .jpeg and then gets posted on Facebook." By identifying "statistical and geometrical and physical regularities" in this life cycle, software can search for inconsistencies to expose manipulation.

In 2008 this research pulled Mr. Farid into another underworld—child pornography. In 2002 the U.S. Supreme Court struck down a ban on "virtual" child porn—computer-generated images that "appear to depict minors but were produced without using any real children." Mr. Farid is sometimes brought in as an outside expert when a defendant claims the material at issue is virtual.

The child-porn industry was nearly defunct by the 1990s, because negatives and videotapes can be confiscated and destroyed. "Then the Internet came," Mr. Farid says, "and all hell broke loose."

Supply can create its own demand. Much like jihadists, deviants formed a global community, finding each other online and sharing what are really crime-scene photos. Like ISIS agitprop, material is continuously copied, cut, spliced, resized, recompressed and otherwise changed, in part to evade detection as it is retransmitted again and again.

Mr. Farid worked with Microsoft to solve both problems—detection and replication. He coded a tool called Photo DNA that uses "robust hashing" to sweep for child porn. "The hashing part is that you reach into a digital image and extract a unique signature. The robust part is if that image undergoes simple changes, the fingerprint shouldn't change. When you change your clothes, cut your hair, as you age, your DNA stays constant," he says. "That's what you want from this distinct fingerprint."

The algorithm matches against a registry of known illegal signatures, or hashes, to find and delete photographs, audio and video. Photo DNA is engineered to work at "internet scale," says Mr. Farid, meaning it can process billions of uploads a day in microseconds with a low false-positive rate and little human intervention.

Monitoring by Photo DNA, which is licensed by Microsoft at no cost and now used in most networks, revealed that the nature of the problem was "not what we thought it was," says Ernie Allen, the retired head of the National Center for Miss-

ing and Exploited Children. Child pornography was far more widely circulated than law enforcement believed. "Hany Farid changed the world," Mr. Allen adds. "His innovation rescued or touched the lives of thousand of kids, and uncovered perpetrators, and prevented terrible revictimization as content was constantly redistributed."

Mr. Farid linked up with the Counter Extremism Project to apply the same robust-hashing method to extremist propaganda. But this effort has encountered resistance. "The pushback from the tech companies has been pretty strong," the project's Mr. Ibsen says dryly.

U.S. law immunizes Internet companies from criminal and civil liability for content that travels over their transoms. Their terms of service forbid abusive content, but they rely on users instead of algorithms to police violations. "It's a very slow and tedious process: You wait for it to get reported, somebody has to review it, they make mistakes," Mr. Farid says. "They take down the Vietnam napalm girl on Facebook."

Liability aside, what about their moral obligations to help prevent death, injury and destruction? "In my mind, we're not asking them even to do something that they haven't said they want to do already. We're saying, hey, would you please do the thing that you promised you would do?" he explains. "I am simply saying, look, for free, you can automate this and make it really efficient and really fast and save you money on the side."

But the "ethos" of Silicon Valley doesn't include becoming the censors of the internet, and tech firms fear a slippery slope. "The concern they have is, OK, first they came for the child porn, then they came for the extremism, next they're going to take the kitten videos," Mr. Farid says. "I think that's a bit of a hysterical leap. We are talking about content with very clear and well-defined harm. These are not abstract notions—'I don't want people to be mean to me.' We're not talking about bullying. We are talking about things with very immediate consequences and very real harm."

One question is how to distinguish support for terrorism from the merely inappropriate or objectionable. What about Islamic State's black-flag brand, or a declaration of a caliphate, or the sermons of Anwar al-Awlaki? Maybe you know it when you see it.

"Is an ISIS fighter saying 'Death to the West' extremism? I don't know. I don't want to have that conversation," Mr. Farid replies. "I'm talking about explicit acts of violence, explicit calls to violence, explicit glorification of violence, depravity, the worst of the worst of the worst."

His point is that tech companies can make judgment calls about the middle ground, wherever it might be, for themselves: "You decide: Yes, no, yes, no, yes, no, and then we'll build a cache and eliminate that content from your networks."

Mr. Farid concedes that there are dangers: "This type of technology is agnostic in what it's looking for. It can be used in ways we would not approve of, such as stifling speech. You can't deny that. This is what we've learned about technology over the years—it can be used for good and for bad. Social media platforms can be good and bad."

There has been some progress. Twitter has deleted hundreds of thousands of handles associated with terrorism since 2015, and late last year Twitter, Facebook, Microsoft and YouTube announced an industry antiterror consortium. But Mr. Farid's robust hashing remains a hard sell.

The irony is that algorithms increasingly govern the world. Networks are perpetually scanned for spam, malware, viruses; Google reads your e-mail to target ads; credit-card companies monitor your financial transactions to prevent fraud. Facebook's Mark Zuckerberg even promises to use algorithms to distinguish truth from falsehood. As a scholar of the differences between the two, Mr. Farid has a few thoughts.

In the backwash of 2016, Mr. Zuckerberg published a 5,800-word manifesto that promised Facebook's artificial intelligence would soon learn to sort real news from hoaxes and misinformation, break up "filter bubbles," and draw a line between free speech and suborning terror. The goal, he wrote, is to preserve "our shared sense of reality."

Mr. Farid is a skeptic: "As somebody who worked for a long time in this space, I think he's underestimating what a hard problem this is." Mr. Zuckerberg "paints this picture like machine learning is going to be fully automatic—basically you'll be able to set criteria on your page, "I don't want to see violence, I don't want to see bad words," and it'll just work.

"Even as a technologist, and despite all the advance of technology, the human brain is astonishing at what it does. Our ability to make these types of assessments that are really hard for these AI algorithms is humbling. I don't think we'll get it in the next five or 10 years."

Meantime, Mr. Farid has developed a technology that could work today to contain a growing threat. While we await the Facebook utopia, perhaps our digital lives—and our real lives—would be healthier if it were widely deployed.

Mr. Rago is a member of *The Wall Street Journal's* editorial board.

Appeared in the March 11, 2017, print edition.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JERRY MORAN TO MONIKA BICKERT

Question 1. Your written testimony emphasized the importance of the credibility of the speaker as it relates to Facebook's efforts to prevent recruitment through "counterspeech." How have your strategic partnerships with non-governmental organization and community groups bolstered Facebook's "counterspeech" efforts?

Answer. We believe that a key part of combating extremism is preventing recruitment by disrupting the underlying ideologies that drive people to commit acts of violence. That's why we support a variety of counterspeech efforts. Although counterspeech comes in many forms, at its core it includes efforts to prevent people from pursuing a hate-filled, violent life or convincing them to abandon such a life.

Our efforts are focused on empowering counterspeech creators and amplifying local voices by building awareness, educating communities, encouraging cohesion, and directly countering hateful narratives. We have partnered with non-governmental organizations and community groups around the world to empower positive and moderate voices. For example, in the U.S., we have worked with EdVenture Partners to develop a peer-to-peer student competition called the Facebook Global Digital Challenge (P2P). This is a semester-long university course during which students build a campaign to combat extremism in their area, launch it, track its success, and then submit the results as part of a global competition. As part of P2P, a team of communications students from the University of Central Oklahoma ran an amazing program called uDefy that reached over one million people in 85 countries using Facebook and other social media platforms. The team behind uDefy encouraged participants to recognize and challenge their own beliefs and stereotypes by taking a four-step pledge: (1) face your truth; (2) get the facts; (3) commit to defy; and (4) spread the word. The goal of the campaign is to channel fear and misconception into truth and understanding one individual at a time. Those who complete the four-step pledge become uDefy ambassadors and take the campaign back to their own campuses. In less than three years, these P2P projects have reached more than 56 million people worldwide through more than 500 anti-hate and extremism campaigns created by more than 5,500 university students in 68 countries.

We have also partnered with the Institute for Strategic Dialogue to launch the Online Civil Courage Initiative, a project that has engaged with more than 100 anti-hate and anti-extremism organizations across Europe. Similarly, we work with Affinis Labs to host hackathons in places like Manila, Dhaka, and Jakarta, where community leaders joined forces with tech entrepreneurs to develop innovative solutions to challenge extremism and hate online.

By fanning out and removing content, and supporting counterspeech efforts, we can limit the audience and distribution of terrorist propaganda.

Question 2. Your written testimony stated: "In the first half of 2017, [Facebook] provided information in response to more than 75 percent of the 1,864 requests for emergency disclosures that [the company] received from U.S. law enforcement agencies." Do you have a company policy when deciding how to respond to the 1,864 requests for emergency disclosures?

(a) Does Facebook do their own assessment as to whether the content constitutes an emergency?

Answer. As part of official investigations, government officials sometimes request data about people who use Facebook. We disclose account records in accordance with our terms of service and applicable law, and we may voluntarily disclose information to law enforcement where we have a good faith reason to believe that the matter involves imminent risk of serious physical injury or death. We have strict processes in place to handle these government requests. We require officials to provide a detailed description of the legal and factual basis for their request, and we push back if the request appears to be legally deficient or is overly broad, vague, or otherwise inconsistent with our policies. More information about the requests we have received from governments around the world can be found at *https://transparency.facebook.com/*.

(b) Do your internal policies account for the 25 percent of requests that are not responded to with information?

Answer. Please see the response to question 2a.

(c) Do you have the resources to deal with these requests?

Answer. Our Law Enforcement Response Team works hard to respond to legitimate law enforcement requests while fulfilling our responsibility to protect people's privacy and security. We have a global team that strives to respond within minutes to emergency requests from law enforcement. Our effort to make our platform safer and more secure is a holistic one that involves a continual evaluation of our personnel, processes and policies, and we make changes as appropriate.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. RON JOHNSON TO MONIKA BICKERT

Question 1. Social media companies are increasingly able to remove terrorist recruitment, incitement, and training materials before it posts to their platforms by relying on improved automated systems. Other than content removal, what else can be done to limit the audience or distribution of these dangerous materials?

Answer. When we find an account that is associated with terrorism, we use artificial intelligence to identify and remove related material that may also support terrorism or terrorists. As part of that process, we utilize a variety of signals, including whether an account is "friends" with a high number of accounts that have been disabled for terrorism, or whether an account shares the same attributes as a disabled account.

Moreover, we believe that a key part of combating extremism is preventing recruitment by disrupting the underlying ideologies that drive people to commit acts of violence. That's why we support a variety of counterspeech efforts. Although counterspeech comes in many forms, at its core these are efforts to prevent people from pursuing a hate-filled, violent life or convincing them to abandon such a life. We have partnered with non-governmental organizations and community groups around the world to empower positive and moderate voices. For example, in the U.S., we have worked with EdVenture Partners to develop a peer-to-peer student competition called the Facebook Global Digital Challenge (P2P). This is a semester-long university course during which students build a campaign to combat extremism in their area, launch it, track its success, and then submit the results as part of a global competition. As part of P2P, a team of communications students from the University of Central Oklahoma ran an amazing program called uDefy that reached over one million people in 85 countries using Facebook and other social media platforms. The team behind uDefy encouraged participants to recognize and challenge their own beliefs and stereotypes by taking a four-step pledge: (1) face your truth; (2) get the facts; (3) commit to defy; and (4) spread the word. The goal of the campaign is to channel fear and misconception into truth and understanding one individual at a time. Those who complete the four-step pledge become uDefy ambassadors and take the campaign back to their own campuses. In less than three years, these P2P projects have reached more than 56 million people worldwide through more than 500 anti-hate and extremism campaigns created by more than 5,500 university students in 68 countries.

We have also partnered with the Institute for Strategic Dialogue to launch the Online Civil Courage Initiative, a project that has engaged with more than 100 anti-hate and anti-extremism organizations across Europe. Similarly, we work with Affinis Labs to host hackathons in places like Manila, Dhaka, and Jakarta, where community leaders joined forces with tech entrepreneurs to develop innovative solutions to challenge extremism and hate online.

By fanning out and removing content, and supporting counterspeech efforts, we can limit the audience and distribution of terrorist propaganda.

Question 2. Terrorist how-to guides are protected by the First Amendment in the United States, but violate the content policies of many social media companies as well as the laws of some international partner nations. What countries have laws that go beyond your company's content policies and can you give examples of how you have worked with those countries to de-conflict those differences?

Answer. A number of countries around the world have laws that limit content that might otherwise be allowed by our Community Standards or U.S. law. In Germany, for example, laws forbid incitement to hatred. In the U.S., on the other hand, even the most vile speech may be legally protected under the U.S. Constitution. There are times when we may have to remove or restrict access to content because it violates a law in a particular country, even though it does not violate our Community Standards. Further, when governments believe that something on the Internet

violates their laws, they may contact companies like Facebook and ask us to restrict access to that content. When we receive such a request, it is scrutinized to determine if the specified content does indeed violate local laws. If we determine that it does, then we make it unavailable in the relevant country or territory. For example, Holocaust denial is illegal in Germany, so if it is reported to us, we will restrict this content for people in Germany.

Question 3. The long-term business interests of social media platforms are aligned with the public safety concerns of this committee: users want to feel safe while engaging with the online community. To this end, Facebook is developing a way to identify users at higher risk of suicide and urgently pass posts from any user in danger to a community operations team, as well as provide that user with a menu of options to reach out to their own friends or other suicide prevention partners. Is Facebook developing any similar tool to identify users at higher risk of terrorist activity? If so, what off-ramp options would Facebook consider offering those users?

Answer. We are using similar automated tools to identify users who are posting content that violates our policies against terrorism, including promoting terror groups, sharing their propaganda, and planning or coordinating violence. We reach out to law enforcement whenever we see a credible threat of imminent harm.

We are eager to partner with government and civil society to develop off-ramp options for users at a higher risk of terrorist activity. A critical part of providing an off-ramp is being able to link people to appropriate, effective, and responsible services. We are exploring ways of partnering with such services.

Response to Written Questions Submitted by Hon. Maria Cantwell to Monika Bickert

Question 1. Facebook prohibits individuals and organizations that promote hate from having a presence on its site. However, I think we can do much more to promptly communicate threats of hate-based violence to the relevant law enforcement agencies and internally police hate-promoting individuals and organizations. Will you commit Facebook to exploring and implementing a more aggressive effort to report hateful images and threats to law enforcement? If not, can you explain why you would not commit to this important request?

Answer. Facebook is opposed to hate speech in all its forms, and we are committed to removing it from our platform any time we become aware of it. We carefully review reports that we receive from the public, media, civil society, and governments, and we remove content that violates our policies. We are committed to improving our approach to addressing these issues, and regularly evaluate our hate speech policies to determine whether they need to be updated. We are also working to enhance our review process so that we are able to respond quickly and accurately to community reporting. We also remove credible threats of physical harm to individuals and specific threats of theft, vandalism, or other financial harm. We have a long history of working successfully with law enforcement to address a wide variety of threats to our platform, and we work with law enforcement when we believe there is a genuine risk of physical harm or direct threats to public safety. Our effort to make our platform safer and more secure is a holistic one that involves a continual evaluation of our personnel, processes, and policies, and we make changes as appropriate.

Question 2. We have strong principles of freedom of speech, but at the same time, we need to balance that freedom with the need to protect against bad actors who would leverage that freedom to plan and promote illegal acts. How can we use artificial intelligence to help us achieve a balance between our American ideal of free speech and the need to protect against extremist acts of terror?

Answer. We already use artificial intelligence (AI) to help us identify threats of real world harm from terrorists and others. We reach out to law enforcement whenever we see a credible threat of imminent harm. The use of AI and other automation to stop the spread of terrorist content is showing promise. Today, 99 percent of the ISIS and Al Qaeda-related terror content we remove from Facebook is content we detect before anyone in our community has flagged it to us, and in some cases, before it goes live on the site. We do this primarily through the use of automated systems like photo and video matching and text-based machine learning. Once we are aware of a piece of terror content, we remove 83 percent of subsequently uploaded copies within one hour of upload.

We believe technology can be part of the fight against terrorism. But deploying AI for counterterrorism is not as simple as flipping a switch. For example, a photo of an armed man waving an ISIS flag might be propaganda or recruiting material, but could be an image in a news story. Ultimately, the use of AI must be reinforced

with manual review from trained experts. To that end, we tap expertise from inside the company and from outside, partnering with those who can help address extremism across the internet.

Question 3. Outside of artificial intelligence, what other technologies could be used to combat potential radicalization on social media platforms? What does the implementation of those technologies look like?

Answer. We are constantly updating our technical solutions, but our current efforts include image matching technology and language understanding. When someone tries to upload a terrorist photo or video, our systems look for whether the image matches a known terrorist photo or video. This means that if we previously removed an ISIS propaganda video, for example, we can work to prevent other accounts from uploading the same video to our site. We also have started experimenting with using AI to understand text that potentially advocates for terrorism. We are working to develop text-based signals to detect praise or support of terrorist organizations. These signals will be incorporated into an algorithm that is in the early stages of learning how to detect similar posts.

We understand that simply working to keep terrorism off Facebook is an inadequate solution to the problem of online extremism, particularly because terrorists are able to leverage a variety of platforms. We believe our partnerships with others—including other companies, civil society, researchers, and governments—are crucial to combating this threat. To this end, we have partnered with our industry counterparts to more quickly identify and slow the spread of terrorist content online. For example, in December 2016, we joined with Microsoft, Twitter, and YouTube to announce the development of a shared industry database of "hashes"—unique digital fingerprints for photos and videos—for content produced by or in support of terrorist organizations. The database now contains more than 60,000 hashes, and the consortium has grown to include thirteen companies.

We believe that computer algorithms and machine learning are necessary but not sufficient to address these problems. That's why we are also using specialized human review, industry cooperation, and counter-speech training. We will also be doubling the number of people who work on safety and security at Facebook by the end of this year—from 10,000 to 20,000 people.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. RICHARD BLUMENTHAL TO MONIKA BICKERT

Question 1. On October 30, 2017, nineteen civil rights groups, including Muslim Advocates, Leadership Conference on Civil and Human Rights, NAACP, Southern Poverty Law Center, and many others, co-signed a letter to Facebook to express concern about the hateful content on the social media platform used to divide the country, and in particular, to promote anti-Muslim, anti-Black, anti-immigrant, and anti-LGBTQ animus.

Ms. Bickert, please provide a copy of Facebook's response to this letter.

Answer. Hate speech and discriminatory advertising have no place on our platform, and we remove such content as soon as we become aware of it. We also have partnerships with academics and experts who study organized hate groups and who share information with Facebook on how organizations are adapting to social media and give feedback on how Facebook might better tackle these problems. We have reached out to the groups referenced in the question and are in a dialogue with them, which has included in-person conversations. We are committed to continuing our dialogue with them and other third parties to ensure that our users feel welcome and safe on our platform.

Question 2. Facebook reports that 99 percent of the ISIS and Al Qaeda-related terror content is detected and removed before it is even flagged on Facebook. However, a recent investigation by ProPublica asked Facebook about its handling of forty-nine posts that they thought might be deemed offensive Facebook's own Community Standards. Facebook acknowledged and apologized that its content reviewers made the wrong call in almost half of them.

Ms. Bickert, what steps is Facebook taking to better counter the use of the platform to spread hateful information and improve the work of its content reviewers?

Answer. We don't allow hate speech, which we define as anything that directly attacks people based on race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, gender identity, or serious disability or disease. However, our policies allow content that may be controversial and at times even distasteful, but which does not cross the line into hate speech. This may include criticism of public figures, religions, professions, and political ideologies. Our challenge is identi-

fying hate speech across different cultures, languages, and circumstances for a community of more than 2 billion people.

Nudity and violence, for example, are fairly easy to spot, but hate speech is often determined by its context. Because of these nuances, we cannot rely on machine-learning or AI to the same degree that we do with other types of content like nudity. Technology can help flag the most blatantly reprehensible language. But it cannot yet understand the context necessary to assess what is or is not hate speech—though we are working on tools to help us improve the accuracy of our enforcement and building new AI to better detect bad content.

We encourage people to report posts and rely on our team of content reviewers around the world to review reported content. Our reviewers are trained to look for violations and enforce our policies consistently and as objectively as possible. We have weekly quality audits of each reviewer, during which we re-review a subset of their work and address any mistakes made. We receive millions of reports of possible content violations every week, so we know that we will unfortunately make many mistakes even if we maintain an accuracy rate of 99 percent. We are always working to make our platform safer and more secure through, among other things, continually evaluating our processes, policies, and training. Enforcement is never perfect, but we will get better at finding and removing improper content.

Question 3. As recently as last year, Facebook reportedly allowed offensive claims that specify a sub-group within a protected class, such as "black children" or "female drivers" but would ban attacks aimed at entire groups, such as "white men." It seems that such policies could easily be gamed to work-around Facebook Community Standards.

Ms. Bickert, Is that still the policy of Facebook? Can you explain the nuances in the new policy and how Facebook is working to make sure such mistakes don't happen again?

Do you think it would be helpful for Facebook to be more transparent about how it applies its standards, or bring in an independent third-party and work with civil rights groups to help it evaluate its current policies?

Answer. No, this is not our policy. Facebook is opposed to hate speech in all its forms, and that includes removing content that targets any of the three groups identified in the question, depending on the context of the post.

We currently define hate speech as anything that directly attacks people based on protected characteristics—race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, gender identity, or serious disability or disease. Such content violates our Community Standards and will be removed. This includes, for example, content that attacks "black children" or "white men." However, there may be other content that is controversial or distasteful, but does not cross the line into hate speech.

We are constantly evaluating—and, where necessary, changing—our content policies to account for shifts in cultural and social norms around the world. For example, we recently updated our hate speech policies to remove violent speech directed at groups of people defined by protected characteristics, even if the basis for the attack may be ambiguous. Under the previous hate speech policy, a direct attack targeting women on the basis of gender, for example, would have been removed from Facebook, but the same content directed at women drivers would have remained on the platform. We have come to see that this distinction is a mistake, and we no longer differentiate between the two forms of attack when it comes to the most violent hate speech. For instance, we would now remove a comment that dehumanized "female drivers" by comparing them to animals. We continue to explore how we can adopt a more granular approach to hate speech, both in the way we draft our policies and the way we enforce on them.

Question 4. Ms. Bickert, Mr. Monje, and Ms. Downs, please provide copies (including images, text, dates and timestamps) of all content identified by your platforms as generated by Russian agents or the Internet Research Agency.

Answer. We have provided this information to the Senate Select Committee on Intelligence and the Senate Judiciary Committee and believe you should have access through those committees.

Question 5. Advocates for preventing gun violence have long been concerned about the "private sale" loophole, which allows individuals to purchase guns without a background check. So, Facebook's announcement in January 2016 that it would ban the private sale of guns and ammo on its site and Facebook was met with great applause. Unfortunately, it soon became evident that new rules have done little, if anything, to stop the flow of guns on the social network. If we are serious about fighting terrorism online, we should be just as serious about really closing this dan-

gerous loophole, which could very well enable violent ideology to be translated into horrific acts.

Ms. Bickert, do you agree that successfully closing this loophole is important?

How would you compare the amount of resources devoted to combatting terrorism online to the amount of resources devoted to ensuring Facebook's prohibition on the private sale of guns and ammo?

In what way is Facebook making sure it applies any relevant technologies, tools, and human resources used to review hate speech to also enforce Facebook's prohibitions on the private sale of guns and ammo?

Answer. We do not allow firearm sales on Facebook, and any time we become aware of content that is facilitating gun sales, we remove it. We allow our users to report such activity. We also look at associated groups and accounts by "fanning out" to identify and remove other content that may violate our policies. We will continue to look for ways to get faster at finding and removing violating content, and we encourage our community to continue to tell us if they see this behavior anywhere on our platform.

Question 6. At least one of your peers in the tech industry has voluntarily initiated an outside assessment of the civil rights impacts of its policies and programs. In response to concerns regarding discrimination on the home-sharing platform, AirBNB hired former U.S. attorney general Eric Holder to help craft an anti-discrimination policy and has promised to pursue technological innovations to guard against future discriminatory events.

