

2019

## Social Media, Manipulation, and Violence

Allyson Haynes Stuart  
*Charleston School of Law*

Follow this and additional works at: <https://scholarcommons.sc.edu/scjilb>



Part of the [Business Organizations Law Commons](#), [Civil Procedure Commons](#), [Communications Law Commons](#), and the [Intellectual Property Law Commons](#)

---

### Recommended Citation

Stuart, Allyson Haynes (2019) "Social Media, Manipulation, and Violence," *South Carolina Journal of International Law and Business*: Vol. 15 : Iss. 2 , Article 7.

Available at: <https://scholarcommons.sc.edu/scjilb/vol15/iss2/7>

This Article is brought to you by the Law Reviews and Journals at Scholar Commons. It has been accepted for inclusion in South Carolina Journal of International Law and Business by an authorized editor of Scholar Commons. For more information, please contact [dillarda@mailbox.sc.edu](mailto:dillarda@mailbox.sc.edu).

## SOCIAL MEDIA, MANIPULATION, AND VIOLENCE

*Allyson Haynes Stuart\**

*For centuries now, inventions heralded as advances in human progress have been exploited by the criminal mind. New technologies, all too soon, can become instruments used to commit serious crimes. The railroad is one example . . . and the telephone another . . . So, it will be with the Internet and social media.<sup>1</sup>*

Many of us lament the ubiquity of social media<sup>2</sup> and the attention it takes from face-to-face activities,

---

*\* Allyson Haynes Stuart is a Professor of Law at Charleston School of Law. She wishes to thank her research assistant, Michael Grabara, and the editorial staff of the South Carolina Journal of International Law and Business. This article follows the 2018 Symposium, "Innovation, Policy, & Technology in Law and Development," where Allyson Haynes Stuart was a featured panelist.*

<sup>1</sup> *Packingham v. North Carolina*, 137 S. Ct. 1730, 1736 (2017) (citations omitted).

<sup>2</sup> Social media refers to “forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos).” *Social Media*, MERRIAM-WEBSTER.COM DICTIONARY <https://www.merriam-webster.com/dictionary/social%20media>. The social media websites (“sites”) I refer to primarily are Facebook, Twitter, Instagram, and YouTube, which are also referred to as third-party platforms. See Jonathan Peters & Brett Johnson, *Conceptualizing Private Governance in A Networked Society*, 18 N.C. J. L. & TECH. 15, 49 (2016) (“Google,

particularly on the part of young people. Social media generally stresses the importance of appearance, and makes most of us believe our lives pale in comparison to everyone else on the planet.<sup>3</sup> It has been linked to depression and low self-esteem in teenagers, and no one is completely sure of the long-term effects of small screen use on the body and brain.<sup>4</sup> But this essay addresses an even darker side to social media: its use for manipulation on scales both small and large and its direct and indirect ties to violence. Increasingly, social media is used as a tool to foment violence, particularly in regions of the world where access to the Internet is otherwise limited. Even when a social media site is not itself an instrument to foment violence, it may have the effect of encouraging and validating the extreme views that result in violence. What is the role of social media sites in containing their use for such nefarious purposes, and how should the law govern them? Should such sites be liable for failing to take down misleading or

---

Facebook, YouTube, and Twitter are examples of third-party platforms that offer a variety of services that enable their users to share content.”).

<sup>3</sup> Note Facebook’s recent “10 Year Challenge,” which encouraged users to post pictures of how little they had changed in 10 years. See Rebecca Jennings, *Why You’re Seeing the 10-Year Challenge Everywhere*, VOX (Jan. 16, 2019, 5:10 PM), <https://www.vox.com/the-goods/2019/1/16/18185256/10-year-challenge-facebook-meme>.

<sup>4</sup> A survey conducted by the Royal Society for Public Health of 14-24 year olds in the United Kingdom found that Snapchat, Facebook, Twitter, and Instagram all led to increased feelings of depression, anxiety, poor body image, and loneliness. See Rachel Ehmke, *How Using Social Media Affects Teenagers*, CHILD MIND INSTITUTE, <https://childmind.org/article/how-using-social-media-affects-teenagers/> (last visited Feb. 26, 2019).

inaccurate content or for allowing hate speech? If not, what are other options for controlling its misuse?

I. UBIQUITY OF SOCIAL MEDIA IN THE U.S. AND AROUND  
THE GLOBE

In a relatively short period of time, our world has been taken over by social media. Facebook is only fifteen years old, but sixty-eight percent of U.S. adults and fifty percent of U.S. teens use the social media platform.<sup>5</sup> Additionally, around forty percent of U.S. adults get their news from Facebook.<sup>6</sup> Thus, social media sites have now surpassed print newspapers as a news source for Americans; twenty percent of U.S. adults surveyed say that they often get their news via social media, as opposed to the sixteen percent who obtain their news from print newspapers.<sup>7</sup>

Globally, social media use has grown astoundingly. According to the Pew Research Center, seventy-five percent of adults in Jordan and an average of sixty-eight percent of adults in the Middle East and North Africa

---

<sup>5</sup> See John Gramlich, *10 Facts About Americans and Facebook*, PEW RESEARCH CENTER (Feb. 1, 2019), <http://www.pewresearch.org/fact-tank/2019/02/01/facts-about-americans-and-facebook/>.

