Atlantic Council
EURASIA CENTER

Democratic Defense Against Disinformation 2.0

Alina Polyakova and Daniel Fried

Facebook turns to artificial intelligence to fight hate and misinformation in Myanmar
'FAKE NEWS VICTIMS' MEET WITH TWITTER AND FACEBOOK
Facebook shuts down Israel-based disinformation campaigns
Google announces plan to combat spread of misinformation
In Ukraine, Russia Tests a New Facebook Tactic in Election Tampering
Google Pledges $300 Million to Clean Up False Ne
Alarming Lessons from Facebook's Push to Stop Fake News in India
FACEBOOK REMOVES FRESH BATCH OF INNOVATIVE, IRAN-LINKED FAKE ACCOUNTS
FAR-RIGHT PROPAGANDA FLOODS FACEBOOK AHEAD EU ELECTIONS
inland is winning the war news. What it's learned may be ucial to Western democracy
Facebook and Twitter disal disinformation campaign w Iran
Inside Facebook's war room: the battle to protect EU elections
FACEBOOK FOUND HOSTING MASSES OF FAR RIGHT EU DISINFORMATION NETWORKS
YouTube to Remove Thousands of Videos Pushing Extreme Views
State Department to take a ste into the digital age in effort to counter disinformation
A.I. Won't Solve the Fake News Problem
Troll Watch: On Eve Of Mic Elections, Misinformation Ta Bigger Political Issues
Facebook deletes fake accounts used in pro-Iranian influence campaign
Twitter And Facebook Crack Down On Iran-Linked Disinformation Scheme
Google, Facebook, Twitter rapped for not doing enough in EU fake news fight
Facebook Bans More Myanmar Military Linked Accounts for Spreading Propagan
Russia-linked hack 'bid to discredit' UK anti-disinformation campaign - Foreign Office
Blame politician for fake news, RT chief tells Whitehall media forum
Disinformation headlines Russian disinformation on YouTube draws ads, lacks warning labels: researchers
Russia's Playbook for Social Media Disinformation Has Gone Global
As Midterms Approac Facebook Ramps Up Disinformation Fig
241m Europeans 'may have received Russian-linked disinformation'
There is no guarantee the truth can spread more successfully than a
RUSSIAN TROLLS PROMOTED ANTI-VACCINATIO PROPAGANDA THAT MAY HAVE CAUSED MEASLE

Democratic Defense Against Disinformation 2.0

Alina Polyakova and Daniel Fried

ISBN-13: 978-1-61977-592-3

This report is written and published in accordance with the Atlantic Council Policy on Intellectual Independence. The author is solely responsible for its analysis and recommendations. The Atlantic Council and its donors do not determine, nor do they necessarily endorse or advocate for, any of this report's conclusions.

June 2019

TABLE OF CONTENTS

INTRODUCTION

This second edition of the paper, *Democratic Defense Against Disinformation*, seeks to capture the rapid development of policy responses to the challenge—especially by governments and social media companies—since initial publication in February 2018. The authors stand by the fundamentals of the earlier analysis and recommendations: that democratic societies can combat and mitigate the challenge of foreign disinformation while working within democratic norms and respect for freedom of expression. The first edition offered a "whole-of-society" vision, with policy suggestions for governments, social media companies, and civil-society groups. Since publication of the first edition, the European Union (EU) and, to a lesser degree, the US government, have taken actions that parallel some of these suggestions. For their part, social media companies have moved from an initial and unsustainable denial of the problem to a stance of willingness to help deal with it, though the depth of this commitment (and the effectiveness of the responses) has yet to be determined.

Collectively, democracies have moved beyond "admiring the problem," meaning a sort of existential despair in the face of a new threat not easily managed. They have now entered a period of "trial and error," in which new ideas and solutions for countering, and building resilience against, disinformation are being tested, though unevenly and with setbacks. Meanwhile, the disinformation challenge has evolved and advanced, as state and nonstate actors deploy new technologies, develop new methods of exploitation, and adapt to responses.

Supply and demand of disinformation. Many of our recommendations address the "supply side" of disinformation; i.e., they recommended, and continue to recommend, policies and actions to limit the influx of disinformation into US and other media ecosystems. But, tools to block disinformation will be imperfect, and some degree of disinformation will be part of the media landscape for the indefinite future. Addressing the "demand side" of disinformation—i.e., reducing general social acceptance of fabrications and distortions—is likely to be more important for sustained societal immunity. It is critical that governments, social media companies, and civil-society groups invest in long-term resilience against disinformation, including raising social awareness of disinformation and encouraging digital literacy education, including how to discern fabricated or deceptive content and sources.

Democratic Defense Against Disinformation 2.0 will review developments over the past year, assess their effectiveness thus far, and offer suggestions for next steps.

Framing Solutions (Against a Moving Target)

Vladimir Putin's Russia was perhaps first among major powers to deploy techniques of full-spectrum, state-sponsored disinformation for the digital age—the intentional spread of inaccurate information designed to influence societies. It will not be the last. Other state actors with perhaps greater capabilities, such as China, and nonstate actors, such as terrorist groups with a higher tolerance for risk, will adapt the disinformation toolkit to undermine democracies or are already doing so. There is nothing new about state propaganda and other means of political subversion ("active measures" was the term of art for Soviet efforts of this kind). But, digital and social media, in combination with more traditional methods, offer new means to achieve traditional ends. Russia's democratic and pro-Western neighbors, especially Ukraine, Georgia, and the Baltic states, have contended with Russian disinformation attacks for years.

Other targets of state-sponsored disinformation campaigns—the United States and some Western European countries—woke up late to the challenge, with the United States doing so only after its 2016 presidential election, in which Russia played a large and malign role. The Department of Justice Special Counsel Report[1] and two independent reports, prepared for the US Senate's Select Committee on Intelligence and published in December 2018, detail Russia's disinformation tactics during and after the 2016 US elections, including by the Russian-government-supported Internet Research Agency (IRA), the now-notorious St. Petersburg troll farm.[2] The February 2018 Department of Justice indictment of thirteen Russian operatives involved in the IRA information operations provides the most in-depth

[1] Robert S. Mueller, III, "Report On The Investigation Into Russian Interference In The 2016 Presidential Election" (US Department of Justice, Washington, DC, 2019), https://www.justice.gov/storage/report.pdf.

[2] Renee DiResta et al., *The Tactics & Tropes of the Internet Research Agency,* New Knowledge, December 17, 2018, https://cdn2.hubspot.net/hubfs/4326998/ira-report-rebrand_FinalJ14.pdf; Philip N. Howard et al., *The IRA and Political Polarization in the United States*, Oxford Internet Institute, August 22, 2018, https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/12/The-IRA-Social-Media-and-Political-Polarization.pdf.

Vladimir Putin's Russia was perhaps first among major powers to deploy techniques of full-spectrum, state-sponsored disinformation for the digital age—the intentional spread of inaccurate information designed to influence societies. *Photo Credit:* The Russian Presidential Press and Information Office

research to date about the internal machinery of the Russian operation.[3]

Exposure is not enough. Disinformation campaigns are not going away. Quite the opposite—other malicious state actors with an interest in undermining democracies, including Iran, North Korea, and China, are learning from Russian tactics. Meanwhile, the tools of attack are evolving and adapting to democratic responses. State-sponsored disinformation campaigns aim to amplify existing social divisions and further polarize democratic societies. As such, they don't stop when the ballot box closes. Still, due to the high level of attention and consequential outcomes, elections provide an ideal high-impact opportunity for this type of influence operation.

Ahead of elections throughout Europe and North America in 2019 and 2020, governments, social media companies, and civil-society groups must learn from each other and accelerate the implementation of best practices to defend against disinformation.

Democracies are learning that means of defense and norms of resilience applicable to traditional propaganda and subversion are inadequate to meet the present danger. Also, disinformation techniques will continue to evolve. For example, innovation in artificial intelligence (AI) is producing "deepfakes" and other "synthetic media" products—video and audio manipulation with the capability to manufacture the appearance of reality, such as nonexistent, but real-looking, remarks by a political leader. As these tools become more low cost and accessible, they will become perfect weapons for information warfare.[4] Such technologies could drive the next great leap in AI-driven disinformation.[5] More generally, disinformation techniques are shifting from the use of simple automated bots to more sophisticated interaction with (and manipulation of) domestic groups, extremist and otherwise, through various forms of impersonation and amplification of organic posts by domestic actors. Thus, it may be increasingly difficult to disentangle foreign-origin disinformation from domestic social media conversations. Rather than trying to break through the noise, the new strategy aims to blend in with the noise—obfuscating manipulative activity and blurring the line between authentic and inauthentic content.

The first edition of *Democratic Defense Against Disinformation* offered recommendations on ways democratic governments and free societies can combat disinformation, while respecting the norms and values of free expression and the rule of law.[6] As democratic countries learned in the struggle against Soviet communism, we need not become them to fight them: democratic societies should not fight propaganda with propaganda,

[3] US Department of Justice, "Internet Research Agency Indictment" (US Department of Justice, Washington, DC, 2018), https://www.justice.gov/file/1035477/download.

[4] + Polyakova, *Weapons of the weak: Russia and the AI-driven asymmetric warfare*, November 15, 2018, https://www.brookings.edu/research/weapons-of-the-weak-russia-and-ai-driven-asymmetric-warfare/.

[5] Alina Polyakova and Spencer Phipps Boyer, *The future of political warfare: Russia, the West, and the coming age of global digital competition*, March 2018, https://www.brookings.edu/research/the-future-of-political-warfare-russia-the-west-and-the-coming-age-of-global-digital-competition.

