A HOOVER INSTITUTION ESSAY

Advanced Persistent Manipulators and Social Media Nationalism

NATIONAL SECURITY IN A WORLD OF AUDIENCES

CLINT WATTS

Aegis Series Paper No. 1812

The tweet went out on Friday, May 5, 2017. "Massive doc dump at /pol/, "Correspondence, documents, and photos from Macron and his team."[1] Included in the tweet was a link to a Pastebin page hosting the hacked contents of French presidential candidate Emmanuel Macron's campaign.[2] The tweet pushed many viewers around the world to the allegedly compromising information at a time when the French media blackout prevented preelection coverage.

The hacking of Macron was not surprising after Russia's influence on the 2016 US election alerted those in subsequent democratic contests to the Kremlin threat. The tweet's source and its language—English, not French or Russian—were the more interesting aspects of the hacked emails. Why would an American working for a fringe Canadian media outlet be the first to signal the disclosure of hacked materials related to the French election?

At the time, Jack Posobiec served as the bureau chief and sole employee of the Washington office of the Canadian outlet *the Rebel*.[3] In his short time in Washington, DC, Posobiec acquired a coveted White House press pass and began reporting heavily for alt-right audiences feverishly supporting the Trump administration. Posobiec noted during an interview that he liked to do "4-D journalism, meaning that I'm willing to break the fourth wall. I'm willing to walk into an anti-Trump march and start chanting anti-Clinton stuff—to make something happen, and then cover what happens . . . so, activism tactics mixed with traditional journalism tactics."[4] Posobiec disseminates his reporting extensively via the social media platforms of YouTube, Periscope, and Twitter.[5]

The #MacronLeaks hashtag served as further confirmation of the rapidly shifting geopolitical landscape that has emerged with the transcendence of social media. A virtual information army, connected via social media more than real-world relationships, worked either wittingly or unwittingly with a nation-state to influence the outcome of a foreign democratic election. The French election represented not the first time, but one in a string of political contests manipulated

by Russia that involved the assistance of foreign social media audiences nudged by the Kremlin.

Future implications of this audience convergence may ultimately be huge for traditional nation-state security. Since the much debated US presidential election of 2016, the world has seen unprecedented employment of social media by authoritarians to either suppress and control their domestic populations or disrupt democratic adversaries. The Kremlin playbook has been nimbly adopted in Southeast Asia. Strongmen in Myanmar, Cambodia, and the Philippines employed social media-powered oppression to quell dissent inside their countries or move certain audiences in a country against another part of the citizenry.[6] Social media-powered nationalist movements threaten Western alliances and democratic unions worldwide and seem poised to alter the world order dramatically.

What does national security mean when threats on social media transcend physical borders? Security threats in the physical world have naturally migrated to the Internet. With the growth of social media, bad actors have rapidly recruited members, financed and coordinated operations, and even moved audiences, wittingly and unwittingly, in pursuit of their objectives. Moving forward, with boundaries and allegiances continually crossing traditional borders, social media platforms will become the battlefields of competing audiences forming a new spectrum of information threats with potential to inflict harm on Western democracies.

Social Media Nationalism: When the Virtual World Drives the Real World

Citizens of Western democracies spend nearly their entire day connected to the virtual world. The rise of mobile phones connected to the Internet has, in only a few years, brought about staggering changes in the daily conduct of human lives. American smartphone users over age eighteen, depending on the research findings, average between two and a half and five hours a day on their devices.[7] At four hours a day, this would roughly equal one fourth of waking hours, and this doesn't even account for the amount of time people use traditional computers for work or pleasure. Apps constitute the majority of mobile device consumption, with social media platforms—particularly Facebook—taking up the largest portion of mobile phone application use.[8] As Americans increasingly spend more time on social media, their virtual connections grow. The increasing number and intensity of virtual connections powered by social media result in many users valuing their virtual relationships more than their physical, real-world relationships.

During the first two centuries of US history, Americans formed associations principally through physical connections. Regions and even some states were known for their ethnic identities, cultural norms, and similar speech and appearance.[9] Cultural

identifiers have naturally shifted over the decades. Liberal views emerged in the urban and coastal regions and more conservative views came to dominate rural areas, the South and the Midwest. Socioeconomic partitioning of America occurred as well. Similar-looking people of similar income levels migrated into similar communities across the United States. These changes occurred over decades and have been encapsulated in the recent spate of political gerrymandering, a process that has mapped communities by political leanings down to the neighborhood or even sub-neighborhood level.

American democracy, at least in its first two centuries, operated well because of the crosscutting physical relationships citizens maintained within civil society. Known in political science as social capital, the network of relationships a constituent had, regardless of race, creed, religion, or economic class, led to an overlapping system of personal and group interests contributing to negotiation between parties and undermining of the "tyranny of the majority"—the ability of a majority of voters to suppress the interests and beliefs of a minority. Labor unions, community associations, political parties, and family gatherings—Americans were usually members of many groups with competing interests that collectively shaped a common national identity.

Two decades ago, as the Internet arose linking people around the world via computers, Robert Putnam warned of the rapid degradation of American social capital. In his book *Bowling Alone*, Putnam said Americans were increasingly isolated. He believed this would lead to the erosion of democracy.[10] Real-world connections and resulting civic associations were in decline, spelling danger for democracies dependent on social and political participation.