Mr. Monje, Ms. Bickert, and Ms. Downs, can you each commit to bringing in an independent entity to conduct a thorough and public audit of the civil rights impact of your policies and programs, including how your platform has been used by hate groups to stoke religious resentment and violence?

Answer. Hate speech and discriminatory advertising have no place on our platform. Our Community Standards prohibit attacks based on protected characteristics, including religion, and we prohibit advertisers from discriminating against people based on religion and other attributes. Facebook has partnerships with academics and experts who study organized hate groups and hate speech. These academics and experts share information with Facebook on how organizations are adapting to social media and give feedback on how Facebook might better tackle these problems. We recently hosted several of these academics at Facebook for multiple days of observation and assessment, during which the academics attended substantive meetings on our content policies and the guidance we provide to our reviewers. Further, in the area of hate speech, there are very important academic projects that we follow closely. Timothy Garton Ash, for example, has created the Free Speech Debate to look at these issues on a cross-cultural basis. Susan Benesch established the Dangerous Speech Project, which investigates the connection between speech and violence. These projects show how much work is left to be done in defining the boundaries of speech online, which is why we will keep participating in this work to help inform our policies at Facebook. We are committed to continuing our dialogue with third parties to ensure that our users feel welcome and safe on our platform.

Question 7. A little over a year ago, Facebook, Twitter, Google, and Microsoft announced a plan to create a joint industry database of "content that promotes terrorism."

Mr. Monje, Ms. Bickert, and Ms. Downs, to what extent does this joint industry database focus on all forms of terror, including the real terror threat presented by white supremacists?

Answer. At last year's EU Internet Forum, Facebook, Microsoft, Twitter, and YouTube declared our joint determination to curb the spread of terrorist content online. Over the past year, we have formalized this partnership with the launch of the Global Internet Forum to Counter Terrorism (GIFCT). The GIFCT is committed to working on technological solutions to help thwart terrorists' use of our services, including through a shared industry hash database, where companies can create "digital fingerprints" for terrorist content and share it with participating companies. The database, which became operational in the spring of 2017, now contains more than 60,000 hashes. It allows the thirteen member companies to use those hashes to identify and remove matching content—videos and images—that violate our respective policies or, in some cases, block terrorist content before it is even posted. Each company has different policies, practices, and definitions as they relate to terrorist content. If content is removed from a company's platform for violating that platform's individual terrorism-related content policies, the company may choose to hash the content and include it in the database.

Facebook's policies do prohibit all forms of terror, including threats by white supremacist organizations.

Question 8. As reported by CNN last August after the events in Charlottesville, only 58 of over 200 Southern Poverty Law Center-designated hate groups with Facebook accounts had been suspended for their hateful actions and rhetoric.

Ms. Bickert, how many of those hate groups with Facebook accounts are now blocked? What further steps are you taking to further review the actions and accounts of these groups?

Answer. Facebook is opposed to hate speech in all its forms. Facebook has partnerships with academics and experts who study organized hate groups, including the Southern Poverty Law Center. These academics and experts share information with Facebook on how organizations are adapting to social media and give feedback on how Facebook might better tackle these problems. That said, we apply our own policies about what constitutes hate speech, and our definition of hate speech may differ from others, including those with whom we partner. We are constantly evaluating—and, where necessary, changing—our content policies, and we currently define hate speech as anything that directly attacks people based on protected characteristics—race, ethnicity, national origin, religious affiliation, sexual orientation, sex, gender, gender identity, or serious disability or disease. Such content violates our Community Standards and will be removed. However, there may be content that is controversial or distasteful, but does not cross the line into hate speech.

Further, our own content policy team includes subject matter experts who are focused on staying ahead of trends in hate speech. Their work is used to inform our Community Operations team, which reviews content that our users and automated tools flag as inappropriate, dangerous, abusive, or otherwise violating our policies—including our hate speech policy.

Managing a global community in this manner has never been done before, and we know we have a lot more work to do. We are committed to improving and to ensuring that hate has no place on Facebook.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. BRIAN SCHATZ TO MONIKA BICKERT

Question 1. Please quantify and explain Facebook's progress in tackling the fake-user account issue. For the most recent full month available and every month in the two years preceding provide:

- number of fake accounts created
- number of fake accounts removed
- number of accounts hacked
- number of hacked accounts restored
- number of duplicate accounts created
- number of duplicate accounts removed
- number of inactive accounts existing
- number of inactive accounts removed
- number of monthly active users
- average number of days a fake account remains on the platform

Please provide the numbers above for Instagram as well.

Answer. Facebook regularly provides information on the number of monthly active users (MAUs), false accounts, and duplicate accounts in its filings with the Securities and Exchange Commission. We define an MAU as a registered Facebook user who logged in and visited Facebook through our website or a mobile device, or used our Messenger application (and is also a registered Facebook user), in the last 30 days as of the date of measurement. MAUs are a measure of the size of our global active user community. As of December 31, 2017, we had 2.13 billion MAUs, an increase of fourteen percent from December 31, 2016.

We regularly evaluate these metrics to estimate the number of "duplicate" and "false" accounts among our MAUs. A duplicate account is one that a user maintains in addition to his or her principal account. We divide "false" accounts into two categories: (1) user-misclassified accounts, where users have created personal profiles for a business, organization, or non-human entity such as a pet (such entities are permitted on Facebook using a Page rather than a personal profile under our terms of service); and (2) undesirable accounts, which represent user profiles that we determine are intended to be used for purposes that violate our terms of service, such as spamming. The estimates of duplicate and false accounts are based on an internal review of a limited sample of accounts, and we apply significant judgment in making this determination. In the fourth quarter of 2017, we estimate that dupli-

cate accounts may have represented approximately ten percent of our worldwide MAUs. We believe the percentage of duplicate accounts is meaningfully higher in developing markets such as India, Indonesia, and the Philippines, as compared to more developed markets. In the fourth quarter of 2017, we estimate that false accounts may have represented approximately three to four percent of our worldwide MAUs. Our estimation of false accounts can vary as a result of episodic spikes in the creation of such accounts. Additional information relating to our estimate of false accounts is included in our filings with the SEC. We do not maintain public statistics on the other types of accounts that are referenced in the question.

We continue to make improvements to our efforts to more effectively detect and deactivate fake accounts to help reduce the spread of spam, false news, and misinformation. We continually update our technical systems to identify, checkpoint, and remove inauthentic accounts, and we block millions of attempts to register fake accounts every day. These systems examine thousands of detailed account attributes and prioritize signals that are more difficult for bad actors to disguise, such as their connections to others on our platform. As with all security threats, we have been incorporating new insights into our models for detecting fake accounts, including information specific to election issues.

Question 2. How does a user find out if they are being impersonated on Facebook or Instagram? Do Facebook and Instagram notify users proactively? Or are users expected to monitor the platforms and report to the company?

Answer. Claiming to be another person violates our Community Standards, and we want to make it harder for anyone to be impersonated on our platform. Users can also report accounts that are impersonating them. We've developed several techniques to help detect and block this type of abuse. At the time someone receives a friend request, our systems are designed to check whether the recipient already has a friend with the same name, along with a variety of other factors that help us determine if an interaction is legitimate. Further, we recently announced new features that use face recognition technology that may help people learn when someone is using their image as a profile photo—which can help stop impersonation. This is an area we're continually working to improve so that we can provide a safe and secure experience on Facebook.

Question 3. What are the average numbers of days or hours that Facebook and Instagram take to investigate impersonation complaints before they are resolved?

Answer. We promptly respond to reports of imposter accounts. Sometimes, these investigations are complex and require, for example, that users upload identification to confirm their identities. In general, the majority of all types of complaints received on Facebook are reviewed within 24 hours.

Question 4. Do Facebook and Instagram have a separate, expedited process for resolving impersonation of minors' accounts?

Answer. We take the issue of safety on our platform very seriously, especially that of our teen users. We want people to connect and share on Facebook, and it's integral that they feel safe in order to do so. We do not tolerate impersonation in any way and we remove profiles that impersonate other people. We have developed several techniques to help detect and block this type of abuse. At the time someone receives a friend request, for example, our systems are designed to check whether the recipient already has a friend with the same name, along with a variety of other factors that help us determine if an interaction is legitimate. It's an area we're continually working to improve so that we can provide a safe and secure experience.

Question 5. According to the current best estimate, approximately 126 million people on Facebook may have been served some piece of content associated with the Internet Research Agency (IRA) between January 2015 and August 2017. How many Instagram users were also served IRA content during the same time period? What was the methodology behind these estimates?

Answer. Using data analysis and modeling, we found that 11.4 million people in the United States saw at least one of the ads associated with the IRA between 2015 and 2017, and that as many as 126 million people in the United States may have seen a piece of IRA content on Facebook. Our data related to Instagram is incomplete, but we believe that as many as 16 million additional people who did not see this content on Facebook saw IRA content on Instagram starting in October 2016.

Question 6. How well did IRA content perform on Facebook and Instagram? Please provide metrics commonly measured on the platforms and benchmark against industry standards. This includes but is not limited to engagement (*i.e.,* time on post), reactions, impressions, and referral traffic for organic content. For ads, please provide the click-through rate and cost per reaction or reach.

Answer. As noted above, we found that 11.4 million people in the United States saw at least one of the ads associated with the IRA between 2015 and 2017, and

that as many as 126 million people in the United States may have seen a piece of IRA content on Facebook. Forty-four percent of total ad impressions were before the U.S. election, and 56 percent of total ad impressions were after the election. Roughly 25 percent of the ads were never shown to anyone. That's because advertising auctions are designed so that ads reach people based on relevance, and certain ads may not reach anyone as a result. Our data related to Instagram is incomplete, but we believe that as many as 16 million additional people who did not see this content on Facebook saw IRA content on Instagram starting in October 2016.

Question 7. Has Facebook shared the content, data, and metadata associated IRA activity above with researchers who are also looking into this? With law enforcement? With other companies? If not, then why not?

Answer. Facebook is providing investigators, including congressional committees, with information it has regarding the scope and nature of Russian information operations on our platform that may be relevant to their inquiries. We have also been working with many others in the technology industry on these issues.

Question 8. Facebook has several advertising tools and properties—including Facebook Events, Facebook Audience Network, and Facebook Canvas. Can you list all of them (including those for advertising on Instagram), and a succinct summary of what each of them do? Which of these were used by the IRA?

Answer. The Facebook family supports multiple advertising types. Each ad has two components: the format (what the ad looks like) and the placement (where it will be displayed). Ads can be placed on Facebook, Instagram, Messenger and Audience Network, which allows ads to be delivered on apps and sites beyond Facebook. Depending on where it is placed, available formats may include video; image; Collection, or displays of items from a product catalog; Carousel, or multiple image or videos within an ad; Slideshow, or video ads that can be seen at slower connection speeds; Canvas, or full-screen ads on mobile devices; lead generation ads, which allow advertisers to collect information from people interested in their business; offer ads that businesses can use to share discounts on their products; Post ads, which allows advertisers to have their Page posts appear beyond their Pages; ads for events; and ads for Page likes. The IRA generally used Page Likes or Page Post ads, typically with still images. The IRA also created some ads to promote events.

Facebook also offers three primary types of targeting, or audiences. Core Audiences are traditional targeting options based on location, demographics (age, gender education, job status, and more), interests, behavior, and connections. Custom Audiences are groups of specific people, like an advertiser's own contacts (Customer File Custom Audiences), visitors to an advertiser's website or app (Website Traffic Custom Audiences), or people who have engaged with an advertiser's content on Facebook services (Engagement Custom Audiences). Finally, Facebook offers Lookalike Audiences, which enables advertisers to find Facebook users that have similar characteristics to another audience. The targeting for the IRA ads that we have identified was relatively rudimentary, targeting broad locations and interests, and did not use Customer File Custom Audiences or Customer File Lookalike Audiences.

Question 9. In 2016, accounts affiliated with RT and Sputnik spent $5.4 million on Facebook advertising. How much did the same accounts spend on Facebook advertising in 2017? Does this include Instagram? How well did they perform? Again, please provide metrics commonly measured on the platforms and benchmark against industry standards. Were there any other Russian-linked accounts that heavily promoted RT or Sputnik content to the U.S. audience?

Answer. We have provided information concerning 2016 spending by RT and Sputnik in response to unique issues regarding the 2016 election. We have not conducted a similar analysis for 2017.

Question 10. Advertisers on both Facebook and Instagram generally pay for the size and quality of the audience that they would like to reach on the platforms. Did Facebook advertise to fake-user accounts? If so, how much revenue or profit did Facebook bring in by advertising to fake-user accounts?

Answer. We regularly evaluate metrics to estimate the number of "false" accounts among our monthly active users. We divide "false" accounts into two categories. The first category includes user-misclassified accounts, where users have created personal profiles for a business, organization, or non-human entity such as a pet (such entities are permitted on Facebook using a Page rather than a personal profile under our terms of service). The second category includes undesirable accounts, which represent user profiles that we determine are intended to be used for purposes that violate our terms of service, such as spamming. We estimate that in the fourth quarter of 2017, false accounts may have represented approximately three to four percent of our worldwide monthly active users. Our estimation of false accounts

can vary as a result of episodic spikes in the creation of such accounts, and additional information relating to our estimate of false accounts is included in our quarterly filings with the Securities and Exchange Commission. We continually update our technical systems to identify, checkpoint, and remove inauthentic accounts, which means that once discovered, these accounts do not remain active or eligible to view ads. We believe that revenue generated by advertising to false accounts is immaterial.

Question 11. Of the 20,000 people Facebook plans to employ by end of 2018 to work on safety and security, how many will be full-time, permanent employees? How many will be contractors? Will the majority of that team be located at the Menlo Park campus? Where is the majority of Facebook's current security and safety team located?

Answer. We have people working around the world on safety and security at Facebook. We use a combination of employees and contractors to make Facebook a place where both expression and personal safety are protected and respected. This allows us to scale globally with coverage across time zones, languages, and markets.

Question 12. Will Instagram have its own security and safety team as well? If so, please provide details.

Answer. Our safety and security teams work across Facebook's family of applications.

Question 13. Facebook recently announced that it will implement additional verification and disclosure requirements for advertisers running election ads for Federal elections. How will the ad onboarding process change for political advertisers on Facebook and Instagram? How will political ads look like to users of Facebook and Instagram? Please provide mock-ups for both the onboarding process and users' view of the ad.

Answer. We support efforts to promote greater transparency in political advertising online and are taking steps to make advertising on Facebook more transparent, increase requirements for authenticity, and strengthen our enforcement against ads that violate our policies. We will require more thorough documentation from advertisers who want to run election-related ads. As part of the documentation process, advertisers may be required to identify that they are running election-related advertising and verify both their entity and location. Once verified, these advertisers will have to include a disclosure in their election-related ads, which reads: "Paid for by." When users click on the disclosure, they will be able to see details about the advertiser, and we will maintain a searchable archive of information. Like other ads on Facebook, they will also be able to see an explanation of why they saw that particular ad. For more information, see *newsroom.fb.com/news/2017/10/update-on-our-advertising-transparency-and-authenticity-efforts.*

Question 14. For political advertisers who do not self-identify, will there be any human controls in addition to the automated tools to identify the ads proactively? Will Facebook and Instagram still publish the ad before its buyer is identified? If not, how long will the advertiser have to wait before the ad is published if they did not self-identify?

Answer. As part of our efforts to promote greater transparency in political advertising online, we'll require more thorough documentation from advertisers who want to run election-related ads. For political advertisers that do not proactively disclose themselves, we are building machine learning tools that will help us find them and require them to verify their identity. Once they are found, we will take appropriate steps to enforce compliance with our policies.

Question 15. Why are Facebook's verification and disclosure requirements for political advertisers only limited to Federal elections?

Answer. We are implementing new verification and disclosure standards on Facebook that will bring greater transparency to political advertising on our platform in general and make it easier for us to enforce our policies. We expect these reforms to be in effect by the 2018 U.S. Federal elections and will progress from there to additional contests and elections in other countries and jurisdictions.

Question 16. The Washington Post reported that Facebook removed the data that Jonathan Albright, a researcher at Columbia University, used to study Russia-linked ads. In response, Facebook stated that this was done in order to correct a bug in its system, and that Facebook policy requires that inactive content is no longer available across its platforms. What is Facebook's data retention policy? Is it consistent across all of Facebook's properties? Does this policy apply to all parties—such as independent researchers, users, advertisers, and data brokers—in the same way?

Answer. Facebook generally retains data for as long as it is commercially reasonable and necessary for our business. We have taken appropriate steps to retain relevant information related to IRA activity on Facebook.

Question 17. Given the importance of collaborating with third-party or independent researchers to prevent further interference by Russia, will Facebook be updating its data retention policy?

Answer. Our effort to make our platform safer and more secure is a holistic one that involves a continual evaluation of our personnel, processes, and policies, and we make changes as appropriate. We have taken appropriate steps to retain relevant information related to IRA activity on Facebook.

Question 18. In terms of dollars and percentage of annual revenue, how much is Facebook now spending on preventing foreign interference with our elections? What was the figure in the election cycle leading up to November 2016? What is the projected spend leading up to November 2018?

Answer. We are determined to do everything we can to protect our platform. We are expanding our threat intelligence team, and more broadly, we are working now to more than double the number of people working on safety and security at Facebook, from 10,000 to 20,000, by the end of 2018. Many of the people we are adding to these efforts will join our ad review team, and we also expect to add at least 3,000 people to Community Operations, which is the team that reviews content that our users and automated tools flag as inappropriate, dangerous, abusive, or otherwise violative of our policies. These investments will help us to enforce our policies, including our authenticity policy, and help us to counter threats from malicious actors, including those who are state-sponsored. We will also significantly expand the number of people who work specifically on election integrity before the 2018 U.S. Federal elections, including people who investigate information operations by foreign actors. The investments that we are making to address election integrity and other security issues will be so significant that we have informed investors that we expect that the amount that we will spend will impact our profitability.

Question 19. Congress will judge success not by Facebook's efforts but by its results. How will Facebook measure success? Will Facebook be conducting an audit after November 2018? When will the results be shared?

Answer. Success would consist of minimizing or eliminating abuse of our platform and keeping our community safe. We have a number of specific goals that we will use to measure our progress in these efforts.

First, we will increase the number of people working on safety and security at Facebook, to 20,000 by the end of this year. We will significantly expand the number of people who work specifically on election integrity, including people who investigate this specific kind of abuse by foreign actors. Those specialists will find and remove more of these actors.

Second, we will work to improve threat intelligence sharing across our industry, including, we hope, by having other companies join us in formalizing these efforts. This is a fight against sophisticated actors, and our entire industry needs to work together to respond quickly and effectively.

Third, we will bring greater transparency to election ads on Facebook by requiring more disclosure from people who want to run election ads about who is paying for the ads and by making it possible to see all of the ads that an advertiser is running, regardless of the targeting. We believe that these efforts will help to educate our community and to arm users, media, civil society, and the government with information that will make it easier to identify more sophisticated abuse to us and to law enforcement.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. TAMMY BALDWIN TO MONIKA BICKERT

Question 1. According to press reports, my home state of Wisconsin was one of the states where voters were targeted by Russian groups with Facebook advertising and political content that employed a series of divisive messages on key issues like race relations and immigration. In your testimony, you state that Facebook "continues to seek more effective ways to combat extremism, crime and other threats to our national security."

Do you believe that these foreign-directed activities intended to sew social discord and influence our elections are threats to our national security?

I request that Facebook provide me a verbal and written briefing regarding the Russia-linked political ads and content targeting Wisconsin, to include: a list and description of the ads and content; the entities responsible for the ads and content;

Facebook's assessment of the intent of such entities; how the ads and content were targeted geographically within the state, with regard to social or political issues, and with regard to audience subgroups; and the timing of the ads and content.

What steps is Facebook taking to ensure that Russia or other foreign governments cannot repeat this effort? And that Facebook users understand the source of this type of advertising and content?

Answer. The foreign interference we saw in the 2016 election is reprehensible and outrageous and opened a new battleground for our company, our industry, and our society. That foreign actors, hiding behind fake accounts, abused our platform and other Internet services to try to sow division and discord—and to try to undermine our election process—is an assault on democracy and our national security, and it violates all of our values. At Facebook, we build tools to help people connect, and to be a force for good in the world. What these actors did goes against everything Facebook stands for. Our goal is to bring people closer together; what we saw from these actors was an insidious attempt to drive people apart.

We're determined to do our part to prevent it from happening again. One improvement that we believe will help to address more subtle kinds of abuse is that our ad review team will do more to assess not just the content, but also the overall context of an ad, including the buyer and intended audience. We will also significantly expand the number of people who work specifically on election integrity before the 2018 U.S. Federal elections, including people who investigate this specific kind of abuse by foreign actors. Additionally, we have begun testing a program where people will be able to click "View Ads" on a Page and view advertisements a Page is running on Facebook, Instagram, and Messenger—whether or not the person viewing it is in the intended target audience for the ad. All Pages will be part of this effort, and we will require that all ads be associated with a Page as part of the ad creation process. We are also taking steps to make political ads on Facebook more transparent, increase requirements for authenticity, and strengthen our enforcement of ads that violate our policies. And, we continue to make improvements to our efforts to more effectively detect and deactivate fake accounts to help reduce the spread of spam, false news, and misinformation. We continually update our technical systems to identify, checkpoint, and remove inauthentic accounts, and we block millions of attempts to register fake accounts every day. These systems examine thousands of detailed account attributes and prioritize signals that are more difficult for bad actors to disguise, such as their connections to others on our platform. As with all security threats, we have been incorporating new insights into our models for detecting fake accounts, including information specific to election issues.

We are determined to do everything that we can to protect our platform. The investments that we are making to address these issues and other security issues will be so significant that we have informed investors that we expect that the amount that we will spend will impact our profitability. We will continue to work the government, and across the tech industry and civil society, to address this important national security matter so that we can do our part to prevent similar abuse from happening again. That's why we have provided all of the ads and associated information to the committees with longstanding, bipartisan investigations into Russian interference, and we defer to the committees to share as appropriate. We believe that Congress and law enforcement are best positioned to assess the nature and intent of these activities.

Question 2. Ms. Bickert, the national security website Just Security recently published troubling evidence that raises doubts about Facebook's ability to prevent, monitor, and remove extremist content. According to the article, in a one month period spanning December 2017–January 2018, a researcher named Eric Feinberg reported dozens of pro-ISIS pages to Facebook. That means the material had gotten past your company's initial means for flagging and removing terrorist content. In 56 percent of those cases, Facebook removed the offending page. But for the other 44 percent of reported pages, Facebook left the content up, noting that it didn't violate community standards. This is despite there being no appreciable difference between content that was removed and content that was retained. For example, pages that Facebook left up included: a photo of gunmen in an urban neighborhood with the caption, "We Will Attack you in Your Home;" an online publication promoting ISIS among the Bangladeshi community; and a photo of Omar Mateen, praising his attack on Orlando's Pulse nightclub.

Retaining this content contradicts Facebook's explicit policies, internal guidelines, and your testimony. Can you please explain this?

Answer. We immediately remove terrorists' accounts and posts that support terrorism whenever we become aware of them. When we receive reports of potential terrorism posts, we review those reports urgently and with scrutiny. After receiving the article mentioned in the question, we reviewed the accounts and content identi-

fied in the article and disabled and removed all those that violated our policies. Managing a global community in this manner has never been done before, and we know we have a lot more work to do. We are committed to improving and to ensuring that hate has no place on Facebook.