[6] Daniel Fried and Alina Polyakova, *Democractic Defense Against Disinformation*, March 5, 2018, https://www.atlanticcouncil.org/images/publications/Democratic_Defense_Against_Disinformation_FINAL.pdf.

nor should they turn to censorship. Freedom of expression and US First Amendment protections do not rob free societies of options. US law prohibits foreign participation in US elections, broadly defined, and permits extensive regulation of commercial advertisements (e.g., outright bans on broad categories, such as smoking). In general, foreign persons, especially outside the United States, do not enjoy full First Amendment protections. Automated (e.g., bot) social media accounts also do not necessarily have First Amendment rights. EU/European options are still broader and include legal and regulatory options for enforcing versions of media fairness and impartiality. And, in some countries this includes authority to ban certain forms of hate speech, though this has drawbacks.

This paper's initial recommendations covered three levels of action, summarized below:

- **Governments,** starting with the United States, European national governments, the EU, and NATO, should introduce and enforce transparency standards, including with respect to foreign-origin political and issue ads on both traditional and social media, and otherwise monitor and notify their publics in real time about the activities of foreign propaganda outlets. To that end, governments should establish "rapid-response task forces" to inform government officials and, as needed, allies, of emerging disinformation campaigns that threaten national security. Governments should also expand institutional capacity, building on the EU's East StratCom, NATO's StratCom Center of Excellence in Riga, the Helsinki Hybrid Center of Excellence, the Department of Homeland Security's task force to counter malign foreign influence, and the US State Department's Global Engagement Center, to identify and expose Russian and other disinformation campaigns.

- **Social media companies** have a responsibility to stop denying and start mitigating the problem of foreign disinformation. The paper specified that they should: identify and label overt foreign propaganda outlets (RT and Sputnik, for example) as state-sponsored content; experiment with labeling and taking down automated and fake accounts; and redesign algorithms to demote, de-rank, or mute known propaganda content and suspect or mislabeled content, based on findings from third-party fact checkers.

- **Civil society groups**, especially the tech-savvy "digital Sherlocks" skilled at identifying disinformation—such as Ukraine's StopFake, Bellingcat, the Atlantic Council's Digital Forensic Research Lab, the Alliance for Security Democracy's Hamilton 68, EU DisinfoLab, and the Baltic Elves—have proven skilled at identifying coordinated disinformation activity driven by inauthentic accounts, often in real time.[7] They, and other innovative startups, are better able than governments to develop the tools to identify emerging disinformation techniques (e.g., synthetic media and deepfakes). Governments, social media companies, and philanthropic organizations should fund such innovators and establish regular lines of communication with them, e.g., through designated points of contact.

The paper's biggest single recommendation was that the United States and EU establish a **Counter-Disinformation Coalition,** a public/private group bringing together, on a regular basis, government and non-government stakeholders, including social media companies, traditional media, Internet service providers (ISPs), and civil society groups. The Counter-Disinformation Coalition would develop best practices for confronting disinformation from nondemocratic countries, consistent with democratic norms. It also recommended that this coalition start with a voluntary **code of conduct** outlining principles and agreed procedures for dealing with disinformation, drawing from the recommendations as summarized above.

In drawing up these recommendations, we were aware that disinformation most often comes from domestic, not foreign, sources.[8] While Russian and other disinformation players are known to work in coordination with domestic purveyors of disinformation, both overtly and covertly, the recommendations are limited to foreign disinformation, which falls within the scope of "political warfare." Nevertheless, it may be that these policy recommendations, particularly those focused on transparency and social resilience, may be applicable to combatting other forms of disinformation.

[7] StopFake, accessed April 3, 2019, https://www.stopfake.org/en/news/; Bellingcat, accessed April 3, 2019, https://www.bellingcat.com/; Digital Forensic Research Lab, accessed April 3, 2019, https://medium.com/dfrlab; Hamilton 68, accessed April 3, 2019, https://securingdemocracy.gmfus.org/hamilton-68/; EU DisinfoLab, accessed April 3, 2018, https://www.disinfo.eu/; "Baltic Elves Fight Kremlin Trolls," Radio Free Europe/Radio Liberty, May 16, 2017, https://www.rferl.org/a/baltic-eleves-vs-trolls/28498222.html.

[8] Jack Nicas, "Alex Jones and Infowars Content Is Removed From Apple, Facebook and YouTube," New York Times, August 6, 2018, https://www.nytimes.com/2018/08/06/technology/infowars-alex-jones-apple-facebook-spotify.html.

PROGRESS REPORT

The initial paper noted that the West's response to disinformation was developing rapidly, albeit from a low baseline, and anticipated significant policy progress soon. This (admittedly easy) prediction has been borne out, though in unexpected ways. The European Union has moved faster than anticipated to outline and seek to operationalize solutions, though it is too early to judge the implementation or impact. The United States—perhaps reflecting the domestic political complexity of addressing Russian disinformation given the controversy over the 2016 US presidential election—has struggled to frame consistent policy responses, though the US administration has started to act at the operational level. Social media companies, stung by their initial (and risible) denial of the disinformation problem and a host of other embarrassing news about their questionable behavior, have sought to be seen as, and possibly to be, part of the solution.

Europe Seeks Solutions

European assessments of the disinformation challenge continue to vary. Prompted by experiences in recent elections, some EU member states are now aware of Russia's aggressive disinformation tactics and determined to deal with them (the Baltic states, United Kingdom, Sweden, Finland, Denmark, and possibly France and others). Some member states that earlier believed that state-sponsored disinformation had nothing to do with them now know that it does (e.g., Spain, after learning of Russian involvement in the Catalonia referendum, and some in Greece in the aftermath of Russian efforts to derail the Greek-North Macedonia agreement resolving the issue of North Macedonia's name). Others appear to remain in a state of denial about the challenge. These differences of assessment are reflected in the EU's internal debates.

Given this background, the EU's progress in framing solutions to the disinformation challenge is impressive. The EU's policy approach to countering disinformation is contained in the following documents, all prepared in 2018:

- "Tackling Online Disinformation: a European Approach," prepared by the EU Commission and published on April 26, 2018;[9]
- a voluntary "Code of Practice on Disinformation" prepared by the EU Commission, published on September 26, 2018, and accepted on October 16 by Facebook, Google, Twitter, and Mozilla, as well as European trade associations representing online platforms and the advertising industry;[10]
- an EU Commission "Progress Report" on implementation of the April recommendations published December 5, 2018;[11] and
- the "Action Plan against Disinformation" jointly prepared by the EU Commission and European External Action Service (the EU "foreign ministry"), also published on December 5, 2018.[12]

Taken together, these documents lay out principles and action items that, if implemented and sustained by political will, would mark the end of Europe's admiring the problem of disinformation and the beginning of its credible efforts to contain and reverse it. The December papers are significantly stronger than the initial April report, which is an encouraging indicator of policy momentum in Brussels and key capitals. The EU appeared to be driven in part by determination to have in place at least preliminary defensive measures prior to the May 2019 European elections, especially given the knowledge that Russian disinformation campaigns were amplifying European extremist groups, while undermining establishment parties and centrists. Aggressive Russian use of disinformation to cover up its use of nerve gas in an attempt to murder a former Russian intelligence officer in Salisbury, United Kingdom (UK), in March 2018 also appears to have contributed to Europe's determination to address the disinformation challenge.

The EU's principles in countering disinformation include fidelity to freedom of expression. Its recommended

[9] European Commission, "Tackling online disinformation: A European Approach" (European Commission, Brussels, 2018), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0236&from=EN.

[10] European Commission, "Code of Practice on Disinformation," (European Commission, Brussels, 2018), https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation.

[11] European Commission, "Report from the Commission on the implementation of the Communication "Tackling online disinformation: a European Approach" (European Commission, Brussels, 2018), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:794:FIN.

[12] European Commission "Action Plan against Disinformation" (European Commission and European External Action Service, Brussels, 2018), https://eeas.europa.eu/sites/eeas/files/action_plan_against_disinformation.pdf.

European Commission progress reports covering implementation of the Code of Practice by social media company signatories indicate a mixed picture.

actions focus on transparency, exposure of disinformation, and encouragement of social media companies (with an implied warning of regulatory mandates) to help by identifying and combatting, rather than accommodating, abuse of their platforms by malign Russian and other disinformation actors. The EU's approach is consistent with the fundamentals of the initial paper.

The major elements of EU-proposed action include the following.

- **Strengthening EU capacity to identify and expose disinformation**. This includes a call to double the budget for strategic communications and increase personnel prior to future European parliamentary elections. The EU's **EastStratCom** unit, based in Brussels, has a mandate to identify and expose Russian disinformation, but has lacked consistent political support and, perhaps as a result, has been inadequately resourced. Happily, this appears to be changing.

- **Establishment of a new EU rapid alert system (RAS) to expose current disinformation campaigns in real time.** This new mechanism is supposed to have a capacity for real-time response to disinformation and designated links to each member state government, as well as less-specified means to exchange information with NATO and Group of Seven (G7) governments. The RAS' mandate will include providing alerts about disinformation campaigns, both publicly and officially. EU experts are thinking of expanding the RAS' links to independent fact-checking and other independent groups, though this will apparently not be part of the RAS' initial capability. The EU also intends to press social media platforms to cooperate with the RAS. The EU "Action Plan" called for the RAS to have an initial operational capacity—probably meaning links to member state governments—by March 2019, two months before the EU parliamentary elections. As of early April 2019, the RAS appeared to exist mostly on paper, though there is hope that this will change.

- **Launch of a voluntary "Code of Practice on Disinformation," agreed to by key social media companies.** The code's language is general and, in some cases, calls only for indirect steps; it has been criticized on that basis. However, the code is the most ambitious effort to date to specify policy norms for social media companies to mitigate the exploitation of their platforms by purveyors of disinformation. Its effectiveness will depend, in part, on whether the EU has political will to insist on action to implement its laudable objectives, given its general, non-binding terms. EU officials appear aware of the challenge, and the EU documents make explicit that if the code proves ineffective, the next step may be formal regulatory action. The policy commitments cover several aspects.