Some reviewers argued Putnam's warnings were overblown, in part because the Internet was steadily connecting people of similar interests and beliefs.[11] By the late 1990s and into the 2000s, many people believed virtual associations would supplant and enhance connections, providing needed bridges between constituents and communities.[12] By 2008, there seemed to be some evidence to support this as presidential candidate Barack Obama, using a grassroots campaign fueled by activism on the World Wide Web, swept past traditional, established political candidates Hillary Clinton and John McCain to win the White House. A newcomer deployed a virtual campaign to power a physical campaign creating connections with and between voters as never seen before.[13]

Since 2008, however, social media have upended the traditional meaning of community and created political and social shifts more rapidly than at any time in human history. Social media, by design, sought to create online communities where people came together and interacted in dialogue. In its infancy—and still to some extent today—crosscutting engagement for good occurred on social media platforms.

The Arab Spring movements of 2010–11 showed the power of social media to connect like-minded citizens and help the oppressed mobilize to overthrow dictators in Tunisia, Egypt, and Libya. These grassroots movements employed social media to harness a decentralized network once suppressed and disenfranchised in the physical world by heavy-handed security forces. More substantially, social media raised international awareness of the campaigns to oust North African authoritarians.[14] Sadly, the social media-powered Arab Spring turned into the Arab Winter. The democracy-inspired hordes filling the streets of Cairo or Tunis lacked strong, physical associations needed to propel democratic change beyond mobilization. The organized and well-resourced, such as the Muslim Brotherhood in Egypt, occupied the political vacuums left in the wake of the social media-networked populist movements that toppled dictators. In the worst cases, authoritarian regimes, like that of Syrian President Bashar al-Assad, employed social media against uprisings, using digital connections to track and suppress political dissidents.[15]

The wonders of social media activism faded as the Syrian civil war hit full stride in 2012. The outside world learned of the plight of the Syrian people via social media platforms, but the world's worst actors came to see opportunity with these digital portals. Unchecked authoritarian regimes like al-Assad's exploited technology to root out dissidents, deployed hackers and proxies to attack internationally, and used social media as an avenue for counter-propaganda to mobilize regional and international partners in support of their regimes.[16] International jihadists took to social media platforms in droves, creating an unprecedented migration to Syria and Iraq. Some recruits joined al-Qaeda affiliate Jabhat al-Nusra, but far more came to call themselves the Islamic State, creating the largest caliphate in modern history. On the ground in the Levant, the Islamic State enacted unimaginable violence. Online, it employed social media to radicalize young adherents abroad, accelerate recruitment into its ranks, and even synchronize operations locally and worldwide. In less than two years, the social media-savvy Islamic State media battalion helped build the largest terrorist fighting force in the world, overtaking al-Qaeda and unleashing a spectacular, international terrorist campaign around the world in the summer of 2016.[17] In eight years, social media went from being the tool of liberty and democracy to the platform of extremists.

The Islamic State proved a harbinger for a trend that now powers movements around the world: virtual communities arising to create and drive physical communities. The Islamic State surged in large part because a digital tail wagged the physical dog. Jihadists created a virtual nation that powered a physical nation. Political movements in the Western world have since followed a similar path.

Today's populist political upheavals, whether the Arab Spring, the Islamic State, or the emergent alt-right, arise in many ways from the formation of social media nations. Users spend hundreds of hours a year online developing and enriching virtual

connections with people they've never or rarely met in the real world. These virtual connections have overtaken the physical connections of their actual communities. In extreme cases, they've led to physical migrations of like-minded people who share similar beliefs and objectives. Entire families of jihadists connected to the Islamic State via social media traveled into war-torn Syria while millions fled devastation. Today, alternative right protests and rallies synchronized via social media occur across Western democracies. Most participants arise from the country where the protest occurs, but it's not uncommon for attendees from abroad to join in these nationalist gatherings.[18]

Whether through Twitter, Facebook, or Instagram, social media nations transcend physical borders. The borders of social media nations are defined primarily by a common language. Forming virtual bonds requires fluid communications, so English, Spanish, Russian, Arabic, or French initially establish the pathways for connections. From there, members of social media nations self-designate with virtual personas. They may note a US state or a country in which they reside, but they are far more likely to denote their allegiance to a prescribed ideology, movement, or campaign. Constituents of social media nations identify with each other through common biographical terms, hashtags, and avatars. Social media nations not only share the same virtual attributes, but the same information sources pandering to their preferred social, religious, or political views. Their information biases create alternative realities that suit the collective preferences of the social media nation and can influence their view of history to support preferred narratives. Social media nations blur the lines between fact and fiction, diverging from a common understanding of what actually happened and opting instead for their preferred versions of events. As their ranks grow and their resources expand, social media nations increasingly fund their own research institutions and initiatives, seeking evidence to support their preferred political agenda, cultural leanings, or science.

Social media nations have only begun to transcend real-world borders. Should they continue to grow, they will further generate political instability, affecting nearly every nation in the world. Describing the United States of America today as two competing political parties provides no clarity on the composition of America. But offering a few keywords can instead quickly distinguish a person's beliefs, ideals, leaders, and goals. #MAGA or #Resistance—those two words alone provide casual observers of social media platforms more perspective on a person, in many ways, than his hometown or professional résumé. America likely consists of at least five large social media nations with smaller subnations overlapping these information bubbles. The alt-right movement today represents one of the largest social media nations, with tens or even hundreds of thousands of members in the United States. Not only are the members a powerful current socially and politically in America, but they tightly intertwine and likely have greater affinity for fellow alt-right members in Europe and Canada than they do for fellow Americans. Similarly, strong virtual connections can be seen among the more

liberal-leaning constituents of North American and European democracies. A separate case could be made that a global oligarchy of wealthy elites has also established its own virtual nation and supporters. Globally, government transparency enthusiasts have created a sizable cohort that at many points overlaps with social movements like the Occupy movement and longer lasting hacker collectives like Anonymous or LulzSec.