Question 3. In the context of extremist content, I would like to learn more about each company's policy for proactively reporting users to law enforcement. I understand your companies evaluate and respond to law enforcement requests for information, but what framework do you use to proactively report terrorist-related content to authorities, including any identifying information of the user? For example, if you use a standard of imminent harm, how do you define and apply it, particularly in a threat environment where terrorist organizations often call on recruits to attack carefully planned targets of opportunity, rather than to launch an immediate, indiscriminate attack?

Answer. We have a long history of working successfully with the DOJ, the FBI, and other government agencies to address a wide variety of threats to our platform, including terrorist threats. We reach out to law enforcement whenever we see a credible threat of imminent harm. We have been able to provide support to authorities around the world that are responding to the threat of terrorism, including in cases where law enforcement has been able to disrupt attacks and prevent harm.

We cooperate with governments in other ways, too. For example, as part of official investigations, government officials sometimes request data about people who use Facebook. We have strict processes in place to handle these government requests, and we disclose account records in accordance with our terms of service and applicable law. We also have law enforcement response teams available around the clock to respond to emergency requests. Further, governments and inter-governmental agencies also have a key role to play in convening and providing expertise that is impossible for companies to develop independently. We have learned much through briefings from agencies in different countries about ISIS and Al Qaeda propaganda mechanisms.

Question 4. I would like to hear from the companies whether they support implementing Mr. Watts's recommendations to: first, fully certify the authenticity of all users—in other words, ensure that each user is a real person; and second, eliminate social bot networks to reduce automated broadcasting of disinformation.

Answer. We have always believed that Facebook is a place for authentic dialogue, and that the best way to ensure authenticity is to require people to use the names they are known by. Fake accounts undermine this objective, and are closely related to the creation and spread of inauthentic communication such as spam and disinformation. We also prohibit the use of automated means to access our platform. We rely on both automated and manual review in our efforts to effectively detect and deactivate fake accounts, including bots, and we are now taking steps to strengthen both. For example, we continually update our technical systems to identify, checkpoint, and remove inauthentic accounts. We block millions of attempts to register fake accounts every day. These systems examine thousands of detailed account attributes and prioritize signals that are more difficult for bad actors to disguise.

Question 5. What are the indicators that you use to identify a Russian disinformation account, whether from the Kremlin's so-called Internet Research Agency or an associated group of hackers or trolls, and what thresholds must be met to disable an account?

Answer. We continually update our technical systems to identify, checkpoint, and remove inauthentic accounts, including accounts used for state-sponsored information operations. We block millions of attempts to register fake accounts every day. These systems examine thousands of detailed account attributes such as location information and connections to others on our platform.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. CATHERINE CORTEZ MASTO TO MONIKA BICKERT

Question 1. During my time as the Attorney General for the State of Nevada, I saw too many instances of sex trafficking cases involving child victims that were dismissed because the conduct occurred online or through social media. So that's why I'm a strong supporter of the Stop Enabling Sex Traffickers Act of 2017 (SESTA), which clarifies the Communications Decency Act (CDA) to allow state Attorneys General to retain their jurisdiction to prosecute those who facilitate human trafficking. We know that trafficking is happening online and on social media, and SESTA is the only current legislative proposal that provides sufficient deterrence to

traffickers by providing the necessary tools for successful prosecutions. As a former prosecutor, I know what it will take to successfully prosecute those who engage in sex trafficking through social media and other websites, and that's why I believe that the House version of SESTA doesn't go far enough to give prosecutors the tools they need to protect sex trafficking victims. I hope that your organizations all agree that victims of sex trafficking deserve meaningful protections and justice.

If so, I'd like to hear whether you will continue to support SESTA over the weaker U.S. House version of the bill.

Answer. Facebook supports SESTA. We look forward to continuing to work with Congress to pass this important legislation.

Facebook is committed to making our platform a safe place, especially for individuals who may be vulnerable. We have a long history of working successfully with governments to address a wide variety of threats to our platform, including child exploitation. When we learn of a situation involving physical abuse, child exploitation, or an imminent threat of harm to a person, we immediately report the situation to first responders or the National Center for Missing and Exploited Children (NCMEC). Further, as part of official investigations, government officials sometimes request data about people who use Facebook. We have processes in place to handle these government requests, and we disclose account records in accordance with our terms of service and applicable law. We also have a global team that strives to respond within minutes to emergency requests from law enforcement.

Our relationship with NCMEC also extends to an effort that we launched in 2015 to send AMBER Alerts to the Facebook community to help find missing children. When police determine that a case qualifies for an AMBER Alert, the alert is issued by the NCMEC and distributed through the Facebook system with any available information, including a photograph of the missing child, a license plate number, and the names and descriptions of the child and suspected abductor. Law enforcement determines the range of the target area for each alert. We know the chances of finding a missing child increase when more people are on the lookout, especially in the critical first hours. Our goal is to help get these alerts out quickly to the people who are in the best position to help, and a number of missing children have been found through AMBER Alerts on Facebook.

Further, we work tirelessly to identify and report child exploitation images (CEI) to appropriate authorities. We identify CEI through a combination of automated and manual review. On the automated review side, we use image hashing to identify known CEI. On the manual review side, we provide in-depth training to content reviewers on how to identify possible CEI. Confirmed CEI is reported to the NCMEC, which then forwards this information to appropriate authorities. When we report content to the NCMEC, we preserve account information in accordance with applicable law, which can help further law enforcement investigations. We also reach out to law enforcement authorities in serious cases to ensure that our reports are received and acted upon.

Question 2. I was glad to hear that the Internet Association supports SESTA, and I'd like to know what else your organization is doing to address concerns about sex trafficking occurring on your platforms and helping us pass this important legislation in the Senate.

Answer. Please see the response to question 1.

Question 3. I am glad that Facebook has acknowledged the practices of discrimination through employment or housing as problematic, by committing to roll out programs that screen for bad ads, and requiring advertisers to certify their compliance with antidiscrimination laws.

Can you provide a brief update on the progress of those programs?

Answer. We have Community Standards that prohibit hate speech, bullying, intimidation, and other kinds of harmful behavior. We hold advertisers to even stricter advertising policies to protect users from things like discriminatory ads. We don't want advertising to be used for hate or discrimination, and our policies reflect that. For example, we make it clear that advertisers may not discriminate against people based on personal attributes such as race, ethnicity, color, national origin, religion, age, sex, sexual orientation, gender identity, family status, disability, and medical or genetic condition. We educate advertisers on our anti-discrimination policy, and in some cases, we require advertisers to certify compliance with our anti-discrimination policy and anti-discrimination laws.

We are committed to getting better at enforcing our advertising policies. We review many ads proactively using automated and manual tools, and reactively when people hide, block, or mark ads as offensive. We are taking aggressive steps to strengthen both our automated and our manual review. Reviewing ads means assessing not just the content of an ad, but the context in which it was bought—such

as the identity of the advertiser and the landing page—and the intended audience. We are changing our ads review system to pay more attention to these signals. We are also expanding our global ads review teams and investing more in machine learning to better understand when to flag and take down ads, such as ads that offer employment or credit opportunity while including or excluding multicultural advertising segments. Enforcement is never perfect, but we will get better at finding and removing improper ads.

Question 4. What metrics are in place that you can provide us to be confident in the facts and figures you provide to address this concern?

In my view, permitting ad targeting on the basis of age, race, religion, or other protected characteristics, especially without a robust process to review ads for compliance with applicable antidiscrimination laws, is likely to facilitate unlawful discrimination.

Answer. Please see the response to question 3.

Question 5. So, if we are to believe that Facebook is serious about combatting discrimination other hateful content, why would it choose to facilitate discriminatory practices for advertisers using Facebook's services?

Answer. Discriminatory advertising has no place on Facebook. Our advertising policies prohibit discrimination, and in some cases, we require advertisers to certify compliance with our anti-discrimination policy and anti-discrimination laws. We use automated and manual review to find and remove discriminatory ads, and we constantly seek to strengthen our ability to enforce our policies and prevent discrimination on Facebook.

Question 6. Given that Facebook, like other tech giants, may not be meeting their own goals for workplace diversity, are you confident that your employment advertising program meets all of its obligations to avoid facilitating unlawful employment discrimination?

Answer. As noted above, discriminatory advertising has no place on Facebook, and we are constantly trying to find ways to improve enforcement of our anti-discrimination policies. To assist us in these efforts, we have met with policymakers and civil rights leaders to listen to their concerns and to gather feedback about ways to improve our enforcement while preserving the beneficial uses of our advertising tools. We are grateful for the collaboration of many experts who have worked with us to develop solutions to combat discriminatory ads. We look forward to finding additional ways to combat discrimination, while increasing opportunity for underserved communities, and to continuing our dialogue with policymakers and civil rights leaders about these important issues.

Question 7. How can you be sure that the program isn't replicating the same biases and blind spots that have impeded your own diversity efforts?

Answer. Please see the response to question 6.

Question 8. Over the past few months, our country has been reckoning with some hard truths about the way that women and minorities are treated in the workplace. And I think this is a moment for all types of organizations, including tech giants like the ones represented here, to take a clear-eyed accounting of their culture and practices, to take responsibility for what hasn't worked, and to renew their commitments to make meaningful improvements. The Equal Employment Opportunity Commission's 2016 report on "Diversity in High Tech" found that women, African Americans, and Hispanics are all represented at significantly lower levels in high tech than in private industry as a whole. And while recent internal studies at Facebook and Google have showed some progress in the hiring of women, there has not been equal improvement in the representation of people of color and other underrepresented groups.

What technically qualifies as diversity to your organization?

Answer. With a global community of over two billion people on Facebook, greater diversity and inclusivity are critical to achieving our mission. Studies have shown that cognitive diversity on teams that are working on hard problems produces better results. Diversity helps us build better products, make better decisions and better serve our community. In order to achieve that, we have developed programming to attract and retain more people from traditionally underrepresented groups which include women, people of color, veterans and people with disabilities.

We are not where we would like to be, but we are encouraged that representation for people from underrepresented groups at Facebook has increased. We've grown Black and Hispanic representation by 1 percent each (2 percent combined) between our first report in 2014 and our most recent report in 2017:

- Black Representation: from 2 percent to 3 percent
- Hispanic Representation: from 4 percent to 5 percent

- Black Non-Tech: from 2 percent to 6 percent
- Hispanic Non-Tech: from 6 percent to 8 percent
- Black Leadership: from 2 percent to 3 percent
- Hispanic Leadership: from 4 percent to 3 percent
- Black and Hispanic Tech have stayed at 1 percent and 3 percent

As of August 2017, the number of women globally increased from 33 percent to 35 percent and the number of women in tech increased from 17 percent to 19 percent. Women made up 27 percent of all new graduate hires in engineering and 21 percent of all new technical hires at Facebook.

We seek to promote diversity in a variety of ways, and we want to highlight three programs in particular. First, we have adopted our Diverse Slate Approach (DSA) to interviewing job candidates. The more people that hirers interview who don't look or think like them, the more likely they are to hire someone from a diverse background. To hardwire this behavior at Facebook, we introduced our DSA in 2015 and have since rolled it out globally. DSA sets the expectation that hiring managers will consider candidates from underrepresented backgrounds when interviewing for an open position.

Second, we are working to reduce unconscious bias. Our publicly available Managing Unconscious Bias class encourages our people to challenge and correct bias as soon as they see it—in others, and in themselves. We've also doubled down by adding two new internal programs: Managing Inclusion, which trains managers to understand the issues that affect marginalized communities, and Be The Ally, which gives everyone the common language, tools, and space to practice supporting others.

Third, we have created Facebook University. We want to increase access and opportunity for students with an interest in software engineering, business, and analytics. Facebook University (FBU) gives underrepresented students extra training and mentorship earlier in their college education. We started FBU in 2013 with 30 students and expect to have 280 in 2018. More than 500 students have graduated from this program, with many returning to Facebook for internships and full-time jobs.

Finally, we have many partnerships to move the numbers nationally such as Black Girls Code, All Star Code, Hack the Hood, The Hidden Genius Project, Level Playing Field Institute, Yes We Code, Streetcode Academy, Dev Color, Dev Bootcamp and Techbridge. And, we now recruit at 300 Universities—including historically black colleges and universities (HBCUs) like Spelman, Morehouse, Howard, NCA&T, and Morgan State (EIR) and the HBCU Faculty Summit.

We're committed to building a more diverse, inclusive Facebook. Much like our approach to launching new products on our platform, we are willing to experiment and listen to feedback.

Question 9. How is your company working to address issues of discrimination in your own workforces?

Answer. Please see the response to question 8.

Question 10. Do you believe those efforts are sufficient?

Answer. Please see the response to question 8.

Question 11. We know that so-called talent pipelines are not the only obstacle to achieving a diverse workforce, and that discrimination and harassment go hand in hand, distorting the operation of workplace meritocracies. This is a moment when many victims of sexual assault and harassment are bravely coming forward about their experiences, allowing us to get a better sense of the true scope and effects of this behavior. Persistent harassment, and the workplace culture that tolerates, ignores, or even encourages such harassment, pushes people out of their workplaces, stalls or derails promising careers, and discourages some from pursuing certain opportunities altogether.

What is your company doing to evaluate the impact of harassment in your workforces?

Answer. Harassment, discrimination, and retaliation in the workplace are unacceptable but have been tolerated for far too long.

At Facebook, we treat any allegations of such behavior with great seriousness, and we have invested significant time and resources into developing our policies and processes. We have made our policies and processes available publicly—not because we think we have all the answers, but because we believe that the more companies are open about their policies, the more we can all learn from one another. These are complicated issues, and while we don't believe any company's enforcement or policies are perfect, we think that sharing best practices can help us all improve, especially smaller companies that may not have the resources to develop their own policies. Every company should aspire to doing the hard and continual work nec-

essary to build a safe and respectful workplace, and we should all join together to make this happen.

Our internal policies on sexual harassment and bullying are available on our Facebook People Practices website (peoplepractices.fb.com), along with details of our investigation process and tips and resources we have found helpful in preparing our Respectful Workplace internal trainings. Our philosophy on harassment, discrimination, and bullying is to go above and beyond what is required by law. Our policies prohibit intimidating, offensive, and sexual conduct even when that conduct might not meet the legal standard of harassment. Even if it's legally acceptable, it's not the kind of behavior we want in our workplace.

In developing our policies, we were guided by six basic principles:

First, develop training that sets the standard for respectful behavior at work, so people understand what's expected of them right from the start. In addition to prescribing mandatory harassment training, we wrote our own unconscious bias training program at Facebook, which is also available publicly on our People Practices website

Second, treat all claims—and the people who voice them—with seriousness, urgency, and respect. At Facebook, we make sure to have HR business partners available to support everyone on the team, not just senior leaders.

Third, create an investigation process that protects employees from stigma or retaliation. Facebook has an investigations team made up of experienced HR professionals and lawyers trained to handle sensitive cases of sexual harassment and assault.

Fourth, follow a process that is consistently applied in every case and is viewed by employees as providing fair procedures for both victims and those accused.

Fifth, take swift and decisive action when it is determined that wrongdoing has occurred. We have a zero-tolerance policy, and that means that when we are able to determine that harassment has occurred, those responsible are fired. Unfortunately, in some cases investigations are inconclusive and come down to one person's word against another's. When we don't feel we can make a termination decision, we take other actions designed to help everyone feel safe, including changing people's roles and reporting lines.

Sixth, make it clear that all employees are responsible for keeping the workplace safe—and anyone who is silent or looks the other way is complicit.

There's no question that it is complicated and challenging to get this right. We are by no means perfect, and there will always be bad actors. Unlike law enforcement agencies, companies don't have access to forensic evidence and instead have to rely on reported conversations, written evidence, and the best judgment of investigators and legal experts. What we can do is be as transparent as possible, share best practices, and learn from one another—recognizing that policies will evolve as we gain experience. We don't have everything worked out at Facebook on these issues, but we will never stop striving to make sure we have a safe and respectful working environment for all our people.

Question 12. How are you working to create a culture where harassment is no longer tolerated?

Answer. Please see the response to question 11.

Question 13. What more could you be doing to be a positive example for other companies and industries?

Answer. Please see the response to question 11.

Question 14. Last October, Facebook announced that it would be improving transparency for all ads run on its platform, including by requiring political advertisers to include a disclaimer telling viewers who paid for an ad, and allowing viewers to see all the ads a page is running, even those that aren't targeting them. Twitter also announced similar measures. Although these policies were announced in response to Russia using social media to interfere in our elections, it seems these transparency measures could help shine a spotlight on other forms of influence campaigns by extremists or terrorists.

Can you provide an update on the status of these measures?

Answer. We recently announced steps to make advertising on Facebook more transparent, increase requirements for authenticity, and strengthen our enforcement against ads that violate our policies. We'll require more thorough documentation from advertisers who want to run election-related ads. We are starting with the 2018 Federal elections in the United States, and will progress from there to additional contests and elections in other countries and jurisdictions. As part of the documentation process, advertisers may be required to identify that they are running election-related advertising and verify both their entity and location. Once verified, these advertisers will have to include a disclosure in their election-related ads,

which reads: "Paid for by." When users click on the disclosure, they will be able to see details about the advertiser, and we will maintain a searchable archive of information. Like other ads on Facebook, they will also be able to see an explanation of why they saw that particular ad. For political advertisers that do not proactively disclose themselves, we are building machine learning tools that will help us find them and require them to verify their identity.

Further, for all ads on Facebook—not just political ads—we have begun testing a program where people will be able to click "View Ads" on a Page and view advertisements a Page is running on Facebook, Instagram, and Messenger—whether or not the person viewing it is in the intended target audience for the ad. All Pages will be part of this effort, and we will require that all ads be associated with a Page as part of the ad creation process. We will roll this feature out to the United States by this summer, ahead of the U.S. midterm elections in November, as well as broadly to all other countries around the same time.

Question 15. When can we expect to see them fully implemented?

Answer. Please see the response to question 14.

Question 16. How are you defining what constitutes a political ad subject to these heightened transparency requirements?

Answer. Our commitment to ad transparency is not limited to political ads. While our most recent announcements have focused on election-related ads—although not necessarily only ads that mention candidates by name—we are bringing greater transparency to all ads by making sure that people can see all of the ads run by any Page, regardless of whether those ads are targeted to them.

Question 17. On January 29, the Director of the Central Intelligence Agency said he expects the Russian government to attempt to influence the 2018 elections in this *country.*

What efforts is Facebook undertaking in the lead up to the 2018 elections to identify and close the platform's remaining vulnerabilities to foreign exploitation?

Answer. The foreign interference we saw in the 2016 election is reprehensible and outrageous and opened a new battleground for our company, our industry, and our society. We're determined to do our part to prevent it from happening again. We are more than doubling the number of people who work on safety and security at Facebook and have already hired thousands more content reviewers. They will be engaged in processes that we are continuously refining, but this significant investment of resources will help us to perform those processes more accurately, quickly, and thoroughly. One improvement that we believe will help to address more subtle kinds of abuse is that our ad review team will do more to assess not just the content, but also the overall context of an ad, including the buyer and intended audience. We will also significantly expand the number of people who work specifically on election integrity before the 2018 U.S. Federal elections, including people who investigate this specific kind of abuse by foreign actors.

Additionally, we have begun testing a program where people will be able to click "View Ads" on a Page and view advertisements a Page is running on Facebook, Instagram, and Messenger—whether or not the person viewing it is in the intended target audience for the ad. All Pages will be part of this effort, and we will require that all ads be associated with a Page as part of the ad creation process. We are also taking steps to make political ads on Facebook more transparent, increase requirements for authenticity, and strengthen our enforcement of ads that violate our policies. And, we continue to make improvements to our efforts to more effectively detect and deactivate fake accounts to help reduce the spread of spam, false news, and misinformation. We continually update our technical systems to identify, checkpoint, and remove inauthentic accounts, and we block millions of attempts to register fake accounts every day. These systems examine thousands of detailed account attributes and prioritize signals that are more difficult for bad actors to disguise. As with all security threats, we have been incorporating new insights into our models for detecting fake accounts, including information specific to election issues.

We are determined to do everything that we can to protect our platform. The investments that we are making to address these issues and other security issues will be so significant that we have informed investors that we expect that the amount that we will spend will impact our profitability.

Question 18. What assistance can Federal, state and local government entities provide in that effort?

Answer. We have a long history of working successfully with the DOJ, the FBI, and other law enforcement to address a wide variety of threats to our platform, including threats emanating from Russia. We deeply respect and value the seriousness, diligence, and support of those organizations, and we would welcome their partnership as we work to address this specific threat. We are particularly encour-

aged by the FBI's creation of a task force dedicated to addressing election interference and we are actively working with that newly-formed body. This is a new kind of threat, and we believe that we will need to work together—across industry and between industry and government—to be successful.

Question 19. In November 2016, National Public Radio reported that Facebook had a few hundred subcontractors across several countries responsible for reviewing malicious content. That amounted to contractors reviewing one tag that's been flagged approximately every 10 seconds. In your testimony you said you now have 7,500 employees reviewing "terror content and other potential violations." You also said you have 180 people focused specifically on countering terrorism.

Can you please describe the content categories you have employees reviewing, how each is defined and how material is categorized?

Answer. Our teams review a variety of content for compliance with our policies. For example, our Business Integrity team focuses on reviewing and removing ads that do not comply with our advertising policies, which prohibit discriminatory practices, ads for certain types of products, misleading or false content, and other activity and content (*see facebook.com/policies/ads*). Our Community Operations team reviews content that our users and automated tools flag as inappropriate, dangerous, abusive, or otherwise violating our Community Standards. Our Community Standards prohibit things like serious threats of harm to public and personal safety, organizations engaged in terrorist or other violent activity, hate speech, bullying and harassment, criminal activity, sexual exploitation, the sale of certain regulated goods, fraud and spam, and other content and activity (*see facebook.com/communitystandards*).

Question 20. Please provide the average number of posts or content flagged per day, the number of employees devoted to each category, and the amount of time each employee has to review the content. If this is not a fair metric of employee time allocation to content review, please provide the metrics requested, as well as an explanation of the content review systems that allow employees to review a vast amount of content quickly.

Answer. All content goes through some degree of automated review, and we use human reviewers to check some content that has been flagged by that automated review or reported by people that use Facebook. Our content reviewers respond quickly to millions of reports each week from people all over the world. We also use human reviewers to perform reviews of content that was not flagged or reported to check the accuracy and efficiency of our automated review systems. The amount of content a reviewer views per day may vary based on the type of content reviewed and the reason for review.

Question 21. What differentiates employees reviewing "terror [content] and other potential violations" from employees focused on "countering terrorism"? What is the role of each?

Answer. Our Community Standards prohibit, among other things, individuals and organizations that are engaged in terrorist activity and also prohibit content that expresses support for terrorism. Our content review teams include more than 7,500 people around the world who work 24 hours a day and in more than 40 of languages to review reports of terrorist content and other content that may violate our Community Standards. Separately, we have more than 180 highly trained people who are exclusively or primarily focused on preventing terrorist content from ever appearing on our platform and quickly and identifying and removing it if it does. This group includes former academics who are experts on counterterrorism, former prosecutors and law enforcement agents, investigators and analysts, and engineers. Among other things, this team helps to build tools and leverages counterterrorism research to detect terrorist activity and prevent the spread of propaganda.

Question 22. Have you found that having 180 employees focused on countering terrorism is sufficient for the vast amount of content posted daily on Facebook?

Answer. While 180 employees focus on countering terrorism as the core part of their job at Facebook, many others in the company share this responsibility as part of their job. This includes the 7,500 content reviewers who remove from our site content that violates our policies, including terrorism policies. We work continuously to make our platform safer and more secure, and our effort to do so is a holistic one that involves not only hiring additional employees when issues arise, but also a continual evaluation of our processes and policies. In addition to our counterterrorism specialists, thousands of reviewers review content that our users and automated tools flag is inappropriate, dangerous, abusive, or otherwise violating our Community Standards—including those prohibiting terrorism. More broadly, we are working now to ensure that we will more than double the number of people working on safety and security at Facebook, from 10,000 to 20,000, by the end of 2018.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JERRY MORAN TO JUNIPER DOWNS

Question 1. Your testimony covered the "counter-narratives" that YouTube is currently utilizing to speak out against terrorism. Last week, I had the pleasure of hearing about the important work that is being done in this effort within Jigsaw's project called "Redirect Method."