- ◆ ***Scrutiny of ad placements***, including restricting placement of ads by demonstrated purveyors of disinformation, with determinations made in cooperation with fact-checking organizations;
- ◆ ***Transparency of political and issue-based advertisements.*** These should be accurately labeled, and their sponsors identified;
- ◆ ***Integrity of services,*** meaning that social media companies have committed to put in place policies to identify and remove fake accounts, including bots;
- ◆ ***Empowering consumers,*** meaning that social media companies have committed to "help people make informed decisions," inform users of potential disinformation, and increase transparency. This is the weakest commitment in the code, long on generalities and short on specifics. It does not, for example, require social media companies to de-rank or mute known deceptive sites; and
- ◆ ***Empowering the research community,*** meaning that social media companies will support independent "good-faith" efforts to research disinformation. This does not constitute a commitment to provide key information, e.g., algorithms about placement of information, but merely "not to prohibit or discourage" such research.

- ■ **European Commission progress reports** covering implementation of the Code of Practice by social media company signatories indicate a mixed picture. The commission has recognized efforts by social media platforms to take down fake accounts, restrict ad purchasing by purveyors of disinformation, identify and block inauthentic behavior, and take other steps to meet the (general) commitments outlined in the code. But, it also noted insufficient information provided by social media companies, and urged specific next steps, including calling on platforms to take more serious actions to address transparency, particularly with respect to political ads. The commission is issuing monthly progress reports to test social media companies' response to their commitments.[13]
- ■ **Improving social resilience against disinformation, including: creating a European network of independent fact checkers; launching a secure online platform addressing disinformation; exploring means of reliable identification of information suppliers; and supporting long-term social media literacy.** Supporting nongovernment, independent fact checkers and researchers—e.g., the existing civil society groups that already possess the skills to expose disinformation campaigns, or journalists' associations—and potentially bringing them into close cooperation with the official rapid alert system, is a sound idea. However, government bodies, even the relatively nimble EastStratCom team, are unlikely to match the best civil society groups in cutting-edge techniques for uncovering disinformation. The EU may also want to empower its researchers to obtain access to social media company data, as is provided in a general way in the Code of Practice. As of late March, little progress was reported in implementing this objective.

In addition, in a wide-ranging article on Europe published March 4, 2019, French President Emmanuel Macron proposed a new "European Agency for the Protection of Democracies," which included providing each EU member state with expertise to protect election processes against cyberattacks and manipulation.[14] While his proposals did not go into detail, and may be only notional, his engagement could boost EU efforts to turn its policy framework into action.

The G7 in the Game

The Canadian G7 Summit in June 2018 produced the "Charlevoix Commitment on Defending Democracy from Threats," which included general language about cooperation to share lessons and information to counter disinformation, including a commitment to

[13] European Commission, "Code of Practice against disinformation: Commission calls on signatories to intensify their efforts," (European Commission, Brussels, 2019), http://europa.eu/rapid/press-release_IP-19-746_en.htm; "Second monthly intermediate results of the EU Code of Practice against disinformation," (European Commission, Brussels, 2019), https://ec.europa.eu/digital-single-market/en/news/second-monthly-intermediate-results-eu-code-practice-against-disinformation. Latest report at time of writing: http://europa.eu/rapid/press-release_STATEMENT-19-2174_en.htm.

[14] Emmanuel Macron, "Renewing Europe," Project Syndicate, March 4, 2019, http://prosyn.org/kCUclh5.

The Canadian G7 Summit in June 2018 produced the "Charlevoix Commitment on Defending Democracy from Threats," which included general language about cooperation to share lessons and information to counter disinformation. *Photo Credit:* Casa Rosada (Argentina Presidency of the Nation)

"Establish a G7 Rapid Response Mechanism [RRM] to strengthen our coordination to identify and respond to diverse and evolving threats to our democracies, including through sharing information..."[15] The Canadian government has now established such a mechanism, reportedly with a staff of twelve and a mandate to analyze disinformation campaigns, with an initial focus, which included the May European elections. The RRM has established the State Department's Global Engagement Center (GEC) as its US government point of contact. Thus, the RRM could link up with the EU's RAS and the State Department's GEC, creating an institutional basis to share information about disinformation campaigns, hopefully in near real time. The RRM has already started to share information about current counter-disinformation projects beyond governments to the wider expert community.

The United States Tries to Get a Grip

The United States has fallen behind the EU, both in terms of conceptual framing and calls for concrete actions to meet the disinformation challenge. Ahead of the 2018 US congressional elections, Dan Coats, the director of national intelligence, warned that "lights were blinking red" on the threat of foreign interference.[16] But, Coats' statement suggests that the US government's priority was election security in the context of the 2018 US congressional elections, rather than addressing disinformation more broadly. While protecting election infrastructure (such as updating or replacing electronic voting machines, hardening security around voter-data storage, and harmonizing information sharing between local, state, and federal election authorities) is critically important, such a policy approach risks

[15] "Charlevoix commitment on defending democracy from foreign threats," January 2, 2019, https://international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2018-06-09-defending_democracy-defense_democratie.aspx?lang=eng.

[16] Dan Coats, "Director Coats: 'The Lights are Blinking Red,'" Video, 3:15, C-SPAN, July 13, 2018, https://www.c-span.org/video/?c4740341/director-coats-the-lights-blinking-red.

US Cyber Command began operations ahead of the 2018 congressional elections, to deter Russian operatives from potential interference. *Photo Credit:* US Army/Steve Stover

becoming episodic and narrow, whereas disinformation campaigns are continuous and consistent.[17]

Disinformation campaigns are ongoing between election cycles, and they need not focus on specific candidates or parties. For this reason, efforts to resist and build resilience against foreign disinformation must be sustainable and involve the whole of government.

The United States has made little progress in addressing this challenge over the last year. Still, there have been notable activities.

US Executive

- The Mueller Report (Volume One) outlined in detail Russia's interference in the 2016 US presidential campaign, including its cyber hacking, intelligence operations, and broad disinformation efforts.

- The Global Engagement Center, a State Department unit within the Public Diplomacy Bureau, initially intended to focus on countering extremist Islamist ideology, has turned to countering state-sponsored disinformation, with an appropriated budget of $120 million. The GEC has begun to disperse significant funding to civil society groups and private-sector partners, including: for research into disinformation and counter-disinformation tactics ($9 million); to journalists, fact checkers, and online influencers ($9 million); to partner organizations to support local counter-disinformation efforts; and to develop new technologies useful for counter-disinformation actions. The GEC is also actively participating in the G7 RRM, the UK-US bilateral coalition, and the Australian-led counter-interference effort.

- The Department of Defense funds the GEC (mandated under a National Defense Authorization Act). Beyond the traditional strategic communications functions conducted by its public-affairs apparatus, the Defense Department's policy arm has a narrow mandate to direct information support activities under the Special Operations/Low Intensity Conflict (SO/LIC) unit, typically in support of US military activities overseas or relations with allies and partners. US European Command (EUCOM) supports the broader US effort to counter Russia's disinformation,

[17] Kate Fazzini, "It's unlikely that hackers can change votes, but here's how they could cause midterm mischief," CNBC, August 20, 2018, https://www.cnbc.com/2018/08/17/how-homeland-security-is-gearing-up-to-protect-the-midterms-from-hackers.html.

and conducts information operations as part of its foreign-presence exercises in Europe, e.g., in Poland.[18]

- The mandate of the Federal Bureau of Investigation's (FBI) Foreign Interference Task Force (FITF), established in October 2017, includes engagement with US technology and social media companies to address the challenge of false personas and fabricated stories on social media platforms (as well as "hard" cybersecurity for voting infrastructure and other potential US election-related targets). At least one social media company has credited the FITF with advancing US government (USG)-social media company discussions to address the threat.

- US Cyber Command began operations ahead of the 2018 congressional elections, to deter Russian operatives from potential interference.[19] Cyber Command, together with the National Security Agency (NSA), reportedly developed information about Russian trolls and their activities, and alerted the FBI and Department of Homeland Security (DHS).[20] The operation followed the Department of Justice's February 2018 and July 2018 indictments of Russian individuals, intelligence officers, and companies involved in the Internet Research Agency and cyber operations against the US elections.[21] Cyber Command has reportedly sent messages to specific individuals active in disinformation operations, de facto outing them and their activities.

- The Department of Homeland Security—the agency that should lead the counter-disinformation efforts in the United States—has an internal working group focused on countering malign influence, but its activities seem more focused on technical election security around critical infrastructure than on broader disinformation.

- The US State Department's Bureau for European and Eurasian Affairs has established a new position—the senior adviser for Russian malign activities and trends (SARMAT)—tasked with coordinating policy on Russian malign influence.[22]

> **"While not a new policy, the Department of the Treasury used existing authorities to impose sanctions on Russian entities tied to disinformation efforts, including those directed at the 2016 US presidential election."**

- The State Department has worked with like-minded European governments to establish an informal consultative group on disinformation efforts. This is a worthy step, though it has not yet generated common actions or standards, or even framed solutions on a conceptual level.

- A USG interagency working group—the Russian Influence Group (RIG)—includes the relevant US government agencies, including DHS, the intelligence community, the State Department, and the Department of Defense. It has generated little high-level action and few recommendations to date and, as of this writing, no USG senior official has been empowered to take the lead on counter-disinformation efforts. Policy issues without senior-level ownership tend to drift.

- While not a new policy, the Department of the Treasury used existing authorities to impose sanctions on Russian entities tied to disinformation efforts, including those directed at the 2016 US presidential election. This included the sanctions designation on December 19, 2019, of entities and individuals tied to the IRA and nine GRU (military intelligence) officers. Material accompanying the Treasury Department's sanctions designations exposed details of Russian

[18] General Curtis Scaparrotti, "EUCOM Posture Statement 2018," (EUCOM, Washington, DC, 2018). https://www.eucom.mil/mission/eucom-posture-statement-2018.

[19] Julian E. Barnes, "U.S. Begins First Cyberoperation Against Russia Aimed at Protecting Elections," New York Times, October 23, 2018, https://www.nytimes.com/2018/10/23/us/politics/russian-hacking-usa-cyber-command.html.