The end result of increased digital connection and the rise of virtual identity may be the degradation of democracies and strengthening of autocracies. As social media users move further to define themselves via their preferences, their allegiance to their physical country may very well wax and wane as their preferred elected leaders control or lose the reins of political power. Social media polarization has accentuated political divides in America, for example, corresponding with governmental gridlock, insufficient compromise, and the pursuit of false information satisfying desired beliefs rather than true information refuting a preferred world.[19] In several extreme cases, virtual divides on social media and the Internet have led to threats and even political action for Texas to leave the United States and for California to break into three different states.[20] These efforts represent minorities among America's population, but these virtually powered phenomena led to real-world activism and measures to divide the nation.

For autocrats, the Internet once threatened their power by opening up their populations to democratic challenges to their rule and connecting oppressed dissidents who could then collectively mobilize. Autocrats caught up during the social media era, learning to harness platforms rather than fear them. The Soviet Union's original design of "active measures," an asymmetric strategy by which it sought to erode enemies through the "force of politics, rather than the politics of force," failed to take hold because it was so difficult for the Kremlin during the analog era to connect with and mobilize selected audiences.[21] Russia's active measures reboot for cyberspace recognized the natural currents unfolding with social media nations. Its social media influence techniques paved new ground, riding the waves and new emerging contours of social media nations. Russia's intelligence services represent the most talented of the current generation of adversaries seeking to connect with and mobilize social media nations. The next generation of threats haunting social media platforms will take the Kremlin's playbook and advance it in ways that threaten both the real and virtual worlds they seek to dominate.[22] Freedoms of speech and press will be used to divide audiences online and mobilize supporters on the ground for whichever cause conducts the most sophisticated and enduring information campaign.

The Rise of Advanced Persistent Manipulators (APM)

A decade ago, as cyberattackers grew in number and type, a new designation surfaced for the most sophisticated attackers. The moniker "advanced persistent threat," now widely known as APT, came to be the greatest fear of governments and large corporations with much to lose and a broad cyberbattleground to protect.[23] APTs were designated as such

because they followed a deliberate approach of reconnaissance, incursion, discovery, capture, and exfiltration. APTs were determined attackers, pursuing their targets on an enduring basis and employing a variety of methods to breach systems. Sustaining these attacks with a range of tools and techniques required resources—the kind only the most sophisticated organized crime groups and nation-states could employ.

Over time, as APTs expanded, less sophisticated cyberattackers and even average people sought to expand their capabilities as their targets improved their cyberdefenses. Demand for more sophisticated cyberweapons led to the creation of markets, on both the open and dark web, for attacker tools. Malware farms popped up rapidly, creating and selling new variants of unseen malicious code called "zero days." This referred to the number of days the malware had been seen in the wild—meaning no cyberprotections were yet available to protect targets. Distributed denial of service (DDoS) attacks became available for rent by anyone with an Internet connection. Today, ransomware, the most pervasive and devastating cybermethod of the last couple years, can be easily acquired and employed by any criminal worldwide.[24]

Hacking people's computers, for all its technical sophistication, feels like child's play compared to the hacking of people's minds that has occurred on social media platforms the past four years. The US government once feared most that a foreign nation might hack into US infrastructure and cripple the power grid. Now the goal of Russia's information warfare may be to infiltrate, influence, and then manipulate an American working inside a nuclear power plant to voluntarily destroy the American power grid. Corporations, which have fretted in recent years about destructive malware attacks similar to that allegedly launched by North Korea on Sony Pictures, should worry equally about a corporate competitor or nation-state recruiting insiders in their organizations via social media to deliver a cyberweapon into their systems, steal their internal communications, or launch a devastating smear campaign on their current or former employers.

APTs, now and in the future, pose a danger to the world's governments, corporations, and citizens through acute breaches. But over the long term, it will be advanced persistent manipulators (APM) who deploy a subtler, more corrosive cancer on the social media battlefield, sowing widespread mistrust of democracies and institutions worldwide. APMs will pursue a series of objectives against their adversaries. The traditional pursuit of information warfare has been the discrediting of adversaries.[25] New social media threats will continue that trend, using ever more advanced techniques to compromise and denigrate their opponents. Discrediting campaigns are and will be used by APMs to suppress internal dissent from political dissidents, to sideline social factions, or to tarnish financial rivals.

The 2016 presidential campaign demonstrated the confluence of multiple actors seeking to influence audiences toward a particular agenda. The Kremlin and political

campaigns used similar techniques to infiltrate and influence audiences for political candidates. Now, many APMs employ the same approach regarding social issues. A second successive objective for APM infiltration will be enlisting allies online who can be employed as agents in both the virtual and physical world. Future information warfare campaigns will seek out issue-based allies who can create an online armada for advancing the APM's agenda in target audiences.

