Accounting for Foreign Disinformation: National Security Regulatory Proposals for Social Media Accounts and False Speech

Jonathan A. Schnader[♦]

I. INTRODUCTION

The ubiquity of social media in modern society carries with it a host of national security concerns, some of which require new approaches, new legal frameworks, and new policies to adequately address them. The current major national security concerns in the social media space are the active measures and disinformation campaigns.[1] Most notable among them are the ones led by Kremlin-backed elements. These campaigns not only lead to interference in Western democratic institutions, but they also cause panic through fake news stories amplified by bots and trolls. All of these efforts ultimately feed into a disinformation feedback loop.

Two examples of successful Russian disinformation campaigns are particularly noteworthy. First, in the Incirlik Air Base incident, false news about terrorists overtaking the United States' base circulated worldwide, causing a small protest outside the gates of the base.[2] Second, the false story

---

♦ Jonathan A. Schnader lives in Washington, D.C. He earned his bachelor's degree in Psychology and Classical Humanities from Miami University of Ohio. He earned his J.D. *cum laude* from Syracuse University College of Law with a Certificate of Advanced Study in National Security and Counterterrorism Law. Following law school, he worked as an Assistant Public Defender in Rochester, NY for five and a half years, handling just under four thousand criminal cases. In addition to being licensed to practice law in New York and Washington D.C., he is a Certified Anti-Money Laundering Specialist. Jonathan graduated with distinction in 2019 from Georgetown University Law Center, where he completed a Master of Laws in National Security. His academics and current practice focus on national security dimensions of several areas, including cybersecurity; artificial intelligence; blockchain and cryptocurrency; intelligence and counterintelligence; and social media.

[1] This article uses the term "disinformation," but some of the cited sources use the terms "misinformation" and "disinformation" interchangeably. For a discussion of "disinformation," that is, intentionally incorrect information, versus, "misinformation" and "mal-information," *see* Alice E. Marwick, *Why do People Share Fake News? A Sociotechnical Model of Media Effects*, 2 GEO. L. TECH REV. 474, 478 (2018).

[2] *See, e.g.*, Clint Watts, *Clint Watts' Testimony: Russia's Info War on the U.S. Started in 2014*, THE DAILY BEAST (Mar. 30, 2017), https://www.thedailybeast.com/clint-watts-testimony-russias-info-war-on-the-us-started-in-2014; Craig Timberg, *Russian Propaganda Effort Helped Spread 'Fake News' During Election, Experts Say*, WASH. POST (Nov. 24, 2016), https://www.washingtonpost.com/business/economy/russian-propaganda-effort-helped-spread-fake-news-during-election-experts-say/2016/11/24/793903b6-8a40-4ca9-b712-716af66098fe_story.html.

about the explosion of the Columbia Chemical plant in Louisiana caused a panic on social media.[3]

While terrorist recruitment on social media platforms and mass-hacks of personal information from social media accounts constitute national security risks, the sophisticated and unified Russian disinformation feedback loop challenges fundamental American values, such as free speech. Mark Zuckerberg has described Facebook as an "idealistic and optimistic company" and a wonderful tool for connecting people.[4] Notwithstanding this positivity, Zuckerberg has also acknowledged his platform's shortcomings:

> [Facebook] didn't do enough to prevent these tools from being used for harm as well. That goes for fake news, foreign interference in elections, and hate speech, as well as developers and data privacy. We didn't take a broad enough view of our responsibility, and that was a big mistake. It was my mistake, and I'm sorry. . . It's not enough to just connect people, we have to make sure those connections are positive. It is not enough to just give people a voice, we have to make sure people aren't using it to hurt people or spread misinformation.[5]

Access to vast troves of data and information by foreign agents allows for those malign actors to conduct orchestrated campaigns against United States interests using social media. The data security issues, election manipulation efforts of Russia, and mass data collection practices "also raise the possibility that regulators, policymakers, consumers, and even the platforms themselves may be significantly underestimating the risks of data-fueled analytics and automated technology."[6]

Another major concern for social media platforms is the antagonism between regulation of the social media space and the First Amendment, which prohibits, in relevant part, the government from abridging freedom of speech. Social media platforms have evolved into not only the modern

---

[3] Adrian Chen, *The Agency*, N.Y. TIMES MAG. (June 2, 2015), https://www.nytimes.com/2015/06/07/magazine/the-agency.html.

[4] *Facebook, Social Media Privacy, and the Use and Abuse of Data: Joint Hearing Before the S. Comm. on the Judiciary, S. Comm. on Commerce, Sci., & Transp.*, 115th Cong. 8 (2018) (testimony of Mark Zuckerberg, Chairman and Chief Executive Officer, Facebook).

[5] *Id.*

[6] Terrell McSweeny, *Psychographics, Predictive Analytics, Artificial Intelligence, & Bots: Is the FTC Keeping Pace?*, 2 GEO. L. TECH. REV. 514, 514-15 (2018).

soapboxes for individuals, but also the virtual bulletin board and outlet for news from all around the world. Consequently, historical notions of traditional media and First Amendment protections of speech continue to be challenged by the universality of social media. These platforms provide accessibility to worldwide information, news, and other data to both normal members of the public as well as terrorists, foreign agents, and hackers.

The purpose of this article is to propose solutions that facilitate the regulation of social media without impinging on cherished First Amendment rights. First, this analysis will discuss the social media traits that present national security issues. Next, this article will turn to the Russian disinformation campaign orchestrated by the Kremlin-backed Internet Research Agency. Then, it will address First Amendment issues with regard to social media. Penultimately, it will evaluate existing or recently proposed regulatory schemes for social media. Finally, a proposed regulatory scheme, implemented under the President's emergency powers pursuant to the International Emergency Economic Powers Act ("IEEPA"), will demonstrate how a regulation can prohibit the intentional dissemination of fake or misleading information, under false pretenses, without trampling First Amendment protections.

II. INHERENT SOCIAL MEDIA CHARACTERISTICS THAT PRESENT NATIONAL SECURITY PROBLEMS

In this context, social media refers to "internet connected platforms and software used to collect, store, aggregate, share, process, discuss or deliver user-generated and general media content, that can influence knowledge, perception and thereby directly or indirectly prompt behaviour as a result of interaction."[7]

As suggested by Professor Ryan Calo, an expert on technology, privacy, and AI issues, some businesses seek to maximize their customer exposure by employing tactics that exploit the vast amount of data available to entities through social media:

> [T]echnology captures and retains intelligence on the consumer's interaction with a given firm. Today, consumer interactions leave a record of the consumer's behavior. A conservative list of information a commercial website might collect could include how many times the consumer

[7] THOMAS ELKJER NISSEN, #THEWEAPONIZATIONOFSOCIALMEDIA 123 (2015).

> has been to the website before; what website the consumer was visiting immediately before arriving; what pages the consumer visited, and for how long; what items the consumer purchased; what items the consumer almost purchased; where the consumer physically was; and what computer or browser the consumer was using. Furthermore, firms might combine the data with public or private information purchased from a third party. Using this compiled and stored information, firms can then run complex algorithms to convert mere behavior into insight (and value).[8]

These *legal* methods of data analytics demonstrate the newfound availability of user data on social media platforms. The ease with which businesses legally gather data about people on a huge scale raises a bright red flag from a national security standpoint: If businesses can use social media information to gather this kind (and volume) of data, it must not be difficult for agents of foreign States to accomplish the same feats—they need only create their own algorithms to learn about their target population. Businesses capitalize on the availability of information for millions of social media users, and the same is true of foreign State agents conducting disinformation campaigns or disseminating propaganda.