[20] David Ignatius, "The U.S. military is quietly launching efforts to deter Russian meddling," Washington Post, February 7, 2019, https://www.washingtonpost.com/opinions/the-us-military-is-quietly-launching-efforts-to-deter-russian-meddling/2019/02/07/4de5c5fa-2b19-11e9-b2fc-721718903bfc_story.html?utm_term=.1cbbaf8bf3ae.

[21] US Department of Justice, "Internet Research Agency Indictment" (US Department of Justice, Washington, DC, 2018), https://www.justice.gov/file/1035477/download; "Indictment" (US Department of Justice, Washington, DC, 2018), https://www.justice.gov/file/1080281/download.

[22] A. Wess Mitchell, "U.S. Strategy Toward the Russian Federation," (US Department of State, Washington, DC, 2018), https://www.state.gov/p/eur/rls/rm/2018/285247.htm.

operation, including establishment of an online English-language website, "USA Really."

US Congress

- The 2019 National Defense Authorization Act (NDAA) added significant (albeit second-order) provisions defining the importance of countering disinformation for US national security.[23]

  - Cementing the role of the GEC by defining its counter-disinformation task within the parameters of US national security, likely securing the center's longer-term funding in future iterations of the NDAA.

  - Defining "malign influence" as "the coordinated, integrated, and synchronized application of national diplomatic, informational, military, economic, business, corruption, educational, and other capabilities by hostile foreign powers to foster attitudes, behaviors, decisions, or outcomes within the United States."

  - Authorizing the establishment of a new position in the National Security Council (NSC) responsible for coordinating the interagency process for countering foreign malign influence. This NSC director-level position now exists and was filled at the time of this writing.

- The Honest Ads Act, introduced in October 2017 and likely to be reintroduced in the current Congress, would require that political ads be identified as such on social media platforms.[24] On one level, the legislation would address only a small number of online ads (those strictly defined as sponsored by a political candidate or campaign). But, by making social media companies liable should they provide a platform for foreign expenditures aimed at influencing US elections (already prohibited under US campaign-finance law), the Honest Ads Act could conceivably curtail Russian-placed and other foreign-placed issue ads with a partisan purpose, including ads placed under hidden or misleading sponsorship. In any case, the legislation has not moved through either chamber of Congress.

  - Possibly to preempt legislation, both Twitter and Facebook have announced that they are implementing many Honest Ads Act requirements. The impact of these announcements is not yet clear and would be limited if these social media companies apply the Act's definitions narrowly.[25]

  - Even if Congress were to pass this legislation, its impact may not be great. Political ads make up a miniscule portion of the overall ads industry. In addition, ads are a secondary tool for spreading disinformation; organic posts, albeit under false identities, are becoming the major Russian disinformation tool.

- The Senate Special Committee on Intelligence (SSCI) commissioned two major reports on the IRA's tactics and techniques, based on data shared by Twitter, Facebook, and Google.[26]

- The Senate introduced the Data Protection Act of 2018, which would have placed standards on what online service providers can do with end-user data. While the bill has not been reintroduced in the new Congress, it laid out the responsibilities of providers in handling user data, and it enjoyed wide support from platforms.[27] This legislation should be reintroduced. The Senate has reintroduced the Defending American Security from Kremlin Aggression Act of 2019 (DASKA); while mostly devoted to sanctions, it also "calls for the establishment of a National Fusion Center to Respond to Hybrid Threats, a Countering Russian Influence Fund to be used in countries vulnerable to Russian malign influence, and closer coordination with allies" [28] (sections 704, 705, and 706).

[23] US Government Publication Office, "National Defense Authorization Act For Fiscal Year 2019" (US Government Publication Office, Washington, DC, 2018), https://www.govinfo.gov/content/pkg/CRPT-115hrpt874/pdf/CRPT-115hrpt874.pdf.

[24] U.S. Congress, Senate, Honest Ads Act, S 1989, 115th Congress, 1st session, introduced in the Senate October 19, 2017, https://www.congress.gov/115/bills/s1989/BILLS-115s1989is.pdf.

[25] Twitter Public Policy (@Policy), 2018, "Twitter is pleased to support the Honest Ads Act. Back in the fall we indicated we supported proposals to increase transparency in political ads," Twitter, April 10, 2018, 11:54 AM, https://twitter.com/policy/status/983734917015199744?lang=en; Mark Zuckerberg, 2018, "With important elections coming up," Facebook, April 6, 2018, https://www.facebook.com/zuck/posts/10104784125525891.

[26] DiResta, *The Tactics & Tropes*.

[27] "Shatz Leads Group Of 15 Senators In Introducing New Bill To Help Protect People's Personal Data Online," U.S. Senator Brian Schatz, December 12, 2018, https://www.schatz.senate.gov/press-releases/schatz-leads-group-of-15-senators-in-introducing-new-bill-to-help-protect-peoples-personal-data-online.

[28] US Congress, Senate, *Defending American Security from Kremlin Aggression Act of 2019*, S 482, 116th Congress, 1st session, introduced in Senate February 13, 2019, https://www.congress.gov/116/bills/s482/BILLS-116s482is.pdf.

Senator Mark Warner and Senator Marco Rubio speak to the press following their public remarks at a July 16 event on the Kremlin's interference in elections. *Photo Credit:* Atlantic Council

- In April 2019, Senators Mark Warner (D-VA) and Deb Fischer (R-NE) introduced the Deceptive Experiences to Online Users Reduction (DETOUR) Act,[29] which seeks to limit tactics used by social media platforms to steer users in various directions. DETOUR appears directed against domestic deceptive online practices, not foreign disinformation. But, the bill suggests that Congress is moving beyond pushing transparency (as in the Honest Ads bill) and toward introduction of more intrusive standards of conduct for social media companies. The precedent could provide a basis for legislation targeting disinformation.

- Congress's main activity on countering disinformation has been to hold hearings with social media companies and their executives. Since 2016, the US Congress has held five open hearings with Google, Facebook, and Twitter executives. These have captured intense media attention and may have generated political pressure on the social media companies to be seen as constructive with respect to disinformation issues.

- Congress has not, however, decided what specific steps it wants the social media companies to take to address issues of data privacy, online advertising transparency, algorithmic transparency with respect to placement of news, or more transparency on how the companies are identifying or de-prioritizing/de-ranking disinformation campaigns, or removing them from their platform.

- This lack of focus does not indicate a lack of understanding of the problem: Senator Warner has developed a draft white paper illustrating proposals that could become the basis for regulation. Like the EU policy papers and Code of Practice, it focuses on transparency (e.g., information about the origin of posts and accounts, mandatory labeling of bots, identification of inauthentic accounts, a degree of algorithmic transparency, and other measures).[30]

- Domestic political issues, such as allegations of partisan bias on the part of social media compa-

[29] US Congress, Senate, *Deceptive Experiences To Online Users Reduction Act*, S 1084, 116th Congress, 1st session, introduced in Senate April 9, 2019, https://www.congress.gov/116/bills/s1084/BILLS-116s1084is.pdf.

[30] Mark R. Warner, *Potential Policy Proposals for Regulation of Social Media and Technology Firms*, Reg Media, July 30, 2018, https://regmedia.co.uk/2018/07/30/warner_social_media_proposal.pdf.

nies, have become distractions in some congressional hearings.

Social Media Companies on the Spot

Twitter, Facebook, and Google have been in the hot seat for failing to do more to address the manipulation of their platforms, their handling of personal data and privacy, their initial lack of cooperation, and denials that the disinformation problem even existed. Over the last three years, the political climate has shifted in significant ways, as has the companies' approach to the problem. Since the 2016 elections, the US Congress held five hearings with social media companies' representatives.[31] The European Parliament has held one.[32] Under pressure, social media companies pivoted to a stance of acknowledging the problem and seeking to be seen as part of the solution. Yet, their efforts fall significantly short of addressing the dual challenges of platform manipulation and data collection and sharing. While disinformation will never be fully "solved" through the actions of social media companies alone, these large content distributors are the first line of defense against information manipulation and, with that, they have a large share of the responsibility for mitigating the problem.

Social media companies have made welcome changes in their policies, advertising rules, and postures vis-à-vis governments and researchers. While companies' approaches vary, they have, on the whole, become more open in sharing disinformation content with researchers, civil society organizations, and targeted governments. They have also established internal teams to identify "coordinated inauthentic behavior"—groups of pages or people working together to mislead others and manipulate the conversation.[33] The companies have announced changes to their algorithms aimed at reducing the visibility of content from automated spammers, suspicious impersonator accounts, and other distracting content. To prevent foreign influence in elections, Google, Facebook, and Twitter have also changed policies around political advertising during elections in the United States and Europe, including tougher certification requirements for individuals or groups seeking to promote candidates and parties.[34]

- "Takedowns" (i.e., removal) of coordinated inauthentic activity—accounts, pages, and content working to actively manipulate discourse on social media—have become a common tool for Facebook and Twitter. In January 2019, Facebook removed a network of pages and accounts from both Facebook and Instagram that were linked to Sputnik and its employees, primarily operating in Eastern Europe and Central Asia.[35] This operation followed a November 2018 takedown of accounts and pages with suspected links to the IRA.[36] So far in 2019, Facebook has also taken down coordinated inauthentic activity in India, Pakistan, Iran, Macedonia, Kosovo, the Philippines, Moldova, Romania, the United Kingdom, Indonesia, and Ukraine.[37] Facebook said it worked with tips from the FBI and pre-released the data for analysis by some independent researchers, but Facebook remains unwilling to share data with most researchers, making it difficult to assess the scope of manipulation on the platform.

[31] Rachel Kaser, "2018: The year Congress and social media collided," The Next Web, https://thenextweb.com//socialmedia/2018/12/28/2018-the-year-congress-and-social-media-collided/.

[32] Prashant S. Rao et al., "Mark Zuckerberg to Meet European Parliament Members Over Facebook's Data Use," New York Times, May 16, 2018, https://www.nytimes.com/2018/05/16/technology/zuckerberg-europe-data-cambridge-analytica.html.