The most damning and harmful objective for society from APMs is reality distortion. The most sophisticated actors on the social media battlefield will manipulate information and amplify falsehoods and inaccurate narratives with computational propaganda. Reality distortion will alter audience perceptions to that preferred by the APM and could possibly cause grave harm to entire societies. In, concert with reality distortion may be the final objective of inciting fear, real or imagined, into target audiences. Fear generation, whether a tactical, opportunistic amplification of terrifying news or an enduring strategic campaign of panic-inducing calamity, weakens the target audience's ability to properly assess information sources. APMs will seek to inject fear into target audiences, hoping information consumers fall back on biases that can be more easily manipulated with determined propaganda.

APM methods will be a mix of old and new techniques, deployed in evolving combinations to achieve their objectives. Digital forgeries have continued to appear in social media smear campaigns, but with the rise of fake audio and video capability, forgeries will take on enhanced sophistication and effectiveness.[26] Hacking teams will continue their tirades as well as harvesting secret information across a swath of APM adversaries.

APMs will create information bases in the form of alternative news outlets and opinion blogs. Some of these information outlets will be overt, but most will be of obfuscated attribution. Extending the reach of APM-created and affiliated news outlets will certainly require employment of computational propaganda—the rapid, repeated broadcast of targeted information via social bots. APMs will mostly rent social bots but may also create them for all social media platforms as they seek to gain audience share. These social bots will be increasingly sophisticated, growing smarter each year as machine learning and emerging artificial intelligence make posts and conversations from bots indistinguishable from real human social media users. Each successive improvement in computational propaganda will make it more difficult for audiences and social media companies to distinguish between fact and fiction, bots and humans.

Much has been said of the technical sophistication of social bots, but talented humans who have mastered the art of influence will be the most important element of emerging APMs. As seen with Russia's Internet Research Agency, effective social media influence comes with the creation of false personas with real propagandists

operating the accounts. Social media trolls will have their own employment skill set. Avatar operators will be tasked with connecting with and moving audiences to defined narratives. The most valuable human operators employed by influence efforts will be provocateurs. These operatives will create real-world incidents and conduct direct recruitment of target audience members in hopes of persuading them to execute actions, unwittingly, on behalf of the influencer. Propagandists and provocateurs provide a mutually reinforcing influence method that only the most sophisticated actors understand and can conduct.

Cyberthreats of the APT era have been predominately defined by four categories: criminals, hacktivists, terrorists, and authoritarian nation-states. The era of APMs will include some of these same actors and an emerging new breed. Authoritarians have led the way in social media influence operations. They will continue this trend, pursuing and employing aggressive social media influence domestically. Regionally, strong states will use social media to influence foreign policy outcomes similar to the Kremlin's deployment of social media influence to achieve its foreign policy objectives in Ukraine and Syria. Extremists of all stripes—religious, social, and political—will continue to be a threat on every social media platform they descend upon. Future social media extremism will come in many forms, with jihadists being one of many rather than the premiere extremist group exploiting platforms. Aggressive activist campaigns will also pursue APM methods and objectives. The best funded may even exploit the more advanced features offered by emerging technology.

The APM threat spectrum will be dominated, however, by the best resourced and equipped actors, who may only now be entering the influence space. Political campaigns have more desire and resources than most nation-states engaging in social media influence. They can hire one or more social media influence operations to advance their messages. Closely related, but equally desiring targeted widespread social media influence, will be lobbyists and corporate public relations firms. Depending on their clients, they will pursue any or all of the objectives and methods discussed above, seeking one-stop shops for influencing their target audiences. Finally, the world's wealthy will have the means and desire to shape their images and influence audiences on social media. The last year has seen disclosures of celebrities buying social media followers from third-party providers, but they likely represent the least aggressive people with resources. Oligarchs and other aristocrats will buy or create their own information outlets for distributing their perspectives, preferred worldviews, and personalized promotional propaganda. The fabulously wealthy will employ all acquirable means to sustain their positions and grow their brands.

Detecting and diagnosing APMs will come from using a methodology similar to cybersecurity's APT attribution. Victim-centric analysis will provide the needed framework for determining which APM is pursuing its agenda across one or more

social media platforms. Investigators, intelligence analysts, and researchers will have to work from objectives (what was the actor trying to do?) back to the methods that were employed. Much like cyberthreats, more sophisticated actors will use more advanced techniques and greater levels of technological sophistication. For example, the parallel to malware proliferation and advancement in cybersecurity will be the growth and sophistication of computational propaganda in information warfare. The type and combination of methods employed will then point over time to the type of actor within the APM spectrum. Certain categories of actors, when studied over a sustained period, will create a set of behavioral and technical signatures which can then be used to attribute the APM across many social media platforms.

Social Media and Governments: At Odds, in Danger, or Both?

Trust—it's the backbone of democracies and social media platforms. Without trust, neither democracy nor social media platforms can survive. Both have suffered massive setbacks with the rise of social media nations and advanced persistent manipulators. Whether they're described as citizens or users, people in today's hyper-connected systems have dwindling trust in the institutions that govern their physical world and with the social media companies that manage their virtual worlds. Both need to regain the trust of their constituents and users, respectively, to sustain their existence.

Democracies must try to govern via physical institutions and retain the trust of citizens who have segmented into competing social media nations that extend far beyond their actual borders. The United States and its Western allies have few institutions, limited regulations and laws, and no clear levers for engaging their APM adversaries among these social media nations. How do they sustain their national security in a virtual world of social media nations sharing allegiances across many borders? How do democracies fend off foreign adversaries that utilize freedoms of speech and press to dismantle their institutions and erode their constituency?