Answer. Yes, these strategies have been researched and developed over many years, and we appreciated your acknowledgement of that work in the hearing.

Question 2. Could you please describe how this project targets the most susceptible audience to "redirect" them to videos debunking recruitment materials?

Answer. The Redirect Method uses Adwords targeting tools and third party curated YouTube videos to confront online radicalization. As you mentioned, the targeting efforts focus on the slice of ISIS' audience that is most susceptible to its messaging, and redirects them toward third party created YouTube videos debunking ISIS recruiting themes. This open methodology was developed in part from interviews with ISIS defectors. Jigsaw initially tested the Redirect Method in an ISIS-focused campaign in Arabic and English. Over the course of 8 weeks, 320,000 individuals watched over half a million minutes of the 116 videos we selected to refute ISIS's recruiting themes. The Redirect Method has recently been deployed in the UK and France. The Redirect Method is open for any institution to use in their work.

Question 3. Does "Redirect Method" or YouTube create the videos that program redirects the audience to? If not, why is that the case?

Answer. Many previous efforts to push back on extremist propaganda have focused on creating new content—writing, videos, etc.—to dispel extremist narratives. Through our research, we found that content that had been created for the sole purpose of dispelling extremist narratives didn't tend to resonate as well as much of the organic content that was already available online. For this reason, The Redirect Method focuses on curation of pre-existing content to push back against extremist propaganda while more effectively reaching the target audience. We work with local scholars and experts to curate the videos in Redirect playlists. In France, our interdisciplinary research team of scholars is based at the Castex Chair of Geostrategy. In the UK, the Institute for Strategic Dialogue (ISD) and Moonshot CVE participate in our curation and research efforts.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. RON JOHNSON TO JUNIPER DOWNS

Question 1. Social media companies are increasingly able to remove terrorist recruitment, incitement, and training materials before it posts to their platforms by relying on improved automated systems. Other than content removal, what else can be done to limit the audience or distribution of these dangerous materials?

Answer. YouTube's Community Guidelines set the rules of the road for content that we allow on the platform. Our policies include prohibitions on hate speech, gratuitous violence, incitement to violence, terrorist recruitment videos, and violent propaganda. We also have robust advertiser-friendly guidelines and demonetize videos that don't comply with those policies, and can age-restrict or place a warning interstitial in front of content that may be shocking.

If our review teams determine that a video does not contain a direct call to violence or incitement to hate but could be inflammatory we may disable some features. Identified borderline content will remain on YouTube behind a warning interstitial, won't be recommended, won't be monetized, and won't have key features including comments, suggested videos, and likes. This new treatment has been positive, with substantial reduction in watch time of those videos.

We disable access to our services for users who repeatedly violate our policies—and, for egregious violations, for the first offense. We also terminate the Google accounts of entities on the U.S. State Department's Foreign Terrorist Organization (FTO) list, regardless of the content they are posting.

In addition to ensuring our policies are effectively enforced, we invest heavily in counterspeech. We see lots of examples of counterspeech working, such as creators stepping up to refute content related to violent extremism. In many cases, these creators are driving even more engagement than the original objectionable content. Exposing susceptible individuals to counterspeech content is universally viewed as a critical component of counterterrorism and other counter radicalization strategies. To that end, we've held over 20 counterspeech workshops around the world, pairing anti-radicalization NGOs with YouTube creators who know how to best engage with and relate to their audiences. In 2016, we launched YouTube #CreatorsforChange,

a global counterspeech initiative aimed at amplifying and multiplying the voices of role models who are tackling difficult social issues such as xenophobia, hate speech, and extremism.

Question 2. Terrorist how-to guides are protected by the First Amendment in the United States, but violate the content policies of many social media companies as well as the laws of some international partner nations. What countries have laws that go beyond your company's content policies and can you give examples of how you have worked with those countries to de-conflict those differences?

Answer. Although we are a U.S.-based company, we respect the law in countries where we operate. Sometimes those laws restrict speech more than our Community Guidelines require. Holocaust denial, for example, while protected by the First Amendment in the United States, is against the law in many European countries. In countries where we conclude the law so requires, we would remove such content from our results.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. MARIA CANTWELL TO JUNIPER DOWNS

Question 1. We have strong principles of freedom of speech, but at the same time, we need to balance that freedom with the need to protect against bad actors who would leverage that freedom to plan and promote illegal acts. How can we use artificial intelligence to help us achieve a balance between our American ideal of free speech and the need to protect against extremist acts of terror?

Answer. YouTube has always used a mix of humans and technology to enforce our policies efficiently and at scale. Humans are a critical component in the enforcement process as very often understanding nuance and context is necessary to determine if content is in violation of the company's policies. That said, as the technology improves in its precision, YouTube leverages these advancements to enhance the enforcement operations.

For example, YouTube has invested heavily in cutting-edge artificial intelligence and machine learning technology designed to help us identify and remove violent extremist and terrorism-related content in a scalable way. Last year, we began deploying these classifiers that detect potential terrorist material and flag it for review by people trained to enforce our policies. These efforts have resulted in some positive progress:

Speed and efficiency: Our machine learning systems are faster and more effective than ever before. Last June, only 40 percent of the videos we removed for violent extremism were identified by our algorithms. Today, that number is 98 percent. Our advances in machine learning let us now take down nearly 70 percent of violent extremism content within 8 hours of upload and nearly half of it in 2 hours.

Accuracy: The efficiency of our systems has improved dramatically due to our machine learning technology. While these tools aren't perfect, and aren't right for every setting, in many cases our systems have proven more accurate than humans at flagging videos that need to be removed.

Scale: With over 400 hours of content uploaded to YouTube every minute, finding and taking action on violent extremist content poses a significant challenge. But since June, our teams have manually reviewed approximately two million videos to improve our machine-learning flagging technology by providing large volumes of training examples. Noteworthy, every subsequent decision on content that has been flagged by this technology serves as an additional input that continues to train and improve the system. We are encouraged by these improvements, and will continue to develop our technology in order to make even more progress. We are also hiring more people to help review and enforce our policies—reaching 10,000 people across Google working to address content that might violate our policies by the end of this year—and will continue to invest in technical resources to keep pace with these issues and address them responsibly. Our commitment to combat these issues is sustained and unwavering.

Question 2. Outside of artificial intelligence, what other technologies could be used to combat potential radicalization on social media platforms? What does the implementation of those technologies look like?

Answer. No single component can solve the problem of extremist content in isolation. In addition to our work on artificial intelligence and machine learning, YouTube uses 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 also 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. Given the higher likelihood that flags from these organizations are actionable, flags from Trusted Flaggers are prioritized for review.

We disable access to our services for users who repeatedly violate our policies—and, for egregious violations, for the first offense. We also terminate the Google accounts of entities on the U.S. State Department's Foreign Terrorist Organization (FTO) list, regardless of the content they are posting.

We also invest heavily in and promote counterspeech to present counternarratives and elevate voices that counter-extremism. For example, our Creators for Change program supports creators who are tackling difficult social issues, including extremism and hate, by building empathy among their influential audiences and acting as positive role models online. Similarly, Google's Jigsaw group, an incubator to tackle some of the toughest global security challenges, has deployed the Redirect Method. This Method uses Adwords targeting tools and third party curated YouTube videos uploaded to disrupt online radicalization.

Question 3. It seems like every week there is a new and more dangerous security breach. It was recently announced that YouTube would only be employing people, rather than relying on the newest technologies, like artificial intelligence, to combat terror-related content. Do you feel like this decision has the potential to open your companies content up to bad actors who do utilize next level technologies?

Answer. Technology will continue to be a part of the how YouTube enforces our policies and protects our services against bad actors. From our use of machine learning classifiers to detect violative content, to video-matching techniques that prevent known bad content from surfacing on the platform, to tools such as the Redirect Method, we will continue enhancing our systems to combat the evolving nature of the threat. We also understand the importance of using methods other than technology to combat terror-related content. 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 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. Late last year we added Ask.fm, Cloudinary, Instagram, Justpaste.it, LinkedIn, Oath, and Snap to the consortium. Industry collaboration is necessary and effective given counter-terrorism research that shows that many terrorist organizations engage in cross-platform abuse, and they especially migrate towards smaller and less-resourced platforms.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. RICHARD BLUMENTHAL TO JUNIPER DOWNS

Question 1. On October 30, 2017, nineteen civil rights groups, including Muslim Advocates, Leadership Conference on Civil and Human Rights, NAACP, Southern Poverty Law Center, and many others, co-signed a letter to Facebook to express concern about the hateful content on the social media platform used to divide the country, and in particular, to promote anti-Muslim, anti-Black, anti-immigrant, and anti-LGBTQ animus.

Ms. Downs, how would YouTube respond to this letter if it had received it?

Answer. At YouTube, we believe everyone deserves to have a voice, and that the world is a better place when we listen, share, and build community through our stories. Our values are based on four essential freedoms that define who we are:

- *Freedom of Expression:* We believe people should be able to speak freely, share opinions, foster open dialogue, and that creative freedom leads to new voices, formats and possibilities.
- *Freedom of Information:* We believe everyone should have easy, open access to information and that video is a powerful force for education, building understanding, and documenting world events, big and small.

- *Freedom of Opportunity:* We believe everyone should have a chance to be discovered, build a business and succeed..
- *Freedom to Belong:* We believe everyone should be able to find communities of support, break down barriers, transcend borders and come together around shared interests and passions.

Using YouTube to promote violence, incite hate, or celebrate violent extremism is not only strictly and specifically prohibited by our terms of service, but is antithetical to our mission. To this end, we believe that hate, 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.

We have worked to enhance not only our technology, but also to expand and deepen our expertise and resources on these issues. In 2017, we grew our Trusted Flagger Program by an additional 50 NGOs, experts on various types of hate and extremism.

Removing content that violates our guidelines or the law is an important part of the solution, but it's equally critical that we foster a better ecosystem for positive content and narratives against hate, extremism, and xenophobia. We are heavily invested in promoting counterspeech on our platform. We have hosted dozens of workshops around the world that teach the core skills needed to produce effective counterspeech. We also launched our flagship counterspeech program, Creators for Change, that is dedicated to amplifying and multiplying the voices of content creators who are tackling important social issues with their channels. YouTube just announced an additional $5M investment into this program.

As a company, we are dedicated to being a part of the solution and we will continue to invest in strategies that prevent the spread of hatred in all its forms.

Question 2. It's now well known how Russian agents used social media platforms to meddle in our elections—sow division and spread disinformation in the United States.

Ms. Downs, I understand that on YouTube, you do not necessarily need to be logged in to view content. However, many users—if not most—are logged in when they are viewing content nonetheless. Will you commit to proactively informing all of those identifiable users if they were victims of Russia's disinformation campaign—as Twitter and Facebook have already started to do? If you cannot, what percent of views on YouTube are anonymous?

Answer. We appreciate your work to promote transparency of these issues. The approximately 1,100 videos we identified as part of our investigation were removed when we disabled the accounts of these users. We have posted notice on the pages where those videos previously appeared, explaining they were removed due to violation of our Company's Terms of Service.

Question 3. Ms. Bickert, Mr. Monje, and Ms. Downs, please provide copies (including images, text, dates and timestamps) of all content identified by your platforms as generated by Russian agents or the Internet Research Agency.

Answer. We have conducted an extensive review of this issue and we provided both electronic and hardcopy versions of the ads associated with accounts we identified as connected to this effort to the Judiciary Committee. We identified limited activity on our platforms, but did identify two Ads accounts with approximately $4,700 of spend. In order to validate our findings, we broadly reviewed all political ads from June 2015 until the election last November that had even the loosest connection to Russia, which substantiated that we had identified the ads connected to this effort.

Our investigation is ongoing, we continue to request and receive leads from peers in our industry, and will be happy to continue cooperating with Congressional investigations on this topic.

Question 4. At least one of your peers in the tech industry has voluntarily initiated an outside assessment of the civil rights impacts of its policies and programs. In response to concerns regarding discrimination on the home-sharing platform, AirBNB hired former U.S. attorney general Eric Holder to help craft an anti-discrimination policy and has promised to pursue technological innovations to guard against future discriminatory events.

Mr. Monje, Ms. Bickert, and Ms. Downs, can you each commit to bringing in an independent entity to conduct a thorough and public audit of the civil rights impact of your policies and programs, including how your platform has been used by hate groups to stoke religious resentment and violence?

Answer. We understand that Airbnb hired outside lawyers to conduct a comprehensive review after Harvard University researchers published a *working paper* that found that users with perceived to be African-American were more likely to be

rejected by Airbnb hosts relative to guests perceived to be white. We applaud the seriousness that Airbnb took in addressing this issue and the alleged violations of civil rights and housing laws on its platform, including hiring Mr. Holder to draft an anti-discrimination policy.

We are committed to preventing the use of our products for unlawful activities, including the violation of civil rights laws. As a platform that hosts content, we deal with difficult questions around many issues, including hate speech, harassment, violence. But over the years, and often in consultation with outside organizations, lawyers, and experts, we have implemented and updated comprehensive policies to deal with these issues of unwanted content on our platforms. We have banned *hate speech* on YouTube and our other hosted platforms, and do not allow these platforms to be used for *harassment or cyberbullying.* And we recently *announced* a new YouTube policy that puts controversial and inflammatory videos behind an interstitial warning, where they will not be monetized, recommended or eligible for comments or user endorsements. These videos will have less engagement and be harder to find. We think this strikes the right balance between access to information without promoting extremely offensive viewpoints.

We have and will continue to engage with outside groups, lawyers, academics, non-governmental organizations and others to address and improve the content on our platforms.

Question 5. A little over a year ago, Facebook, Twitter, Google, and Microsoft announced a plan to create a joint industry database of "content that promotes terrorism." Mr. Monje, Ms. Bickert, and Ms. Downs, to what extent does this joint industry database focus on all forms of terror, including the real terror threat presented by white supremacists?

Answer. In December 2016, Facebook, Microsoft, Twitter and YouTube announced a coalition to launch a shared database of hashes of terrorist videos and images to prevent the spread of terrorist content between services. We expanded this partnership in July 2017, with the *launch* of the Global Internet Forum to Counter Terrorism (GIFCT). GIFCT's first meeting was held in August 2017, where representatives from the tech industry, government and non-governmental organizations came together to focus on three key areas: technological approaches, knowledge sharing, and research.

The GIFCT is committed to working on technological solutions to help thwart terrorists' use of our services, and has built on the groundwork laid by the EU Internet Forum, particularly through a shared industry hash database, where companies can create "digital fingerprints" for terrorist content and share it with participating companies. The database now contains more than 50,000 hashes. It allows member companies to use those hashes to identify and remove matching content—videos and images—that violate our respective policies or, in some cases, block terrorist content before it is even posted.

Each platform must make difficult decisions about how to balance issues surrounding free speech. YouTube has been working for years to combat extremist and hateful content on our platforms, and has long had policies that prohibit *terrorist content,* including terrorist recruitment, violent extremism, incitement to violence, and *instructional content* that could be used to facilitate substantial bodily injury or death.

On YouTube, 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.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. BRIAN SCHATZ TO JUNIPER DOWNS

Question 1. Please share a detailed update of YouTube's investigation into Russia's interference of the 2016 election via YouTube. Please include any other relevant investigations as they relate to other U.S. Federal or state elections.

Answer. We conducted an extensive review of this issue, which spanned across nearly twenty of our products, including YouTube.

As a result of this investigation, we found 18 YouTube channels likely associated with this campaign that made videos publicly available, in English, and with content that appeared to be political. There were 1,108 such videos uploaded, representing 43 hours of content and totaling 309,000 U.S. views from June 2015 to November 2016. These videos generally had very low view counts; only around 3 percent had more than 5,000 views. Upon confirmation with our internal systems validating the identity of these accounts, we suspended these YouTube channels at an account level.

Question 2. Google announced last year that it will identify the names of advertisers running election-related campaigns on Search, YouTube, and the Google Display Network within Election Ads. When will this be implemented? How will the ad onboarding process change for political advertisers on YouTube? How will political ads look to YouTube's users? Please provide mock-ups for both the onboarding process and users' view of the ad.

Answer. We've announced several measures to enhance transparency within election advertising, which we'll roll out prior to the November midterm elections:

- *Transparency Report.* We'll release a transparency report for election ads, where we'll share data about who is buying election-related ads on our platforms and how much money is being spent.
- *Creative Library.* We'll also introduce a publicly accessible library of election ads purchased on AdWords and YouTube (with information about who bought each ad).
- *In-ad disclosures.* We'll identify the names of advertisers running election-related campaigns on our platforms, including YouTube.
- *Verification program.* We'll reinforce our existing protections by requiring that advertisers proactively identify who they are and where they are based during the onboarding process before they can run any election-related ads. As they do, we'll verify that they are permitted to run U.S. election campaigns through our own checks.

We are in the process of implementing this change to our systems. While we do not have mock-ups we are able to share at this time, we are happy to provide information once we are closer to finalizing our work.

Question 3. Columbia University researcher Jonathan Albright has identified several series of AI-generated videos that consist of a slideshow and an auto-generated voice narrating an article. His research has focused on fake news and misinformation content. In one group of channels ("A Tease . . ."), a new video is created and uploaded every three minutes. Some of the videos from "A Tease . . ." channels appear to be taken down. What specific YouTube policy violation prompted this removal? Why did YouTube remove only a portion of these videos but not all? Does YouTube plan to take down the rest of them? How will YouTube prevent more of these from being uploaded in the future?

Answer. We have strict policies that prohibit spam, including posting post large amounts of untargeted, unwanted or repetitive content. We've developed proprietary technologies to help us fight spam, which are effective at capturing and preventing the widespread, and often automated, dissemination of low-quality information. If we detect large scale automated behavior around account creation, we can terminate the accounts. We also have policies against videos with *misleading titles and metadata.* Among other things, metadata added in an attempt to mislead viewers or game search algorithms will lead to the removal of videos and could lead to further action on the channel. In 2017, we removed over 130,000 videos for violation of this specific policy.

Question 4. Out of the total number of videos uploaded to YouTube per day, how many are generated primarily by computers, with little or no human input?

Answer. We have advanced systems to detect and terminate accounts that upload high volumes of spam, fraud, and other low-quality content. These systems are trained to analyze a variety of signals at account creation and at content upload. As our systems adapt, so do the behaviors of those who seek to abuse and game our systems. Given the evolving nature of the threat, our enforcement methods must and do evolve to respond to them. No matter what challenges emerge, our commitment to combat them will be sustained and unwavering. We're committed to getting this right and are increasing both human and engineering resources to tackle this ever-evolving landscape.

Question 5. Can you list the number of videos that violate YouTube Community Guidelines and the average amount of time YouTube takes to address the violation? Please break out by all of the violation categories currently listed today (1/30/2018).

Answer. We understand that people want a clearer view of how we're tackling problematic content. Our *Community Guidelines* give users notice about what we do not allow on our platforms and we want to share more information about how these are enforced. Over the next few months we will be creating a regular transparency report where we will provide more aggregate data about the flags we receive and the actions we take.

When it comes to violent extremist content, we have removed over 160,000 videos and terminated approximately 30,000 accounts since June 2017. Machine learning

is helping our human reviewers remove nearly five times as many videos than they were previously. Today, 98 percent of the videos we remove for violent extremism are detected by our machine-learning algorithms. Our advances in machine learning let us now take down nearly 70 percent of violent extremist content within eight hours of upload and nearly half of it in two hours, and we continue to accelerate that speed.

Question 6. For each of these categories, what percentage of videos are primarily generated by computers with little or no human input?

Answer. As noted in the answer above, we have advanced systems to detect and terminate accounts that upload high volumes of spam, fraud, and other low-quality content. We are continuously investing in improving these systems to counteract the ever-evolving nature of the threat.

Question 7. Once a video is taken down, does YouTube delete the data permanently, or does it just shield the content from public view? What is YouTube's data retention policy?

Answer. Videos removed for policy violations are generally retained for a period of time to allow for user appeal. We may also retain some data about videos for a period of time to comply with applicable laws. When a user deletes a video, we permanently remove the video and user-identifiable metadata associated with the video after a recovery window, which allows the user to recover the video in the case of accidental deletion.

Question 8. It was announced at the hearing that Google will employ 10,000 people this year to address content that might violate its policies. Will these employees be focused on any specific set of policies? Will they be full-time, permanent employees of Google? How many will be dedicated to YouTube? How many hours of content do you expect a single person to be responsible for per day (with or without the help of AI)?

Answer. These employees will primarily sit on the Trust & Safety teams across YouTube and Google, which work with our in-house legal and policy departments on escalations and also oversees vendors we hire to help us scale our operations. The new hires will consist of engineers and content reviewers, among others, who will work across Google to address content that violates any of our policies. Many of these reviewers will be dedicated solely to YouTube content moderation and they will be made up of a mix of full-time employees and contractors.

Question 9. In terms of dollars and percentage of annual revenue, how much is YouTube now spending on preventing foreign interference with our elections? What was the figure in the election cycle leading up to November 2016? What is the projected spend leading up to November 2018?

Answer. We've been tackling malicious actions directed at our users or services, including those originating from government-backed actors, since long before the 2016 elections. For more than a decade, we've offered our Safe Browsing tool, which helps protect users from phishing, malware, or other attacks; today it is used on more than three billion devices worldwide. Additionally, when we detect that a user's account has been targeted by a government-backed attacker, we show a warning that includes proactive steps the user can take to increase the security of his or her account.

Our existing advertising safeguards include policies that prohibit foreign nationals from buying U.S. election ads. In 2016, we tightly restricted which advertisers can serve ads to audiences based on their political leanings. (We offered two categories—left-leaning and right-leaning—starting in August 2016.) And we're continuing to invest in enhancements to our safeguards, which will include a transparency report and creative library that feature election ads bought using our front-end advertising systems, including those appearing on YouTube.

Election-related advertising is not a large business for Google relative to other advertising verticals; nevertheless, we understand the importance of election advertising and are making investments in launching transparency tools in 2018 that are industry-leading. Across the world, we have a global team of thousands of policy experts, reviewers, product managers, and data scientists focused on creating, maintaining, and enforcing our policies, including those related to these issues. It is difficult to estimate, however, the overall dollar and revenue-percentage value of these investments as many of the resources we're leveraging are widely shared among numerous technical infrastructural and product development teams.

Question 10. Congress will judge success not by YouTube's efforts but by its results. How will YouTube measure success? Will YouTube be conducting an audit after November 2018? When will results be shared?

Answer. We found very limited activity connected to this effort on our platform in our investigation into potential election interference with respect to the 2016

Presidential election, however we are committed to preventing the misuse of our platforms for the purpose of interfering with democratic elections. At the end of last year, we announced we would take the *following steps:*

- *Transparency Report.* In 2018, we'll release a transparency report for election ads, which will share data about who is buying election-related ads on our platforms and how much money is being spent.
- *Creative Library.* We'll also introduce a publicly accessible database of election ads purchased on AdWords and YouTube (with information about who bought each ad). That means people will not only be able to learn more about who's buying election-related ads on our platforms; they'll be able to see the ads themselves, regardless of to whom they were shown.
- *In-ad disclosures.* Going forward, we'll identify the names of advertisers running election-related campaigns on Search, YouTube, and the Google Display Network via our "Why This Ad" icon.
- *Verification program.* U.S. law restricts entities outside the U.S. from running election-related ads. We'll reinforce our existing protections by requiring that advertisers proactively identify who they are and where they are based before running any election-related ads.