<sup>6</sup> See *id.*

<sup>7</sup> In 2017, the portion of U.S. adults who got news via social media was about equal to the portion who got news from print newspapers. See Elisa Shearer, *Social Media Outpaces Print Newspapers in the U.S. as a News Source*, PEW RESEARCH CENTER (Dec. 10, 2018), <http://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/>.

now use social media.<sup>8</sup> Worldwide, an average of sixty percent of adults use social media.<sup>9</sup> In some countries, social media adoption has risen dramatically in the last few years. For example, “only forty-nine percent of Lebanese adults used social media in 2015, but in 2017, that number increased to seventy-two percent.”<sup>10</sup> Similarly, “just two years ago only fifty-one percent of South Korean adults were on social media, compared to almost sixty-nine percent in 2017.”<sup>11</sup>

In certain countries, social media sites have a monopoly status. In an extreme example, fewer than five percent of Myanmar’s population had mobile phones until 2014, when the government loosened restrictions and first allowed SIM cards to be sold at affordable levels.<sup>12</sup> After that, a much greater number of people bought smartphones—and those phones were all

---

<sup>8</sup> See Jacob Poushter, Caldwell Bishop & Hanyu Chwe, *Social Network Adoption Varies Widely by Country*, PEW RESEARCH CENTER (June 19, 2018), <http://www.pewglobal.org/2018/06/19/3-social-network-adoption-varies-widely-by-country/>.

<sup>9</sup> See *id.*

<sup>10</sup> *Id.*

<sup>11</sup> *Id.*

<sup>12</sup> See Euan McKirdy, When Facebook Becomes ‘The Beast’: Myanmar Activists Say Social Media Aids Genocide, CNN (Apr. 6, 2018, 8:05 PM), <https://www.cnn.com/2018/04/06/asia/myanmar-facebook-social-media-genocide-intl/index.html>; See Megan Specia & Paul Mozur, *A War of Words Puts Facebook at the Center of Myanmar’s Rohingya Crisis*, N.Y. TIMES (Oct. 27, 2017), <https://www.nytimes.com/2017/10/27/world/asia/myanmar-government-facebook-rohingya.html>.

preinstalled with Facebook.<sup>13</sup> The use of Facebook in Myanmar went from about two million in 2014 to over thirty million today.<sup>14</sup> There, Facebook “is the Internet” and is the way most people get their news.<sup>15</sup> That can have dire results when the news people receive is not necessarily true and is being used for improper purposes.<sup>16</sup>

## II. THE DARK SIDE OF SOCIAL MEDIA USE

There are myriad benefits from social media, like its use during the revolutions of the Arab Spring to put pressure on governments and spread truth in the face of propaganda.<sup>17</sup> But there are three primary problems:

---

<sup>13</sup> See Specia & Mozur, *supra* note 12.

<sup>14</sup> See *id.*

<sup>15</sup> See Barbara Ortutay, *Facebook Admits Not Doing Enough to Prevent Myanmar Violence*, WASH. TIMES (Nov. 6, 2018), <https://www.washingtontimes.com/news/2018/nov/6/facebook-admits-not-doing-enough-to-prevent-myanmar/>.

<sup>16</sup> See *infra* Incitement to Violence.

<sup>17</sup> See Kevin Gregg, “Text ‘Revolution’ to Vote”: *Social Media’s Effect on Popular Consent and Legitimacy of New Regimes*, 31 B.U. INT’L L.J. 315, 334 (2013) (describing the significant role of camera-phone videos posted on Facebook in Tunisia’s popular revolution); Kitsuron Sangsuvan, *Balancing Freedom of Speech on the Internet Under International Law*, 39 N.C. J. INT’L L. & COM. REG. 701, 755 (2014) (“[T]he history books will also note that 2011 marked the beginning of a new age when mass protest, revolution and armed conflict was not only facilitated by, but made possible through, digital communication networks and social networking sites. Facebook, Twitter, and YouTube had transformed civil society’s engagement with, and in, warfare.”); Philip N. Howard et al., *Opening Closed Regimes: What Was the Role of Social Media During the Arab Spring?*, PROJECT ON INFORMATION TECHNOLOGY AND POLITICAL ISLAM (2011) (unpublished manuscript),

misleading or inaccurate stories used to influence elections, misleading or inaccurate stories used to foment violence, and the more tangential effect on violence that results when people with extreme views find amplification and validation of those views on social media.

#### A. ELECTIONS

News had certain built-in safeguards when traditional media outlets controlled its distribution. Journalists and newspapers have reputations to protect, in addition to incentives to avoid liability, and so they follow safeguards that require source-vetting, fact-checking, and backup evidence.<sup>18</sup> There is a level of

---

([https://deepblue.lib.umich.edu/bitstream/handle/2027.42/117568/2011\\_Howard-Duffy-Freelon-Hussain-Mari-Mazaid\\_PITPI.pdf?sequence=1&isAllowed=y%20](https://deepblue.lib.umich.edu/bitstream/handle/2027.42/117568/2011_Howard-Duffy-Freelon-Hussain-Mari-Mazaid_PITPI.pdf?sequence=1&isAllowed=y%20)) (“Our evidence shows that social media was used heavily to conduct political conversations by a key demographic group in the revolution – young, urban, relatively well-educated individuals, many of whom were women. Both before and during the revolutions, these individuals used Facebook, Twitter, and YouTube to put pressure on their governments.”).