Social media companies face the inverse challenge. They must secure their users from a range of APMs, each seeking different objectives by, with, and through social media nations. Social media companies face a morass of overlapping laws and regulations from more than one hundred countries. Each requires different standards and disclosures from the social media companies and imposes costs to their operations while seeking to ensure privacy protections across many jurisdictions. How do social media platforms corral and please social media nations, govern them fairly and transparently, while detecting and disrupting a range of APMs?

Self-regulation or Government Regulation?

Democracies struggle to regulate social media companies for a few reasons. First, many democratic legislators have only rudimentary understanding of the social

media platforms. Social media technology remains complicated and lawmakers have little guidance from precedent how any rule they impose might affect social media companies and users—an ecosystem of major economic growth for Western countries. EU countries have pursued more stringent data privacy rules on social media companies, but simple, limited laws in the United States such as the Honest Ads Act have moved at a glacial pace.

In contrast, self-regulation by social media companies seems to happen only after a bad actor perpetrates an attack. Social media companies have no incentive to hold back their profits for potential problems that may arise. Absent nation-state laws, self-regulation, if not undertaken by the industry collectively, puts the first company to self-regulate at a significant disadvantage. In contrast, if a major social media platform with a wide range of services works to establish the regulatory structure in conjunction with the government, it may be able to further cement its monopoly by crafting terms with which only the behemoth tech companies could possibly comply. This potential scenario began playing out during the April 10, 2018, Senate hearing with Mark Zuckerberg, when Senator Lindsey Graham of South Carolina asked the Facebook CEO what federal regulations he might recommend.

As of April 10, 2018, Facebook, along with Twitter, preemptively endorsed and supported the Honest Ads Act, which in 2017 sought the same disclosures for political and issue advertisements on social media as those on traditional print, radio, and television ads. Curiously, the social media companies more quickly made preemptive changes than did Congress, despite strong public support for the measure. It's unclear why something as simple as the Honest Ads Act would not be quickly passed by legislatures. This signals that the US Congress will be unlikely to regulate the social media companies any time soon.

Terms of Service: The Weapon of Social Media

Terrorists' exploitation of social media platforms led to creative ways to curb their online behavior. Absent clear laws and seeking to avoid censoring free speech, social media companies fell back on their terms of service for squelching the violent rants and images of jihadist terrorists. Russian disinformation has made this even more difficult. The Internet Research Agency, indicted by the Special Counsel's office in February 2018, sought to manipulate audiences via social media platforms by staying largely within the terms of service issued by each social media platform.[27] Nation-state actors, political campaigns, and public relations firms seek to cultivate social media nations, nudging them on issues rather than pushing directly for violence. Advocating for social or political issues does not violate the terms of service for social media companies.

Further terms-of-service restrictions with regard to speech and behavior will be difficult for social media companies to implement evenly across dozens of physical

states and could hurt their business models. The biggest hurdle in disrupting APMs will be identity verification. The most challenging debate in social media is anonymity versus authenticity. Should social media companies preserve account anonymity to prevent the persecution of people for their views and speech? Or should social media companies protect the safety of users and sustain the trust of users by ensuring real humans operate social media accounts?

The design of social media platforms affects how companies can tackle this challenge. Reddit and Twitter sought to be open platforms. Instituting identity verification would assist in disrupting APMs but may adversely affect their business models. Twitter has improved its social bot removal but remains plagued by bots' growth and expansion. Twitter has also undertaken a new and novel approach to curb "behaviors that distort and detract from the public conversation on Twitter."[28] Twitter, after identifying corrosive conversations, will examine whether an account's email address has been confirmed or if the same entity signs up for multiple accounts to thwart coordinated attacks.[29] The proposals appear to be a step forward for thwarting manipulators on their platform, but the changes are too recent to assess now.

Facebook has been the best and first to detect and check the spread of Russian disinformation. The platform's robust understanding of its users allows for better identity verification.[30] Moving forward, curbing the manipulators of social media nations can best be done by expanding verification mechanisms on social media platforms. Twitter's blue checkmark shouldn't be limited but, rather, expanded, as it will reinforce users' trust that what they are seeing on the platform is authentic. Similarly, permitting social bots to proliferate on platforms does far more damage to users and social media nations than any positive gains that might be achieved through the inauthentic replication of accounts and their chatter. Eliminating manipulative social bots will help democracies and also assist in restoring user trust.

Restoring the Integrity of Information

False information and its dubious sources have been the bane of both social media companies and democratic governments. APMs seek to use the freedoms of speech and the press to achieve their objectives of influencing audiences, maligning adversaries, and distorting reality. Democratic governments are poorly positioned to counter disinformation and misinformation on social media. Any government regulation or even funding of research into disinformation provides fuel to adversarial social media nations propelling conspiracies and may even empower an APM's narratives against a democratic government. Attempts by social media companies to curb false information proliferating on their platforms have found mixed success. Tagging "fake news" failed quickly at Facebook and other places as APMs can create false news faster than anyone can detect and screen it.[31] Facebook has since sought to

have users rate news outlets, but this will likely lead to further confirmation bias as users will rate highly those outlets that confirm their views.[32] Google News Initiative committed $300 million to improving "the accuracy and quality of news appearing on its platform" by teaming up with traditional news publishers.[33] These new methods and initiatives have only just begun and can't yet be assessed for their effectiveness.