Additionally, the structure of social media differs from traditional forms of media in significant ways: First, virtually any user can create their own content and broadcast or distribute that content for consumption on a social media platform.[9] Second, a user's existing networks and connections on the platform facilitate the transmission of the content.[10] Third, algorithms built into a platform will adjust the content that appears for a user, ultimately displaying content that aligns with a user's past activity and preferences.[11] In other words, social media promotes then forwards content preferred by the user, to the user. That strategy vastly differs from that of traditional media, which requires the reader to sift through content and make a choice. In the context of social media, the content shown to the user is chosen by the platform rather than the user him/herself. If disinformation broadcasted on a social media platform aligns with users' views, they will be more likely to

---

[8] Ryan Calo, *Digital Market Manipulation*, 82 G. WASH. L. Rev. 995, 1003-04 (2014).
[9] Marwick, *supra* note 1, at 503.
[10] *Id.*
[11] *Id.*

see that disinformation and subsequently share it with other like-minded users.

The primary goal of social media platforms is to connect individuals, not to determine the truth of a statement or discern the factual quality of a post. Considering its goal in facilitating interaction between connected users, social media makes sharing opinions, stories, and thoughts so easy; all it takes is a few clicks of the mouse to transmit an idea, re-share an online article, post a picture, or live-stream an event happening in real time. For better or for worse, the effect of this easy access is that vast quantities of information about people or events can be found on social media platforms, and naturally, more traditional media outlets have understandably attempted to integrate information gleaned from social media sites into the news. According to Danish Military Analyst, Thomas Nissen, who has written extensively on martial uses of social media:

> Media agendas (and not least the media's sources) are to a high degree now informed by social network media, making them a hugely powerful tool. This, however, also presents an unprecedented challenge for the media in terms of source criticism or validation of the attribution and validity of the information picked up from social network media.[12]

While traditional media outlets pluck stories and information from social media, social media users digest and interpret news from traditional media sources, and republish or share them on the same or different social media platforms.

Much of the traditional media content trickles down through the filter of a person's social feed on a social media platform, and "this stream is affective; among social media participants, news is 'collaboratively constructed out of subjective experience, opinion, and emotion,'" and in this "social [space], the traditional journalistic value of objectivity no longer makes sense: virtually every story is augmented with someone's opinion."[13] Because the social media platform prioritizes information so that the content itself aligns with its users, those users will be increasingly receptive to information because it mirrors their own views. Terrorist recruiters take advantage of the social media prioritization of information based on user

---

[12] NISSEN, *supra* note 7, at 99.

[13] Marwick, *supra* note 1, at 504 (citing Zizi Papacharissi, *Toward New Journalism(s) Affective News, Hybridity, and Liminal Spaces*, 16 JOURNALISM STUD., 27 (2015)).

preference: Extremist recruiters zero in on social media users who align with specific viewpoints or join particular ideological groups that suggest a susceptibility to being cajoled into espousing terrorist ideologies. From there, recruiters befriend those users, and forward information and stories that attract them, then isolate them in encrypted messaging applications.[14] Terrorist recruiters use this approach strategy because users with similar viewpoints are more closely situated in the social media space because their content preferences result in connections (hashtags and "likes," for instance). Indeed, just as people that "like" musical theater may be grouped together and easily viewed in a list, so too are people who "like" semi-automatic rifles or "like" the thought of banning them. Thus, easily viewed ideological or preferential groupings on the platforms provide a simple and effective way to target a specific user based on his/her interests, which are conveniently listed in the social media interface on the computer screen.

III. RUSSIAN INFORMATION CAMPAIGNS USING SOCIAL MEDIA

In order to protect against future Russian disinformation campaigns or election meddling, it is crucial to understand the Russian strategy employed to undermine United States national security. Recently, Special Counsel Robert S. Mueller III released the 448 page Report on the Investigation into Russian Interference in the 2016 Presidential Election, which includes two volumes comprehensively detailing, in relevant part, the scope of the Russian disinformation campaign.[15] The Kremlin's aggressive and sophisticated social media propaganda campaign generates fake news as a part of a self-reinforcing news cycle, while simultaneously flooding the internet with false information. This not only directs people to its narrative, but it also sows mistrust in western institutions and democratic systems.

> Russia appears to actively synchronize social media products with those of various other information outlets, including Russian-branded TV broadcasts and web news, proxy civil society agencies, and web outlets. However, the Kremlin's web campaign that relies on anonymous web comments and non-attributed social media content

[14] *See id.* at 506-07. Notably, social media terrorist recruitment is an important topic with vast scholarly attention and substantial legislation, but it is not the focus here.

[15] U.S. DEP'T OF JUSTICE, REPORT ON THE INVESTIGATION INTO RUSSIAN INTERFERENCE IN THE 2016 PRESIDENTIAL ELECTION (2019). This article uses numerous sources, but it is important to note that Special Counsel Mueller's report provides a deep examination the Russian disinformation campaign.

> disseminated by bots and trolls offers Russia the opportunity to target unsuspecting audiences with malign and often fake-news content.[16]

Mark Zuckerberg declared Facebook's commitment to protecting the veracity of information flowing through Facebook, saying that "[i]t is not enough to just give people a voice; we need to make sure that voice isn't used to harm other people or spread misinformation."[17] According to the testimony of Jack Dorsey to the Senate Select Committee on Intelligence, "Twitter's built and measured by how we help encourage more healthy debate, conversations, and critical thinking. Conversely, abuse, malicious automation, and manipulation detracts from it."[18]

Recently, various experts testified before Congress to discuss Russian active measures and disinformation operations through social media and their effects on United States democratic processes. Clint Watts, fellow at the Foreign Policy Research Institute, noted:

> While Russia certainly seeks to promote Western candidates sympathetic to their worldview and foreign policy objectives, winning a single election is not their end goal. Russian active measures hope to topple democracies through the pursuit of five complementary objectives: [o]ne, undermine citizen confidence in democratic governance; [t]wo, foment and exacerbate divisive political fissures; [t]hree, erode trust between citizens and elected officials and their institutions; [f]our, popularize the Russian policy agenda within foreign populations; [a]nd five, create general distrust or confusion over information sources by blurring the lines between fact and fiction, a very pertinent issue today in our country.[19]

---

[16] TODD HELMUS, ET AL., RUSSIAN SOCIAL MEDIA INFLUENCE, UNDERSTANDING RUSSIAN PROPAGANDA IN EASTERN EUROPE 25 (2018).

[17] *Hearing*, *supra* note 4.

[18] *Open Hearing on Foreign Influence Operations' Use of Social Media Platforms (Company Witnesses): Hearing Before the S. Select Comm. on Intelligence*, 115th Cong. 22 (2018) (testimony of Jack Dorsey, Chief Executive Officer, Twitter, Inc.).

[19] *Disinformation: A Primer in Russian Active Measures and Influence Campaign, Panel I: Hearing Before the S. Select Comm. on Intelligence*, 115th Cong. 30 (2017) (statement of Clint Watts, Robert A. Fox Fellow, Foreign Policy Research Institute).