[33] Nathaniel Gleicher, "Coordinated Inauthentic Behavior Explained," Facebook Newsroom, December 6, 2018, https://newsroom.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior. Other national governments, such as the UK and Germany, have also held public hearings with social media companies' representatives.

[34] Laura Kayali, "Twitter tightens political ad rules ahead of Eu election," Politico, February 19, 2019, https://www.politico.eu/article/twitter-expands-political-ads-rules-in-the-eu-ahead-of-parliament-election/. At the time of writing, Twitter and Google introduced new policies around political advertising only. Facebook also included issue ads.

[35] Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior from Russia," Facebook Newsroom, January 17, 2019, https://newsroom.fb.com/news/2019/01/removing-cib-from-russia/.

[36] "More Information About Last Week's Takedowns," Facebook Newsroom, November 13, 2018, https://newsroom.fb.com/news/2018/11/last-weeks-takedowns/.

[37] Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior and Spam From India and Pakistan," Facebook Newsroom, April 1, 2019, https://newsroom.fb.com/news/2019/04/cib-and-spam-from-india-pakistan/; Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior from Iran, Russia, Macedonia and Kosovo," Facebook Newsroom, March 26, 2019, https://newsroom.fb.com/news/2019/03/cib-iran-russia-macedonia-kosovo/; Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior From the Philippines," Facebook Newsroom, March 28, 2019, https://newsroom.fb.com/news/2019/03/cib-from-the-philippines/; Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior From Moldova," Facebook Newsroom, February 13, 2019, https://newsroom.fb.com/news/2019/02/cib-from-moldova/; Nathaniel Gleicher, "Removing Coordinated Inauthentic Behavior From the UK and Romania," Facebook Newsroom, March 7, 2019, https://newsroom.fb.com/news/2019/03/removing-cib-uk-and-romania/; Nathaniel Gleicher, "Taking Down Coordinated Inauthentic Behavior in Indonesia," Facebook Newsroom, January 31, 2019, https://newsroom.fb.com/news/2019/01/taking-down-coordinated-inauthentic-behavior-in-indonesia/; Nathaniel Gleicher, "Removing More Coordinated Inauthentic Behavior From Russia", Facebook Newsroom, May 6, 2019, https://newsroom.fb.com/news/2019/05/more-cib-from-russia/.

- Facebook also announced in early 2019 that it will establish an independent "oversight board" of experts to review Facebook's content decisions.[38] In essence, the board will complement Facebook's existing algorithmic review with human review, which was also recommended in the first edition of this report. It remains unclear, however, how the membership will be selected, which content the board will review, how independence will be ensured, and what information Facebook will share publicly. Ahead of the European Parliamentary elections in May 2019, the company opened a "war room" in Dublin to monitor manipulation on Facebook, Instagram, and WhatsApp (the company set up similar command posts ahead of the 2018 US elections and elections in India).[39]

- Twitter has focused on improving algorithmic review of accounts suspected of abuse, harassment, hate speech, or manipulation. In the first half of 2018, Twitter identified and challenged 232.4 million accounts that the company identified as engaging in spammy or manipulative behavior.[40] Twitter has also been transparent and willing to share data with researchers on a regular basis. The challenge that Twitter and other social media companies face today is less about automated accounts, however, and more about impersonators and organic disinformation shared by real people. The company also established an election-integrity data archive, which stores and makes available content on information operations from state actors, including Russia, Iran, and Venezuela.[41] And Twitter barred RT and Sputnik from advertising on its platform.[42]

- Google and YouTube, in contrast, have been less transparent about their activities to limit false and manipulated information on their platforms. In 2017, the then-chairman of Alphabet, Eric Schmidt, said that RT and Sputnik would be de-ranked from appearing at the top of Google result searches.[43] As of this writing, however, that policy change has not been implemented.

Facebook CEO Mark Zuckerberg. Since 2016, the US Congress has held five open hearings with Google, Facebook, and Twitter executives. These have captured intense media attention and may have generated political pressure on the social media companies to be seen as constructive with respect to disinformation issues. *Photo credit:* Anthony Quintano at https://flickr.com/photos/22882274@N04/41118890174

- In March 2018, Google announced that it would commit $300 million to support quality journalism and combat false information on its platform by, among other things, supporting youth education and partnering with research organizations.[44] In February 2019, the company published a white paper on its ongoing efforts to fight disinformation.[45] Identifying

[38] Nick Clegg, "Charting a Course for an Oversight Board for Content Decisions," Facebook Newsroom, January 28, 2019, https://newsroom.fb.com/news/2019/01/oversight-board/.

[39] Adam Satariano, "Facebook Opens a Command Post to Thwart Election Meddling in Europe", New York Times, May 5, 2019, https://www.nytimes.com/2019/05/05/technology/facebook-opens-a-command-post-to-thwart-election-meddling-in-europe.html.

[40] *Transparency* Report, Twitter, https://transparency.twitter.com/en/platform-manipulation.html.

[41] *Elections Integrity*, Twitter, https://about.twitter.com/en_us/values/elections-integrity.html#data.

[42] "Announcement: RT and Sputnik Advertising," Twitter Blog, October 26, 2017, https://blog.twitter.com/en_us/topics/company/2017/Announcement-RT-and-Sputnik-Advertising.html.

[43] Halifax International Security Forum, "PLENARY 6: SATELLITE ARMIES: THE RACE IN SPACE and HALIFAX CHAT – featuring Ms. Theresa Hitchens, Gen. John Hyten, Ms. Julie Perkins The Boeing Company..." Facebook, Video, 18 November, 2017, https://www.facebook.com/Halifaxtheforum/videos/1642381182492613/

[44] Philipp Schindler, "The Google News Initiative: Building a stronger future for news," The Keyword, March 20, 2018, https://www.blog.google/outreach-initiatives/google-news-initiative/announcing-google-news-initiative/.

[45] *How Google Fights Disinformation*, Google, February 2019, https://storage.googleapis.com/gweb-uniblog-publish-prod/documents/How_Google_Fights_Disinformation.pdf.

SIgn at YouTube headquarters in San Bruno, California. Unlike FacebookYouTube has received little scrutiny from lawmakers, despite reporting that it has promoted radical content. *Photo credit:* © BrokenSphere/Wikimedia Commons

and preventing manipulation by malicious actors remains a challenge, according to the paper, because "[a]lgorithms cannot determine whether a piece of content on current events is true or false, nor can they assess the intent of its creator just by reading what's on a page." Google's answer has been to prohibit impersonation in its terms of use, augment algorithmic review with human review, and invest in better detection of "spammy" behavior.

- YouTube is quickly catching up with Facebook in terms of the number of active users, which could soon make the video channel the most popular social media site in the world.[46] Compared to Facebook, however, YouTube has received little scrutiny from lawmakers, despite reporting that YouTube has promoted radical content.[47] In 2017, Google announced a change in YouTube's terms of service, which would de-rank, label, and remove from promotion potentially extremist or offensive videos, making them harder to find.[48] In January 2019, YouTube announced that it would change its recommendations system to no longer recommend conspiracy theories or highly misleading content—however, the change will affect less than 1 percent of YouTube content.[49] Google and YouTube frame these moves as part of a company effort to counter extremism and reduce the prevalence of harmful content. Certainly, limiting recommendations of low-quality content is an important part of a broader set of actions aimed at identifying and tackling foreign disinformation. However, as this paper has argued, content controls are both more controversial and probably less effective than dealing with fake and inauthentic online presences. YouTube's content-control-policy changes do not address how the company plans to deal with foreign manipulation or coordinated inauthentic behavior, likely making the impact on disinformation campaigns negligible.

Social media companies' actions to curb manipulation of their platforms vary, but the trend is positive, a welcome change from initial denials. But, to varying degrees, the companies still seem reticent to fully cooperate with lawmakers and researchers. For example, the two independent research organizations

[46] "Most popular social networks worldwide as of April 2019, ranked by number of active users (in millions)," Graph, Statista, April 2019, https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.

[47] Zeynep Tufekci, "YouTube, the Great Radicalizer," New York Times, March 10, 2018, https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html.

[48] Daisuke Wakabayashi, "YouTube Sets New Policies to Curb Extremist Videos," New York Times, June 18, 2017, https://www.nytimes.com/2017/06/18/business/youtube-terrorism.html.

[49] "Continuing our work to improve recommendations on YouTube," YouTube Official Blog, January 25, 2019, https://youtube.googleblog.com/2019/01/continuing-our-work-to-improve.html.

commissioned by SSCI to analyze the IRA data provided by Facebook, Twitter, and Google noted that the companies provided incomplete and, in some cases, unusable data. The firms have a long way to go to convince users and regulators of their commitment to tackle disinformation and manipulation.

One core problem remains: social media companies' business models are premised on increasing ad revenues, and sensationalist content sells best. The online advertising ecosystem that supports this business model has ballooned into a multibillion-dollar business, without much attention from regulators.

In addition, the tools social media companies have developed to maximize ad revenue, including microtargeting, rely on increasing amounts of precise personal data to develop sophisticated social graphs of US and other societies. Under pressure to compete for data, these companies will continue pushing to get access to users' data. Facebook, for example, ran a research program that paid young users a nominal sum ($20) in exchange for detailed access to their phone-use habits.[50] Social media companies' objectives are primarily commercial, but their methodology (and the data access it provides) has served the purposes of the IRA, Russian military intelligence, and other purveyors of disinformation.

Data protection is, thus, not just a privacy issue. Access to data provides a commercial suite of tools to deliver precise disinformation to bespoke slices of the population. In other words, social media companies may be providing valuable research for purveyors of disinformation, including Russian intelligence. None of the social media companies have pushed to seriously address this core problem, preferring to instead focus on filtering content, because it risks undercutting profit margins. Yet, this is exactly where the industry will have to change its approach and policies if it is serious about solving the disinformation problem.