Social media companies, since the US presidential election of 2016, still struggle with false information and the outlets that spread it. Again, a better solution that social media companies can pursue is the development of an information rating agency. The rating agency would reside outside of the social media companies and would be funded by them collectively. The agency would produce a rating icon for news-producing outlets displayed next to their news links in social media feeds and search engines. The icon would provide users an assessment of the news outlet's ratio of fact versus fiction and reporting versus opinion. The opt-out or opt-in rating system appearing in social media feeds would not infringe on freedom of speech or freedom of the press but would inform users as to the veracity of content and its disposition.[34] Users wanting to consume information from outlets with a poor rating wouldn't be prohibited, and if they were misled about the truth they would have only themselves to blame. Cumulatively, the public, the mainstream media, and social media companies would all benefit from the rating system without restricting or regulating individual freedoms. Information outlets that perform well and receive strong ratings would garner more clicks, advertising revenue, and even subscriptions.

Social Media Intelligence Center—Proactive Rather Than Reactive

Countering APMs seeking to destroy democracies and ruining social media platforms requires a new mindset: intelligence-led social media security. Social media companies repeatedly fail to detect bad actors on their platforms because their approach remains too reactive and technical, overly reliant on searching for knowns and unable to anticipate the unknowns. Democratic governments have been equally behind, unaware of the seeds of dissension spread among their constituents and displaying a bit of hubris in the face of the evidence of what was coming.

As with counterterrorism and cybersecurity, social media security must change its mindset and pursue a proactive, intelligence-led approach to get in front of the APMs wrecking governments and companies. Rather than focusing on technical vulnerability assessments and system signatures, social media companies must move beyond better threat intelligence to map out the biggest and most prolific APMs operating on their systems. Again, the solution for social media companies and democracies is an independent effort: a social media intelligence center (SMIC) able to bring together technologists with social engineers to anticipate—rather than react to—what bad actors will attempt to do on social media platforms. The goal will be

to protect the platform and, in so doing, protect customers by sustaining their trust in social media companies. Social media companies might consider hiring reformed propagandists or public relations experts (social media white hackers) who strengthen the resiliency of each platform, its current features, and its future innovations.

A SMIC would provide the threat intelligence needed for government and social media companies to protect themselves against APMs. The SMIC would spot and assess APMs across all platforms, connecting threat behavior with personas. The intelligence center would identify the most prolific and capable APMs, attributing them to known entities and studying their methods for information attacks. Even further, the SMIC's intelligence analysts would identify or create policy changes, tools, and processes for tracking and disrupting bad actors across multiple platforms. The SMIC would help all social media companies, exchanging threat information between the platforms and significantly aiding smaller social media companies with limited resources for defending against APM infestations. The SMIC would issue technical signatures and warnings to both the private and public sector through alerts and intelligence reports. Public education materials could be issued to help users understand and avoid threats they encounter on social media. A diverse, interdisciplinary staff with the social media equivalent of white hat hackers in cyberspace might reside alongside geopolitical and intelligence specialists as well as skilled technicians in machine learning and data analytics. A consolidated SMIC, rather than the partitioned security efforts of social media companies, could better detect and disrupt APMs in a far more efficient manner.

A parallel goal for the intelligence-led approach should be to avoid impeding innovation and competitiveness. Security professionals have a tendency, if allowed to dominate government or industry, to overcorrect toward protections that can cripple the functionality or appeal of goods and services. Social media companies have been an important economic driver for democracies and their technology innovations require freedom of thought and maneuver. Sustaining innovation while providing security will require a delicate balance. Otherwise, security concerns could lead to the death of platform engagement.

Next Step: Government and Social Media Summit, Milestones, and Timeline

Government and social media platforms have battled each other on legislative floors while they've been battered by APMs in cyberspace. The next step moving forward is to bridge the divide between the two, as they are ultimately on the same side. Protecting real nations and virtual ones comes by developing solutions rather than simply passing blame. The goal in the next six months should be to move past hearings and testimony to develop a clear timeline and milestones for protecting constituents and users on the social media battlefield. If government and social media platforms don't move closer together, social media subnations residing in democracies

and the APMs that manipulate them will continue pushing them further apart. Neither governments nor social media companies will survive if they stay on their current trajectory.

Providing their citizens with information, an understanding of fact and fiction, and a distinguishable difference between real and virtual is the only way democracies can survive and thrive. The divergent pathways of democratic governments and social media companies compound and accelerate the movement and polarization of social media nations. In the United States, and increasingly in the West, time is fleeting. Failure by either party to secure social media platforms from APMs and dissuade social media nations from overtaking physical nations can and will lead to permanent damage to democracies.

NOTES

1 Jack Posobiec (@JackPosobiec), "Correspondence, documents, and photos from Macron and his team," Twitter, May 5, 2017, accessed August 31, 2018, https://twitter.com/jackposobiec/status/860567072010620929?lang=en.

2 Ibid.

3 For background, see "Meet the Rebel's NEW Washington, DC Correspondent, Jack Posobiec," TheRebel.media, April 3, 2017, accessed August 31, 2018, https://www.therebel.media/meet_the_rebel_s_new_washington_dc_correspondent_jack_posobiec.