<sup>18</sup> See Washington Post Staff, *About Us: Policies and Standards*, WASHINGTON POST (Jan. 1, 2016) [https://www.washingtonpost.com/news/ask-the-post/wp/2016/01/01/policies-and-standards/?noredirect=on&utm\\_term=.9f3bf843b3c4](https://www.washingtonpost.com/news/ask-the-post/wp/2016/01/01/policies-and-standards/?noredirect=on&utm_term=.9f3bf843b3c4) (“The Post has a multi-level structure for the review and editing of stories that may include fact-checking. These include assignment editors (department heads, their deputy editors and assistant editors) who collaborate with reporters on the origination of stories and typically provide initial review when a story is submitted by a reporter; multiplatform editors (also called copy editors) who often provide initial review on breaking news stories and routinely provide second-level review on print and other less time-sensitive stories;

reliability in traditional news outlets that is not inherent in stories on social media. Additionally, the news sources themselves on social media may not always be what they appear. Russia, via its government Internet Research Agency, uses misleading website names to suggest its stories are coming from the U.S. rather than Russia.<sup>19</sup> As they say, “On the Internet, nobody knows you’re a dog.”<sup>20</sup>

The public has known for some time about Russia’s interference in the 2016 U.S. Presidential election on social media, but in December 2018 two expert reports revealed a trove of details.<sup>21</sup> The Senate commissioned

and senior editors who have overall oversight of the daily and weekend report for digital publication throughout the day as well as The Post’s print editions. Editors who oversee digital platforms also may be involved in the presentation of stories as well as headlines, news alerts and newsletters. The number of editors who review a story prior to publication and the extent of their involvement varies depending on a range of factors, including complexity, sensitivity, and the pressure of time.”).

<sup>19</sup> The website usareally.com is funded by the sponsors of the Russian “troll factory” accused of interference in the 2016 US Presidential election. *See Russia Protests After Journalist Detained, Interrogated at DC Airport*, ASSOCIATED PRESS (Nov. 11, 2018), <https://www.washingtontimes.com/news/2018/nov/11/alexander-malkevich-usa-really-editor-detained-dc/>.

<sup>20</sup> See Michael Cavanaugh, ‘Nobody Knows You’re A Dog’: As Iconic Internet Cartoon Turns 20, Creator Peter Steiner Knows The Joke Rings As Relevant As Ever, WASHINGTON POST (Jul. 31, 2013), [https://www.washingtonpost.com/blogs/comic-riffs/post/nobody-knows-youre-a-dog-as-iconic-internet-cartoon-turns-20-creator-peter-steiner-knows-the-joke-rings-as-relevant-as-ever/2013/07/31/73372600-f98d-11e2-8e84-c56731a202fb\\_blog.html?noredirect=on&utm\\_term=.6e128dc23f88](https://www.washingtonpost.com/blogs/comic-riffs/post/nobody-knows-youre-a-dog-as-iconic-internet-cartoon-turns-20-creator-peter-steiner-knows-the-joke-rings-as-relevant-as-ever/2013/07/31/73372600-f98d-11e2-8e84-c56731a202fb_blog.html?noredirect=on&utm_term=.6e128dc23f88).

<sup>21</sup> One report was prepared by social media analysts New Knowledge and the other by an Oxford University team working with analytical firm

the expert reports as part of its bipartisan investigation of Russian interference, and the reports are based largely on data about the Russian operations provided to the Senate by Facebook, Twitter, and the other companies whose platforms were used.<sup>22</sup> The reports describe the ongoing efforts of the Russian government's Internet Research Agency, based in St. Petersburg, Russia, to erode trust in U.S. democratic institutions and to divide Americans by race and extreme ideology.<sup>23</sup> Part of that effort included targeting African-Americans.<sup>24</sup> On Facebook and Instagram, Russians specifically targeted black American communities using Gmail accounts with American-sounding names to recruit and sometimes pay unwitting American activists to stage rallies and spread content.<sup>25</sup> One tactic was using Facebook ads targeted at users who had shown interest in particular topics, including black history, the Black Panther Party, and

---

Graphika. See Mark Hosenball, *Russia Used Social Media for Widespread Meddling in U.S. Politics: Reports*, REUTERS (Dec. 17, 2018, 2:18 PM), <https://www.reuters.com/article/us-usa-trump-russia-socialmedia/russia-used-social-media-for-widespread-meddling-in-u-s-politics-reports-idUSKBN1OG257?il=0>.

<sup>22</sup> See Renee DiResta et al., *The Tactics & Tropes of the Internet Research Agency*, NEW KNOWLEDGE, <https://disinformationreport.blob.core.windows.net/disinformation-report/NewKnowledge-Disinformation-Report-Whitepaper.pdf>.

<sup>23</sup> See Hosenball, *supra* note 21.

<sup>24</sup> See Scott Shane & Sheera Frenkel, *Russian 2016 Influence Operation Targeted African-Americans on Social Media*, N.Y. TIMES (Dec. 17, 2018), <https://www.nytimes.com/2018/12/17/us/politics/russia-2016-influence-campaign.html>.

<sup>25</sup> See DiResta et al., *supra* note 22, at 34.

Malcolm X.<sup>26</sup> The most popular of the Russian Instagram accounts was @blackstagram, with 303,663 followers, while on YouTube, the largest share of Russian material covered the Black Lives Matter movement and police brutality, with channels called “Don’t Shoot” and “BlackToLive.”<sup>27</sup> The expert reports revealed that, while much attention has been focused on Facebook, the Internet Research Agency in fact created social media accounts under fake names on virtually every available platform,<sup>28</sup> with the goal of supporting Donald J. Trump as a candidate in the Republican primary, then in the general election, and as president since his election.<sup>29</sup>

In April of 2019, Special Counsel Robert Mueller released his long-awaited investigative report of Russian interference in the 2016 election. The report further detailed the actions of the Russian troll farm Internet Research Agency and its social media interference dating back to 2014, including the use of fake Facebook accounts.<sup>30</sup>

---

<sup>26</sup> See *id.*

<sup>27</sup> See *id.* at 30.

<sup>28</sup> See Shane & Frenkel, *supra* note 24.

<sup>29</sup> See Craig Timberg & Tony Romm, *New Report on Russian Disinformation Prepared for the Senate, Shows the Operation’s Scale and Sweep*, WASH. POST (Dec. 17, 2018), [https://www.washingtonpost.com/technology/2018/12/16/new-report-russian-disinformation-prepared-senate-shows-operations-scale-sweep/?noredirect=on&utm\\_term=.aca9ee116e41](https://www.washingtonpost.com/technology/2018/12/16/new-report-russian-disinformation-prepared-senate-shows-operations-scale-sweep/?noredirect=on&utm_term=.aca9ee116e41).