Civil Society Stays the Course

Since the launch of the first report, civil society has continued to be instrumental in raising awareness and understanding the breadth of the challenge. Civil society groups and researchers are becoming savvier and more sophisticated in their analysis of social media data, tracking of disinformation campaigns, and exposure of such campaigns.

> **"One core problem remains: social media companies' business models are premised on increasing ad revenues, and sensationalist content sells best."**

- Graphika, Oxford University's Computational Propaganda Project, and New Knowledge were enlisted by the Senate Intelligence Committee to study millions of posts released to the committee by major platforms.

- The Atlantic Council's Digital Forensic Research Lab has been given access to data ahead of major takedowns by Facebook and has been able to offer analysis of the pages and increase understanding of the information operations.

- AlgoTransparency is a program, developed by a former YouTube employee, that tracks YouTube's recommended-videos algorithm, which accounts for seven hundred million view hours per day.[51] The project aims to "inform citizens on the mechanisms behind the algorithms that determine and shape our access to information." As the project finds, most such videos tend to lean toward conspiracy theories, radical content, and misinformation. YouTube's recent change in its recommendations system seems to reflect input from AlgoTransparency.

- Public Editor is a start up that provides credibility scores for news articles, journalists, public figures, and news sites by guiding volunteers to complete online media literacy tasks identifying argumentative fallacies, psychological biases, inferential mistakes, and other misleading content.[52]

[50] Josh Constine, "Facebook pays teens to install VPN that spies on them," Tech Crunch, https://techcrunch.com/2019/01/29/facebook-project-atlas/.

[51] "About," Algo Transparency, https://algotransparency.org/methodology.html?candidat=Marine%20Le%20Pen&file=ytrecos-presidentielle-2017-06-10.

[52] Public Editor, http://publiceditor.io.

NEXT STEPS

The US Executive Should Get Serious

- **Get organized**
  - ◆ Designate a **lead agency** or office, and a lead senior official, for counter-disinformation policy and operations. The mandate should be broadly established to cover disinformation generally, rather than focusing on a specific actor (such as Russia).
    - As noted in the previous edition of this paper, DHS could take the lead in devising counter-disinformation policy, working with and through the State Department as the United States seeks to build common action with European and other governments.
    - Given the broad nature of the challenge, the USG could establish an **interagency "Counter-Disinformation Center**," perhaps a smaller version of the National Counter Terrorism Center established after the terrorist attacks of September 11, 2001, as an operating arm for monitoring of foreign disinformation, with a mandate for real-time liaison with foreign governments, social media companies, civil society, and other relevant actors. DASKA legislation includes a proposal for a similar "fusion cell."
      - ✱ A key task for such a center should be to develop a baseline and outflowing set of metrics and indicators for when and how the government should respond to an identified state-sponsored disinformation attack against the United States, and who within the government should do so. Not all disinformation efforts warrant a USG response, as there is risk of further amplifying disinformation content. Identifying when the benefit of exposure outweighs the risk of inadvertent amplification is critical. A common set of guidelines is sorely needed.
      - ✱ US government employees and officials should be regularly informed of detected disinformation attacks and, where appropriate, trained on how to handle and respond to such attacks when their agencies are targeted.
      - ✱ What will not do is the current situation, in which no one in the USG seems to have responsibility for countering foreign disinformation. On the bright side, vacuums of policy leadership—one way to describe the situation within the US administration with respect to countering disinformation—tend to be filled when the politics of an issue start generating pressure, providing incentives to step up first.

- **Work with friends, starting with the EU**
  - ◆ Stand up a transatlantic **Counter-Disinformation Coalition**, as recommended in the first edition of this paper, bringing in like-minded governments (e.g., the United States, the EU, other willing European national governments, and possibly G7 member states), social media, relevant tech companies, and civil society groups.
    - The coalition would serve as a platform to share experience and information with respect to foreign disinformation, and develop, in non-binding fashion, best practices for dealing with it in ways consistent with common commitments to freedom of expression.
    - The coalition could draw from the EU's Code of Practice as the basis for a broader set of expectations for social media companies' behavior. It could also serve as a place to develop consistent approaches to issues of data privacy, transparency, and identification (and possibly removal) of bots, foreign trolls, impersonators, and other malign actors.
    - The USG and EU should also engage private cybersecurity companies for support in detection of disinformation. Some of these are led by, or employ, former US intelligence officers whose technical expertise can help spot Russian and other disinformation techniques, while also evaluating social media companies' counter-disinformation actions.
    - There are organizational variants for an informal coalition.
      - ✱ A G7 institutional base, building on the Charlevoix Commitment from the Canadian presidency. The G7's Rapid Response Mechanism, already functioning, could be its core.

Ahead of the 2018 US congressional elections, Dan Coats (front right), the director of national intelligence, warned that “lights were blinking red” on the threat of foreign interference *Photo Credit:* Official White House Photo by Shealah Craghead

    - While the Organization for Economic Cooperation and Development (OECD) has limited political weight, it has resources and a broad membership beyond the G7 and transatlantic community.

    - Alternatively, the existing NATO Strategic Communications Centre of Excellence in Riga or the European Centre of Excellence for Countering Hybrid Threats could play a role as a resource hub for such a coalition.[53]

  - As a potential longer-term objective, practices of cooperation and informal groups could grow into formal arrangements. Former Danish Prime Minister Anders Fogh Rasmussen has proposed establishment of a digital “Alliance of Democracies,” with extensive norm-and-rule-setting authorities like those of the World Trade Organization.[54]

  - Whatever the short- and long-term institutional arrangements, the United States and EU should start now to pool assessments of social media companies’ performance in terms of their counter-disinformation commitments—including through the Code of Practice, via testimony to Congress, and in public statements—and use their combined leverage to encourage greater transparency and implementation.

- **Establish a USG rapid alert system (RAS)** to inform the public, allied governments, and social media companies of emerging disinformation campaigns that threaten national security.

  - The European rapid alert system (if established along lines forecast by the EU Action Plan) can help the USG judge the potential of this idea. Some of the challenges can be anticipated: given US politics and traditions, issues will arise around a US RAS’s mandate (e.g., the definition and attribution of disinformation, as opposed to media content the USG doesn’t like) and its composition, credibility, and independence. Issues regarding

[53] *NATO Strategic Communications Centre of Excellence*, https://www.stratcomcoe.org/; *European Centre of Excellence for Countering Hybrid Threats*, https://www.hybridcoe.fi/.

[54] Anders Fogh Rasmussen, “The West’s dangerous lack of tech strategy,” Politico, March 11, 2019, https://www.politico.eu/article/opinion-the-wests-dangerous-lack-of-tech-strategy.

the current administration's credibility on media independence and freedom will also have to be addressed.

- The authors judge the effort to establish a US RAS worthwhile, challenges notwithstanding. An independent board could be established, possibly by legislation, with Senate confirmation of its members. The precedent of a European RAS could give the idea legitimacy.

- **Follow the money**
  - The USG should continue to **impose sanctions** on foreign official, or officially-controlled or directed, purveyors of disinformation and their sponsors, and to identify and prosecute violations of federal elections laws (prohibitions on foreign contributions).
    - On September 12, 2018, the Trump administration issued Executive Order 13848, which provides for sanctions imposed against persons found to have interfered in US elections. While, in part, simply an effort by the administration to preempt stronger legislation (i.e., the "DETER" Act introduced by Senators Marco Rubio (R-FL) and Chris Van Hollen (D-MD)), it provides a useful vehicle, should the administration use it.
    - US sanctions laws restrict US citizens from financial dealings with or "providing material support" to foreign persons under sanctions. Enforcement of these and federal election laws could limit the ability of Russian or other foreign purveyors of disinformation to work with US agents.
  - The USG should continue to **share findings** about Russian disinformation methods, financing, and relevant individuals and organizations with the EU and affected member states, the G7, and other friends and allies—seeking, among other things, coordinated sanctions.

- **Develop "forward-defense" and asymmetric options.** US Cyber Command's reported actions targeting Russian disinformation actors raise the question of whether the USG toolkit for countering disinformation should include asymmetric options for deterrence or retaliation, including in the cyber realm. This policy question leads to broader issues of cybersecurity and cyber deterrence, which are generally beyond the scope of this paper, but worth noting.

- **In approaching regulation (and associated legislation), start with low-hanging policy fruit, then determine how far to proceed**.
  - **Regulation of advertisement and sponsored content** seems relatively straightforward, as precedent exists for limits on commercial speech. Social media companies can be mandated to post accurate information about sponsors of ads, rather than euphemistic or misleading self-descriptions. ("Americans for Puppies" is not an acceptable label for an ad sponsor if the ultimate placer of the ad is the St. Petersburg Internet Research Agency.)
    - Regulation and limitation of foreign ads are useful, but may be of limited impact, especially given the shift to organic content.
  - More complex would be **mandatory identification of bots**. This would raise major domestic commercial issues, given that bots play a role in domestic, non-political advertising and are used for benign purposes, e.g., bots distribute content from major media organizations such as the *New York Times*. Individuals also use bots to schedule their tweets or re-post their own content. Still, identification could be mandated for foreign-origin bots or for bots disguised as persons, following principles of transparency and integrity, or for specific behaviors.
  - Still more complicated would be **regulatory mandates to disclose or remove inauthentic accounts and impersonators**, e.g., a Russian troll masquerading as "Jimmy" from Minneapolis, or a purported domestic "news organization" that was, in fact, controlled from Russia or Iran. We favor removal of inauthentic accounts and labeling of anonymous accounts if they engage in misleading presentation. Using a pseudonym online may be legitimate for human-rights activists operating in authoritarian countries but could also provide cover for foreign deception. (Russian human-rights activists, however, have told us that the marginal protection of anonymity in Russia is insignificant; this is likely true in China as well.) In any case, there is a difference between an anonymous online persona and a deceptive one, and this paper recommends steps to limit the latter.
  - As a further (and yet more complicated) step, transparency regulation could mandate an **online sign-in system for access to part, or all, of the Internet.** All identified human beings, but only identified human beings, would have access

to social media accounts. Some sites could be placed off limits to minors. This could (but need not necessarily) be combined with requirements that online persons also be identified, i.e., no more anonymity.