4 Andrew Marantz, "The Far-Right American Nationalist Who Tweeted #MacronLeaks," *New Yorker*, May 7, 2017. accessed August 31, 2018, https://www.newyorker.com/news/news-desk/the-far-right-american-nationalist-who-tweeted-macronleaks.

5 Ibid.

6 Tom Miles, "U.N. Investigators Cite Facebook Role in Myanmar Crisis," Reuters, March 12, 2018, accessed August 31, 2018, https://www.reuters.com/article/us-myanmar-rohingya-facebook/u-n-investigators-cite-facebook-role-in-myanmar-crisis-idUSKCN1GO2PN; Joel Rosenblatt, "Facebook Resists Inquiry into Cambodia Voter-Manipulation Claim," Bloomberg, April 29, 2018, accessed August 31, 2018, https://www.bloomberg.com/news/articles/2018-04-30/facebook-resists-inquiry-into-cambodia-voter-manipulation-claims; Lauren Etter, "What Happens When the Government Uses Facebook as a Weapon?" Bloomberg, December 7, 2017, accessed August 31, 2018, https://www.bloomberg.com/news/features/2017-12-07/how-rodrigo-duterte-turned-facebook-into-a-weapon-with-a-little-help-from-facebook.

7 There are many surveys and estimates of mobile phone use. For an example, see Sarah Perez, "U.S. Consumers Now Spend 5 Hours per Day on Mobile Devices," *TechCrunch*, March 3, 2017, accessed August 31, 2018, https://techcrunch.com/2017/03/03/u-s-consumers-now-spend-5-hours-per-day-on-mobile-devices.

8 Ibid. Also see Shannon Greenwood, Andrew Perrin, and Maeve Duggan, "Facebook Usage and Engagement is on the Rise, While Adoption of Other Platforms Holds Steady," Pew Research Center, November 11, 2016, accessed August 31, 2018, http://www.pewinternet.org/2016/11/11/social-media-update-2016.

9 A thorough examination of subcultures in the United States is by Colin Woodard, *American Nations: A History of the Eleven Rival Regional Cultures of North America* (New York: Penguin Group, 2011).

10 See generally Robert Putnam, *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon & Schuster, 2001).

11 For an example of the counters to Robert Putnam's arguments and subsequent research to evaluate the impact of the Internet, see Andrea L. Kavanaugh and Scott Patterson, "The Impact of Community Computer Networks on Social Capital and Community Involvement," *American Behavioral Scientist* 45, no. 3 (November 1, 2001): 496–509.

12 Many studies have examined the relationship of the Internet and social media to social capital formation. One example of this is Miki Caul Kittilson and Russell J. Dalton, "Virtual Civil Society: The New Frontier of Social Capital?" *Political Behavior* 33, no. 4 (December 2011): 625–44. See also Thierry Penard and Nicolas Poussing, "Internet Use and Social Capital: The Strength of Virtual Ties," *Journal of Economic Issues* 44, no. 3 (September 2010): 569–95.

13 Claire Cain Miller, "How Obama's Internet Campaign Changed Politics," *New York Times*, November 7, 2008, accessed August 31, 2018, https://bits.blogs.nytimes.com/2008/11/07/how-obamas-internet-campaign-changed-politics.

14 For a comparative discussion of the effects of social media on the Arab Spring, see Taylor Dewey, Juliane Kaden, Miriam Marks, Shun Matsushima, and Beijing Zhu, "The Impact of Social Media on Social Unrest in the Arab Spring," Stanford University report for the Defense Intelligence Agency, March 20, 2012, accessed August 31, 2018, https://publicpolicy.stanford.edu/publications/impact-social-media-social-unrest-arab-spring.

15 "Social Media: A Double-edged Sword in Syria," Reuters, July 13, 2011, accessed August 31, 2018, https://www.reuters.com/article/us-syria-social-media-idUSTRE76C3DB20110713?sp=true.

16 Zack Whittaker, "Surveillance and Censorship: Inside Syria's Internet," CBS News, November 12, 2013, accessed August 31, 2018, https://www.cbsnews.com/news/surveillance-and-censorship-inside-syrias-internet.

17 Amarnath Amarasingam and Charlie Winter, "ISIS's Perverse, Bloody Interpretation of Ramadan," *Atlantic*, May 26, 2017, accessed August 31, 2018, https://www.theatlantic.com/international/archive/2017/05/ramadan-isis-attack-muslim/528336.

18 Jessie Singal, "Undercover with the Alt-Right," *New York Times*, September 19, 2017, accessed August 31, 2018, https://www.nytimes.com/2017/09/19/opinion/alt-right-white-supremacy-undercover.html.

19 "Once Considered a Boon to Democracy, Social Media Have Started to Look Like Its Nemesis," *Economist*, November 4, 2017, accessed August 31, 2018, https://www.economist.com/briefing/2017/11/04/once-considered-a-boon-to-democracy-social-media-have-started-to-look-like-its-nemesis.

20 See the Texas secessionist movement's virtual and physical manifestations at Aaron Mak, "Here Are Some of the Social Media Posts that Russia Used to Meddle in the 2016 Election," *Slate*, November 1, 2017, accessed August 31, 2018, http://www.slate.com/articles/technology/technology/2017/11/here_are_the_facebook_posts_russia_used_to_meddle_in_the_2016_election.html. And for the California state-splitting campaign, see John Myers, "Radical Plan to Split California into Three States Earns Spot on November Ballot," June 12, 2018, accessed August 31, 2018, http://www.latimes.com/politics/la-pol-ca-california-split-three-states-20180612-story.html.