<sup>30</sup> See Dustin Volz & Allan Cullison, ‘*Putin Has Won*’: Mueller Report Details the Ways Russia Interfered in the 2016 Election, Wall St. J. (Apr.

Facebook plays a prominent role in an additional social media election scandal involving Cambridge Analytica and one of the largest data leaks in social media history.<sup>31</sup> Cambridge Analytica was a voter profiling company started in 2014 by Steve Bannon and Robert Mercer, using a psychographics platform developed by Christopher Wylie and Aleksandr Kogan, that mapped personality traits based on what people liked on Facebook.<sup>32</sup> The company's goal was to develop detailed psychological profiles of every American voter so that campaigns could tailor their pitches from person to person.<sup>33</sup> To that end, the company created an application for Facebook with a personality quiz that asked 120 questions about personality and behavior.<sup>34</sup> Then, it scored people on traits like openness, extroversion, and agreeableness, and mixed those results with polls, voter records, and online activity in order to create personality models for voters.<sup>35</sup> Users of the app

---

19, 2019), <https://www.wsj.com/articles/putin-has-won-mueller-report-details-the-ways-russia-interfered-in-the-2016-election-11555666201>.

<sup>31</sup> See Carole Cadwalladr, Nicholas Confessore & Matthew Rosenberg, *How Trump Consultants Exploited the Facebook Data of Millions*, N.Y. TIMES (Mar. 17, 2018), <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html>.

<sup>32</sup> See *id.*

<sup>33</sup> See Scott Detrow, *What Did Cambridge Analytica Do During the 2016 Election?*, NAT'L PUB. RADIO (Mar. 20, 2018), <https://www.npr.org/2018/03/20/595338116/what-did-cambridge-analytica-do-during-the-2016-election>.

<sup>34</sup> See *id.*

<sup>35</sup> See *id.*

— about 270,000 people — technically consented to having their data harvested when they took the quiz (including education, location, the groups and pages they liked, their relationship status, and where they worked).<sup>36</sup> But the app also gathered information from the friends of those users, amounting to about “50 million raw profiles” that were provided to Cambridge Analytica.<sup>37</sup> The company then used that data for targeted advertisements and other election-related efforts on behalf of conservative candidates.<sup>38</sup> Most notably, the data was used to support Donald Trump in the 2016 election, when Cambridge Analytica designed target audiences for digital advertisements and fundraising appeals, modeled voter turnout, bought television advertisements, and determined where then-candidate Trump should travel to increase his support.<sup>39</sup>

The U.S. is not alone in experiencing election interference via social media use.<sup>40</sup> The democracy advocacy group Freedom House found that at least

---

<sup>36</sup> See *id.*; Rosenberg, *supra* note 31.

<sup>37</sup> See Rosenberg, *supra* note 31. The app violated Facebook rules both in gathering data for a purpose other than the research purpose that was disclosed, and for violating a loophole in Facebook’s architecture that allowed the app to harvest data of users who had not given permission.

<sup>38</sup> See *id.*

<sup>39</sup> See *id.*

<sup>40</sup> See Jackie Snow, *Last Year, Social Media Was Used to Influence Elections in at Least 18 Countries*, MIT TECH REV. (Nov. 14, 2017), <https://www.technologyreview.com/the-download/609478/last-year-social-media-was-used-to-influence-elections-in-at-least-18-countries/>.

seventeen other countries also had their elections manipulated through social media throughout 2016.<sup>41</sup>

*B. INCITEMENT TO VIOLENCE*

Facebook has also been tied to violence in countries where the social media platform has been used to foment action against disfavored groups.<sup>42</sup> In Myanmar, Facebook has an outsized importance because of the way the country's Internet use developed.<sup>43</sup> Until 2014, fewer than 5% of individuals there used mobile phones because they were prohibitively expensive.<sup>44</sup> That year, changes in government policies lowered the price of SIM cards, allowing a much greater number of people to afford them.<sup>45</sup> Importantly, those mobile phones were usually programmed to include Facebook and its messaging application.<sup>46</sup> Users had little experience with the Internet before having mobile phones, therefore for many in Myanmar, Facebook "is the Internet."<sup>47</sup> The social media platform has been used to target the disenfranchised Rohingya Muslim minority, which has been the target of a sustained campaign of violence and

---

<sup>41</sup> *See id.*

<sup>42</sup> *See id.*

<sup>43</sup> *See* McKirdy, *supra* note 12.

<sup>44</sup> *See id.*

<sup>45</sup> *See id.*

<sup>46</sup> *See id.*

<sup>47</sup> *See* Ortutay, *supra* note 15.

abuse by the Myanmar military that the United Nations calls “ethnic cleansing”.<sup>48</sup>

Ultra-nationalist Buddhists have used Facebook to publish hate speech and foment violence against the Rohingya. This includes posting fake pictures of the Muslims burning their own homes and of decaying bodies they said were Buddhist victims of Rohingya attacks, and using the Facebook messenger service to forward messages warning of incoming attacks by Muslims.<sup>49</sup> In September 2017, the Buddhists warned of a fabricated “jihad,” which incited a “call to arms” against the Muslims.<sup>50</sup> The purge of Muslims and subsequent refugee crisis have seen 700,000 Rohingya forced from their homes.<sup>51</sup>

Sri Lanka is another example where there have been instances of hate speech used against the Muslim minority.<sup>52</sup> In March 2018, parts of Sri Lanka’s Central Province experienced a wave of anti-Muslim riots, which the government blamed on hate speech against the Muslim community that was spread over Facebook, WhatsApp, Instagram, and Viber.<sup>53</sup> There are also

---

<sup>48</sup> See Specia & Mozur, *supra* note 12; McKirdy, *supra* note 12.