- Any effort to verify the identity of individuals may raise objections on grounds of free speech, or unintended discriminatory impact if online sign-in requires documentation to verify identity. Additional concerns around online surveillance would also have to be taken into account.

◆ **Mandating standard terms of service** for social media companies would ease implementation of regulatory limitations. Precedent exists; the Credit Card Accountability Responsibility and Disclosure (CARD) Act of 2009 required credit-card companies to standardize the terms used in their terms of service (e.g., fixed rates).[55] In the context of social media companies, this could mean requiring platforms to agree on both **common definitions of impersonator and inauthentic accounts and standards on removing bots,** and, though this is more questionable, what qualifies as hate speech, credible versus weak content, and violent or harmful content (e.g., anti-vaccination propaganda).

◆ Social media platforms have created a commodity from aggregated personal information. This has the unintended consequence of turning social media companies into potential research engines for foreign intelligence agencies or other malign actors. The issue of corporate responsibility for safeguarding personal data, based on privacy protocols, thus takes on a new dimension. The policy challenge is whether to mandate privacy protocols and/or what some scholars have termed **"information fiduciaries"** to help set corporate standards for safeguarding personal data beyond areas already limited, e.g., health and banking.[56] This would complicate the business model of social media companies, so at this stage the recommendation is for further exploration.

◆ Another difficult regulatory issue, also affecting social media companies' business models, has to do with social media platforms' **algorithmic bias toward sensational content**, driven by commercial preference for user engagement but easily exploited by those with intention to inflame and distort. Government mandates with respect to overall social media algorithms are apt to be contentious, (e.g., what would a "fairness doctrine" look like when applied to social media?) and could easily become the object of partisan politics (e.g., Senator Ted Cruz (R-TX) has publicly advocated regulation to counter what he claims is liberal bias in social media).

In 2017, the then-chairman of Alphabet, Eric Schmidt, said that RT and Sputnik would be de-ranked from appearing at the top of Google result searches. As of this writing, however, that policy change has not been implemented. *Photo credit:* Hecker/MSC

◆ Thus, we prefer not to start with a mandated makeover at this stage, but to start with targeted fixes, such as:

- Systems in place to **de-rank, rather than remove, false or deceptive sites or posts** while at

[55] US Congress, *Credit Card Accountability Responsibility and Disclosure Act of 2009*, HR 627, 111th Congress, May 22, 2009, https://www.congress.gov/111/plaws/publ24/PLAW-111publ24.pdf

[56] For example, see Jack Balkan and Jonathan Zittrain, "A Grand Bargain to Make Tech Companies Trustworthy," The Atlantic, October 3, 2016, https://www.theatlantic.com/technology/archive/2016/10/information-fiduciary/502346/.

> **"Democratic societies will need to invest in building resilience to (digital) disinformation, even as they take steps to mandate norms that diminish vulnerabilities to disinformation exploitation."**

the same time focusing on elevating high-quality content. ("Deceptive" could be defined to refer not to content, but to deceptive provenance, e.g., RT, Sputnik, or known propaganda outlets.)

- System improvements around **"data voids"** or search terms for which there is little available content and thus the void can be filled by malicious actors (e.g., Russian outlets strategically launching a barrage of content on a specific search query to crowd out any other sources).[57] This is a particular problem for search engines including Google and YouTube.

◆ Facebook is reportedly introducing such changes to its algorithms to de-rank questionable content.[58] Google, however, is reportedly not reviewing ranking of individual sites.

■ We remain **skeptical about government-mandated content control**. Taking down pornography, hate speech, and terrorist-related sites is challenging, but seems within the realm of the possible for democratic societies with robust norms and laws protecting free expression. Broader content restrictions (against false information, for example), however, are more difficult. While arguments can be made for restricting or eliminating dangerous content, e.g., how to build weapons of mass destruction or anti-vaccine propaganda, the application of content restrictions against disinformation, whether foreign or domestic, will constantly bump against norms of free expression.

■ **Remember resilience**

◆ Public policy and regulatory efforts only go so far by themselves. Democratic societies will need to invest in building resilience to (digital) disinformation, even as they take steps to mandate norms that diminish vulnerabilities to disinformation exploitation. The "supply side" of disinformation and the "demand side" go together: social media platforms' transparency measures, such as labeling and identification of bots, impersonator trolls, etc., work best when people recognize and shun them by habit.

◆ Building social resilience will be a generational challenge. Successful campaigns that have changed individuals' behavior included effective public policy, (sometimes) buy-in from industry, and mobilization of civil society groups. These have taken decades, a great deal of trial and error, and struggle. The US anti-smoking campaign is one example of successful societal change coupled with regulation and changes in industry approaches (the latter only under sustained social and regulatory pressure).[59] The digital revolution is young—Facebook, Google, and Twitter are still young companies—and democracies in the beginning stages of grappling with the disinformation challenges of the twenty-first century. More trial and error will follow, but democracies must stay the course.

◆ There are bright spots. For example, Lie Detectors is a European project that sends journalists into classrooms to teach schoolchildren about the difference between false and real news.[60] It is showing impactful early results, while empowering youths with digital literacy skills. If instituted at scale, such efforts could begin to yield results quickly.

The US Congress Should Fill the Gap

■ **Hold additional (and better prepared) public and private hearings** with social media companies' chief

[57] Michael Golebiewski and Danah Boyd, "Data Voids: Where Missing Data Can Easily Be Exploited," Data & Society, May 11, 2018, https://datasociety.net/output/data-voids-where-missing-data-can-easily-be-exploited/.

[58] Elizabeth Dwoskin, "How Facebook is trying to stop its own algorithms from doing their job," Washington Post, April 10, 2019, https://www.washingtonpost.com/technology/2019/04/10/how-facebook-is-trying-stop-its-own-algorithms-doing-their-job/?utm_term=.1e2bfb3f9fc4.

[59] Alina Polyakova and Geysha Gonzalez, "The U.S. anti-smoking campaign is a great model for fighting disinformation," Washington Post, August 4, 2018, https://www.washingtonpost.com/news/democracy-post/wp/2018/08/04/the-u-s-anti-smoking-campaign-is-a-great-model-for-fighting-disinformation/?utm_term=.68e671b5e3e3.

[60] Lie Detectors, https://lie-detectors.org/.

technology officers (CTOs) or equivalent, including on algorithmic bias, manipulation of search engine optimization (SEO) (i.e., how malicious actors game the search engine results), data collection, and privacy practices.

- ◆ Such hearings should be designed to encourage (and, if necessary, force) companies to inform policymakers and the public, in clear language, how their data and behaviors are tracked, collected, stored, and shared with third parties, and how the profit incentive drives such practices. If company representatives are later revealed to have misled congressional leaders, they should once again be called to testify and held accountable.

- ■ **Authorize and appropriate funds to "build capacity** of civil society, media, and other nongovernmental organizations," countering Russian and other sources of foreign disinformation (from DASKA Sec 705(b)), in coordination with the EU, NATO, and other bodies. Funding is already available to the State Department's Global Engagement Center; this should be increased.

- ■ **Develop in-house expertise** on disinformation. If Congress and the executive branch establish regulatory frameworks to combat disinformation (e.g., as suggested above, with respect to ads, standards, and expectations for removal of bots, identification and/or removal of impersonation accounts, privacy/information fiduciary standards, de-ranking of falsified or deceptive organic content), Congress's capacity for detailed analysis, independent from social media companies, will be critical.

- ■ **Prepare legislation—on a step-by-step basis—to support a regulatory framework for social media companies.** This paper has recommended (above) a layered approach to regulation, starting with simpler steps. In similar fashion, it recommends "light" legislation (and, probably, discrete legislative packages) to build a regulatory framework, rather than an attempt at comprehensive regulatory legislation, given the issues of freedom of expression—and the difficulty of legislating at an evolving threat from a relatively low knowledge base in Congress.

  - ◆ The **Honest Ads Act**, introduced in the last Congress, is a solid step toward setting transparency standards around online advertising (not just narrowly defined political ads). This analysis suggests that these standards be established evenly across the tech industry, not just for social media firms. This act, revised and strengthened along the above lines, could be a vehicle for this effort.

  - ◆ The **DETOUR Act**, noted above, suggests that Congress is prepared to consider legislation to regulate what has been standard social media company practice to steer (or manipulate) users. This suggests Congress is getting ready to consider legislation to support a regulatory framework to deal with disinformation.

  - ◆ This paper also recommends legislation to provide a framework for regulation to address **transparency** (especially with respect to bots), **integrity and authenticity of service** (i.e., targeting deceptive and impersonator accounts, whether individuals or false-front organizations), and **common terms of service** across the social media industry.

  - ◆ Legislation to address the problem of **algorithmic bias** toward the sensational, or to de-rank deceptive content, appears more complex, and should be addressed separately from (and probably after) initial efforts.

  - ◆ Congress could also mandate that media outlets determined by the Department of Justice to be acting as agents of foreign governments be de-ranked in searches and on newsfeeds and be barred from buying ads. RT, for example, was required to register under the Foreign Agents Registration Act (FARA). Governmental assessments and FARA determination should be one of many variables considered in rankings for search engines.

  - ◆ Congress should explore establishing a federal statute that would limit companies' collection of personal data about individuals. Such a statute would specify that any personal data collected would be specific to the stated purpose of the technology. Such data collection limitation would make microtargeting and exploitation of individuals' personal data more difficult while also reducing the ability of malicious actors to influence. The California Consumer Privacy Act of 2018[61] could serve as precedent for a federal mandate.

  - ◆ The issue of social media companies' use of personal information (the issue of "**information fiduciaries**") should await additional study.

[61] Dipayan Ghosh, "What You Need to Know About California's New Data Privacy Law," Harvard Business Review, July 11, 2018, https://hbr.org/2018/07/what-you-need-to-know-about-californias-new-data-privacy-law.

