NEW KNOWLEDGE

Statement for the record from Renee DiResta, Director of Research, New Knowledge

Honorable Committee Members – My name is Renee DiResta, and I research influence operations and social network manipulation. I appreciate the opportunity to submit this written and verbal testimony to your committee.

Over the past decade, disinformation, misinformation, and social media hoaxes have evolved from a nuisance into high-stakes information war. Our frameworks for dealing with them, however, remain the same -- we discuss counter-messaging and counter-narratives, falling into the trap of treating this as a problem of false stories rather than as an attack on our information ecosystem. We find ourselves in the midst of an arms race, in which responsibility for the integrity of public discourse is largely in the hands of private social platforms, and determined adversaries continually find new ways to manipulate evolving feature sets and circumvent new security measures. It is critical to acknowledge that computational propaganda and disinformation is not about arbitrating truth, nor is it a question of free speech. Information Warfare is a cybersecurity issue, it is an ongoing national security issue, and it must be addressed through a collaboration between governments responsible for the safety of their citizens and private industry responsible for the integrity of their products and platforms.

Propaganda and malign narratives have existed for a very long time, but today's influence operations, which co-opt popular social platforms, are materially different – the propaganda is shared by our friends, often in the form of highly effective, shareable, immediately graspable memes. It is efficiently amplified by algorithms, and the campaigns achieve unprecedented scale. To conduct an operation, adversaries leverage the entire media ecosystem to push a narrative and manufacture the appearance of popular consensus. The operation is planned on one platform, such as a messaging or chat board. Content is created, tested, and hosted on others, such as Reddit, Pinterest and YouTube. It's then pushed to platforms like Twitter and Facebook, with standing audiences of hundreds of millions of people, and targeted at those most likely to be receptive to it. The platform's trending algorithms are gamed to make the content go viral - this often delivers the added benefit of mainstream media coverage, increasing attention via traditional media channels including television. If an operation is successful and the content gets wide distribution, or a manipulative Page or Group gains enough followers, the recommendation engine and search engine will continue to serve up the content on an ongoing basis.

We are here because the Internet Research Agency (IRA) employed this playbook, conducting an operation that leveraged our social networks to spread propaganda and disinformation directly to American citizens. Their operation likely began sometime in 2013, continued throughout the 2016 election cycle, and even increased on Instagram in 2017. While many accounts were shut down in 2017 as the tech companies began their investigations, Twitter accounts and Facebook pages associated with the IRA remain active. The IRA content on Facebook and Instagram alone had 293 million engagements; Facebook itself estimates 146 million users across the two platforms were affected. The Internet Research Agency's disinformation campaign was conducted on all the major platforms in the social network ecosystem. The presence of manipulated content on Facebook and Twitter is well-documented. In the case of Alphabet, YouTube, G+, Gmail, and Google Voice were all leveraged to either host content or to support personas. Reddit, Tumblr, and Medium have confirmed that they were misused; Twitter's Vine video app was co-opted as well, and IRA meme boards were discovered on Pinterest. Games and music apps were created and pushed to teenagers to download. Even popular game Pokémon Go was incorporated into the operation. Outside of social platforms, a number of websites were created to host original written content, many of which looked very much like citizen journalism-style blogs and think tanks. Topics ran the gamut, from social issues to concerns about wars, the environment, corporate greed, GMOs, energy policy, and immigration. Twitter accounts were created to masquerade as local news stations. White House petitions were either created or co-opted to engineer a perception of social consensus. Dozens of Facebook Events were promoted, and activists were contacted personally via Messenger, to take the operation to the streets.

*"The IRA content on Facebook and Instagram alone had **293 million engagements**. Facebook itself estimates **146 million users** across the two platforms were affected."*

The Internet Research Agency's campaign pressed on a variety of socially divisive issues, but the primary focus was on racial tension. Despite YouTube's claim that the content attributed to the IRA on its platform was "not targeted to the US or to any particular sector of the US population", it appears that the overwhelming majority of the videos were related to issues of importance to the black community, particularly officer-involved shootings. Hundreds of thousands of Americans liked Facebook Pages with names like Blacktivist, Heart of Texas, and Stop All Invaders. The percentage of explicitly political content that mentioned candidates by name was small –

approximately 10% – but the political content targeting both right-leaning and left-leaning Americans was unified in its negativity toward the candidacy of Secretary Clinton. In pages targeting the left, this included content intended to depress voter turnout among black voters, or to paint Secretary Clinton in a negative light as compared to candidates Jill Stein or Senator Bernie Sanders. Only the social networks that hosted this campaign are in a position to gauge its full impact in changing voter attitudes on their respective platforms. However, independent of its impact, the fact that it was attempted, went undetected, and achieved such significant reach is sufficient cause for concern.

Although this hearing was convened because of the Internet Research Agency's interference in the 2016 election, Russia was not the first to target American citizens with propaganda using the social ecosystem. In 2014, the co-opting of social communication infrastructure rose to mainstream awareness in the United States as ISIS established a virtual caliphate, using every social app imaginable to push propaganda boldly and transparently, using the features of our of social ecosystem in precisely the way they were meant to be used: to build an audience and connect with followers. This was a visible indication that the tools built to enable marketers and messengers and friends to communicate could be co-opted and misused; the ensuing debate about what to do about the problem made it apparent to anyone watching that no one was in charge, and that American companies, American civil society organizations, and the American government were deeply divided on how to respond to the threat. That confusion continues even as the threat expands beyond extremists and state actors: The Wall Street Journal recently revealed that a private intelligence company, Psy-Group, openly marketed their ability to conduct similar types of influence operations to impact the 2016 election.