<sup>49</sup> See Specia & Mozur, *supra* note 12.

<sup>50</sup> See McKirdy, *supra* note 12.

<sup>51</sup> See *id.*

<sup>52</sup> See Amantha Perera, *Social Media – the New Testing Ground for Sri Lanka’s Freedom*, INTER PRESS SERV. NEWS AGENCY (July 18, 2018), <http://www.ipsnews.net/2018/07/social-media-new-testing-ground-sri-lankas-freedom/>.

<sup>53</sup> See *id.*

examples of people trolling journalists and using threats to prevent them from telling the truth.<sup>54</sup> Between five and six million Sri Lankans use Facebook, and research shows that social media is the primary platform of political interaction for Sri Lankans between eighteen and thirty-four.<sup>55</sup> Misinformation, disinformation, and threats to journalists have risen sharply.<sup>56</sup>

In India, there are infamous examples of social media weaponization.<sup>57</sup> In July 2018, a mob lynched five people after rumors spread, via WhatsApp messages, that these people were child abductors.<sup>58</sup> The victims were poor agricultural workers from a nearby district, who were surrounded and attacked by a crowd of about three thousand people.<sup>59</sup> Earlier that week, three more people were killed in another part of India after social media rumors spread of child abduction and human organ harvesting.<sup>60</sup> Mobs have killed people in at least three other instances, all related to fake rumors circulated mostly through WhatsApp groups.<sup>61</sup>

---

<sup>54</sup> *See id.*

<sup>55</sup> *See id.*

<sup>56</sup> *See id.*

<sup>57</sup> *See generally* Swati Gupta & Bard Wilkinson, *WhatsApp India: Five lynched after online child kidnap rumors*, CNN (July 3, 2018, 4:37 AM), <https://www.cnn.com/2018/07/02/asia/india-lynching-whatsapp-intl/index.html> (providing examples of recent weaponization of social media).

<sup>58</sup> *See id.*

<sup>59</sup> *See id.*

<sup>60</sup> *See id.*

<sup>61</sup> *See id.* (discussing similar mob attacks in India).

Finally, to a lesser extent, conspiracy theorists have touted their beliefs online in the U.S. and sparked violence. In response to Alex Jones's assertion that the 2012 Sandy Hook Elementary mass shooting was a hoax perpetrated by gun control advocates, Mr. Jones's supporters harassed the grieving parents of child victims.<sup>62</sup> And after the circulation of the bizarre Pizzagate conspiracy theory, alleging that Hillary Clinton and other Democrats were secretly running a child-sex ring, one supporter brought an assault rifle to a Washington, D.C. pizza restaurant, vowing to save the children locked in the basement.<sup>63</sup>

### C. FOMENTING HATE

Aside from direct incitement of violence, social media links to violence in more insidious ways. There is a tie between recent shootings in the U.S. and a history of social media radicalization. Before Cesar Sayoc allegedly sent fourteen pipe bombs through the mail to prominent Democrats around the country, he had become radicalized online and was "sucked into a vortex

---

<sup>62</sup> See Kevin Roose, *Cesar Sayoc's Path on Social Media: From Food Photos to Partisan Fury*, N.Y. TIMES (Oct. 27, 2018), <https://www.nytimes.com/2018/10/27/technology/cesar-sayoc-facebook-twitter.html>.

<sup>63</sup> See Spencer S. Hsu, 'Pizzagate' Gunman Sentenced To Four Years In Prison, As Prosecutors Urged Judge To Deter Vigilante Justice, WASH. POST (June 22, 2017), [https://www.washingtonpost.com/local/public-safety/pizzagate-gunman-sentenced-to-four-years-in-prison-as-prosecutors-urged-judge-to-deter-vigilante-justice/2017/06/22/a10db598-550b-11e7-ba90-f5875b7d1876\\_story.html?utm\\_term=.2d464811e9a1](https://www.washingtonpost.com/local/public-safety/pizzagate-gunman-sentenced-to-four-years-in-prison-as-prosecutors-urged-judge-to-deter-vigilante-justice/2017/06/22/a10db598-550b-11e7-ba90-f5875b7d1876_story.html?utm_term=.2d464811e9a1).

of partisan furor.”<sup>64</sup> Beginning in 2016, Sayoc started posting right-wing news stories, pro-Donald Trump images, and stories about Muslims and the Islamic State on Twitter and Facebook.<sup>65</sup> Weeks before the 2018 bombings, Sayoc had posted violent fantasies and threats against several people to whom pipe bombs were sent, including Democratic representatives, prominent liberals, and Trump critics.<sup>66</sup> On social media he found a home for his more strident views, including conspiracy theories involving the Clintons, Barack Obama, George Soros, and illegal immigration.<sup>67</sup> Many subjects of his tweets were targeted with mailbombs.<sup>68</sup>

Robert Bowers, the suspected gunman who killed 11 people at a Philadelphia synagogue, was an avid user of the social media site Gab.<sup>69</sup> Unlike mainstream social

---

<sup>64</sup> See Nicole Chavez, *Cesar Sayoc Was a DJ, Bodybuilder and Pizza Delivery Man Before He Became a Bomb Suspect*, CNN (Oct. 27, 2018, 9:39 PM), <https://www.cnn.com/2018/10/27/politics/cesar-sayoc-mail-bomb-suspect/index.html>.