A salient example of a coordinated disinformation campaign purportedly perpetrated by Kremlin-backed actors is the Columbia Chemical scare, in which numerous social media accounts alleged that a chemical plant in Louisiana had exploded. Senator Marco Rubio described the false story as "not some simple prank . . . but a highly coordinated disinformation campaign involving dozens of fake accounts that posted hundreds of tweets for hours, targeting a list of figures precisely chosen to generate maximum attention."[20]

In his statement to the Senate Committee on the Judiciary and the Senate committee on Commerce, Science and Transportation, Mark Zuckerberg described the nature of the Kremlin-backed disinformation and election meddling campaign. He explained, "What we found was that bad actors had used coordinated networks of fake accounts to interfere in the election, promoting or attacking specific candidates and causes, creating distrust in political institutions, or simply spreading confusion."[21] It is important to note that fake accounts artificially increased the visibility of disinformation narratives, thus driving the Russian disinformation machine. The nexus between fake accounts and social media posts will ultimately underpin this article's regulatory proposals in its conclusion.

A. Disinformation Operations

"The Kremlin has built a complex production and dissemination apparatus that integrates actors at varying levels of attribution to enable large-scale and complex information operations."[22] Essentially, the feedback loop of disinformation has three levels of what Helmus et al. refer to as the "grayscale of deniability": the first level consists of "white outlets" overtly attributable to the Russian state, which include "a constellation of Russian state-controlled, state affiliated, and state censored media and think tanks, such as [Russia Today] and Sputnik News," among others.[23] The second level, known as "gray" outlets because of their "uncertain attribution," include conspiracy websites, as well as far left or far right websites, news aggregators, and data dump websites.[24] The third level, considered to be the level of "covert attribution, referred to as 'black' in the grayscale of

---

[20] *Id.* at 46 (statement of Sen. Marco Rubio, Member, Senate Select Committee on Intelligence).
[21] *Hearing*, *supra* note 4, at 11-12.
[22] HELMUS ET AL., *supra* note 16, at 11.
[23] *Id.*
[24] *Id.* (citing Andrew Weisburd et al., *Trolling for Trump*: *How Russia is Trying to Destroy Our Democracy*, WAR ON THE ROCKS (Nov. 6, 2016), https://warontherocks.com/2016/11/trolling-for-trump-how-russia-is-trying-to-destroy-our-democracy).

deniability, produce content on user-generated media, such as YouTube, but also add fear-mongering commentary to amplify content produced by others and supply exploitable content to data dump websites."[25]

B. Trolls, Bots, Honeypots

"*Trolls*, *bots*, and *honeypots* all refer to fake social media accounts used for various purposes, but trolls and honeypot accounts are operated by humans, while bot accounts are automated."[26] One threat posed by "bots" is that they "magnify the number of hits [a computer] might get to a particular social media site."[27] Similarly, "you can create more personas in Twitter . . . which makes it look like there are more people than there really are."[28] Clint Watts describes how large numbers of bots and fake social media accounts receive internet-wide visibility by "amplif[ying] your appearance":

> [S]o what they do is they launch those simultaneously as they begin the engagement or push of false news stories, usually from [Russia Today] and Sputnik News. They do that in unison, which games the social media system such that such a high volume of content being pushed at the same time raises that into the trends that you'll see. If you look at Facebook or Twitter or whatever it might be, you'll see the top ten stories that are out right now. It pushes that up there. As soon as it pushes that into that top ten feed, mainstream media outlets then are watching that and they start to examine that content.[29]

According to Jack Dorsey, although Russian-linked accounts that tweeted about the election amounted to "less than two one-hundredths of a percent (0.016%) of the total accounts on Twitter at the time. Of all election-related tweets that offered during that period, these malicious accounts constituted approximately one percent (1.00%), totaling 2.12 million Tweets."[30] Clint Watts, in his prepared statement to the Senate Select Committee on Intelligence outlined how Kremlin-backed actors essentially

---

[25] *Id.*
[26] *Id.*
[27] *Hearing*, *supra* note 19, at 48 (statement of Sen. Mark Warner, Vice Chairman, Senate Select Committee on Intelligence).
[28] *Id.* (statement of Clint Watts, Robert A. Fox Fellow, Foreign Policy Research Institute).
[29] *Id.*
[30] *Hearing*, *supra* note 18.

digested massive amounts of public data from social media accounts, which they used to then construct fake social media accounts that were virtually indistinguishable from a regular American social media user. The Kremlin-backed actors then used those fake American personae to promote the Russian foreign policy goals, including creating chaos and confusion during elections.[31]

The creation of fake social media accounts in the form of automated bots or manned trolls gives visibility to foreign agents orchestrating social media disinformation campaigns, and therefore any regulation of social media should begin by addressing the creation of social media accounts.

IV. SOCIAL MEDIA AND FIRST AMENDMENT CONCERNS: PROTECTED VS. UNPROTECTED SPEECH

The First Amendment prohibits Congress from "abridging the freedom of speech, or of the press."[32] A fundamental principle of the First Amendment is that "all persons have access to places where they can speak and listen, and then, after reflection, speak and listen once more."[33] Indeed, "[w]hile in the past there may have been difficulty in identifying the most important places (in a spatial sense) for the exchange of views, today the answer is clear. It is cyberspace—the 'vast democratic forums of the Internet' in general, and social media in particular."[34] Justice Kennedy, in the *Packingham* decision, eloquently describes the importance of social media in the modern context, explaining how banning access to social media generally violates the First Amendment:

> North Carolina with one broad stroke bars access to what for many are the principle sources for knowing current events, checking ads for employment, speaking and listening in the modern public square, and otherwise exploring the vast realms of human thought and knowledge. These websites can provide perhaps the most powerful mechanisms available to a private citizen to make his or her voice heard.[35]

---

[31] *Hearing*, *supra* note 19, at 48 (statement of Clint Watts, Robert A. Fox Fellow, Foreign Policy Research Institute).

[32] U.S. CONST. amend. I.

[33] Packingham v. North Carolina, 137 S. Ct. 1730, 1735 (2017).

[34] *Id.* (quoting Reno v. ACLU, 521 U.S. 844, 868 (1997)).

[35] *Id.* at 1737.

The Constitutional free-speech analysis

> requires a court to determine whether the law (1) regulates a category of speech that is unprotected under the First Amendment or enjoys something less than full protection, giving the government the regulatory authority, and whether the law (2) is a content-based restriction–which are presumed invalid under strict scrutiny–or a content-neutral restriction–which are subject to intermediate scrutiny, a less speech-protective test.[36]

Some speech, however, is not entitled to heightened constitutional protection, like true threats, fraud, child pornography, libel, incitement, defamation, and imminent threats the government has the power to prevent.[37] The Supreme Court uses the following derivative of the "clear and present" danger test:[38]

> Speech advocating the use of force or crime can only be proscribed where (1) the speech is "directed to inciting or producing imminent lawless action"–a requirement of intent; and (2) the advocacy is also "likely to incite or produce such action." Importantly, when the Court examines the strength of the government interest proffered today, it "unmistakably insists that any limit on speech be grounded in realistic, factual assessment of harm."[39]

Public expression has evolved from traditional soap-box or street corner speeches to posts on social media, and that shift requires an evaluation of how to protect constitutional rights while guarding national security interests.

---

[36] Louis W. Tompros et al., *The Constitutionality of Criminalizing False Speech Made on Social Networking Sites in a Post-Alvarez, Social Media-Obsessed World*, 31 HARV. J. L. & TECH 65, 89 (2017).

[37] *Id.* (citing United States v. Alvarez, 567 U.S. 709, 717 (2012)).

[38] Schenk v. United States, 249 U.S. 47, 52 (1919).

[39] Tompros et al., *supra* note 36, at 92 (first citing Brandenburg v. Ohio, 395 U.S. 444, 447 (1969) and then citing United States v. Williams, 553 U.S. 285, 321-22 (2008) (Souter, J., dissenting)).