We also have increased our longstanding support to non-profits and journalists dedicated to ensuring the integrity of our election systems. For example, we've recently contributed nearly $750,000 to the bipartisan "Defending Digital Democracy" project, led by the Belfer Center for Science and International Affairs at Harvard Kennedy School. And we are deeply committed to helping people participate in the election by providing users with timely and comprehensive information they need to make their voice heard.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. TAMMY BALDWIN TO JUNIPER DOWNS

Question 1. In the context of extremist content, I would like to learn more about each company's policy for proactively reporting users to law enforcement. I understand your companies evaluate and respond to law enforcement requests for information, but what framework do you use to proactively report terrorist-related content to authorities, including any identifying information of the user? For example, if you use a standard of imminent harm, how do you define and apply it, particularly in a threat environment where terrorist organizations often call on recruits to attack carefully planned targets of opportunity, rather than to launch an immediate, indiscriminate attack?

Answer. Google discloses information to government entities when the threat of loss of life or serious physical injury is brought to our attention by governmental entities, and when we learn of the threat ourselves or from other sources. Emergency disclosures are handled twenty-four hours a day, every day of the year and consistent with the law.

Evaluating whether there is a credible threat presented in any particular case requires evaluation of the facts as we know them at the time. This can include looking to information on Google's platform as well as from other sources. Public sources of information can also assist in making the determination as well as having an understanding of the threat environment more generally.

Question 2. I would like to hear from the companies whether they support implementing Mr. Watts's recommendations to: first, fully certify the authenticity of all users—in other words, ensure that each user is a real person; and second, eliminate social bot networks to reduce automated broadcasting of disinformation.

Answer. We have not seen the same degree of social media bots that have been reported on other platforms. Our systems rely on a host of inputs about historical use and pattern recognition across various services in an effort to detect if an account creation or login is likely to be abusive. The system operates to block "bad" account creation or to close groups of such accounts. We prevent users from creating a large number of Google Accounts in a short time period if our systems detect that the user might be abusive. If we detect suspicious conduct, we also require verification, aimed at detecting if a bot is attempting to access or create an account. We have also developed robust protections over the years to address attempts to manipulate our systems by bots or other schemes, like link farms. (Our webmaster guidelines provide more information about this: *https://support.google.com/webmasters/answer/35769.*) We use both algorithmic and manual methods, and we deploy these across our products including Search and YouTube.

Question 3. What are the indicators that you use to identify a Russian disinformation account, whether from the Kremlin's so-called Internet Research Agency or an associated group of hackers or trolls, and what thresholds must be met to disable an account?

Answer. We developed a list of actors we know or suspect were involved in this effort from (1) our research of publicly available information, (2) the work of our security team, and (3) leads we received from others in the industry and Jigsaw. In addition, we reviewed all ads from June 2015 until the election in November 2016 that were categorized as potentially political by our systems and had even the loosest connection to Russia, such as a Russian I.P. address or billing address or were paid for using Russian currency. When we identified accounts that we believed were associated with this effort, we removed them.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. CATHERINE CORTEZ MASTO TO JUNIPER DOWNS

Question 1. During my time as the Attorney General for the State of Nevada, I saw too many instances of sex trafficking cases involving child victims that were dismissed because the conduct occurred online or through social media. So that's why I'm a strong supporter of the Stop Enabling Sex Traffickers Act of 2017 (SESTA), which clarifies the Communications Decency Act (CDA) to allow state Attorneys General to retain their jurisdiction to prosecute those who facilitate human trafficking. We know that trafficking is happening online and on social media, and SESTA is the only current legislative proposal that provides sufficient deterrence to traffickers by providing the necessary tools for successful prosecutions. As a former prosecutor, I know what it will take to successfully prosecute those who engage in sex trafficking through social media and other websites, and that's why I believe that the House version of SESTA doesn't go far enough to give prosecutors the tools they need to protect sex trafficking victims. I hope that your organizations all agree that victims of sex trafficking deserve meaningful protections and justice.

If so, I'd like to hear whether you will continue to support SESTA over the weaker U.S. House version of the bill.

Answer. This is a very important issue and we are glad that Congress has taken it up. There's no question that Backpage is a bad actor, and we strongly agree that they and other bad actors should be prosecuted for their crimes under the full extent of the law. We firmly support both increased enforcement of existing laws against sex trafficking, and strengthening of those laws where appropriate to combat this heinous activity. As you know, we are members of the Internet Association, which has endorsed both the House and Senate legislation. We will continue to work with any lawmakers interested in addressing this important problem, while ensuring that these efforts do not inadvertently hinder the ability of websites to remove illegal content and to fight sex trafficking.

Question 2. I was glad to hear that the Internet Association supports SESTA, and I'd like to know what else your organization is doing to address concerns about sex trafficking occurring on your platforms and helping us pass this important legislation in the Senate?

Answer. Google has made it a priority to tackle the horrific crime of sex trafficking. We have donated over $20 million to organizations on the front lines of ending modern day slavery and human trafficking. We have developed and built extensive technology to connect victims with the resources they need. And we have helped pioneer the use of technologies that identify trafficking networks to make it easier and quicker for law enforcement to arrest these abusers.

Google has a zero tolerance policy for any advertising related to sex trafficking and prostitution. We work tirelessly to ensure bad actors are not able to exploit our products and use an industry-leading combination of sophisticated technology and manual review to detect and remove bad ads. We've invested millions of dollars in building these systems to scale. We also work with law enforcement and NGOs that are focused on stopping human trafficking.

Additionally, we have always had strict policies against child endangerment, and we partner closely with regional authorities and experts to help us enforce these policies and report to law enforcement through the National Center for Missing and Exploited Children (NCMEC). Key principles in Section 230 of the Communications Decency Act enable us to develop new strategies around user flags, machine learning, and automated enforcement to stop sex trafficking and other forms of child endangerment on our services.

We will continue to work to prevent this type of activity from occurring on our platform, and we welcome efforts by Congress to help us do so.

Question 3. Over the past few months, our country has been reckoning with some hard truths about the way that women and minorities are treated in the workplace. And I think this is a moment for all types of organizations, including tech giants like the ones represented here, to take a clear-eyed accounting of their culture and practices, to take responsibility for what hasn't worked, and to renew their commitments to make meaningful improvements. The Equal Employment Opportunity Commission's 2016 report on "Diversity in High Tech" found that women, African Americans, and Hispanics are all represented at significantly lower levels in high tech than in private industry as a whole. And while recent internal studies at Facebook and Google have showed some progress in the hiring of women, there has not been equal improvement in the representation of people of color and other underrepresented groups.

What technically qualifies as diversity to your organization?

Answer. At Google we take a broad and intersectional view of diversity—from race and gender, to sexual orientation, gender identity, age, disability, socio-economic background and more. That said, we are actively working to improve the representation of underrepresented groups in our workforce.

We believe deeply in diversity and inclusion and it underpins Google's business. We believe that if we tap the full range of human experience, capability and contribution, we will move faster, increase innovation and creativity, and can tackle more and more of the world's problems. Increasing diversity not only makes good business sense, it's also the right thing to do.

Google's core mission is to organize the world's information and make it universally accessible and useful. Our goal for diversity is a natural extension of this mission—to increase access to opportunity, by breaking down barriers and empowering people through technology. Products will only get better and more useful if we invite all segments of society, and people from all over the world, to influence and create technology.

Question 4. How is your company working to address issues of discrimination in your own workforces?

Answer. We've worked hard over many years to create and foster a fair and inclusive Google, and we absolutely do not tolerate discrimination, or any actions that create a hostile work environment.

This work starts with creating and building a fair and inclusive culture, that keeps discrimination from happening in the first place.

Once Googlers get here, we strive to ensure that our work environment is fair and inclusive, so they can grow and flourish. And we keep close tabs on our programs and process to ensure they yield fair and equitable outcomes. When we learn that something is amiss, we take action early and follow up to make sure we continue to create a great place to work.

We have a culture that empowers Googlers to quickly raise up issues of concern and where they feel they've been mistreated, or discrimination has happened. If Googlers are concerned by any inappropriate behavior they experience, or see, in the workplace, we ask them to please report it.

There are many avenues to do this, including by anonymous means, and we review each complaint. We do a lot to make Googlers aware of those channels—ranging from reaching out directly to HR to a third-party helpline if anyone wants to stay anonymous.

Question 5. Do you believe those efforts are sufficient?

Answer. Google's approach to our products and our business is that we can always do better, and our approach to diversity and inclusion, and how we oppose discrimination, is no different. We have always been transparent about our commitment to diversity, inclusion, equity and compliance in our workforce. More importantly, we're also transparent about our challenges and key learnings in this arena. We believe the best way to be a positive example is by making progress in our own workforce and culture, and that's a key priority of ours.

Question 6. I've seen that Facebook works to make their labor diversity information public, can you provide a status on your labor figures, or commit to sharing those with the Committee and the public?

Answer. As you may know, Google was the first large tech company to publish workforce diversity data in 2014 (all of this information is available at google.com/diversity). We are committed to sharing our numbers every year, and 2018 is no different. We plan on once again releasing those numbers once again later this year.

As stated earlier, we believe it is important to be transparent about our challenges and key learnings in this arena. Our original decision to release our workforce (diversity) numbers led to other companies following suit. Google stands firm in its commitment to foster dialogue and to drive impact on this important issue.

Question 7. We know that so-called talent pipelines are not the only obstacle to achieving a diverse workforce, and that discrimination and harassment go hand in hand, distorting the operation of workplace meritocracies. This is a moment when many victims of sexual assault and harassment are bravely coming forward about their experiences, allowing us to get a better sense of the true scope and effects of this behavior. Persistent harassment, and the workplace culture that tolerates, ignores, or even encourages such harassment, pushes people out of their workplaces, stalls or derails promising careers, and discourages some from pursuing certain opportunities altogether.

What is your company doing to evaluate the impact of harassment in your workforces? How are you working to create a culture where harassment is no longer tolerated?

Answer. Google has clear policies on appropriate behavior by employees of the company. Harassment has never been tolerated. We review all concerns, and take action when necessary. Here's a snapshot of what some of this work has looked like over the past few years:

- In 2015, we launched the Respect@ program as a way for Googlers to raise concerns, share experiences and get support. The program is a way for Googlers to learn about the standards of behavior we expect, the different ways for reporting unacceptable behavior, and the process we undertake to investigate complaints.
  - Respect@ is supported by the senior-most leadership of the company, and is championed by an Executive Oversight Committee that currently consists of 18 VPs, across functions and regions.
  - As part of Respect@, we also created go/saysomething, an internal online resource to provide a discreet way to report inappropriate behavior. If Googlers are concerned by any inappropriate behavior they experience, or see, in the workplace, we ask them to please report it. There are many avenues to do this, including by anonymous means, and we review and investigate each complaint.
  - Since launching Respect@, complaints/investigations have increased significantly, as people became aware of the avenues open to them
- In 2015, we launched the first annual Internal Investigation Report, so that Googlers can see the number and type of complaints we receive, as well as the outcomes of these complaints. We have published these reports every year since.

Question 8. What more could you be doing to be a positive example for other companies and Industries?

Answer. Google has always been transparent about its commitment to diversity, inclusion, equity and compliance in our workforce. More importantly, we're also transparent about our challenges and key learnings in this arena. We believe the best way to be a positive example is by making progress in our own workforce and culture, and that's a key priority of ours.

Question 9. Last October, Facebook announced that it would be improving transparency for all ads run on its platform, including by requiring political advertisers to include a disclaimer telling viewers who paid for an ad, and allowing viewers to see all the ads a page is running, even those that aren't targeting them. Twitter also announced similar measures. Although these policies were announced in response to Russia using social media to interfere in our elections, it seems these transparency measures could help shine a spotlight on other forms of influence campaigns by extremists or terrorists.

Can you provide an update on the status of any improvements YouTube is making?

Answer. At the end of last year, we announced we would take the *following steps across Google's platforms, including on YouTube:*

- *Transparency Report.* In 2018, we'll release a transparency report for election ads, which will share data about who is buying election-related ads on our platforms and how much money is being spent.
- *Creative Library.* We'll also introduce a publicly accessible database of election ads purchased on AdWords and YouTube (with information about who bought each ad). That means people will not only be able to learn more about who's buying election-related ads on our platforms; they'll be able to see the ads themselves, regardless of to whom they were shown.

- *In-ad disclosures.* Going forward, we'll identify the names of advertisers running election-related campaigns on Search, YouTube, and the Google Display Network via our "Why This Ad" icon.
- *Verification program.* U.S. law restricts entities outside the U.S. from running election-related ads. We'll reinforce our existing protections by requiring that advertisers proactively identify who they are and where they are based before running any election-related ads.

Question 10. If applicable, when can we expect to see them fully implemented?

Answer. We are in the process of implementing these steps and plan to have them completed in time to be helpful to users in understanding how ads are purchased during the November elections.

Question 11. If applicable, how are you defining what constitutes a political ad subject to these heightened transparency requirements?

Answer. We will apply the new requirements to political advertisements that either constitute "express advocacy" or contain a reference to a clearly identified candidate, as each of those terms is defined by the Federal Election Commission.

Question 12. On January 29, the Director of the Central Intelligence Agency said he expects the Russian government to attempt to influence the 2018 elections in this country.

What efforts is YouTube undertaking in the lead up to the 2018 elections to identify and close the platform's remaining vulnerabilities to foreign exploitation?

Answer. As mentioned above, at the end of last year, we announced we would take the *following steps* across Google's platforms, including on YouTube:

- *Transparency Report.* In 2018, we'll release a transparency report for election ads, which will share data about who is buying election-related ads on our platforms and how much money is being spent.
- *Creative Library.* We'll also introduce a publicly accessible database of election ads purchased on AdWords and YouTube (with information about who bought each ad). That means people will not only be able to learn more about who's buying election-related ads on our platforms; they'll be able to see the ads themselves, regardless of to whom they were shown.
- *In-ad disclosures.* Going forward, we'll identify the names of advertisers running election-related campaigns on Search, YouTube, and the Google Display Network via our "Why This Ad" icon.
- *Verification program.* U.S. law restricts entities outside the U.S. from running election-related ads. We'll reinforce our existing protections by requiring that advertisers proactively identify who they are and where they are based before running any election-related ads.

We also have increased our longstanding support to non-profits and journalists dedicated to ensuring the integrity of our election systems. For example, we've recently contributed nearly $750,000 to the bipartisan "Defending Digital Democracy" project, led by the Belfer Center for Science and International Affairs at Harvard Kennedy School. And we are deeply committed to helping people participate in the election by providing users with timely and comprehensive information they need to make their voice heard.

Question 13. What assistance can Federal, state and local government entities provide in that effort?

Answer. We are committed to working with Congress, law enforcement, others in our industry, and the NGO community to strengthen protections around elections, ensure the security of users, and help combat disinformation, and we welcome tips and other information from knowledgeable government agencies on these matters.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. ROGER WICKER TO CARLOS MONJE, JR.

Question 1. Has Twitter placed any restrictions on the U.S. Government's use of publicly available information on your platform? If yes, please describe what those restrictions are, why they have been imposed and on which U.S. Government agencies?

Answer. Twitter is a public platform. When users choose to share information by posting it to their public profile, the information is available to anyone who visits those users' profiles. With respect to Twitter's application programming interface ("API") (public and commercial), through which we provide developers and other third parties access to subsets of public Twitter content, all users of our developer

products must comply with Twitter's developer terms and policies. Those policies include long-standing provisions that prohibit, among other things, the use of Twitter data for surveillance purposes or for purposes in contravention of the Universal Declaration of Human Rights. While Twitter works closely with its developer community to address questions and investigate instances of potential abuse, each developer is responsible for compliance with Twitter's applicable policies.

Twitter maintains strong working relationships with law enforcement. We publish guidelines for law enforcement personnel that explain our policies and the process for submitting requests for information. We regularly respond to law enforcement requests, have a dedicated 24/7 response team for that purpose, and have developed a user-friendly online submission form to streamline responses to law enforcement agencies through properly scoped valid legal process. There are also a number of news alert products that are available and used by law enforcement, including the Federal Bureau of Investigation.

Question 2. Are U.S. Government agencies or intelligence organizations permitted to search for or monitor—either directly with Twitter or through third-party aggregators—counterterrorism information or specific Twitter accounts that are likely affiliated with terrorist organizations within the publicly available content found on Twitter's platform? If not, why? Please explain.

Answer. The answer to Question 2 has been provided in response to Question 1.

Question 3. Are companies (such as casinos) allowed to monitor—either directly with Twitter or through third-party aggregators—specific Twitter accounts that have made public threats against their venues or staff?

Answer. The answer to Question 2 has been provided in response to Question 1.

Question 4. Does Twitter have any policies that prohibit the use of its data, by any public or private third-party, for counterterrorism analyses focused on terrorist organizations? If so, can you please explain the purpose of that policy and the parameters of it?

Answer. The answer to Question 2 has been provided in response to Question 1.

Question 5. During the hearing, Mr. Monje testified that Twitter works with law enforcement through the "proper legal process". Please describe the legal process to which Mr. Monje was referring and how it applies to law enforcement's use of Twitter's aggregate user data.

Answer. As we noted above in response to Question 1, Twitter maintains strong working relationships with law enforcement. We publish guidelines for law enforcement personnel that explain our policies and the process for submitting requests for information. See *https://help.twitter.com/en/rules-and-policies/twitter-law-enforcement-support.* We regularly respond to law enforcement requests, have a dedicated 24/7 response team for that purpose, and have developed a user-friendly online submission form to streamline responses to law enforcement agencies through valid legal process. Before launching this system to all U.S. law enforcement agencies, we conducted a pilot with the Federal Bureau of Investigation. We have begun rolling out this tool for global use.

In addition, we have offered and conducted training sessions to law enforcement officials to familiarize them with our policies and procedures. In 2017, we have attended and provided training at a national conference for investigators of crimes against children, training events for FBI legal attachés posted to U.S. embassies abroad, and other conferences with the participation of federal, state and local law enforcement. We continue to build upon and invest in our law enforcement outreach and training. And we welcome feedback from law enforcement experts and professionals about how we can improve our systems.

We regularly and directly engage with law enforcement officials on a wide range of issues, including extremist content online. We receive and respond to "Internet Referral Unit" reports of extremist content. Our recently published Transparency Report for the first half of 2017 details the statistics of those responses. See *https://blog.twitter.com/official/en_us/topics/company/2017/New-Data-Insights-Twitters-Latest-Transparency-Report.html.* In addition, we receive briefings from government experts on terrorist use of online platforms, which help inform our proactive efforts.

Law enforcement requests to Twitter must comply with applicable laws in the jurisdiction where they are issued. For Federal law enforcement, this includes the Electronic Communications Privacy Act 18 U.S.C. 2510 et seq. These requirements only apply to data sought from Twitter. If law enforcement is able to access publicly available information from the Twitter service they may do so subject to any other legal or policy restrictions that may apply to their conduct (*e.g.,* Department of Justice guidance). If law enforcement seeks access to Twitter data via our API directly or through a third party developer, they must do so in a manner that complies with the applicable Twitter terms and policies for our API.

Question 6. Does a U.S. Government agency have to obtain a warrant (or go through a similar legal process as discussed in Question #5) to search publicly available information found on Twitter? If yes, why? If no, does Twitter allow U.S. Government agencies to gain access to publicly available information on its platform through third-parties that have purchased aggregate user data from Twitter?

Answer. The answer to Question 6 has been provided in response to Question 1.

Question 7. Twitter's platform allows users to "follow" other users. In your view, what is the difference between "following" someone and "surveilling" someone?

Answer. A user may follow another account holder via the Twitter service. An account holder may view their list of followers at any time. The account holder may take a range of actions in response to receiving a "follow" from another user. They may decide to follow that user in return. They may also choose to block that follower, preventing the follower from viewing in their timeline content posted that account holder or receiving notifications of posts by the account holder Twitter users may also choose to make their account private so as to restrict new followers to those that they have expressly allowed as followers. These choices, including the ability to restrict followers and the transparency inherent in our platform, are important aspects of the Twitter service.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. JERRY MORAN TO CARLOS MONJE, JR.

Question 1. How is the battle against terrorist content different from misinformation and abuse?

Answer. Twitter is committed to combating terrorist content, abusive activity (including malicious automated activity) and the spread of misinformation on our platform. Because bad actors often rely on the same methods to propagate such content, to some extent, we deploy similar tools to detect and stop all malicious activity on Twitter, including the proliferation of terrorist content. At the same time, we recognize that each of these areas presents unique challenges to deploying our technology at scale. For example, with terrorist content, we can more readily identify the signals of bad actors intending to disseminate terrorist propaganda and can efficiently find and suspend many of these accounts using machine learning. In contrast, for abuse and misinformation, the context of conversations and the content itself are often needed to determine whether something crosses a policy line.

However, Twitter's approach to addressing the spread of malicious automation and inauthentic accounts on our platform is to focus wherever possible on identifying problematic behavior, and not on the content itself. Those who are seeking to influence a wide audience often find ways to try to artificially amplify their messages across Twitter. As with spam, these behaviors frequently provide more precise signals than focusing on content alone.

Accordingly, we monitor various behavioral signals related to the frequency and timing of Tweets, Retweets, likes, and other such activity, as well as to similarity in behavioral patterns across accounts, in order to identify accounts that are likely to be maliciously automated or acting in an automated and coordinated fashion in ways that are unwelcome to our users. We monitor and review unsolicited targeting of accounts, including accounts that mention or follow other accounts with which they have had no prior engagement. For example, if an account follows 1,000 users within the period of one hour, or mentions 1,000 accounts within a short period of time, our systems are capable of detecting that activity as aberrant and as potentially originating from suspicious accounts.

Twitter is continuing its effort to detect and prevent malicious automation by leveraging our technological capabilities and investing in initiatives aimed at understanding and addressing behavioral patterns associated with such accounts. For example, in early 2017, we launched the Information Quality initiative, an effort aimed at enhancing the strategies we use to detect and stop bad automation, improve machine learning to spot spam, and increase the precision of our tools designed to prevent such content from contaminating our platform.

In 2017, we have made significant improvements to reduce external attempts to manipulate content visibility. These improvements were driven by investments in methods to detect malicious automation through abuse of our API, limit the ability of malicious actors to create new accounts in bulk, detect coordinated malicious activity across clusters of accounts, and better enforce policies against abusive third-party applications.

In addition, we have developed new techniques for identifying patterns of activity inconsistent with legitimate use of our platform (such as near-instantaneous replies to Tweets, nonrandom Tweet timing, and coordinated engagement), and we are cur-

rently implementing these detections across our platform. We have improved our phone verification process and introduced new challenges, including reCAPTCHA (utilizing an advanced risk-analysis engine developed by Google), to give us additional tools to validate that a human is in control of an account. We have enhanced our capabilities to link together accounts that were formed by the same person or that are working in concert. And we are improving how we detect when accounts may have been hacked or compromised.

With our improved capabilities, we are now detecting and blocking approximately 523,000 suspicious logins each day that we believe to be generated through automation. In December 2017, our systems identified and challenged more than 6.4 million suspicious accounts globally per week—a 60 percent increase in our detection rate from October 2017. Over three million of those accounts were challenged upon signup, before their content or engagements could impact other users. Since June 2017, we also suspended more than 220,000 malicious applications for API abuse. These applications were collectively responsible for more than 2.2 billion Tweets in 2017. We plan to continue building upon our 2017 improvements, including through collaboration with our peers and investments in machine-learning capabilities that help us detect and mitigate the effect on users of fake, coordinated, and malicious automated account activity.

We have also observed the expansion of malicious activity on our platform from automated accounts to human-coordinated activity, which poses additional challenges to making our platform safe. We are determined to meet those challenges and have been successful in addressing such abusive behavior in other contexts. We are committed to leveraging our technological capabilities in order to do so again by carefully refining and building tools that respond to signals in the account behavior.