<sup>65</sup> See Julie Turkewitz & Kevin Roose, *Who Is Robert Bowers, the Suspect in the Pittsburgh Synagogue Shooting*, N.Y. TIMES (Oct. 27, 2018), <https://www.nytimes.com/2018/10/27/us/robert-bowers-pittsburgh-synagogue-shooter.html>.

<sup>66</sup> See Roose, *supra* note 62, at A20.

<sup>67</sup> See *id.*

<sup>68</sup> See Benjamin Weiser, *Mail Bomb Suspect Accused of Targeting Clinton, Obama and Other Democrats to Plead Guilty*, N.Y. TIMES, Mar. 15, 2019, at A21 <https://www.nytimes.com/2019/03/15/nyregion/mail-bomber-cesar-sayoc.html>.

<sup>69</sup> See Becky Metrick, *Pittsburgh Synagogue Shooting Suspect used Social Media Platform Created by Pa. Native*, PA. REAL-TIME NEWS, (Mar. 27, 2019, 9:27 AM), <https://www.pennlive.com/news/2019/03/an->

media, Gab is infamous for allowing racist, anti-Semitic, and bigoted comments.<sup>70</sup> Bowers frequently used the site to criticize and threaten Jews; he referred to Jews as children of Satan, and the enemy of white people.<sup>71</sup> He accused Jews of aiding migrant caravan “invasions.”<sup>72</sup> Five minutes before police were alerted to the shooting, Bowers posted on the site: “[I] can’t sit by and watch my people get slaughtered. Screw your optics, I’m going in.”<sup>73</sup>

Scott Paul Beierle killed two people and wounded five others in a Tallahassee yoga studio before turning the gun on himself.<sup>74</sup> Evidence shows that he, like Sayoc and Bowers, bore a grudge against a certain group and found kinship for that hatred on social media.<sup>75</sup> Beierle

allentown-school-administrator-shared-photos-of-minors-having-sex-and-police-looked-the-other-way-lawsuit-says.html.

<sup>70</sup> See Alina Selyukh, *Feeling Sidelined by Mainstream Social Media, Far-Right Users Jump to Gab*, NAT’L PUB. RADIO, (Mar. 27, 2019, 9:25 AM),

<https://www.npr.org/sections/alltechconsidered/2017/05/21/529005840/feeling-sidelined-by-mainstream-social-media-far-right-users-jump-to-gab>.

<sup>71</sup> See Saeed Ahmed & Paul P. Murphy, *Here’s What We Know So Far About Robert Bowers, the Pittsburgh Synagogue Shooting Suspect*, CNN, (Oct. 28, 2018, 7:14 AM), <https://www.cnn.com/2018/10/27/us/synagogue-attack-suspect-robert-bowers-profile/index.html>.

<sup>72</sup> See *id.*

<sup>73</sup> *Id.*

<sup>74</sup> See Dakin Adone & Artemis Moshtaghian, *Gunman in Florida Yoga Studio Shooting made Misogynistic Comments on Youtube*, CNN, (Mar. 27, 2019, 9:36 AM), <https://www.cnn.com/2018/11/03/us/tallahassee-shooting-yoga-studio/index.html>.

<sup>75</sup> See Anna North, *How Mass Shooters Practice Their Hate Online*, VOX MED. (Nov. 3, 2018, 9:19 PM)

made numerous racist and misogynistic comments in Youtube videos and identified with “involuntary celebrities.”<sup>76</sup> He made violent comments about women who rejected him, and he expressed sympathy toward Elliot Roger, who killed 6 people and injured others in Santa Barbara after posting a manifesto blaming the cruelty of women for his own virginity.<sup>77</sup>

In probably the worst example yet of online extremism leading to offline violence, an Australian man fatally shot 51 Muslim worshippers and injured 39 others at two mosques in Christchurch, New Zealand.<sup>78</sup> Just before the attack, he published a link on a right-wing forum to a 74-page white supremacist manifesto.<sup>79</sup> He then used a camera mounted to his head to livestream the attack on Facebook.<sup>80</sup> The New York Times reported that “[t]he gunman appeared to pair the shooting with the typical trolling tactics of the internet’s most far-right instigators, playing to a community of

---

<https://www.vox.com/identities/2018/10/31/18039294/scott-beierle-tallahassee-shooting-pittsburgh-gab>.

<sup>76</sup> *See id.*

<sup>77</sup> *See id.*

<sup>78</sup> Charlotte Graham-McLay, *Death Toll in New Zealand Mosque Shooting Rises to 51*, N.Y. TIMES, (May 2, 2019), <https://www.nytimes.com/2019/05/02/world/asia/new-zealand-attack-death-toll.html>.

<sup>79</sup> Kevin Roose, *A Mass Murder of, and for, the Internet*, N.Y. TIMES (Mar. 15, 2019), <https://www.nytimes.com/2019/03/15/technology/facebook-youtube-christchurch-shooting.html>.

<sup>80</sup> *Id.*

like-minded supporters online who cheered him on in real time as they watched bodies pile up.”<sup>81</sup>

While social media did not cause these individuals to take the action they did, it provided an echo chamber for their disturbed views, validated those views, and encouraged hate.<sup>82</sup> Bowers seemed to see himself as having an audience on social media for whom he was performing in slaughtering people.<sup>83</sup> The New Zealand shooter quoted Bowers (“screw your optics”) in a posting before he began his own horrific live performance.<sup>84</sup> The Internet did not create hate, but it has given hate a worldwide platform.

### III. WHAT CAN BE DONE?

Social media sites themselves are increasingly addressing the problems of manipulation and violence.<sup>85</sup> In addition, many countries have opted for government

---

<sup>81</sup> Daniel Victor, *In Christchurch, Signs Point to a Gunman Steeped in Internet Culture*, N.Y. TIMES (Mar. 15, 2019), <https://www.nytimes.com/2019/03/15/world/asia/new-zealand-gunman-christchurch.html?action=click&module=RelatedCoverage&pgtype=Article&region=Footer>.