V. POTENTIAL PARADIGMS FOR REGULATING SOCIAL MEDIA

Social media platforms strongly favor the free expression of ideas, and any attempt to regulate them will be met by intense opposition. For example, "although social media companies recognize that terrorist content should be removed, stricter regulation could 'ruin' social media platforms by deterring normal users from posting objectionable but non-extremist content."[40] Indeed, "[t]his type of 'chilling effect' prevents speakers from exercising their rights to expression, which, although objectionable, may provide valuable contributions to public discourse and debate."[41] However, some commentators believe that "existing governmental measures are inadequate to the extent they allow terrorist activity on the Internet."[42] Mark Zuckerberg acknowledged that the ubiquity of social media platforms, Facebook in particular, creates the need for "some regulation."[43]

A. Proscribing "False Speech" on Social Media

> Prosecution and ultimate convictions for 'social media crimes' in the United States are few and far between, but in other nations like the United Kingdom, people are regularly prosecuted and convicted for such offenses,[44] including 'two people [who] were sentenced to four years in prison for spreading false information through posts on Facebook during the 2011 riots.'[45]

There are states that do attempt to regulate some forms of speech. For instance, New York's paradigm for criminalizing false speech falls under a

---

[40] Paulina Wu, *Impossible to Regulate: Social Media, Terrorists, and the Role for the UN*, CHI J. INT'L L. 281, 300 (2015) (citing Ben Flanagan & Asma Ajroudi, *ADMS: Facebook Shuns 'Full Regulation' Despite ISIS Threat*, AL ARABIYA NEWS (Nov. 18, 2014), https://english.alarabiya.net/en/media/2014/11/18/ADMS-Facebook-shuns-full-regulation-despite-ISIS-threat).

[41] *Id.* (citing Frederick Schauer, *Fear, Risk and the First Amendment: Unraveling the Chilling Effect*, 58 B.U. L. REV. 685, 691-92 (1978)).

[42] Susan Klein & Crystal Flinn, *Social Media Compliance Programs and the War against Terrorism*, 8 HARV. NAT'L SEC. J. 53, 72 (2017).

[43] *Facebook: Transparency and Use of Computer Data: Hearing Before H. Comm. on Energy & Commerce*, 115th Cong. 33 (2018) (testimony of Mark Zuckerberg, Chairman and Chief Executive Officer, Facebook).

[44] Becky Evans, *5,000 People Investigated by Police for Something They Said on Facebook or Twitter as 'Social Network Crime' Soars 800%*, DAILY MAIL ONLINE (Dec. 27, 2012), http://www.dailymail.co.uk/news/article-2253692/Facebook-Twittercrime-sees-fold-increase-police-deal-5-000-cases-involving-websites.html.

[45] Tompros et al., *supra* note 36, at 92.

"Falsely Reporting an Incident" statute in the Penal Law. The misdemeanor of "Falsely Reporting an Incident in the Third Degree" states:

> A person is guilty of falsely reporting an incident in the third degree when, knowing the information reported, conveyed or circulated to be false or baseless, he or she: 1. Initiates or circulates a false report or warning of an alleged occurrence or impending occurrence of a crime, catastrophe or emergency under circumstances in which it is not unlikely that public alarm or inconvenience will result.[46]

The penalties under New York law for a conviction of Falsely Reporting an Incident in the Third Degree, a Class A Misdemeanor, are a mandatory state surcharge ($200), $50 DNA fee, up to $1000 fine, and up to one year in jail. The purpose of the statute is to prevent behavior akin to Justice Holmes' famous example of yelling "fire" in a crowded theater.[47]

Other states have similar false report paradigms for proscribing false speech.[48] There still exists a question of whether laws like New York state's Falsely Reporting an Incident statute are constitutional.[49] Indeed, the New York Court of Appeals, New York's highest court, struck down a cyberbullying statute that prohibited "any act of communicating or causing a communication to be sent by mechanical or electronic means . . . with the intent to harass, annoy, threaten, abuse, taunt, intimidate, torment, humiliate, or otherwise inflict significant emotional harm on another person."[50] The court held that while the government did have a compelling interest to "[protect] children from harmful publications or materials," the statute had "alarming breadth . . . [which] would criminalize a broad spectrum of speech outside the popular understanding of cyberbullying."[51] The concern about applying false reporting statutes as a paradigm to control false or misleading information on social media is that "broad false reporting statutes like the one in New York may counterproductively restrict" well intentioned, but false speech that in theory should be protected under the First Amendment.[52]

---

[46] N.Y. PENAL LAW § 240.50 (McKinney 2013).
[47] Schenk v. United States, 249 U.S. 47, 52 (1919).
[48] Tompros et al., *supra* note 36, at 84-86.
[49] *Id.* at 101-04.
[50] People v. Marquan, 19 N.E.3d 480, 484 (N.Y. 2014) (citation omitted).
[51] *Id.* at 486.
[52] Tompros et al., *supra* note 36, at 108.

B. The Ban on Providing Material Support to Foreign Terrorist Organizations as Applied to Social Media Platforms

Existing law "prohibits the willful provision of anything of value to a group designated as an Foreign Terrorist Organization ("FTO") *if* the provider knows that such organization has either been so designated, or knows that it engages in terrorism."[53] In *Holder v. Humanitarian Law Project*, the Supreme Court found that the statutory ban on providing material support to terrorism was constitutional as applied to United States citizens and domestic organizations who "wanted to assist the lawful political and humanitarian ends of two designated FTOs."[54] The Court held that the statute passed strict scrutiny because of the compelling government interest in national security.[55] The Court also determined that it was narrowly tailored, reasoning that "foreign organizations that engage in terrorist activity are so tainted by their criminal conduct that any contribution to such an organization facilitates that conduct" and even "promot[ing] peaceable, lawful conduct . . . can further terrorism by foreign groups in multiple ways."[56] Rachel VanLandingham suggests that:

> Even speech that is nowhere near incitement nor a true threat becomes criminal when uttered on the behalf of, to, or even simply in coordination with a foreign terrorist group. . . [T]he knowing coordination of value-providing speech, or other such conduct, with a terrorist group is sufficient to criminalize it.[57]

VanLandingham undertakes an exhaustive application of the ban on providing material support to terrorists to social media platforms, concluding that a social media platform may have difficulty determining whether "a particular user is actually an FTO or someone working in coordination with such a group," and even due diligence on the platform's part may not reveal the answer.[58] Consequently, platforms tend to "suppress all content that indicates support of an FTO in order to remain clear of §

---

[53] Rachel E. VanLandingham, *Jailing the Twitter Bird: Social Media, Material Support to Terrorism, and Muzzling the Modern Press*, 39 CARDOZO L. REV. 1, 4 (2017) (describing 18 U.S.C. § 2339B (2012)).

[54] *Id.* at 32 (citing Holder v. Humanitarian Law Project, 561 U.S. 1, 9 (2010)).

[55] *Humanitarian L. Project*, 561 U.S. at 31-32.

[56] VanLandingham, *supra* note 54, at 35 (citing *Humanitarian L. Project*, 561 U.S. at 30).

[57] *Id.* at 35 (citing *Humanitarian L. Project*, 561 U.S. at 43).

[58] *Id.* at 43.