21 "Soviet Active Measures in the 'Post-Cold War' Era 1988–1991," US Information Agency report for US House of Representatives Committee on Appropriations, June 1992, accessed September 1, 2018, http://intellit.muskingum.edu/russia_folder/pcw_era/index.htm#Content.

22 A more detailed explanation of this phenomenon and its root causes is available in Clint Watts, *Messing with the Enemy: Surviving in a Social Media World of Hackers, Terrorists, Russians, and Fake News*, chap. 9 (New York: Harper Collins, 2018).

23 For background on the term APT and what it means, see "Advanced Persistent Threats: A Symantec Perspective," Symantec, accessed September 1, 2018, https://www.symantec.com/content/en/us/enterprise/white_papers/b-advanced_persistent_threats_WP_21215957.en-us.pdf.

24 Bruce Sterling, "Ransomware: The Basics," *Wired*, May 13, 2017, accessed September 1, 2018, https://www.wired.com/beyond-the-beyond/2017/05/ransomware-the-basics.

25 For a recent example of cyber-enabled information warfare employed for the traditional purpose of discrediting an adversary, see Andrew Higgins, "Foes of Russia Say Child Pornography Is Planted to Ruin Them," *New York Times*, December 9, 2016, accessed September 1, 2018, https://www.nytimes.com/2016/12/09/world/europe/vladimir-putin-russia-fake-news-hacking-cybersecurity.html.

26 "Far-right Smear Campaign against Antifa Exposed by Bellingcat," BBC, August 24, 2017, accessed September 1, 2018, https://www.bbc.com/news/blogs-trending-41036631.

27 See United States of America v. Internet Research Agency LLC et al., United States District Court for the District of Columbia, accessed September 1, 2018, https://www.justice.gov/file/1035477/download.

28 Del Harvey and David Gasca, "Serving Healthy Conversation," Twitter Blog, May 15, 2018, accessed September 1, 2018, https://blog.twitter.com/official/en_us/topics/product/2018/Serving_Healthy_Conversation.html.

29 Ibid.

30 For a comparison of Facebook and Twitter and the data they have on users to assist in account verification, see Matt Burgess, "Here's What Twitter and Facebook Know about You," *Wired*, May 19, 2017.

31 Steven Musil, "Facebook Dumps 'Disputed Flags' on Fake News for Context," CNet, December 20, 2017, accessed September 1, 2018, https://www.cnet.com/news/facebook-dumps-disputed-flags-on-fake-news-for-context.

32 See Facebook's new policy for surveying selected users to gauge their familiarity and trust with different news sources: Adam Mosseri, "Helping Ensure News on Facebook Is from Trusted Sources," Facebook, January 19, 2018, accessed September 1, 2018, https://newsroom.fb.com/news/2018/01/trusted-sources.

33 "Google Launches News Initiative to Combat Fake News," *The Star Online* (Petaling Jaya, Malaysia), March 21, 2018, accessed September 1, 2018, https://www.thestar.com.my/tech/tech-news/2018/03/21/google-launches-news-initiative-to-combat-fake-news.

34 Clint Watts and Andrew Weisburd, "Can the Michelin Model Fix Fake News?" *The Daily Beast*, January 22, 2017, accessed September 1, 2018, https://www.thedailybeast.com/can-the-michelin-model-fix-fake-news.

The preferred citation for this publication is: Clint Watts, “Advanced Persistent Manipulators and Social Media Nationalism: National Security in a World of Audiences,” Hoover Working Group on National Security, Technology, and Law, Aegis Series Paper No. 1812 (September 18, 2018), available at https://www.lawfareblog.com/advanced-persistent-manipulators-and-social-media-nationalism-national-security-world-audiences.

About the Author

CLINT WATTS

Clint Watts is the author of *Messing with the Enemy: Surviving in a Social Media World of Hackers, Terrorists, Russians and Fake News*.

He is a distinguished research fellow at the Foreign Policy Research Institute, senior fellow at the Center for Cyber and Homeland Security, and non-resident fellow at the Alliance for Securing Democracy.

Working Group on National Security, Technology, and Law

The Working Group on National Security, Technology, and Law brings together national and international specialists with broad interdisciplinary expertise to analyze how technology affects national security and national security law and how governments can use that technology to defend themselves, consistent with constitutional values and the rule of law.

The group focuses on a broad range of interests, from surveillance to counterterrorism to the dramatic impact that rapid technological change—digitalization, computerization, miniaturization, and automaticity—are having on national security and national security law. Topics include cybersecurity, the rise of drones and autonomous weapons systems, and the need for—and dangers of—state surveillance. The group's output will also be published on the Lawfare blog, which covers the merits of the underlying legal and policy debates of actions taken or contemplated to protect the nation and the nation's laws and legal institutions.

Jack Goldsmith and Benjamin Wittes are the cochairs of the National Security, Technology, and Law Working Group.

For more information about this Hoover Institution Working Group, visit us online at http://www.hoover.org/research-teams/national-security-technology-law-working-group.

Hoover Institution, Stanford University
434 Galvez Mall
Stanford, CA 94305-6003
650-723-1754

Hoover Institution in Washington
The Johnson Center
1399 New York Avenue NW, Suite 500
Washington, DC 20005
202-760-3200

HOOVER INSTITUTION