<sup>82</sup> *See id.*

<sup>83</sup> *See* Turkewitz & Roose, *supra* note 65.

<sup>84</sup> Roose, *supra* note 79.

<sup>85</sup> *See, e.g.*, Alex Hern, *WhatsApp to Restrict Message Forwarding After India Mob Lynchings*, GUARDIAN (July 20, 2018), <https://www.theguardian.com/technology/2018/jul/20/whatsapp-to-limit-message-forwarding-after-india-mob-lynchings>.

action to address online hate.<sup>86</sup> The action of the sites has been criticized as doing too little, while the action of governments has been criticized as doing too much.<sup>87</sup> Either way, limits on “bad” speech risk censorship of “good” speech.

#### A. GOVERNMENT ACTION

One solution to problematic use of social media sites is for governments to control any such use for improper means.<sup>88</sup> However, government action too often amounts to outright censorship.<sup>89</sup> After the March 2018 anti-Muslim riots in Sri Lanka, the Sri Lankan government responded with a weeklong shutdown of Facebook, WhatsApp, Instagram, and Viber.<sup>90</sup> In addition to censorship, government action often targets social media speech critical of its own actions.<sup>91</sup> The Chinese government is infamous for Internet censorship, imposing a system of Internet filters known as the Great

---

<sup>86</sup> See e.g., Saritha Rai, *India Seeks Access to Private Messages in WhatsApp Crackdown*, BLOOMBERG (Feb. 13, 2019, 6:50 AM), <https://www.bloomberg.com/news/articles/2019-02-13/india-targets-facebook-evil-in-backlash-against-u-s-giants>.

<sup>87</sup> See generally DiResta et al., *supra* note 22.

<sup>88</sup> See, e.g., Howard, *supra* note 17.

<sup>89</sup> See *id.*

<sup>90</sup> See Perera, *supra* note 52.

<sup>91</sup> See Mirae Yang, *The Collision of Social Media and Social Unrest: Why Shutting Down Social Media Is the Wrong Response*, 11 NW. J. TECH. & INTELL. PROP. 707, 709 (2013) (discussing Egypt’s shutdown of Internet and cell phone service in response to civil unrest).

Firewall to block content and shut out foreign technology companies.<sup>92</sup>

Now, India is accused of following that model and proposing strong, new measures that would allow Indian officials to demand that online platforms like Facebook and Twitter remove posts or videos that they deem libelous, invasive of privacy, hateful, or deceptive.<sup>93</sup> Those sites would be responsible for the content their users share.<sup>94</sup> In addition, the Indian government has cracked down on WhatsApp, by pressing the service to provide access to encrypted messages.<sup>95</sup>

In the European Union, data protection laws strongly protect users of social media, and in some instances require companies to remove material from their sites or scrub links to particular stories.<sup>96</sup> Under the “right to be

---

<sup>92</sup> See Vindu Goel, *India Proposes Chinese-Style Internet Censorship*, N.Y. TIMES (Feb. 14, 2019), <https://www.nytimes.com/2019/02/14/technology/india-internet-censorship.html?smid=tw-nytimesbits&smtyp=cur>.

<sup>93</sup> See *id.*

<sup>94</sup> See Rai, *supra* note 86 (draft published on the government’s website suggests the guidelines would hold these services responsible for a broad range of content, including information found to be “blasphemous, defamatory, obscene, pornographic, pedophilic, libelous, invasive of another’s privacy, hateful, or racially, ethnically objectionable, disparaging, relating or encouraging money laundering or gambling, or otherwise unlawful in any manner whatever”).

<sup>95</sup> See *id.*; See also Goel, *supra* note 92.

<sup>96</sup> See Regulation 2016/679, 2016 O.J. (L 119) (General Data Protection Directive Regulation or “GDPR”); Directive 95/46/EC, art. 12, 1995 O.J. (L 281).

forgotten,” citizens have a right to their own personal information, and the right to demand that data be deleted when it is no longer necessary for the legitimate purposes for which it was gathered.<sup>97</sup> Google, which is the search engine used by a whopping 95% of Europeans, was first ordered to take down links in Spain in May 2014,<sup>98</sup> and has since been asked to delete links to 2.9 million websites.<sup>99</sup> The right to be forgotten or “right to erasure” was enshrined in the General Data Protection Regulation (GDPR), which went into effect in 2018.<sup>100</sup>

Even if we wanted this kind of control over Internet speech in the U.S., such action would likely be deemed unconstitutional. Protection of free speech in the U.S. is “unparalleled elsewhere in the world.”<sup>101</sup> The U.S. Supreme Court has specifically acknowledged the importance of free speech on social media websites such

---

<sup>97</sup> *Id.* (Under the GDPR, an EU citizen has the right to demand that an organization erases her personal data if, among other reasons, the data is no longer relevant to the reason it was collected or if she withdraws consent to the data being used. This right is limited, for example, by the organization’s right to freedom of expression and if the data is in the public interest.)

<sup>98</sup> See Case C-131/12, *Google Spain SL v. Agencia Espanola de Proteccion de Datos (AEPD)*, 2014 E.C.R. 317.

<sup>99</sup> See *id.*; *Google Wins Bout in Fight Against ‘Right to Be Forgotten’*, LIVEMINT (Jan. 11, 2019, 6:49 AM), <https://www.livemint.com/Companies/tUyyXERKejWs08dnovzFJM/Google-can-limit-right-to-be-forgotten-to-EU-Top-court-ad.html> (France has unsuccessfully attempted to globalize the EU’s orders that Google delete links.)