2339B's reach" which may be an overbroad self-regulation, resulting in the censorship of some permissible speech under the First Amendment.[59] An alternate to self-imposed regulation is a reporting requirement on the part of social media platforms that could both offset any criminal liability under the Material Support statute for allowing terrorist communications, and help law enforcement address terrorist recruitment efforts. As described by Klein and Flinn, "[t]hese social media sites must be encouraged to discover offending posts and report them to federal law enforcement authorities to avoid what on a practical level constitutes complicity with terrorist organizations."[60]

C. Conspiracy to Defraud the United States

In 2018, Special Counsel Robert S. Mueller III unsealed an indictment against the purportedly Kremlin-backed Internet Research Agency ("IRA") and a number of other Russian nationals alleged to be foreign agents for their part in influencing the United States presidential elections through social media.[61] Because the federal criminal code lacks a specific statute proscribing social media manipulation, Special Counsel Mueller used a classic federal criminal statute as his basis for the indictment: Conspiracy to Defraud the United States. This statute criminalizes, in relevant part, two or more persons who "conspire either to commit any offense against the United States, or to defraud the United States, or any agency thereof in any manner or for any purpose."[62] Under the statute's umbrella also falls "any conspiracy for the purpose of impairing, obstructing, or defeating the lawful function of any department of government"[63] through "deceit, trickery, or at least by means that are dishonest."[64]

Notably, no federal statute specifically applies to the fake social media accounts created by the IRA. Consequently, Mueller had to take the additional step of connecting the social media disinformation disseminated by the IRA to some other theory of criminality. He creatively linked the disseminated disinformation to the Foreign Agent Registration Act ("FARA"),[65] which requires agents of foreign principals to register "so that

---

[59] *Id.* at 44; *see also*, Klein & Flinn., *supra* note 42, at 69-70 ("As long as these sites continue to openly provide fora for the distribution of terrorist material, each one of them provides material support to an FTO, which, if done knowingly, would be in direct contravention of 18 U.S.C. § 2339B.").

[60] Klein & Flinn, *supra* note 42, at 70.

[61] Indictment, United States v. Internet Research Agency, LLC., Case 1:18-cr-00032-DLF, (D.D.C. 2018), https://www.justice.gov/file/1035477/download.

[62] 18 U.S.C. § 371 (2012).

[63] Haas v. Henkel, 216 U.S. 462, 479 (1910).

[64] Hammerschmidt v. United States, 265 U.S. 182, 188 (1924).

[65] 22 U.S.C. § 612.

the U.S. government and the people of the United States are informed of the source information and the identity of persons attempting to influence U.S. public opinion, policy, and law."[66] Mueller alleged that the IRA, in attempting to influence the elections and use money to support a particular candidate or agenda, failed to register as an agent of a foreign principal pursuant to FARA. Mueller also linked the IRA's conduct to defrauding the Federal Election Commission, which is tasked with "providing the American public with accurate data about . . . entities supporting federal candidates," and likewise prohibits "foreign nationals from making any contributions, expenditures . . . or disbursements for electioneering communications."[67] Lastly, under the theory of defrauding the United States, Mueller coupled the Russian social media disinformation campaign with the fraudulent representations made by IRA agents on visa applications, as well as misrepresentations they made about their identities and purposes in the United States.[68]

D. Computer Fraud and Abuses Act ("CFAA")

The CFAA prohibits knowing, unauthorized access to specifically designated computer systems, government networks, and protected computers, and prohibits the dissemination of information contained in those systems.[69] The statute thus lays out two ways to commit the crime of improperly accessing a protected computer: (1) obtaining access without authorization; and (2) obtaining access with authorization but then using that access improperly.[70] In a recent civil case, plaintiff Ticketmaster overcame a motion against them to dismiss the civil CFAA complaint in which the defendant business used automated bots to purchase large numbers of tickets online from Ticketmaster with the goal of reselling them.[71] The court held that "the proper inquiry is whether Ticketmaster has sufficiently pled that Defendants accessed Ticketmaster's computers . . . in excess of the authorization they did have."[72] The court went on to explain that "each use of a bot to purchase a ticket was a use in excess of authorization because an individualized cease-and-desist letter sent to [Defendant] Prestige . . .

[66] Indictment, United States v. Internet Research Agency, LLC., Case 1:18-cr-00032-DLF at 11, (D.D.C. 2018), https://www.justice.gov/file/1035477/download.
[67] *Id.*
[68] *Id.* at 12.
[69] *See* 18 U.S.C. § 1030.
[70] Musacchio v. United States, 136 S. Ct. 709, 713 (2016).
[71] Ticketmaster, LLC v. Prestige Entm't West, Inc., 315 F. Supp. 3d 1154, 1172 (C.D. Cal. 2018)
[72] *Id.* at 1169.

explicitly prohibited Prestige and other Defendants from using bots to access Ticketmaster's website."[73] Therefore, the Defendants accessed the Ticketmaster website in a manner explicitly forbidden to them.[74]

Importantly, "the phrase 'exceeds authorized access' in the CFAA does not extend to violations of use restrictions . . . [because] [i]f Congress wants to incorporate misappropriation liability into the CFAA, it must speak more clearly."[75] In summary of the aforementioned principles, the Ninth Circuit established two instructive rules to guide understanding of criminal liability under the CFAA in the realm of social media.

> First, a defendant can run afoul of the CFAA when he or she has no permission to access a computer or when such permission has been revoked explicitly. Once permission has been revoked, technological gamesmanship or the enlisting of a third party to aid in access will not excuse liability. Second, a violation of the terms of use of a website–without more–cannot establish liability under the CFAA.[76]

Hypothetically speaking, under the Ninth Circuit analysis, a social media platform, once it identifies a malign actor, foreign agent, bot account, etc., could not only deactivate the account under the policies in their terms of service, but could also send a specific cease-and-desist notice to the actor or user. Should that same actor or user create a new account or circumvent the platform's safeguards tailored to prevent their access, federal prosecutors could charge them under the CFAA.

E. Social Media Self-Regulation

The increasing use of social media across the globe has prompted numerous providers to implement rules for use of their platforms, using terms of service and privacy policies as self-imposed regulation.

> Social network media have also started to act as a form of "gatekeepers" themselves with the evolving legal role of "terms of use" and the platform administrator's arbitrary

[73] *Id.*
[74] *Id.* at 1172.
[75] United States v. Nosal, 676 F.3d 854, 863 (9th Cir. 2012).
[76] Facebook v. Power Ventures, Inc., 844 F.3d 1058, 1067 (9th Cir. 2016).

> decisions about which posts, videos and imagery to hold on their sites and which to erase or which accounts should be entirely closed.[77]

The social media platforms, through their terms of service, have wide latitude to control access and content.

Facebook recently implemented new Artificial Intelligence ("AI") algorithms and tools to prevent disinformation and election meddling. Referencing French and German presidential elections, as well as the Alabama Special Election in 2017, Mark Zuckerberg pronounced that "the AI tools that we deployed in those elections were able to proactively take down tens of thousands of fake accounts that may have been trying to [influence those elections]."[78] Clearly, the terms of service provide a functional means of policing fake accounts linked to foreign agents or terrorist recruiters.

Interestingly, Facebook does not allow hate groups. According to Mark Zuckerberg, "If there is a group that their primary purpose or a large part of what they do is spreading hate, we will ban them from the platform."[79] Indeed, Facebook acknowledged that it generally needed to adjust its algorithms to prevent those interested in violence or bad activities from being connected with other like-minded individuals.[80] To demonstrate its commitment to removing terrorist related content, Facebook hired a former federal prosecutor, Monika Bickert, in the position of "global policy management" to lead its anti-terrorism efforts. Bickert has implemented a zero-tolerance policy for such material.[81]

Notwithstanding the efforts of Facebook to prevent terrorist propaganda, self-regulation by the social media platforms themselves is insufficient. For instance, Twitter asserts that an algorithm to seek out and reliably flag terrorist-related information is difficult to create and deploy.[82] Several social media platforms have been criticized for either refusing to remove terrorist-

---

[77] NISSEN, *supra* note 7, at 123.