- Explore legislation that could be applicable to social media, based on precedent in other industries.

    - The issue of **data portability**, or the user's ability to take their data across platforms, was resolved in the **telecom sector** through legislation. Individuals have the right to own, control, and move their data between telecom providers. The same should be true for social media companies.

    - Regarding **network manipulation** for malicious purposes: the social media sector shares similarity with the **financial sector**, in that both industries operate networks of sensitive data that flow globally. Financial institutions are legally required to investigate and identify illicit financial flows (or the malicious misuse of their processes) and report any finding to federal agencies. A similar set of compliance practices, with supporting federal agencies, could be considered for social media companies.

Europe Should Make it Real

- **Use the RAS.** The rapid alert system was intended to be online and active in advance of the May European parliamentary elections, but reports of its effectiveness are mixed. If successful, the RAS could (and hopefully will) develop operational ties with regular and social media, and with civil society groups and tech companies involved in counter-disinformation efforts. While this paper focuses on foreign-origin disinformation, information sharing about disinformation generated within EU Member States is valuable. Making the RAS effective as soon as possible will be an important indicator of the credibility of the EU's promising start at tackling disinformation.[62]

- **Establish and use the fact-checkers network**, as called for under the EU Action Plan, though this element was relatively less developed. The issue of "arbiters of truth" (e.g., empowering independent groups to render editorial judgement) is complex, and even fraught. Still, a networks of fact checkers and independent journalists can be on alert to identify new disinformation sites as they emerge, track the credibility of outlets based on their track record of publishing false or misleading content, and share this information with governments and industry. Organizations focused on social media analysis, such as Graphika, New Knowledge, and DFRLab, can work to identify and monitor coordinated manipulation activity and demonstrable deception of identity, which may be easier to establish.

- **Provide adequate resources and political backing to EastStratCom**. This innovative institution needs to be supported without ambivalence. Recent indications are promising and need to be maintained. In the best case, EastStratCom would have the resources to work with independent civil society groups to identify disinformation campaigns in real time and disseminate that knowledge through the RAS to member states and to European publics.

- **Continue to monitor social media companies' implementation of the Code of Practice (with the option to proceed to regulation)**, especially with respect to control over foreign bots, political ads, fake accounts closed, and cooperation with civil society researchers.

    - The current arrangement under the Code of Practice could be termed **audited or monitored self-regulation**. Should this prove inadequate, next steps could include a move to formal EU Commission review and approval of social media industry codes, or to a **full-fledged regulatory system**.

    - In any case, the United States and EU should consult closely about common regulatory standards; establishing a pattern of consultation for green-field regulatory standards will be far easier before, rather than after, respective US and EU regulatory standards are established.

    - Establishment of an **"algorithmic ombudsman"**—a credible figure designated by the EU (or USG) to work with social media companies to reform their algorithmic bias toward sensational content, and the related vulnerability to exploitation by purveyors of disinformation—has been circulating among various experts seeking a middle ground between self-regulation by social media companies and their full regulation by governments. While sympathetic to the intent behind this concept, the authors prefer to tackle the challenge through incremental regulation, on the grounds that regulatory standards either make sense or do not, and the intermediary of an ombudsman will not alter those categories.

- **Empower independent civil society researchers.** The EU should expand and simplify its grant-making

[62] "Rapid Alert System" European External Actions Service, March 2019, https://eeas.europa.eu/sites/eeas/files/ras_factsheet_march_2019_0.pdf

> **“Facebook, Twitter, and Google should continue to improve transparency requirements for content produced by state-funded media outlets, including due diligence around the parent companies behind the content.”**

system to support civil society groups, both inside and outside the EU, with expertise in identifying disinformation campaigns. These groups are on the front lines of counter-disinformation efforts, and those operating inside their home countries are likely to be best placed for these efforts.

Social Media Companies Should Step Up

In a March 30, 2019, op-ed in the *Washington Post*, Facebook Chief Executive Officer Mark Zuckerberg expressed support for Internet regulation, including: content standards (not defined in detail); “election integrity,” including political ad transparency; privacy and data protection based on current EU standards (from the EU’s General Data Protection Regulation); and data portability. These are useful areas to explore, and a welcome indication of changing thinking at Facebook.

The regulatory challenges covered above include far more than Zuckerberg’s relatively modest suggestions. Moreover, social media companies need to do more than cede responsibility to government regulators. As initial steps, they could do the following.

- **Institutionalize cross-platform coordination processes** for tracking, monitoring, and taking down disinformation campaigns, which often work across platforms simultaneously to amplify their content. For example, when Facebook deletes or mutes suspicious accounts, other platforms should de-rank or delete their corollaries or amplifiers. Some coordination already exists, but, as manipulators increasingly work to amplify content across platforms simultaneously, much more nuanced coordination will be needed.

- **Reassess anonymity on these platforms**. To that end, the companies should make publicly available an accurate report on the number of anonymous or unverified accounts on their platforms. Such a report should also assess how, and if, ending anonymous accounts could negatively affect vulnerable civil society or human-rights groups operating in authoritarian countries. A next step could be introduction of “authentic spaces” accessible only to verified human beings—not bots, cyborgs, or impersonators. This would mean use of online sign-in systems capable of identity verification for access, either to some or all social media platforms.

- **Start algorithmic reform.** As an initial example, and without waiting for regulatory mandates, Google could set the industry standard by seeking to prevent overt authoritarian-state-sponsored media outlets from appearing at the top of any search results. RT and Sputnik continue to appear in top search results for current event searches on a variety of topics. Google, and others, should not censor or delete this content. Rather, social media platforms should use search algorithms to prioritize content from independent media over identified propaganda, which should not be regarded as either relevant or authoritative. Social media companies should be transparent about their efforts.

- **Facebook, Twitter, and Google should continue to improve transparency requirements** for content produced by state-funded media outlets, including due diligence around the parent companies behind the content. Laudably, in February 2019, Facebook suspended pages run by Maffick Media, a company largely owned by a subsidiary of RT.[63] Unlike YouTube, which publishes disclaimers on videos funded by the Russian government and other state-media outlets, Facebook and Twitter do not have a similar policy. They should.

- **Large tech companies, not just social media platforms, should work closely with innovative nonprofits**, such as AlgoTransparency, to identify problems in their algorithms and correct them, as YouTube is starting to do. To that end, private-sector firms should establish independent funding mechanisms (e.g., foundations) to fund civil society initiatives and innovators who are using new technologies, such as AI and machine learning, to identify problems before they arise.

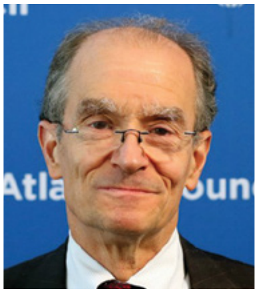
[63] Donnie O’ Sullivan et al., “Russia is backing a viral video company aimed at American millennials,:” CNN Business, February 18, 2019, https://www.cnn.com/2019/02/15/tech/russia-facebook-viral-videos/index.html.

ABOUT THE AUTHORS

DR. ALINA POLYAKOVA
Director, Project on Global Democracy and Emerging Technologies
Brookings Institution

Alina Polyakova is the director of the Project on Global Democracy and Emerging Technologies at the Brookings Institution and Professor of European Studies at the Paul H. Nitze School of International Studies at Johns Hopkins University. She was previously a David M. Rubenstein Fellow for Foreign Policy at Brookings. She specializes in European politics, far-right populism and nationalism, and Russian foreign policy. She is the editor and co-author of the Atlantic Council's report series, *The Kremlin's Trojan Horses*, which examines Russian political influence in Western Europe. Dr. Polyakova specializes in Russian foreign policy, European politics, and far-right populism. Her recent book, *The Dark Side of European Integration* (ibidem-Verlag and Columbia University Press, 2015) examines the rise of far-right political parties in Western and Eastern Europe. She has also written extensively on Russian political warfare, Ukraine, and transatlantic relations for the *New York Times, Wall Street Journal, Foreign Affairs, Foreign Policy*, and the *American Interest.*

Prior to joining Brookings, Dr. Polyakova served as director of research and senior fellow for Europe and Eurasia at the Atlantic Council. She is a term member of the Council on Foreign Relations and a Swiss National Science Foundation senior research fellow. She has also been a fellow at the Fulbright Foundation, Eurasia Foundation, Woodrow Wilson International Center for Scholars, National Science Foundation, Social Science Research Council, International Research and Exchanges Board (IREX), and a senior research fellow and lecturer at the University of Bern. Dr. Polyakova holds a doctorate from the University of California, Berkeley.



AMBASSADOR DANIEL FRIED
Distinguished Fellow, Future Europe Initiative and Eurasia Center
Atlantic Council

Ambassador Daniel Fried is a distinguished fellow with the Atlantic Council's Future Europe Initiative and Eurasia Center. Ambassador Fried has played a key role in designing and implementing US policy in Europe after the fall of the Soviet Union. Prior to joining the Atlantic Council, Ambassador Fried served as the US Department of State's coordinator for sanctions policy from 2013 to 2017. Previously, he served as special envoy for the closure of the Guantanamo detention facility and was assistant secretary of state for European and Eurasian affairs under the Bush Administration, as well as special assistant to the president and senior director for European and Eurasian affairs at the National Security Council. From November 1997 until May 2000, he served as ambassador to Poland, where he had developed much of his earlier career. Ambassador Fried has focused on designing and implementing US policy to advance freedom and security in Central and Eastern Europe, NATO enlargement, and the Russia-NATO relationship. Ambassador Fried holds a BA with magna cum laude honors from Cornell University and earned his MA at Columbia University's School of International and Public Affairs.

Atlantic Council

Board of Directors

Atlantic Council

The Atlantic Council is a nonpartisan organization that promotes constructive US leadership and engagement in international affairs based on the central role of the Atlantic community in meeting today's global challenges.

Atlantic Council

1030 15th Street, NW, 12th Floor,
Washington, DC 20005

(202) 463-7226, www.AtlanticCouncil.org