Those tools have also been successful at detecting and removing terrorist content on our platform. For example, as of September 2017, 95 percent of account suspensions for promotion of terrorist activity were accomplished using our existing proprietary detection tools—up from 74 percent in 2016. These tools focus on indicia of violating activity beyond the content of the Tweet. Although they have proved successful, our efforts to address terrorist content on our platform do not end with investments in our proprietary detection tools. We recognize that the spread of terrorist and extremist content online is not unique to Twitter, and we are committed to collaborating with our industry peers to address this shared thread. Accordingly, in June 2017, we launched the Global Internet Forum to Counter Terrorism (the "GIFCT"), a partnership among Twitter, YouTube, Facebook, and Microsoft. The GIFCT facilitates, among other things, information sharing; technical cooperation; and research collaboration, including with academic institutions. In September 2017, the members of the GIFCT announced a multimillion dollar commitment to support research on terrorist abuse of the Internet, and how governments, tech companies, and civil society can respond effectively. We are looking to establish a network of experts that can develop these platform-agnostic research questions and analysis that considers a range of geopolitical contexts. The GIFCT opened a call for proposals in December 2017, and we look forward to sharing further details of the initial projects this year.

The GIFCT has created a shared industry database of "hashes"—unique digital "fingerprints"—for violent terrorist imagery or terrorist recruitment videos or images that have been removed from our individual services. The database allows a company that discovers terrorist content on one of its sites to create a digital fingerprint and share it with the other companies in the forum, who can then use those hashes to identify such content on their services or platforms, review against their respective policies and individual rules, and remove matching content as appropriate, or even block extremist content before it is posted in the first place. The database now contains more than 40,000 hashes. Instagram, Justpaste.it, LinkedIn, Oath, and Snap have also joined this initiative, and we are working to add several additional companies in 2018. Twitter also participates in the Technology Coalition, which shares images to counter child abuse.

As part of our work with the GIFCT, we have hosted more than 50 small companies at workshops through the Tech Against Terrorism initiative, our partners under the UN CounterTerrorism Executive Directorate. Twitter believes that this partnership will provide a unique opportunity for us to share our knowledge and technical expertise with smaller and emerging companies in the industry and for all industry actors to harness the expertise that has been built up in recent years.

We have also focused on NGO outreach and, since 2013, and have participated in more than 100 Countering Violent Extremism training and events around the world, including in Beirut, Bosnia, Belfast and Brussels, and summits at the White House, at the United Nations, London, and Sydney. Twitter has partnered with

groups like the Institute of Strategic Dialogue, the Anti-Defamation League and Imams Online to bolster counterspeech that offers alternatives to radicalization. As a result of that work, NGOs and activists around the world are able to harness the power of our platform in order to offer positive alternative narratives to those at risk and their wider communities.

Finally, in addressing abuse directed at users on the platform, context matters. A turn of phrase can be playful or offensive, depending on the circumstance, topic, and author. This means we need more nuanced and creative approaches to our machine learning models in order to address abusive activity at scale. One example where we have made progress is in our improving ability to action reports of abuse by witnesses (instead of by victim directly). By looking at various signals, including the relationship and activity between the reported abuser and reported victim, we can better identify, escalate, and take action against instances of abuse.

Question 2. Can you walk through your track record of removing terrorist content?

Answer. As noted above, we have made considerable inroads against the proliferation of terrorist content on our platform. For example, in February 2016 when we first started sharing metrics for our enforcement efforts, we announced that, since the middle of the preceding year, we had suspended more than 125,000 accounts for threatening or promoting terrorist acts. See *https://blog.twitter.com/official/en_us/a/2016/combating-violent-extremism.html.* By August 2016, we announced that we had suspended an additional 235,000 accounts for violating Twitter policies related to the promotion of terrorism. *https://blog.twitter.com/official/en_us/a/2016/anupdate-on-our-efforts-to-combat-violent-extremism.html.* We also announced at that time that our daily suspension records increased by more than 80 percent compared to the previous year, and that our response time for suspending reported accounts decreased dramatically.

We made additional improvements the following year. As we noted in our September 2017 Transparency Report, for the reporting period between January 1 and June 30, 2017, we suspended nearly 300,000 accounts for violations of Twitter policies prohibiting the promotion of terrorism. Of those suspensions, 95 percent were accomplished using our proprietary tools—up from 74 percent in 2016. Approximately 75 percent of those accounts were suspended before posting their first Tweet. In total, between August 1, 2015 and June 30, 2017, we suspended nearly 1 million accounts for violating Twitter rules and policies prohibiting the promotion of violence or terrorist content.

Question 3. What is the next challenge on the Common Vulnerabilities and Exposures (CVE) front? How do we empower smaller platforms that the terrorists are moving to?

Answer. As noted above, we plan to continue building upon our 2017 improvements, including through collaboration with our peers and investments in machine-learning capabilities that help us detect and mitigate the effect on users of fake, coordinated, and malicious automated account activity.

We have observed the expansion of malicious activity on our platform from automated accounts to human-coordinated activity, which poses additional challenges to making our platform safe. We are determined to meet those challenges and have been successful in addressing such abusive behavior in other contexts. We are committed to leveraging our technological capabilities in order to do so again by carefully refining and building tools that respond to signals in the account behavior.

Those tools have also been successful at detecting and removing terrorist content on our platform. For example, as of September 2017, 95 percent of account suspensions for promotion of terrorist activity were accomplished using our existing proprietary detection tools—up from 74 percent in 2016. These tools focus on indicia of violating activity beyond the content of the Tweet.

Although they have proved successful, our efforts to address terrorist content on our platform do not end with investments in our proprietary detection tools. We recognize that the spread of terrorist and extremist content online is not unique to Twitter, and we are committed to collaborating with our industry peers to address this shared thread. Accordingly, in June 2017, we launched the Global Internet Forum to Counter Terrorism (the "GIFCT"), a partnership among Twitter, YouTube, Facebook, and Microsoft. The GIFCT facilitates, among other things, information sharing; technical cooperation; and research collaboration, including with academic institutions. In September 2017, the members of the GIFCT announced a multi-million dollar commitment to support research on terrorist abuse of the Internet and how governments, tech companies, and civil society can respond effectively. We are looking to establish a network of experts that can develop these platform-agnostic research questions and analysis that considers a range of geopolitical contexts. The

GIFCT opened a call for proposals in December 2017, and we look forward to sharing further details of the initial projects this year.

The GIFCT has created a shared industry database of "hashes"—unique digital "fingerprints"—for violent terrorist imagery or terrorist recruitment videos or images that have been removed from our individual services. The database allows a company that discovers terrorist content on one of its sites to create a digital fingerprint and share it with the other companies in the forum, who can then use those hashes to identify such content on their services or platforms, review against their respective policies and individual rules, and remove matching content as appropriate, or even block extremist content before it is posted in the first place. The database now contains more than 40,000 hashes. Instagram, Justpaste.it, LinkedIn, Oath, and Snap have also joined this initiative, and we are working to add several additional companies in 2018.

As part of our work with the GIFCT, we have hosted more than 50 small companies at workshops through the Tech Against Terrorism initiative, our partners under the UN CounterTerrorism Executive Directorate. Twitter believes that this partnership will provide a unique opportunity for us to share our knowledge and technical expertise with smaller and emerging companies in the industry and for all industry actors to harness the expertise that has been built up in recent years.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. RON JOHNSON TO CARLOS MONJE, JR.

Question 1. Social media companies are increasingly able to remove terrorist recruitment, incitement, and training materials before it posts to their platforms by relying on improved automated systems. Other than content removal, what else can be done to limit the audience or distribution of these dangerous materials?

Answer. Twitter has been at the forefront of developing a comprehensive response to the evolving challenge of preventing terrorist exploitation of the Internet. We initially focused on scaling up our own, in-house proprietary spam technology to detect and remove accounts that promote terrorism. In early 2016, the technological tools we had at our disposal detected approximately one-third of terrorism-related accounts that we removed at that time. In 2017, 95 percent of account suspensions for promotion of terrorist activity were accomplished using our existing proprietary detection tools—up from 74 percent in 2016. Approximately 75 percent of those accounts were suspended prior to sending their first Tweet. In total, since 2015, we have suspended nearly a million accounts that we determined violated our terms of service. In December 2016, for example, we took steps toward a hash-sharing agreement with Facebook, Microsoft, and YouTube, intended to further curb the spread of terrorist content online. Pursuant to this agreement, the four companies created an industry database of "hashes"—unique digital "fingerprints"—for violent terrorist imagery or terrorist recruitment videos or images that we have removed from our services. By sharing this information with each other, we may use the shared hashes to help identify potential terrorist content on our respective hosted consumer platforms.

In June 2017, we launched the Global Internet Forum to Counter Terrorism (the "GIFCT"), a partnership among Twitter, YouTube, Facebook, and Microsoft. The GIFCT facilitates, among other things, information sharing; technical cooperation; and research collaboration, including with academic institutions.

In September 2017, the members of the GIFCT announced a multimillion dollar commitment to support research on terrorist abuse of the Internet and how governments, tech companies, and civil society can respond effectively. We are looking to establish a network of experts that can develop these platform-agnostic research questions and analysis that considers a range of geopolitical contexts. The GIFCT opened a call for proposals last month, and we look forward to sharing further details of the initial projects early in 2018.

The GIFCT has created a shared industry database of "hashes"—unique digital "fingerprints"—for violent terrorist imagery or terrorist recruitment videos or images that have been removed from our individual services. The database allows a company that discovers terrorist content on one of its sites to create a digital fingerprint and share it with the other companies in the forum, who can then use those hashes to identify such content on their services or platforms, review against their respective policies and individual rules, and remove matching content as appropriate, or even block extremist content before it is posted in the first place. The database now contains more than 40,000 hashes. Instagram, Justpaste.it, LinkedIn, Oath, and Snap have also joined this initiative, and we are working to add several

additional companies in 2018. Twitter also participates in the Technology Coalition, which shares images to counter child abuse.

As part of our work with the GIFCT, we have hosted more than 50 small companies at workshops through the Tech Against Terrorism initiative, our partners under the UN CounterTerrorism Executive Directorate. Twitter believes that this partnership will provide a unique opportunity for us to share our knowledge and technical expertise with smaller and emerging companies in the industry and for all industry actors to harness the expertise that has been built up in recent years.

We have also focused on NGO outreach and, since 2013, and have participated in more than 100 Countering Violent Extremism training and events around the world, including in Beirut, Bosnia, Belfast and Brussels and summits at the White House, at the United Nations, London, and Sydney. Twitter has partnered with groups like the Institute of Strategic Dialogue, the Anti-Defamation League and Imams Online to bolster counterspeech that offers alternatives to radicalization. As a result of that work, NGOs and activists around the world are able to harness the power of our platform in order to offer positive alternative narratives to those at risk and their wider communities.

Question 2. Terrorist how-to guides are protected by the First Amendment in the United States, but violate the content policies of many social media companies as well as the laws of some international partner nations. What countries have laws that go beyond your company's content policies and can you give examples of how you have worked with those countries to de-conflict those differences?

Answer. The Twitter Rules prohibit violent threats and the promotion or incitement of violence, including terrorism. Twitter is committed to removing such content swiftly from the platform. In addition, our Hateful Conduct policy is designed to protect users from harassment on the basis of protected categories, such as race, ethnicity, national origin, gender identity, age and religion. Examples of hateful conduct that we do not tolerate include targeting users with: (1) harassment; (2) wishes for the physical harm, death, or disease of individuals or groups; (3) references to mass murder, violent events, or specific means of violence in which or with which such groups have been the primary victims; (4) behavior that incites fear about a protected group; and (5) repeated and/or or non-consensual slurs, epithets, racist and sexist tropes, or other content that degrades someone. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories. We have also updated our policies to clearly prohibit users who 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 established channels for law enforcement agencies to request removal of content that may be illegal under the requesting jurisdiction's laws. Specifically, Twitter publishes global guidelines for law enforcement personnel that explain our policies and the process for submitting requests for content removal. See *https://help.twitter.com/en/rules-andpolicies/twitter-law-enforcement-support.* We accept requests from law enforcement agencies in countries in which Twitter operates, and we evaluate each request and, if appropriate, we will take action against the content at issue within the jurisdiction from which the removal request originated. As part of our commitment to transparency, since 2012, Twitter has published biannual Transparency Reports, reflecting the number of requests that we have received for user information and content removal on a per-country basis. See *https://transparency.twitter.com.*

Those reports indicate the number of requests that we have received and the number of requests with which we have complied.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. MARIA CANTWELL TO CARLOS MONJE, JR.

Question 1. We have strong principles of freedom of speech, but at the same time, we need to balance that freedom with the need to protect against bad actors who would leverage that freedom to plan and promote illegal acts. How can we use artificial intelligence to help us achieve a balance between our American ideal of free speech and the need to protect against extremist acts of terror?

Answer. Twitter recognizes that freedom of speech is a fundamental human right and we are committed to providing a platform where users feel safe to share their views and opinions. Twitter has a history of facilitating civic engagement and political freedom, and we intend for Twitter to remain a vital avenue for free expression here and abroad. But we cannot foster free expression without ensuring trust in our platform. We are determined to take the actions necessary to prevent the manipulation of Twitter, and we can and must make sure Twitter is a safe place.

Twitter's enforcement activity is designed with care to avoid, as much as possible, having an inadvertent negative impact on free expression. We do this by focusing our detection mechanisms primarily on signals and behavior, rather than content. We also do this by determining the proper use for automation, which is not always to take direct action on accounts or content.

Twitter relies on automation, artificial intelligence, and machine learning models to identify content for review by Twitter's teams and to detect potentially malicious activity on the platform. For example, when we learned that bad actors who shared terrorist propaganda on the platform were attempting to avoid permanent suspension from the platform by creating serial accounts, Twitter designed automated techniques to improve our detection rates for accounts that engage in such activity. Our efforts have been successful and our detection and suspension rates have increased as a result.

Question 2. Outside of artificial intelligence, what other technologies could be used to combat potential radicalization on social media platforms? What does the implementation of those technologies look like?

Answer. Keeping Twitter safe includes maintaining the quality of information on our platform. Our users look to us for useful, timely, and appropriate information. To preserve that experience, we are always working to ensure that we surface for our users the highest quality and most relevant content first. We are taking active steps to stop malicious accounts, abusive conduct, and terrorist and extremist content from spreading, and we are determined that our strategies will keep ahead of the tactics of bad actors.

For example, Twitter is continuing its effort to detect and prevent malicious automation by leveraging our technological capabilities and investing in initiatives aimed at understanding and addressing behavioral patterns associated with such accounts. In early 2017, we launched the Information Quality initiative, an effort aimed at enhancing the strategies we use to detect and stop bad automation, improve machine learning to spot spam, and increase the precision of our tools designed to prevent such content from contaminating our platform.

We have also made significant improvements to reduce external attempts to manipulate content visibility. These improvements were driven by investments in methods to detect malicious automation through abuse of our API, limit the ability of malicious actors to create new accounts in bulk, detect coordinated malicious activity across clusters of accounts, and better enforce policies against abusive third-party applications.

We have also introduced changes to our Twitter Rules, including how we correspond with those who violate them, and to our rules' enforcement process. We recently unveiled clarifications and updates to rules regarding hateful display names, hateful imagery, violent groups, and content that glorifies violence, which we began enforcing in December 2017. We made the various updates available prior to enforcement in order to provide our users and the general Twitter community with sufficient time to review and understand them. Twitter is continually working to make the platform a safe place for our users. For example, we are introducing changes to our Twitter Rules, including how we correspond with those who violate them, and to our rules' enforcement process.

As with most technology-based threats, the best approach is to share information and ideas to increase our collective knowledge. Working with the broader community, we will continue to test, to learn, to share, and to improve, so that our product remains effective and safe.

Another important tool against radicalization online is fostering alternative narratives from credible voices within communities. Twitter has invested in groups like WISE Muslim Women (NYC), Imams Online (UK) and Hedayah (UAE).

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. RICHARD BLUMENTHAL TO CARLOS MONJE, JR.

Question 1. On October 30, 2017, nineteen civil rights groups, including Muslim Advocates, Leadership Conference on Civil and Human Rights, NAACP, Southern Poverty Law Center, and many others, co-signed a letter to Facebook to express concern about the hateful content on the social media platform used to divide the country, and in particular, to promote anti-Muslim, anti-Black, anti-immigrant, and anti-LGBTQ animus.

Mr. Monje, how would Twitter respond to this letter if it had received it?

Answer. Keeping Twitter safe includes maintaining the quality of information on our platform. Our users look to us for useful, timely, and appropriate information. To preserve that experience, we are always working to ensure that we surface for

our users the highest quality and most relevant content first. We are taking active steps to stop malicious accounts, abusive conduct, and terrorist and extremist content from spreading, and we are determined that our strategies will keep ahead of the tactics of bad actors.

We have introduced changes to our Twitter Rules, including how we correspond with those who violate them, and to our rules' enforcement process. We recently unveiled clarifications and updates to rules regarding hateful display names, hateful imagery, violent groups, and content that glorifies violence, which we began enforcing in December 2017. We made the various updates available prior to enforcement in order to provide our users and the general Twitter community with sufficient time to review and understand them.

Question 2. Ms. Bickert, Mr. Monje, and Ms. Downs, please provide copies (including images, text, dates and timestamps) of all content identified by your platforms as generated by Russian agents or the Internet Research Agency.

Answer. We can provide the set of data on the Internet Research Agency that we have previously provided to other congressional committees through a secure data transfer.

Question 3. At least one of your peers in the tech industry has voluntarily initiated an outside assessment of the civil rights impacts of its policies and programs. In response to concerns regarding discrimination on the home-sharing platform, AirBNB hired former U.S. attorney general Eric Holder to help craft an anti-discrimination policy and has promised to pursue technological innovations to guard against future discriminatory events.

Mr. Monje, Ms. Bickert, and Ms. Downs, can you each commit to bringing in an independent entity to conduct a thorough and public audit of the civil rights impact of your policies and programs, including how your platform has been used by hate groups to stoke religious resentment and violence?

Answer. We agree that the decisions we make can have tremendous implications for civil rights, and we take that responsibility very seriously. We strive to create a platform that is conducive to robust democratic debate. Twitter is built around the idea of giving voice to people who may otherwise not be heard. From the #MeToo movement, to #BlackLivesMatter, to countless other campaigns through Twitter's history, we repeatedly see and are awed by the power of people who use Twitter to drive social change, encourage diverse perspectives, and share their stories. It is critical that our platform remains a welcoming place for these voices.

We share your concern regarding hate groups stoking religious resentment and violence, and we have taken significant steps over the past year to address such activity on our platform. We introduced changes to our Twitter Rules, including how we correspond with those who violate them, and to our rules' enforcement process. We recently unveiled clarifications and updates to rules prohibiting hateful display names, hateful imagery, violent groups, and content that glorifies violence, which we began enforcing in December 2017. In developing these changes, we worked closely with external advisers from around the world in the form of our Trust and Safety Council. The Twitter Trust and Safety Council provides input on our safety products, policies and programs. It includes safety advocates, academics, researchers, grassroots advocacy organizations, and community groups. A full list of our Council members is available here: *https://about.twitter.com/en_us/safety/safety-partners.html.*

Question 4. A little over a year ago, Facebook, Twitter, Google, and Microsoft announced a plan to create a joint industry database of "content that promotes terrorism."

Mr. Monje, Ms. Bickert, and Ms. Downs, to what extent does this joint industry database focus on all forms of terror, including the real terror threat presented by white supremacists?

Answer. In June 2017, we launched the Global Internet Forum to Counter Terrorism (the "GIFCT"), a partnership among Twitter, YouTube, Facebook, and Microsoft. The GIFCT facilitates, among other things, information sharing; technical cooperation; and research collaboration, including with academic institutions. In September 2017, the members of the GIFCT announced a multimillion dollar commitment to support research on terrorist abuse of the Internet and how governments, tech companies, and civil society can respond effectively. We are looking to establish a network of experts that can develop these platform-agnostic research questions and analysis that consider a range of geopolitical contexts. The GIFCT opened a call for proposals last month and we look forward to sharing further details of the initial projects early in 2018.

The GIFCT has created a shared industry database of "hashes"—unique digital "fingerprints"—for violent terrorist imagery or terrorist recruitment videos or im-

ages that have been removed from our individual services. The database allows a company that discovers terrorist content on one of its sites to create a digital fingerprint and share it with the other companies in the forum, who can then use those hashes to identify such content on their services or platforms, review against their respective policies and individual rules, and remove matching content as appropriate, or even block extremist content before it is posted in the first place. The database now contains more than 40,000 hashes. Instagram, Justpaste.it, LinkedIn, Oath, and Snap have also joined this initiative, and we are working to add several additional companies in 2018. Twitter also participates in the Technology Coalition, which shares images to counter child abuse.

We have introduced changes to our Twitter Rules, including how we correspond with those who violate them, and to our rules' enforcement process. We recently unveiled clarifications and updates to rules regarding hateful display names, hateful imagery, violent groups, and content that glorifies violence, which we began enforcing in December 2017. We made the various updates available prior to enforcement in order to provide our users and the general Twitter community with sufficient time to review and understand them.

For example, pursuant to our violent extremist groups policy, users are prohibited from engaging with our platform to make specific threats of violence or wish for the serious physical harm, death, or disease of an individual or a group of individuals. The policy makes clear that this prohibition includes, but is not limited to, threatening or promoting terrorist acts. We also indicate that users may not use our platform to affiliate with organizations that—whether by their own statements or by their activity both on and off the Twitter platform—use or promote violence against civilians to further their causes. We consider extremist groups to (1) identify as such through their stated purpose, publication, or actions; (2) have engaged in (or currently engage in) violence and/or the promotion of violence as a means of furthering their cause; and (3) target civilians in their acts and/or promotion of violence.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. BRIAN SCHATZ TO CARLOS MONJE, JR.

Question 1. Please quantify and explain Twitter's progress in tackling the fake-user account issue. For the most recent full month available and every month in the two years preceding provide:

- number of fake accounts created
- number of fake accounts removed
- number of accounts hacked
- number of hacked accounts restored
- number of duplicate accounts created
- number of duplicate accounts removed
- number of inactive accounts existing
- number of inactive accounts removed
- number of monthly active users
- average number of days a fake account remains on the platform

Answer. In December 2017, Twitter challenged more than 6.4 million accounts per week and prevented more than 523,000 suspicious logins per day. These actions, in addition to protections built into our account signup process, are part of a range of enforcement actions designed to catch and prevent the creation and use of fake accounts. Given our increasing focus on proactive detection (*i.e.,* detections focused on identifying fake or suspicious accounts at signup and preventing the creation of those accounts in the first place), attempting to quantify the precise number of accounts created or removed on a monthly basis would under-represent both the scale of the challenges we face and the scope of Twitter's activities to defend against them.

Twitter takes a wide range of actions to protect the security of user accounts on Twitter. For example, we partner with vendors that provide real-time information regarding data breaches on other websites and services that may impact consumers who also have Twitter accounts (since many consumers reuse passwords across services). In these cases, we take proactive steps to require users to secure their accounts (for instance, through e-mail reconfirmation and password changes)—even if there is no evidence of any malicious activity on Twitter itself. We also offer a range of security features, including two-factor authentication that users can take advantage of to help ensure that their accounts remain safe even if their passwords are

compromised. We also maintain a Vulnerability Rewards Program (also known as a "bug bounty") that provides financial incentives for security researchers to report vulnerabilities in our services so that we can quickly identify and remediate security risks before they affect our users.

The Twitter Rules prohibit the creation of "duplicate or substantially similar accounts" for spamming purposes, and we aggressively enforce these rules as part of our overall anti-spam and Information Quality initiatives. Accordingly, it is not possible to disaggregate specifically duplicative accounts from the overall volume of accounts challenged or suspended on the basis of spamming or malicious activity on Twitter.