[78] *Facebook: Transparency and Use of Computer Data: Hearing Before H.R. Comm. on Energy and Commerce*, 115th Cong. 28, 28 (2018) (testimony of Mark Zuckerberg, Chairman and Chief Executive Officer, Facebook).

[79] *Id.*

[80] *Id.*

[81] Scott Higham & Ellen Nakashima, *Why the Islamic State Leaves Tech Companies Torn between Free Speech and Security*, WASH. POST, (July 16, 2015), https://www.washingtonpost.com/world/national-security/islamic-states-embrace-of-social-media-puts-tech-companies-in-a-bind/2015/07/15/0e5624c4-169c-11e5-89f3-61410da94eb1_story.html?utm_term=.0ed9e5c1fb61.

[82] Klein and Flinn, *supra* note 42, at 71 (citing Higham & Nakashima, *supra* note 82).

related content or failing to do so in a timely fashion.[83] Moreover, at least one lawmaker does not believe Facebook, can "be trusted to regulate itself."[84]

The social media space exists as a place, idealistically, to share ideas, connect with friends and family, and bring the world closer together. However, the same characteristics of social media that make the platforms an excellent place to exchange ideas also make them an excellent place for foreign agents to exploit United States persons. The challenge going forward will be balancing cherished rights like First Amendment free speech against the need to regulate social media to protect United States national security interests.

The most notable, recent threat to national security borne out of social media is the Russian disinformation campaign, discussed above. The Russian disinformation campaign not only interfered with the American democratic process. It also, in some instances, caused hysteria and widespread loss of faith in government and the press, *inter alia*. But, any regulation of social media must achieve a precarious balance between national security concerns and First Amendment issues.

VI. A PROPOSED SOCIAL MEDIA REGULATORY SCHEME

While Congress could enact a statute addressing the concerns of social media in the national security domain and propose corresponding legislation,[85] the more expedient legal mechanism for regulating social media lies with the President and the executive branch.

Through the International Emergency Economic Powers Act (IEEPA), Congress granted the President the power

> to deal with any unusual and extraordinary threat, which has its source in whole or substantial part outside the United

---

[83] *Id.*

[84] Dave Paresh, *U.S. Lawmaker Says Facebook Cannot be Trusted to Regulate Itself*, REUTERS, (Nov. 14, 2018), https://www.reuters.com/article/us-usa-facebook-congress/u-s-lawmaker-says-facebook-cannot-be-trusted-to-regulate-itself-idUSKCN1NJ38R; *see also* Eli Sanders, *Facebook's Attempt at Regulating Itself Isn't Good Enough, Attorney General Says*, THE STRANGER, (Dec. 11, 2018), https://www.thestranger.com/slog/2018/12/11/36659524/facebooks-attempt-at-regulating-itself-isnt-good-enough-attorney-general-says.

[85] Although the Federal Trade Commission ("FTC") could "flex and stretch its existing authorities and resources to meet" the growing challenges posed by technological innovations, particularly in a social media space, "it would be far better for Congress to strengthen the agency and the protections afforded consumers for their data necessary authorities and resources [sic]." McSweeny, *supra* note 6, at 530.

> States, to the national security, foreign policy, or economy of the United States . . . the President may, under such regulations as he may prescribe, by means of instructions, licenses, or otherwise – investigate, block . . . regulate, direct and compel, nullify, void, prevent or prohibit, any acquisition, holding, withholding, use, transfer . . . or dealing in, or exercising any right, power, or privilege with respect to, or transactions involving, any property in which any foreign country or a national thereof has any interest by any person, or with respect to any property, subject to the jurisdiction of the United States.[86]

Congress' express authorization under the IEEPA, coupled with the President's foreign relations power,[87] means that the President's power is at its zenith with regard to regulating national security, foreign policy, and economic matters, so long as the President declares the object of regulation to be an "unusual and extraordinary threat" and also declares a state of emergency.[88]

The President has the authority, therefore, to proclaim a state of emergency with regard to social media if the threat of social media manipulation constitutes an unusual and extraordinary threat. In general, the IEEPA's language refers to property interests and commercial transactions. The legal keystone allowing the IEEPA to govern social media regulation is intangible property jurisprudence, which suggests that people and businesses have proprietary interests in their social media accounts.[89] Requiring any regulation be based on monetary harm would further vest the regulations with the authority of the IEEPA. Consequently, social media accounts implicate the IEEPA, and their regulation falls under the umbrella of the executive if the President declares an emergency with regard to social media.

---

[86] 50 U.S.C. §§ 1701-1702(a)(1)(B) (2012).

[87] *See, e.g.*, United States v. Curtiss-Wright Export Corp., 299 U.S. 304, 319 (1936) (holding "the President is the constitutional representative of the United States with regard to foreign nations.").

[88] Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579 (1952).

[89] *See* In re CTLI, 528 B.R. 359, 374 (Bankr. S.D. Tex. 2015) (concluding in relevant part "the social media accounts were property of the estate"); Salonclick LLC v. SuperEgo Mgmt. LLC, 2017 WL 239379 at 4 (S.D.N.Y. Jan. 18, 2017) (holding that domain names and social media accounts were "property" capable of being trespassed upon); *see also* Thyroff v. Nationwide Mut. Ins., 8 N.Y.3d 283, 291 (2007) (stating that, in a case of conversion of an intangible property, "it cannot be seriously disputed that society's reliance on computers and electronic data is substantial, if not essential.").

A. Goals of the Proposed Regulations

The goal of the regulation will be to protect against emerging national security threats, the most serious of which are "State-backed disinformation campaigns" similar to the disinformation campaign orchestrated by the Kremlin-backed IRA. Any regulation of social media will brush against the First Amendment, so the proposed regulations must not chill speech unnecessarily.

The regulations' specific purpose would be to regulate fake accounts and bots, which facilitate the spread of fake information. Because even fake speech is protected under the First Amendment,[90] the regulation cannot simply prohibit the spread of false information. To prevent an overbroad prohibition, the regulation needs several layers of specific intent and a fraud-like or impersonation element, which does not enjoy the same level of First Amendment protection.[91] The regulation also must be worded in such a way that the prohibited behavior is "closer to conduct than speech" as to avoid prohibiting "pure speech."[92] Consequently, it makes sense to anchor the free-speech regulations dealing with conduct to the creation and misuse of social media accounts.

The regulation should include criminal and administrative violations, much like the CFAA: One directed at a social media *user* and one directed at the social media platform itself. The regulations should be aimed at global social media platform companies, the threshold for regulation being either the number of users or the amount of network traffic. Those platforms should be required to implement technical measures to seek out bots and fake accounts. They also ought to submit reports to a regulatory body that reviews the reports and issues rules and guidelines. Should a platform disagree with or challenge a rule or regulation, a hearing court headed by an Administrative Law Judge ("ALJ") could hear the challenge. The ALJ would also preside over any administrative action brought by the government against social media platforms.

---

[90] United States v. Alvarez, 567 U.S. 709, 727 (2012) (striking down the Stolen Valor Act as a content-based restriction on First Amendment grounds when defendant pleaded guilty to falsely claiming that he had received the Medal of Honor, reasoning "the remedy for speech that is false is speech that is true").