As the internet has evolved, we've seen the consolidation of users into large standing audiences on a small handful of social networks. This infrastructure has been a phenomenal tool for small businesses to reach customers, and for the previously voiceless to find a voice. But like any tool without appropriate safeguards, it can be misused. These platforms employ gameable algorithms and facilitate personalized targeting that is enabled by the ongoing collection of extensive amounts of personal data. As a result, social networks continue to be the most effective vector to manipulate public sentiment and cause lasting damage to our democratic process. To combat this evolving threat, we have to address those structural weaknesses and design an effective deterrence strategy.

Individually, several social platforms have begun to take steps to reduce the spread of disinformation, by disrupting economic incentive structures, reducing the spread of clickbait headlines, and reducing the granularity of targeting criteria that were used to push malicious content directly to subsets of the American people. Political ad content

on Facebook and Twitter is somewhat public now; we look forward to this database being searchable via API, better equipping researchers and journalists to understand our paid political conversation. Several platforms are beginning to take source quality into account, which may help curb the ability of manipulative propaganda websites to reach their audience. These steps, several of which were inspired by prior hearings in this chamber, are a good start. But as platform features and protections change, determined adversaries will develop new tactics.

In addition to the ongoing exploitation of social divisions, targeting of elections, and disinformation about geopolitical events (such as the conflict in Syria), campaigns targeting U.S. industry have emerged and are thriving. Influence operations are increasingly appealing to a variety of actors: ideological true believers, non-state extremists, economically-motivated enterprises, and State Actors. The last of these requires a whole-of-government defense strategy, as it's unlikely that commercial platforms will be able to compete with the sophistication of well-resourced and motivated hostile foreign government experienced in bypassing common security checks.

*"In addition to the ongoing exploitation of social divisions, targeting of elections, and disinformation about geopolitical events, campaigns targeting U.S industry have **emerged and are thriving**"*

The gaps in our ability to combat this type of information warfare became apparent while attempting to address ISIS propaganda: the U.S. government was legally constrained in its ability to respond, and the social platforms proved slow to act as extremist content, assisted by platform targeting algorithms, easily made its way into the social feeds of Americans. That wake-up call fell on deaf ears as our adversaries prioritized, deployed, and perfected their influence operation capabilities. They were able to exploit gaps in our intelligence community's authorities and take advantage of our commitment to civil liberties; this left social platforms in the impossible position of having to individually respond to this global threat, which has resulted in the implementation of inadequate solutions and self-serving defensive policies on the part of those private companies.

Several years after the threat emerged, the U.S. Government and the tech industry respectively took small steps towards combating this threat by establishing the Global Engagement Center and the Global Internet Forum to Counter Terrorism. The focus of

the latter is still solely terrorism, although the Global Engagement Center's mandate has expanded to countering foreign propaganda. The DOJ and NSA & Cyber Command's recent announcements that they will prioritize the mitigation and prosecution of this activity is a positive sign. However, addressing this asymmetric threat requires a 21st century Information Operations Doctrine, the implementation of a global real-time detection and deterrence strategy, and the cooperation of private industry, press, law enforcement, and the intelligence community.

The evolution of social media propaganda and influence techniques will bring serious threats. We should anticipate an increase in the misuse of less popular and less resourced social platforms, and an increase in the use of peer-to-peer messaging services. We believe that future campaigns will be compounded by the employment of witting or unwitting U.S. Persons through whom these state actors will filter their propaganda, in order to circumvent detection by social platforms and law enforcement. We should anticipate the incorporation of new technologies, such as videos and audio produced by artificial intelligence, to supplement these operations, making it increasingly difficult for citizens to trust their own eyes.

*"This will be one of the **defining threats of our generation**. Influence operations exploit divisions in our society using vulnerabilities in our information ecosystem. They take advantage of America's commitment to **freedom of speech** and free flow of ideas."*

This will be one of the defining threats of our generation. Influence operations exploit divisions in our society using vulnerabilities in our information ecosystem. They take advantage of America's commitment to freedom of speech and the free flow of ideas. The social media platforms cannot, and should not, be the sole defenders of democracy and public discourse. In that light, here are several recommendations we are proposing toward achieving the goal of restoring integrity to the information ecosystem:

First, to address the most pressing short-term issue, we recommend immediate government action to identify and eliminate malign influence campaigns and to educate the public in preparation for the 2018 elections.

Second, this domestic defense must be complimented by an updated global IO doctrine and international detection and deterrence strategy, with the goal of mitigating foreign influence targeting our allies, including the clear delegation of responsibility of this activity within the U.S. Government. Taking example from the U.S. Government's cyber security response over the past decade, we must implement legislation that defines and criminalizes foreign propaganda that targets not just our political process but also addresses the targeting of commercial industry and social issues. Empowering law enforcement with updated legal tools to investigate and prosecute sophisticated foreign propaganda is essential to combatting this threat in the age of information warfare.

Third, the private tech platforms must be held accountable to ensure that they are doing their utmost to manage and mitigate the pervasiveness of disinformation and manipulative narratives in our privately-owned public squares. A number of regulatory frameworks are on the table, including mandating that automated accounts be labeled, limiting high-frequency advertising practices, and curtailing and reporting inauthentic accounts. Regardless of which is chosen, and whether these policies are implemented voluntarily by platforms (self-regulatory action) or via formal regulation, the incorporation of oversight is key.

Finally, given that this asymmetric persistent threat impacts our social, geo-political, and economic spheres, and given the sophistication of its tradecraft, we need new structures for cooperation and information sharing between the public and private sectors. Formal partnerships between security companies, researchers, and government will be essential to defending our values, our democracy, and our society.

Thank you.