Twitter does not proactively remove inactive accounts from the platform; however, we take a range of steps to ensure that inactive accounts remain secure. Consistent with our overall approach to account security, we may take steps to proactively lock down dormant or inactive accounts if we receive information that suggests they are at potential risk of compromise. We also employ real-time detections built to detect malicious or abnormal activity from inactive accounts which suddenly become active; an account which was previously inactive but suddenly begins producing a high volume of automated content would likely be challenged by our automated systems for detecting spamming or malicious activity.

As noted above, Twitter's efforts to combat fake accounts generally take two forms: (1) measures to challenge accounts at signup and prevent the creation of fake accounts; and (2) measures to detect, challenge, and prevent spamming or malicious activity in real time. We continue to invest heavily in improving our signup process to prevent the creation of new fake accounts. However, fake accounts which already exist on the platform are challenged and remediated consistent with their activity. A "fake" account which is completely inactive on Twitter would likely not be caught by our detections until or unless it begins to engage in spamming or malicious activity.

In terms of our monthly active users ("MAUs"), we have reported the following numbers over the last two years: 330 million in the fourth quarter of 2017; 330 million in third quarter of 2017; 326 million in second quarter of 2017; 327 million in the first quarter of 2017; 318 million in fourth quarter of 2016; 317 million in third quarter of 2016; 313 million in second quarter of 2016; and 310 million in the first quarter of 2016.

We report the number of active users on the platform quarterly as part of our earnings report, which you can find here: *https://investor.twitterinc.com/results.cfm.* As part of our commitment to transparency, since 2012, Twitter has published bi-annual Transparency Reports, reflecting the number of requests that we have received for user information and content removal on a per-country basis. See *https://transparency.twitter.com/*.

Question 2. How does a user find out if they are being impersonated on Twitter? Does Twitter notify users proactively? Or are users expected to monitor the platform and report to Twitter?

Answer. The Twitter Rules prohibit impersonation accounts. In response to reports—from either the user who is being impersonated or their authorized representatives—Twitter takes action against accounts that deceptively impersonate another user or account. Users and non-users alike can report impersonation accounts through a dedicated form in our Help Center or directly from the impersonated account's profile on the platform.

We are determined to expedite the suspension process for accounts deemed to be impersonating other users. Once we receive a report of potential user impersonation, we investigate the reported accounts to determine if the accounts are in violation of the Twitter Rules, which prohibit such profiles. Accounts determined to be in violation of our impersonation policy, or those not in compliance with our parody, commentary, and fan account policy, are either suspended or asked to update their profile so they no longer violate our policies.

In addition, Twitter strictly prohibits the purchasing and selling of account interactions on our platform. We advise our users that, by purchasing followers, Retweets, and likes, they are often purchasing bots, fake, or hacked accounts. Accounts found to have purchased, sold, or promoted the selling of followers, Retweets, and likes are in violation of the Twitter Rules and may be subject to suspension. See *https://help.twitter.com/en/rules-and-policies/twitter-rules.* Twitter has initiated a process of reviewing such activity and accounts, including accounts that appear to impersonate an actual, existing user. Where we determine that an account is fake (or that it impersonates an existing user), we immediately suspend the account or require that it complete a series of challenges before it can resume engaging with the platform in order to verify that it belongs to an actual user.

Question 3. What is the average number of days or hours that Twitter takes to investigate impersonation complaints before they are resolved?

Answer. We address impersonation reports as we receive them. Depending on the source of the report, we may request additional information before we can take action. Because our response takes place on a case-by-case basis and involves a fact-specific inquiry and manual review by Twitter personnel, there is not a uniform frequency or pattern to such enforcement actions or an average response time that we are able to provide.

Question 4. Does Twitter have a separate, expedited process for resolving impersonation of minors' accounts? Does Twitter know the age of its users?

Answer. We take seriously all reported impersonations and any other such violations of the Twitter Rules. We address all such reports with the same expediency and strive to suspend as quickly as possible those accounts that have been found to violate our policy against impersonation. Children younger than 13 are not allowed to create accounts or otherwise use our platform.

Question 5. Even a relatively small number of fake users can have an outsized impact in misleading voters. During the 2016 general election, there were 529 accounts that pushed incorrect, "text-to-vote" tweets. A similar phenomenon happened during the Virginia gubernatorial election in 2017. How many users saw these tweets? How many users interacted with these tweets? Who were behind these two voter suppression campaigns? If it is still unknown, is Twitter working with law enforcement to identify the accounts' creators?

Answer. During the period leading up to the 2016 election, Twitter labeled as "restricted pending deletion" a total of 918 such Tweets from 529 Twitter accounts. Assigning that label to a Tweet requires the user to delete the Tweet before the user is permitted to continue using the account and engage with the platform. So long as the Tweet has that label—and until the user deletes the tweet—the Tweet remains inaccessible to and hidden from all Twitter users. The user is blocked from Tweeting unless and until he or she deletes the labeled Tweet.

Twitter's review indicates that the 918 labeled Tweets (a) were viewed 222,111 times (an average of 242 views per tweet (also known as "impressions")); (b) liked by 10 users; (c) Retweeted by 801 users; and (d) received 318 replies.

In addition to labeling the individual Tweets, Twitter permanently suspended 106 accounts that were collectively responsible for 734 "vote-by-text" Tweets. Twitter's review of the suspended accounts' history indicates that those 734 Tweets (a) were viewed 162,656 times; (b) liked by 75 users; (c) Retweeted by 603 users; and (d) received 153 replies.

Twitter identified an additional 286 Tweets from 239 Twitter accounts with the relevant voting-related content upon which Twitter did not take any action. Twitter determined that those Tweets propagated the content in order to refute the message and alert other users that the information is false and misleading. Those Tweets generated significantly greater engagement across the platform compared to the 918 Tweets that Twitter labeled and the 106 accounts that Twitter suspended. Specifically, the 286 refuting tweets (a) were viewed 1,634,063 times (an average of 5,714 impressions per Tweet); (b) liked by 358 users; (c) Retweeted by 11,620 users; and (d) Received 611 replies.

During the period leading up to the Virginia gubernatorial election, Twitter received reports about an account that posted similar Tweets. We suspended the user upon receiving those reports. Our action against this account is consistent with the approach we took against illegal voter suppression Tweets during the 2016 election. Here, however, and well before our manual review of the account's activity resulted in its permanent suspension, Twitter's automated spam detection systems identified malicious behavior originating from this account and took action to hide that user's Tweets from appearing in searches and counting toward trends. Those automated systems, which we continue to invest in as part of our Information Quality initiative, help us address emerging malicious behavior even before a human reviewer can assess the content.

Question 6. Have there been any other voter suppression campaigns in elections following the 2016 general election—in the United States or abroad?

Answer. Other than the voter suppression Tweets discussed in Question 5, Twitter is not aware of similar voter suppression campaigns.

Question 7. Twitter has found 3,814 accounts so far linked with the Internet Research Agency (IRA) in a "relevant time period" to the 2016 election. Of the 3,814 IRA-linked accounts, how many of these were automated bots? How many were trolls? How many were impersonations of real American users' accounts? If so, do you plan to notify the accounts' owners?

Answer. We were able to determine that 307 of the 3,814 accounts that we have previously identified as linked to the IRA appear to be automated accounts. As we reported in our January 19 and January 31, 2018, blog posts, we notified U.S.-based users with an active e-mail address if they had directly engaged with or actively followed one of the 3,814 IRA-linked accounts we had identified. See *https://blog.twitter.com/official/en_us//company/2018/2016-election-update.html.* In total, approximately 1.4 million Twitter users received a notification from Twitter.

In addition, the Twitter Rules prohibit impersonation accounts. In response to reports—from either the user who is being impersonated or their authorized representatives—Twitter takes action against accounts that deceptively impersonate another user or account. Users and non-users alike can report impersonation accounts through a dedicated form in our Help Center or directly from the impersonated account's profile on the platform.

We are determined to expedite the suspension process for accounts deemed to be impersonating other users. Once we receive a report of potential user impersonation, we investigate the reported accounts to determine if the accounts are in violation of the Twitter Rules, which prohibit such profiles. Accounts determined to be in violation of our impersonation policy, or those not in compliance with our parody, commentary, and fan account policy, are either suspended or asked to update their profile so they no longer violate our policies.

Question 8. What was the "relevant time period" under which Twitter found the 3,814 IRA-linked accounts? What was the justification for using this time period?

Answer. Our search for IRA-linked accounts was not limited to a particular time period. We conducted a thorough review of data available to us in order to identify any account we could reasonably link to accounts we had previously determined to be associated with the IRA either by our independent research or through information provided to us by third parties (as discussed in greater detail below in response to Question 9).

Question 9. The IRA has been targeting American users well before the 2016 election—for instance during Russia's invasion of Crimea in 2014. When did Twitter first detect IRA activity on the platform? When did Twitter first detect IRA activity specifically targeting American users? From the first time Twitter detected IRA activity to the most recent time period available, how many accounts has the IRA created on Twitter?

Answer. Twitter first identified and suspended IRA-linked accounts in June 2015 following a June 2, 2015, *New York Times* article about the IRA and its well-known online Kremlin propaganda activities. *The Agency: From a Nondescript Office Building in St. Petersburg, Russia, an Army of Well-Paid "Trolls" Has Tried to Wreak Havoc All Around the Internet—and in Real-Life American Communities,* NYTimes.com, available at *https://www.nytimes.com/2015/06/07/magazine/the-agency.html?_r=0.* Twitter suspended a total of 467 those accounts within days of the article's publication. Those accounts were suspended for violating Twitter's spam rules.

On August 8, 2017, a third-party security firm provided Twitter with a report listing Russian-linked accounts that were suspected of being associated with the IRA. Twitter commenced its review of that list immediately. Based on that review, between August 8 and August 10, 2017, Twitter suspended 473 accounts listed in the report for engaging in spam activity prohibited under the Twitter Rules.

Also on August 8, 2017—but separately from the security firm report—Twitter's Information Security team received from Facebook a list of e-mail addresses, which Facebook indicated were connected to the IRA. Twitter reviewed that list and identified 22 accounts that matched the e-mail addresses that Facebook provided. All 22 accounts had already been suspended or were subsequently suspended. On August 10, 2017, Facebook shared with Twitter account-related information for one additional Facebook account. Twitter identified 181 accounts that were linked or related to the 23 accounts that Facebook shared with us, bringing the total of Russian-related accounts under examination to 204. As of August 22, 2017, all but 14 of those accounts had already been suspended or set in read-only mode pending phone number verification. Following a manual review of all 204 accounts, three were determined to be non-automated, legitimate users; those accounts remain active on the platform.

Finally, in connection with our retrospective review of Russian interference in the 2016 U.S. election through activity on our platform, we identified additional accounts linked to the IRA, bringing the total number of such accounts to 2,752. And, as we reported in our January 19, 2018, update to Congress, through our continued analysis, we identified an additional 1,062, for a total of 3,814 IRA-linked accounts. All 3,814 IRA-linked accounts were suspended for Terms of Service violations, and

all but a few compromised accounts that have subsequently been restored to their legitimate account owners remain suspended.

Question 10. Between September 1 and November 15, 2016, 175,993 tweets from IRA-linked accounts received a total of 351,632,679 impressions within the first seven days after posting. And these tweets were retweeted 4,509,781 times. Out of the total number of impressions, how many were via organic reach vs. promoted or paid reach? How do these metrics compare to Twitter's typical performance benchmarks? Please give specific numbers, for instance, the number of impressions that an average tweet got between September 1 and November 15, 2016, for both organic reach vs. promoted or paid reach.

Answer. None of the Tweets in question were promoted.

The number of impressions on a Tweet can vary substantially depending on the content in question, its proliferation on Twitter and through other online channels, and the accounts responsible for posting or sharing it. For example, a user who has many followers will generally receive more impressions per Tweet than a user with a smaller number of followers. Similarly, a news organization embedding a Tweet in an article would likely generate an increase in the number of impressions—compared to Tweets with similar characteristics which were not embedded—given the additional impressions from off-platform views. Due to the variability of accounts, content, and sharing patterns on Twitter, we do not have a general or baseline measure of Tweet performance across different contexts.

Question 11. How many users did the aforementioned 351,632,579 impressions reach exactly? What technological tools (*e.g.,* bots, third-party applications) were used to drive up the number of impressions?

Answer. Impression counts do not drive content ranking or content display and are not publicly visible in the Twitter product. We are not aware of specific attempts to manipulate the number of impressions on content—and believe that such attempts, if made, would not have appreciable impact on the Twitter platform.

Due to the variability of accounts, content, and sharing patterns on Twitter, there is not a one-to-one correlation between impressions and Twitter users. Further, impressions will include the number of times a Tweet is viewed across a range of products (such as the Twitter website, mobile apps, or Tweets embedded on an external website), including by logged out users about whom Twitter has limited information. We do not have a way of measuring impressions on Tweets from any non-Twitter applications or tools, and impressions on content via third-party applications would not contribute to the overall impression count on a Tweet.

Question 12. Twitter removed 935,000 accounts for terrorism promotion. How many of these were automated accounts?

Answer. We do not track suspensions for promotion of terrorism in a manner that specifically flags automation. As we have previously noted in our blog posts on this subject, however, our proprietary spam-fighting tools offer significant assistance in the fight against the dissemination of terrorist propaganda online. We are able to leverage our spam-fighting tools to stop the spread of this content given the prevalence of spam signals such as indicia of automation and attempts at ban evasion.

Question 13. Please share more details about the medium of content in tweets identified with active measures. What percentage of these tweets are text only? Contain static image? Video? How many of these tweets embed an external hyperlink?

Answer. Of the 175,993 Tweets posted by the 3,814 IRA-linked accounts during the election time period: approximately 92,000 embedded external hyperlinks; 114,000 were text only; 58,000 contained static images; and 3,000 contained videos.

Question 14. Once an account is flagged by an algorithm for removal, what is the average amount of time before the account is removed from public view? After the account is removed from public view, what is the average amount of time before the data is deleted? Why is the data deleted rather than just removed from public view?

Answer. Twitter systems are designed to take different steps with respect to different types of malicious activity on the platform. Our systems cast a wide net to detect and label malicious accounts or malicious activity, and we may take additional steps to confirm the accuracy of those processes before removing the content from public view. Those additional steps are designed to minimize false positives and inadvertent action against users who we ultimately determine not to be in violation of our policies.

In other circumstances, where our automated tools operate with high precision, we may take more immediate action. Such circumstances include, but are not limited to, instances where we detect child sexual exploitation or malware on our platform. As a general matter, content removal on Twitter reflects a careful balancing of platform protection and individual user rights, and it requires constant evaluation and assessment to refine and improve upon our existing methodologies. Such

frequent iterations and reexamination of our actions are critical to enhancing and improving our detection tools.

Data from accounts that have been suspended from the Twitter platform are retained in order to inform and improve existing detection systems and for responding to requests from law enforcement.

Question 15. What is Twitter's data retention policy, as it relates to the suspension of accounts for violating Twitter's terms of service? Does this policy apply to all parties—such as independent researchers, users, advertisers, and data brokers—in the same way?

Answer. Data from accounts that have been suspended from the Twitter platform and are no longer publicly visible are retained in Twitter's internal systems for safety and security purposes, including to inform and improve existing detection systems.

Through our API, we give developers and other third parties access to subsets of public Twitter content. Access to this publicly available data through our API is conditioned on acceptance of our policies, including the requirement that developers not use the API to undertake activities with respect to content that users have removed from the platform. Examples of situations this policy is designed to address include a parent deciding to remove pictures of their children if they have safety concerns or a college student removing Tweets as they prepare to apply for jobs. This is a long-standing Twitter policy.

Question 16. Given the importance of collaborating with third-party or independent researchers to prevent further interference by Russia, will Twitter be updating its data retention policy?

Answer. Twitter is committed to addressing how information spreads online, crosses between platforms and services, and raises the attention of voters, elected officials, and the media. Consistent with our commitment to transparency, we recognize that our efforts at addressing this issue must be part of a broader discussion about how important societal conversations take place online and how Russia has leveraged digital services, including Twitter, to interfere with U.S. elections. Indeed, cooperation to combat this challenge is essential. We cannot defeat this novel, shared threat alone. As with most technology-based threats, the best approach is to share information and ideas to increase our collective knowledge. Working with the broader community, we will continue to test, to learn, to share, and to improve, so that our product remains effective and safe.

Twitter looks forward to continuing to work closely with third party researchers consistent with its commitment to transparency and improvement in these critical areas for democracy and elections. Last year, Twitter offboarded Russia Today and Sputnik as advertisers on our platform and dedicated the $1.9 million those accounts had spent on advertising globally on the platform to research in these areas.

Question 17. Do hashes, as mentioned during the hearing, exist only for static images? What about video content? If hashes for video content are not yet fully deployed, please share the timeline to do so.

Answer. Hashes are deployed for static images, in close partnership with industry groups, to fight terrorist content as well as child sexual exploitation online. We continue to work with peer companies and industry groups to expand on hash sharing partnerships, including through the potential use of hashes for video content. It is critical that these programs are deployed urgently, but also with careful cross-industry collaboration and buy-in, to maximize the potential for shared success in fighting these challenges.

Question 18. Last year, Twitter shared that 220,000 malicious applications were suspended for abuse of the Twitter application programming interface (API). How does Twitter define API abuse? Who created, managed, or plugged in the malicious applications to the Twitter API?

Answer. Please see the answer to question 19 below.

Question 19. How often does Twitter monitor the network of applications that use the Twitter API? Is suspension done on a rolling basis? Is it solely up to Twitter employees (rather than its users or developer community) to identify malicious applications?

Answer. Any developer signing up for access to Twitter's API is required to agree to the Developer Agreement and Policy prior to obtaining access. *See https://t.co/devpolicy.* The Developer Agreement, Developer Policy, Automation Rules (*https://t.co/automation*), Display Requirements (*https://developer.twitter.com/en/developer-terms/display-requirements*), and Geo Guidelines (*https://developer.twitter.com/en/developer-terms/geo-guidelines*) collectively make up the body of rules that govern developers' use of our platform. We also make available to developers additional guidance regarding how to interpret and implement these guidelines, either via our

Developer website (*e.g. https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases*) or via the Twitter Community forums (*https://twittercommunity.com/c/dev-platform-feedback/rules-and-policies*).

At this time, subject to limits on use as well as automated and human review for policy compliance, any user with a valid Twitter account can register for access to Twitter's API. As we noted in our January 19, 2018, blog post, we are currently working on introducing an improved developer onboarding process to better manage the use cases for developers building on Twitter's platform. See *https://blog.twitter.com/developer/en_us/topics/tools/2017/introducing-twitter-premium-apis.html.*

Twitter enforces the terms of the Twitter API through multiple channels. Through Twitter's compliance team, we investigate and address instances of potential violations of Twitter's policies with respect to Twitter's commercial data products. We also work with third parties and other stakeholders to investigate and address other reported abuses of our APIs. As with any community, our enforcement mechanisms are best served by a combination of affirmative steps and reactive investigations that we take to address concerns raised by community members in order to protect our users and customers.

Question 20. In terms of dollars and percentage of annual revenue, how much is Twitter now spending on preventing foreign interference with our elections? What was the figure in the election cycle leading up to November 2016? What is the projected spend leading up to November 2018?

Answer. As we stated in our February 8, 2018, shareholder letter, Twitter continues to invest considerable resources in our Information Quality efforts. *See http://files.shareholder.com/downloads/AMDA-2F526X/5990710870x0x970892/F9B4F616-659A-454B-89C6-28480DA53CCA/Q4_2017_Shareholder_Letter.pdf*

Based on the understanding we have gained from our retrospective review, we have also established an internal, cross-functional team dedicated to addressing election-related instances of abuse on Twitter, as we discussed with the Senate Commerce Committee in January.

The election team will address this challenge in a number of ways. Among other things, to detect and promptly address impersonation attempts, the team will verify major party candidates for all statewide and Federal offices, as well as all major national party accounts. In addition to monitoring and enforcing the Twitter Terms of Service and Twitter Rules, the election team will cooperate and communicate with Federal and state election officials to swiftly escalate and address in real time attempts at election interference. And consistent with Twitter's commitment to curtailing malicious automation, spam, and false accounts on our platform, the election team will focus on deploying our proprietary tools specifically to detect and stop malicious election-related activity.

Question 21. Congress will judge success not by Twitter's efforts but by its results. How will Twitter measure its success? Will Twitter be conducting an audit after November 2018? When will the results be shared?

Answer. Twitter will continue to work closely with Congress, our industry peers, civil society, experts, and law enforcement agencies to consider these challenges and novel threats for Twitter, the Internet, and society as a whole. We are committed to addressing, mitigating, and ultimately preventing any future attempts to interfere in elections and the democratic process, and to doing so in the most transparent way possible. We look forward to continuing to provide information to the Committee about malicious activity we detect on our platforms and the measures we take to address such activity.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. TAMMY BALDWIN TO CARLOS MONJE, JR.

Question 1. Mr. Monje, you note in your testimony that Twitter's effective anti-spam technology—which the company will be using to address networks of malicious bots attempting to interfere with the 2018 U.S. midterms—is "in-house proprietary." At the same time, you tout Twitter's record of and support for industry collaboration. Wouldn't that collaboration—as well as our shared fight against Russian active measures—be improved by Twitter sharing its anti-spam technology with other social media companies? Do you share that technology with others?

Answer. We agree that collaboration with our industry peers and civil society is critically important to addressing common threats and that it has been successful in meeting shared challenges. In June 2017, for example, we launched the Global Internet Forum to Counter Terrorism (the "GIFCT"), a partnership among Twitter, YouTube, Facebook, and Microsoft. The GIFCT facilitates, among other things, in-

formation sharing; technical cooperation; and research collaboration, including with academic institutions. In September 2017, the members of the GIFCT announced a multimillion dollar commitment to support research on terrorist abuse of the Internet and how governments, tech companies, and civil society can respond effectively. We are looking to establish a network of experts that can develop these platform-agnostic research questions and analysis that consider a range of geopolitical contexts. The GIFCT opened a call for proposals in December 2017 and we look forward to sharing further details of the initial projects this year.

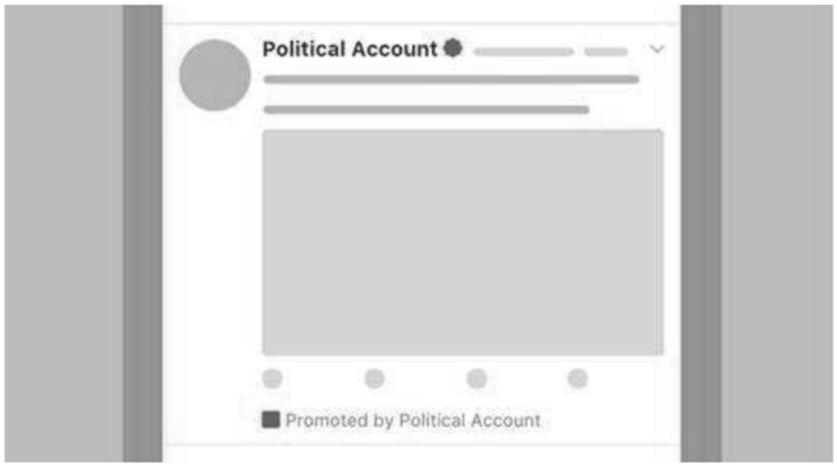
The GIFCT has created a shared industry database of "hashes"—unique digital "fingerprints"—for violent terrorist imagery or terrorist recruitment videos or images that have been removed from our individual services. The database allows a company that discovers terrorist content on one of its sites to create a digital fingerprint and share it with the other companies in the forum, who can then use those hashes to identify such content on their services or platforms, review against their respective policies and individual rules, and remove matching content as appropriate, or even block extremist content before it is posted in the first place. The database now contains more than 40,000 hashes. Instagram, Justpaste.it, LinkedIn, Oath, and Snap have also joined this initiative, and we are working to add several additional companies in 2018. Twitter also participates in the Technology Coalition, which shares images to counter child abuse.