[91] United States v. Chappell, 691 F.3d 388, 396-97 (4th Cir. 2012) (noting "significantly, no [Supreme Court] Justice thought it advisable to drape a broad cloak of constitutional protection over actionable fraud, identity theft, or the impersonation of law enforcement officers"); *see also Alvarez*, 567 U.S. at 719 (stating that "falsity alone may not suffice to bring the speech outside the First Amendment," reasoning that "[t]he statement must be a knowing or reckless falsehood").

[92] *Chappell*, 691 F.3d at 396.

The proposed regulation thus considers First Amendment concerns and aims to not overburden social media platforms or users by over-regulating with onerous rules.

B. Proposed Regulations

Social Media and National Security Administration Regulations ["SMANSAR"]

§1: Authority

Pursuant to the President's foreign affairs power and the IEEPA, the President, by finding that foreign State-backed disinformation campaigns and social media manipulation constitute an "unusual and extraordinary threat," and having declared a state of emergency[93] with respect to such social media manipulation, hereby creates the following regulations.

§2: Purpose

To proscribe the dissemination of false information on social media platforms through false pretenses, such as fake accounts and bots, with the goal of curbing foreign State-backed propaganda and disinformation campaigns influencing American free exercise of constitutional rights and interfering with democratic institutions; to create an oversight body for global social media platforms to ensure their proactive measures to prevent the creation of fake accounts and bots; to promote information and idea sharing between the government and global social media platforms and their users; and to foster discourse, discussion, and free speech concomitant to the evolution of technology and social networks.

§3: Definitions

a) "Global social media platform" means an internet or cyberspace-based public or quasi-public forum for discussing and sharing information, socializing, and creating a network among people, with at least ten million users.[94]

b) "Dissemination" means to post, broadcast, release, make public, tweet, retweet, publish, or otherwise distribute information with the purpose of exposing other people to that information.

---

[93] This section presumes that the President declared a state of emergency and that social media constituted an unusual and extraordinary threat.

[94] It seems important that only social media platforms with widespread and global reach should be regulated.

c) "False pretenses" means, with the intent to deceive, misrepresent, or trick any other user or users or global social media platforms; to use the account(s), login(s), and/or credentials of fake or invented users; or to impersonate another real person, business, organization, or persons, businesses, organizations or use the account(s) login(s), and/or credentials of other real person or persons, business or businesses, organization or organizations; or to use software applications in order to run automated scripts in fake account(s) or login(s) or an account(s) or login(s) not belonging to that person that augment, bolster, or mask that person's cyberspace presence.[95]

d) "User" means a natural person, business, organization, corporation, or other verifiable, formal, official identity or entity, that employs the services of a global social media platform.

e) "Harm" means any degree of monetary loss, depreciation of value, physical harm, or measurable emotional distress or fear.

§4: The Regulatory Body

a) This section shall create the Social Media and National Security Administration ("SMANSA"), a regulatory body, under the auspices of the Department of Homeland Security ("DHS");

b) SMANSA shall:

1. review all reports submitted by global social media platforms as well as all complaints submitted by users and global social media platforms;

2. create a mechanism for user submission of complaints, catalogue those complaints, and review them in a timely manner;

3. review allegations of all violations of SMANSAR and be responsible for bringing administrative and/or civil action for alleged violations, but shall refer alleged criminal violations to the Department of Justice National Security Division;

4. be responsible for issuing additional rules and regulations under SMANSAR it finds germane to SMANSAR's purposes, considering information from user complaints and global social media platform reports submitted pursuant to §5, subject to a six-month notice and comment period.[96]

---

[95] The intent of this definition is to capture "bots." The proscriptions in these proposals focus on the creation of accounts to avoid prohibiting "pure speech."

[96] Although not expressly addressed here, SMANSA would have to comply with rulemaking requirements of the Administrative Procedures Act. 5 U.S.C. § 553 (2012).

§5: Reports for Global Social Media Platforms

Global social media platforms shall:

a) submit reports bi-annually to SMANSA explaining proactive measures taken to prevent foreign State-backed disinformation campaigns, including a general description of those measures and policies, and/or any accompanying technical measures undertaken, except that this section shall not require global social media platforms to disclose specific trade secrets, algorithms, or other proprietary information;

b) issue reports for any apparent violations by users, discovered by global social media platforms under these rules, within forty-eight hours.[97]

§6: Social Media Disinformation

Any person who intentionally disseminates information known to be false or misleading on a global social media platform, under false pretenses, causing any degree of harm to another user, users, or global social media platform,[98] shall be guilty of social media disinformation.[99]

A person charged with social media disinformation violation may be subject to administrative, civil or criminal penalties, or a combination thereof.

§7: Foreign Social Media Disinformation Campaign

Any person, acting on behalf of, or in furtherance of the interests of a foreign State, a non-State terrorist or criminal organization, or any foreign State agent or instrumentality, or foreign State organ, who commits the

[97] This section derives from the "suspicious activity reports" required under the Bank Secrecy Act for combatting Anti-Money Laundering. *See* 12 C.F.R. § 21.11 (2020).

[98] The "harm" element was ultimately added to the social media disinformation violation for three reasons: it creates a measurable element of harm that will help the regulation survive First Amendment analysis by a Court; it connects the false pretenses element more closely to fraud, which is unprotected speech under First Amendment jurisprudence; and finally, it further strengthens the regulatory attachment to IEEPA, which requires an economic nexus for the President to implement regulation under its framework. Additionally, the monetary harm element may be easy to prove if a person impersonates another user because, arguably, the impersonated person is suffering a trespass on their proprietary social media account, the deprivation of which is a monetary loss.

[99] This violation has several levels of *mens rea*: intent to disseminate information; knowledge of the information's falsity; and intent to deceive (under the false pretenses definition). A person's conduct must be extremely specific to fall under this section, and ignorance or even recklessness are insufficient to create liability. This prohibition should be a lesser criminal violation, with consequences like fines, community service, supervised release, or at worst, two years or less of imprisonment. The violation also protects global social media platforms by specifically naming them as a potential victim.

violation of social media disinformation pursuant to §6 shall be guilty of conducting a foreign social media disinformation campaign.[100]

§8: Global Social Media Platform Responsibility Violation[101]

A Global Social Media Platform shall be guilty of an administrative or civil violation and subject to a correspondent administrative or civil penalty, if the social media platform:

a) fails or neglects to submit reports in accordance with these rules and regulations; or

b) fails or neglects to implement reasonable safeguards to prevent the use of false pretenses on the platform as outlined in §3; or

c) fails to implement or abide by, or monitor user compliance with these rules or regulations.

§9: Terms of Service Requirements

Global Social Media Platforms must require a user's compliance with these rules as a part of their terms of service.[102]

§10: Creation of an Administrative Hearing Body:

a) The SMANSAR Court will preside over administrative actions brought by SMANSA for violations arising under these rules.

b) Administrative actions will be heard by administrative law judges, appointed by the President of the United States, and adjudicated in accordance with the Administrative Procedures Act.[103]

C. Limitations of the Proposed Regulations

The above proposed regulations are not without limitations. Global social media companies will likely protest any kind of oversight. Although the regulations do not seem to be particularly onerous or restrictive, they do burden social media platforms with the task of self-policing (which they would pay for) and liability for non-compliance.

---

[100] This violation is the primary crime and national security threat that the regulations are intended to prevent. Thus, the penalties for this crime should be severe. The "in furtherance" prong is meant to be difficult to achieve, in order to prevent overbroad application of the crime.

[101] Importantly, this violation puts no onus on the social media platform to stop false speech itself.

[102] One additional way to regulate the creation of fake accounts would be to require "know your customer" protections, commonly used in an anti-money laundering context. Global social media platforms could require proof of identity, such as submission of identifying information (driver's license numbers, pictures of a person holding up their driver's license, etc.).

[103] 5 U.S.C. § 500 *et seq.* (2012).

Moreover, budgetary constraints always play a role in how an agency functions. The regulatory body would exist as another piece of the DHS patchwork of responsibilities, and while DHS seems like the logical place for this kind of regulatory body, it would likely have to stretch its resources even further to cover the management of SMANSA.

Of chief concern is that the regulation would criminalize discussion, creativity, art, political discourse, satire, and many forms of otherwise protected speech. However, the language of the violations prohibits only a narrow type of conduct, and the addition of an element of harm may also limit potential regulatory overreach.

D. First Amendment Analysis

SMANSAR would likely survive First Amendment analysis based on current jurisprudence. Indeed, to say something "presents a First Amendment issue is not necessarily to say that it constitutes a First Amendment violation."[104]

The analysis under the First Amendment is twofold: "(1) does the rule regulate a category of speech that is unprotected under the First Amendment or enjoys something less than full protection, giving the government regulatory authority" and "(2) is [the law] a content-based restriction--which are presumed invalid under strict scrutiny--or a content-neutral restriction– which are subject to intermediate scrutiny, a less speech-protective test?"[105]

First, turning to unprotected speech under the First Amendment, the regulations include a fraud element requiring misrepresentation and actual, demonstrable harm, which would render the proscribed fake speech unprotected in terms of First Amendment jurisprudence.[106] The framework likely falls within the purview of governmental regulation because it prohibits a specific type of fraud rather than a broad ban on the false speech at issue in *Alvarez*,[107] and will thus be subject to the highly deferential

---

[104] Members of City Council of L.A. v. Taxpayers for Vincent, 466 U.S. 789, 803-04 (1984) (quoting Metromedia, Inc. v. San Diego, 453 U.S. 490, 561 (1981) (Burger, C.J., dissenting)).

[105] Tompros et al., *supra* note 36, at 89.

[106] *See, e.g.*, Illinois ex rel. Madigan, v. Telemarketing Assoc., Inc., 538 U.S. 600 (2003). In *Illinois ex rel. Madigan*, the Court found that the government may impose disclosure requirements for fundraising practices, and that "[s]tates may maintain fraud actions when fundraisers make false or misleading representations designed to deceive donors about how their donations will be used." 538 U.S. at 624.

[107] *See, e.g.*, United States v. Alvarez, 567 U.S. 709, 718-20 (2012). *See also* United States v. Chappell, 691 F.3d 388, 400 (4th Cir. 2012) (concluding that "[t]he First Amendment is a central and essential part of our constitutional life . . . [f]alsely identifying oneself as a policeman in order to get out of a speeding ticket is simply not the kind of expressive conduct the Framers . . . had in mind.").

rational basis standard.[108] However, even assuming that the speech does not fall into the category of unprotected speech, it could be colored as such because it tends "to incite an immediate breach of the peace." For instance, the Russian disinformation cycle induced panic in the United States by disseminating stories about made-up events like the Columbia Chemical plant explosion scare or the protest outside of Incirlik Air Force Base.[109]

Secondly, and more likely, the false speech proscribed in SMANSAR is subject to a heightened level of scrutiny (most likely strict scrutiny) because it targets a specific content—false speech. Still, even in light of the jurisprudential presumption of invalidity that accompanies content-based restrictions, SMANSAR provisions are narrowly tailored—that is, they "further a compelling state interest by the least restrictive means."[110] Indeed, the proposed rules address compelling government interests relating to national security[111]—namely Russian social media disinformation campaigns aimed at toppling democratic institutions—which are particularly salient in light of the rapid expansion of social media. Additionally, the rules proscribe a narrow form of fake speech—fake speech made to deceive others, while impersonating a real person, or using a fake account (or bot account), resulting in some form of harm. In other words, if a person broadcasts fake information from a legitimate account, his/her conduct or speech lives outside of the umbrella of SMANSAR rules. Notably, there exists a distinction between anonymous speech,[112] which is protected, and speech made from a fake or impersonated identity, because impersonation is a type of fraud when used to gain some benefit.[113] These

---

[108] Tompros et al., *supra* note 36, at 89. *See also Alvarez*, 567 U.S. at 722, (rejecting "the notion that false speech should be in a general category that is presumptively unprotected"). However, if the false speech is "made to effect a fraud or secure moneys or other valuable considerations, say offers of employment, it is well established that the Government may restrict speech without affronting the First Amendment." *Id.* at 723 (citing Virginia Bd. of Pharmacy v. Virginia Citizens Consumer Council, 425 U.S., 748, 771 (1976)).

[109] Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942).

[110] Holder v. Humanitarian Law Project, 561 U.S. 1, 46 (2010) (internal citation and quotations omitted).

[111] *See, e.g.*, *id.* at 28 (stating "[e]veryone agrees that the Government's interest in combating terrorism is an urgent objective of the highest order.").

[112] John Wiley & Sons, Inc. v. Doe Nos. 1-30, 284 F.R.D. 185, 189 (S.D.N.Y. 2012) (stating that "internet users have a limited First Amendment privacy interest in anonymous internet usage, including the use of peer to peer file copying networks to download, distribute or make available for distribution copyrighted material in electronic form," but "in the file-sharing context, First Amendment protection is limited and subject to other considerations") (internal quotations omitted) (citing Sony Music Entertainment Inc. v. Does 1-40, 326 F. Supp. 2d 556, 564 (S.D.N.Y. 2004)).

[113] United States v. Chappell, 691 F.3d 388, 392 (4th Cir. 2012) (stating that the "Virginia impersonation statute has a plainly legitimate sweep" because it protects "unsuspecting citizens from

rules would not prohibit anonymous speech itself, but would make it difficult to use fake information to create a social media account for the sole purpose of remaining anonymous.

VII. CONCLUSION

The existence of foreign State disinformation campaigns represents a real and burgeoning national security threat to the United States, as demonstrated by the Russian IRA's efforts to the influence the 2016 presidential election by harnessing social media to sow distrust, enmity, and discord in the United States. Russia pursued a similar goal throughout the Cold War through the use of traditional media, but with the recent rise of social media to the global level, Russia shifted its operations to this cheaper and more ubiquitous method for reaching millions of people. Lawmakers and social media entrepreneurs alike agree that social media needs *some* kind of regulation to secure its societal benefits—chiefly, the "positive connections" made between people and the free-flow of ideas. Without some kind of regulation, the social media space is vulnerable to attack by foreign State agents who attempt to undermine the principles, like the First Amendment, that make the United States the greatest incubator of discourse and ideas. We must implement regulation to shield our open fora from malign foreign influence while remaining cognizant that any rules must capture disinformation campaigns without overburdening free speech. The proposals in this article carefully consider First Amendment protections and concerns while managing the risks to national security inherent in the social media fabric.

---

those who falsely pretend to be law enforcement officials" to gain some benefit and thus "serves the Commonwealth's critical interest in public safety"). Impersonation cases can thus be distinguished from the false speech in *Alvarez*, because impersonation statutes require the actor to have an intent to harm or gain from the false statements. In *People v. Golb*, the New York Court of Appeals struck down a harassment statute on First Amendment grounds, but did not apply First Amendment analysis for the counts of criminal impersonation in the second degree, upholding nine out of fourteen counts. The Court found that "injury to reputation" satisfies the injury element of the statute, but also noted that an "email sent in another person's name does not prove the requisite intent to cause injury, either to reputation or otherwise." 23 N.Y.3d 455, 465-66 (2014).