Because each platform is unique, there are many elements of our coordinated work that do not translate easily across platforms. Although we share with other companies our approach to addressing shared threats, including certain signals that we use to identify malicious content, solutions applicable to the Twitter platform are not always applicable to other companies. We describe our tools as "in-house and proprietary" to distinguish them from tools that are developed by and licensed from third-party vendors.

Question 2. Mr. Monje, ensuring Americans know the source of political advertising on social media is one of the best ways to combat interference in U.S. elections by foreign actors. Put simply, we should apply to social media the same rules that apply to TV and print media. Please tell us more about Twitter's Ads Transparency Center, including the status of implementation and any hurdles you foresee.

Answer. Twitter's approach to greater transparency in political advertising centers on two components: a new electioneering policy and an industry-leading Transparency Center. We expect to roll out the new policy in the U.S. during the first quarter of 2018. To make it clear when a user is viewing or engaging with content considered to be an electioneering ad, our policy will require that advertisers that meet the definition of electioneering to identify their campaigns as such. We will also change the interface of such ads and include a visual political ad indicator (*see, e.g.,* Fig. 1 below).

Fig. 1: Template for New Electioneering Ad



Twitter's definition of electioneering ads will be derived from the FEC regulations' definition of that term, which includes any broadcast, cable, or satellite communication that refers clearly to a candidate for Federal office, is published 60 days before a general election or 30 days before a primary, convention, or caucus, and is targeted to the relevant electorate (if the candidate is running for Congress).

The goal of the Transparency Center is to offer the public increased visibility into all advertising on the platform and to provide users with tools to share feedback with us. With respect to electioneering ads and the Transparency Center, we intend to better enable users and outside parties to conduct their own research or evaluation regarding particular ads. Electioneering ads information accessible through the Transparency Center will include, among other things, the identity of the organization funding the campaign, all ads that are currently running or have run on Twitter, campaign spend, and targeting demographics for specific ads or campaigns. We plan to launch the Transparency Center as soon as feasible after rolling out our electioneering policy in the first quarter of 2018, and we are continuing to refine the tools we will make available in conjunction with launching the Transparency Center to ensure the best experience for our users.

Question 3. In the context of extremist content, I would like to learn more about each company's policy for proactively reporting users to law enforcement. I understand your companies evaluate and respond to law enforcement requests for information, but what framework do you use to proactively report terrorist-related content to authorities, including any identifying information of the user? For example, if you use a standard of imminent harm, how do you define and apply it, particularly in a threat environment where terrorist organizations often call on recruits to attack carefully planned targets of opportunity, rather than to launch an immediate, indiscriminate attack?

Answer. Twitter actively works on establishing and maintaining close relationships with law enforcement by providing ongoing training opportunities and through recurring meetings that allow for urgent, proactive outreach in the event Twitter becomes aware of imminent harm related to content on the platform. In these circumstances, it is of paramount importance that relationships allow for immediate connection at any time, day or night. This includes regularly working with the F.B.I. on domestic issues and U.S. legal attachés across the globe to assist in vetting complex international situations. The circumstances leading to proactive reporting will depend on the nature of the issue. Typically, when we become aware of content on one of Twitter's products that contains a serious immediate threat or an actual depiction of live physical violence, whether of harm to one's self or to other persons, Twitter will proactively contact law enforcement or other appropriate authorities so they can properly assess the nature of the situation. As noted in Twitter's Law Enforcement Guidelines, our emergency request protocols are covered 24/7, every day of the year. Twitter evaluates emergency disclosure requests on a case-by-case basis in compliance with relevant law (*e.g.,* 18 U.S.C. § 2702(b)(8)). If we receive information that provides us with a good faith belief that there is an exigent emergency involving the danger of death or serious physical injury to a person, we may provide information necessary to prevent that harm, if we have it. Where law enforcement believes identifying information is warranted, we can immediately provide such information through our emergency response protocols.

Question 4. I would like to hear from the companies whether they support implementing Mr. Watts's recommendations to: first, fully certify the authenticity of all users—in other words, ensure that each user is a real person; and second, eliminate social bot networks to reduce automated broadcasting of disinformation.

Answer. We have dramatically improved our ability to identify and disrupt social bot networks, as explained elsewhere in these responses. Using malicious automation is a violation of our terms of service and we action accounts with increasing effectiveness.

Twitter is committed to defending the voices of our users, including those who rely on anonymous or pseudonymous accounts to do their work safely. Journalists, activists, political dissidents, whistleblowers, and human rights practitioners have been imprisoned, tortured and worse on the basis of their personally identifiable information online. This decision to protect their identity was made in consultation with leading NGOs working on the front lines of these issues worldwide. We seek to protect them on Twitter.

Question 5. What are the indicators that you use to identify a Russian disinformation account, whether from the Kremlin's so-called Internet Research Agency or an associated group of hackers or trolls, and what thresholds must be met to disable an account?

Answer. We relied on a number of different sources in order to identify accounts linked to the IRA. Twitter first identified and suspended IRA-linked accounts in June 2015 following a June 2, 2015, *New York Times* article about the IRA and its well-known online Kremlin propaganda activities. *The Agency: From a Nondescript Office Building in St. Petersburg, Russia, an Army of Well-Paid "Trolls" Has Tried to Wreak Havoc All Around the Internet—and in Real-Life American Communities,* NYTimes.com, available at *https://www.nytimes.com/2015/06/07/magazine/the-agency.html?_r=0.* Twitter suspended a total of 467 those accounts within days of the article's publication. Those accounts were suspended for violating Twitter's anti-spam rules.

On August 8, 2017, a third-party security firm provided Twitter with a report listing Russian-linked accounts that were suspected of being associated with the IRA. Twitter commenced its review of that list immediately. Based on that review, between August 8 and August 10, 2017, Twitter suspended 473 accounts listed in the report for engaging in spam activity prohibited under the Twitter Rules.

Also on August 8, 2017—but separately from the security firm report—Twitter's Information Security team received from Facebook a list of e-mail addresses, which Facebook indicated were connected to the IRA. Twitter reviewed that list and identified 22 accounts that matched the e-mail addresses that Facebook provided. All 22 accounts had already been suspended or were subsequently suspended. On August 10, 2017, Facebook shared with Twitter account-related information for one additional Facebook account. Twitter identified 181 accounts that were linked or related to the 23 accounts that Facebook shared with us, bringing the total of Russian-related accounts under examination to 204. As of August 22, 2017, all but 14 of those accounts had already been suspended or set in read-only mode pending phone number verification. Following a manual review of all 204 accounts, three were determined to be non-automated, legitimate users; those accounts remain active on the platform.

Finally, in connection with our retrospective review of Russian interference in the 2016 U.S. election through activity on our platform, we identified additional accounts linked to the IRA, bringing the total number of such accounts to 2,752. And, as we reported in our January 19, 2018, update to Congress, through our continued analysis, we identified an additional 1,062, for a total of 3,814 IRA-linked accounts. All 3,814 IRA-linked accounts were suspended for Twitter Terms of Service violations, and all but a few compromised accounts that have subsequently been restored to their legitimate account owners remain suspended.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. CATHERINE CORTEZ MASTO TO CARLOS MONJE, JR.

Question 1. During my time as the Attorney General for the State of Nevada, I saw too many instances of sex trafficking cases involving child victims that were dismissed because the conduct occurred online or through social media. So that's why I'm a strong supporter of the Stop Enabling Sex Traffickers Act of 2017 (SESTA), which clarifies the Communications Decency Act (CDA) to allow state Attorneys General to retain their jurisdiction to prosecute those who facilitate human trafficking. We know that trafficking is happening online and on social media, and SESTA is the only current legislative proposal that provides sufficient deterrence to traffickers by providing the necessary tools for successful prosecutions. As a former prosecutor, I know what it will take to successfully prosecute those who engage in sex trafficking through social media and other websites, and that's why I believe that the House version of SESTA doesn't go far enough to give prosecutors the tools they need to protect sex trafficking victims. I hope that your organizations all agree that victims of sex trafficking deserve meaningful protections and justice.

If so, I'd like to hear whether you will continue to support SESTA over the weaker U.S. House version of the bill.

Answer. The answer to Question 1 has been provide in response to Question 2 below.

Question 2. I was glad to hear that the Internet Association supports SESTA, and I'd like to know what else your organization is doing to address concerns about sex trafficking occurring on your platforms and helping us pass this important legislation in the Senate.

Answer. Human trafficking or the facilitation of such activities have no place on Twitter. Twitter is deeply committed to working together with Congress, law enforcement, victims groups, and NGOs to combat such heinous crimes. To that end, we take a multifaceted approach to this issue.

We do not tolerate child sexual exploitation on Twitter. When we are made aware of links to images of or content promoting child sexual exploitation, they will be removed from the site without further notice and reported to The National Center for Missing & Exploited Children (or "NCMEC"). We also permanently suspend accounts promoting or containing links to child sexual exploitation or engaging in child sex trafficking. Furthermore, we take measures to discourage repeat signups from these users.

More broadly, sharing explicit sexual images or videos of someone online without their consent is a violation of their privacy and one of the most serious violations of the *Twitter Rules.* We will suspend any account we identify as the original poster of intimate media that has been produced or distributed without the subject's consent. We will also suspend any account dedicated to posting this type of content.

Twitter works with NGOs and victims' groups which focus on counter-child sex trafficking measures, including Love146, Thorn and NCMEC—all of which serve on our Twitter Trust Council.

Twitter is also a member of and serves as the current chair of the Technology Coalition, powered by leaders in the Internet services sector. Formed in 2006, the Coalition's vision is to eradicate online child sexual exploitation. The group's strategy is to sponsor the development of technology solutions that disrupt the ability to use the Internet to exploit children or distribute child pornography. The Technology Coalition works with the NCMEC and its sister agency, the *International Centre for Missing & Exploited Children* (the "ICMEC"), to identify and propagate technology solutions that create effective disruption.

In addition, the Technology Coalition seeks and creates platforms for collaboration with the private and public sectors for the creation of standards, the sharing of best practices, and similar initiatives that advance the fight against online sexual exploitation of children.

The Technology Coalition's efforts are structured with a view toward balancing the privacy interests of Internet users with its mission to eradicate online child sexual exploitation.

We also have close working relationships with law enforcement and expeditiously review and action legal requests. *Guidelines* intended for law enforcement authorities seeking information about Twitter accounts are posted on our website. Information concerning requests to *withhold content on Twitter* is available. More general information is available in our *Privacy Policy, Terms of Service,* and *Twitter Rules.*

Twitter has been an active participant in Congress' recent efforts to address human trafficking. Our team has held dozens of meetings with lawmakers to engage in meaningful dialogue about how we can work together to meet this challenge.

Question 3. Over the past few months, our country has been reckoning with some hard truths about the way that women and minorities are treated in the workplace. And I think this is a moment for all types of organizations, including tech giants like the ones represented here, to take a clear-eyed accounting of their culture and practices, to take responsibility for what hasn't worked, and to renew their commitments to make meaningful improvements. The Equal Employment Opportunity Commission's 2016 report on "Diversity in High Tech" found that women, African Americans, and Hispanics are all represented at significantly lower levels in high tech than in private industry as a whole. And while recent internal studies at Facebook and Google have showed some progress in the hiring of women, there has not been equal improvement in the representation of people of color and other underrepresented groups.

What technically qualifies as diversity to your organization?

Answer. Twitter takes a holistic approach to the way we define diversity. At Twitter, diversity includes the hiring, retention and advancement of individuals with protected characteristics (*e.g.,* age, gender, race, sexual orientation, etc.). However, these are not the only characteristics that impact our diversity profile. We value the uniqueness of our people. This means that we challenge ourselves to hire talented people who have different ideas, perspectives, and approaches to the work. In our experience, diversity allows employees to feel comfortable sharing the intersectionality of who they are as individuals, which leads to the exchange and development of cutting-edge ideas.

Question 4. How is your company working to address issues of discrimination in your own workforces?

Answer. At Twitter we value inclusion and diversity. For that reason, we have a multi-tiered approach to addressing discrimination in the workplace. Some of the methods we use include, but are not limited to, clear and accessible policies against discrimination; consistent messaging to employees that discrimination is not tolerated at Twitter; an active Inclusion & Diversity team; annual training for our people

managers to ensure they are equipped to promptly respond to any issue that may be perceived as discriminatory; a dedicated hotline for employees to report concerns of discrimination; and an Employee Relations ("ER") Team that is responsible for prompt, thorough and neutral investigation of discrimination complaints.

We also sponsor Business Resource Groups ("BRGs"), which are an excellent support system within Twitter to foster awareness, respect, and inclusion within the workplace. These groups include Blackbirds (African American), Twitter Alas (Hispanic/LatinX), Twitter Women, and Twitter Open (LGBTQ). BRGs serve as a sounding board around strategic diversity objectives within the organization to help create a more inclusive work environment.

In addition, Twitter's Diversity Advisory Council offers suggestions and advice on strategies. The Council provides a forum for sharing best practices across the tech industry and in the field of diversity and inclusion. We are continuously assessing the work environment at Twitter, conducting periodic pulse surveys of employees and hosting lunches for women and people of color.

Question 5. Do you believe those efforts are sufficient?

Answer. Twitter's current anti-discrimination practices are targeted to foster a fair, inclusive, and healthy environment, and they exceed the minimum legal standards in every jurisdiction in which we operate. To that end, while we believe that our efforts are sufficient, we will continue to work to exceed minimal requirements and strive to be an industry leader in this area.

Question 6. I've seen that Facebook works to make their labor diversity information public, can you provide a status on your labor figures, or commit to sharing those with the Committee and the public?

Answer. We commit to sharing labor diversity information with the Committee and the public. The latest report can be found here: *https://blog.twitter.com/en_us/topics/company/2017/building-a-more-inclusive-twitter-in-2016.html.*

Question 7. We know that so-called talent pipelines are not the only obstacle to achieving a diverse workforce, and that discrimination and harassment go hand in hand, distorting the operation of workplace meritocracies. This is a moment when many victims of sexual assault and harassment are bravely coming forward about their experiences, allowing us to get a better sense of the true scope and effects of this behavior. Persistent harassment, and the workplace culture that tolerates, ignores, or even encourages such harassment, pushes people out of their workplaces, stalls or derails promising careers, and discourages some from pursuing certain opportunities altogether.

What is your company doing to evaluate the impact of harassment in your workforces?

Answer. Twitter conducts periodic and annual reviews and analysis of Employee Relations data.

Question 8. How are you working to create a culture where harassment is no longer tolerated?

Answer. Twitter maintains and enforces a clear and accessible policy against sexual harassment. We foster an environment that encourages employees to report concerns and maintain proper complaint procedures for reporting concerns. Employee Relations promptly investigates all allegations of sexual harassment. We provide annual sexual harassment training for all managers and provide team and individualized training on an as-needed basis.

Question 9. What more could you be doing to be a positive example for other companies and industries?

Answer. Building and maintaining a diverse workforce and fostering a culture of inclusion is a top priority for our company. Recognizing that our platform serves as a powerful tool to educate and facilitate conversations around these topics, we recently launched the Twitter handle, @TwitterTogether, which allows Twitter users, industry peers, and the public to learn about internal and external activities happening at Twitter Inc.

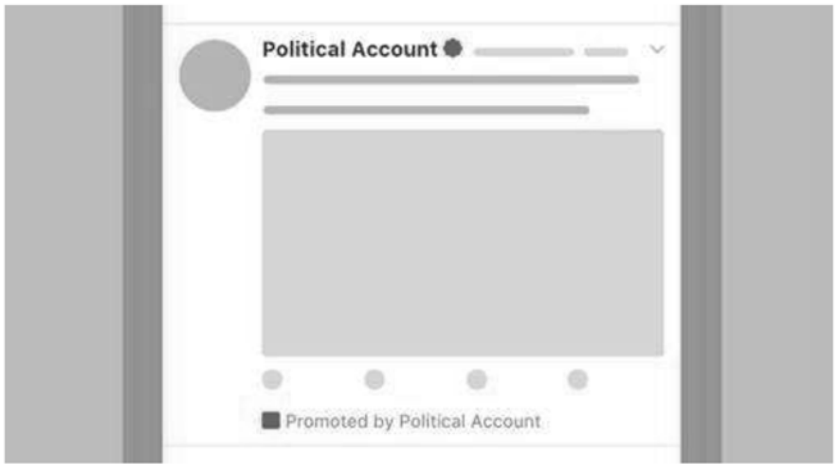
Question 10. Last October, Facebook *announced* that it would be *improving transparency* for all ads run on its platform, including by requiring political advertisers to include a disclaimer telling viewers who paid for an ad, and allowing viewers to see all the ads a page is running, even those that aren't targeting them. Twitter also announced *similar measures.* Although these policies were announced in response to Russia using social media to interfere in our elections, it seems these transparency measures could help shine a spotlight on other forms of influence campaigns by extremists or terrorists.

Answer. Can you provide an update on the status of these measures?

Twitter's approach to greater transparency in political advertising centers on two components: a new electioneering policy and an industry-leading Transparency Center. We expect to roll out the new policy in the U.S. during the first quarter of 2018.

To make it clear when a user is viewing or engaging with content considered to be an electioneering ad, our policy will require that advertisers that meet the definition of electioneering to identify their campaigns as such. We will also change the interface of such ads and include a visual political ad indicator (*see, e.g.,* Fig. 1 below).

Fig. 1: Template for New Electioneering Ad



The goal of the Transparency Center is to offer the public increased visibility into all advertising on the platform, and to provide users with tools to share feedback with us. With respect to electioneering ads and the Transparency Center, we intend to better enable users and outside parties to conduct their own research or evaluation regarding particular ads. Electioneering ads information accessible through the Transparency Center will include, among other things, the identity of the organization funding the campaign, all ads that are currently running or have run on Twitter, campaign spend, and targeting demographics for specific ads or campaigns.

Question 11. When can we expect to see them fully implemented?

Answer. We plan to launch the Transparency Center as soon as feasible after rolling out our electioneering policy in the first quarter of 2018, and we are continuing to refine the tools we will make available in conjunction with launching the Transparency Center to ensure the best experience for our users.

Question 12. How are you defining what constitutes a political ad subject to these heightened transparency requirements?

Answer. Twitter's definition of electioneering ads will be derived from the FEC regulations' definition of that term, which includes any broadcast, cable, or satellite communication that refers clearly to a candidate for Federal office, is published 60 days before a general election or 30 days before a primary, convention, or caucus, and is targeted to the relevant electorate (if the candidate is running for Congress).

Question 13. On January 29, the Director of the Central Intelligence Agency said he expects the Russian government to attempt to influence the 2018 elections in this country.

What efforts is Twitter undertaking in the lead up to the 2018 elections to identify and close the platform's remaining vulnerabilities to foreign exploitation?

Answer. Our efforts to detect and stop malicious activity on our platform continue, particularly in the context of elections. Based on the understanding we have gained from our retrospective review of activity on our platform during the period leading up to the 2016 election, we have established an internal, cross-functional team dedicated to addressing election-related instances of abuse on Twitter, as we discussed with the Committee during the January 17, 2018, hearing.

The election team will address this challenge in a number of ways. Among other things, to detect and promptly address impersonation attempts, the team will verify

major party candidates for all statewide and Federal offices, as well as all major national party accounts. In addition to monitoring and enforcing the Twitter Terms of Service and Twitter Rules, the election team will cooperate and communicate with Federal and state election officials to swiftly escalate and address in real time attempts at election interference. And consistent with Twitter's commitment to curtailing malicious automation, spam, and false accounts on our platform, the election team will focus on deploying our proprietary tools specifically to detect and stop malicious election-related activity.

Question 14. What assistance can Federal, state and local government entities provide in that effort?

Answer. Twitter was pleased to learn that the Federal Bureau of Investigation ("FBI") has launched a Task Force to assist companies and the public in identifying foreign manipulation efforts through social media platforms. We believe that Federal law enforcement agencies are uniquely positioned to access, synthesize, and comprehend disparate sources of intelligence, and to alert the public, Congress, and social media companies of their findings in a way that provides broader picture of the activity.

RESPONSE TO WRITTEN QUESTION SUBMITTED BY HON. JERRY MORAN TO CLINT WATTS

Question. What is the benefit to social media companies employing human verification systems? How prevalent are the use of "social bots" by extremists?

Answer. Social bots allow the replication of anonymous accounts to promote and distribute falsehoods or manipulated truths at such high volumes that they alter the perceptions of reality and blur the line between fact and fiction. Human verification systems will prevent the computer generation of bots by ensuring real people, not computer programs that replicate accounts, are behind the communications on their platforms. This will limit the creation of social bots and increase real human communication on social media platforms. Social media companies should also want to implement these controls as it improves the integrity of their systems and ensures the authenticity of accounts on the platform.

As for extremist use of bots, I am less familiar. The Islamic State did try to construct its own applications, but I'm not aware of any terrorist group employing social bots on a large scale. I'd recommend contacting J.M. Berger, co-author of ISIS: The State of Terror who may have greater insight into terrorists' use of social bots.

RESPONSE TO WRITTEN QUESTIONS SUBMITTED BY HON. TAMMY BALDWIN TO CLINT WATTS

Question 1. Mr. Watts, what lessons are other authoritarians and adversarial nations learning from Russia's malign activities on social media? Do you see others copying the Russian playbook?

Answer. Authoritarian regimes have begun using the Kremlin playbook largely to suppress internal political dissent. The two most recent and poignant examples of this phenomenon are the violent oppression and uprooting of the Rohingya population in Myanmar. Media reports suggest a large portion of the content on Facebook contains false information and smears against this minority Muslim population in Myanmar, which has created a refugee crisis in Bangladesh. The Myanmar government and many of its security services members now seek to rewrite history by denying the oppression of the Rohingya or that the group even exists.[1]

The Philippines leader Rodrigo Duterte also uses Facebook to suppress internal domestic challenges to his rule and promote his regime's authoritarian actions.[2] The most startling development has been U.S. political campaigns and associated political action groups, which have adopted similar tactics and in some cases have hired public relations companies to replicate the same methods employed by the Kremlin. In short, absent regulation and political leadership, everyone, in America and abroad, will be using these tactics on their political opponents.

---

[1] See Hannah Beech, "No Such Thing as Rohingya': Myanmar Erases a History." *New York Times,* 2 December 2017 available at: *https://www.nytimes.com/2017/12/02/world/asia/myanmar-rohingya_denial_history.html.*

[2] *See* Lauren Etter, "What Happens When the Government Uses Facebook As A Weapon?", *Bloomberg,* 7 December 2017 available at *https://www.bloomberg.com/news/features/2017–12–07/how-rodrigo-duterte-turned-facebook-into-a-weapon-wtih-a-little-health-from-facebook.*

Question 2. If so, how does that impact U.S. interests and how would you assess the current administration's efforts to respond?

Answer. The U.S. in nearly all theaters and countries around the world has seen its public image under attack and tarnished. As an asymmetric approach, nefarious social media influence against the U.S. is the most effective and least expensive method for undermining America's power around the world, affecting our military, economic and diplomatic strength and harming American companies. With regards to the current administration's response, I've briefed nearly every element of the U.S. Federal Government on the threat of Russian influence operations in social media, and I've not been able to glean what their strategy might be. I'm uncertain why there has not been concerted action to counter the Kremlin despite so much of U.S. public discussion and debate on the topic. I cannot imagine how an administration could do less to respond on such a critical national security issue.

This page intentionally left blank.

This page intentionally left blank.

This page intentionally left blank.