<sup>100</sup> See GDPR Art. 17, Part 1, Right to Erasure (2018) <https://www.gdpreu.org/the-regulation/list-of-data-rights/right-to-erasure/>.

<sup>101</sup> Kitsuron Sangsuvan, *Balancing Freedom of Speech on the Internet Under International Law*, 39 N.C. J. INT’L L. & COM. REG. 701, 716 (2014).

as Facebook.<sup>102</sup> The Court interprets the First Amendment to require that any content-based restrictions on speech will be presumed invalid, and the government has the burden of showing those restrictions are narrowly tailored to serve a significant governmental interest.<sup>103</sup> Although the First Amendment provides less protection to obscenity, defamation, and threats of violence, it would certainly protect speech that is true but “no longer necessary” as the GDPR would condemn.<sup>104</sup>

In addition, U.S. law protects social media websites from any liability based on what others post to those sites. Since passage of the Communications Decency Act of 1996 (CDA), social media and other sites have been immunized from liability based on content provided by a third party.<sup>105</sup> Section 230 of the CDA protects Internet providers and Internet-related services

---

<sup>102</sup> See *Packingham v. North Carolina*, 137 S. Ct. 1730, 1735 (2017) (“While in the past there may have been difficulty 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. . . . One of the most popular of these sites is Facebook . . .”).

<sup>103</sup> See *Ashcroft v. ACLU*, 542 U.S. 656, 660 (2004).

<sup>104</sup> See *R.A.V. v. City of St. Paul, Minn.*, 505 U.S. 377, 383–84 (1992) (noting that “these areas of speech can, consistently with the First Amendment, be regulated because of their constitutionally proscribable content (obscenity, defamation, etc.)—not that they are categories of speech entirely invisible to the Constitution, so that they may be made the vehicles for content discrimination unrelated to their distinctively proscribable content. Thus, the government may proscribe libel; but it may not make the further content discrimination of proscribing only libel critical of the government.”).

<sup>105</sup> See 47 U.S.C. § 230 (2018).

— such as social media platforms<sup>106</sup> — by prohibiting those platforms from being treated as “the publisher or speaker of any information provided by another information content provider.”<sup>107</sup> The Section establishes federal immunity to any cause of action that would make a social media site liable for information originating from a third-party user, even when the platform has actual knowledge of harmful content.<sup>108</sup>

Because of this immunity, any legal challenges to the way social media sites control their content would likely

---

<sup>106</sup> See *Batzel v. Smith*, 333 F.3d 1018, 1031 (9th Cir. 2003) (including a listserv email newsletter sent by a website operator within the definition of interactive computer service); *Carafano v. Metrosplash.com, Inc.*, 339 F.3d 1119, 1125 (9th Cir. 2003) (holding that a matchmaking and dating website is an interactive computer service); *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 n.2 (4th Cir. 1997) (concluding that America Online fell within the definition of interactive service provider); *Klayman v. Zuckerberg*, 753 F.3d 1354, 1357 (D.C. Cir. 2014) (Facebook qualifies as interactive computer service because it “provides information to ‘multiple users’ by giving them ‘computer access...to a computer server’...namely the servers that host social networking websites.” (quoting 47 U.S.C. § 230(f)(2) (2012))).

<sup>107</sup> 47 U.S.C. § 230 (2018).

<sup>108</sup> See *id.* at § 230(c)(1); *Klayman*, 753 F.3d at 1358 (The immunity is contingent on three conditions: (1) the party seeking immunity is a “provider or user of an interactive computer service”; (2) the claim treats the party seeking immunity “as the publisher or speaker” of the disputed content; and (3) the claim is based on content produced “by another information content provider.”; *Zeran*, 129 F.3d at 330 (Alternatively, providers cannot claim immunity when an action arises out of content the ICSP produced itself. However, a service’s capability to control or remove content posted on their social media website does not void its immunity.) This incentivizes services to willingly regulate and improve their content while keeping their protection intact.

require legislative changes.<sup>109</sup> Recently, Congress has focused on problems posed by unchecked use of social media. Legislators have introduced several bills in Congress and in states like California that would increase privacy protections for users of social media in light of the Cambridge Analytica scandal.<sup>110</sup> Specifically, Congress seems increasingly willing to weaken the immunity granted to social media sites under the CDA.<sup>111</sup> In April 2018, a new law went into effect that limits the immunity provided under CDA Section 230 for online services that knowingly host third-party content to promote or facilitate sex trafficking.<sup>112</sup> The law removes CDA immunity for those online platforms with respect to state criminal charges if the conduct underlying the state violation would constitute a

---

<sup>109</sup> *See id.*

<sup>110</sup> *See* Customer Online Notification for Stopping Edge-Provider Network Transgressions (“CONSENT Act”), S. 2639, 115th Cong. (2018) (requiring websites to notify customers about the collection, use, and sharing of sensitive customer proprietary information of the customer, and to obtain opt-in consent to use, share, or sell the sensitive information of customer); Social Media Privacy Protection and Consumer Rights Act of 2019, S. 189, 116th Cong. (2019) (requiring online platform operators to inform a user, prior to a user creating an account or otherwise using the platform, that the user’s personal data produced during online behavior will be collected and used by the operator and third parties, along with the option to specify privacy preferences, and requiring notification within 72 hours violation of a user’s data security); Assembly 288, 2018-2019 Sess. (Cal. 2019) (requiring social media companies to allow users to have their personally identifiable information permanently removed from the company’s database and records upon closing of their accounts).

<sup>111</sup> *See* 47 U.S.C. § 230, *supra* note 105.

<sup>112</sup> *See* Allow States and Victims to Fight Online Sex Trafficking Act of 2017 (“FOSTA”), H.R. 1865, 115th Cong. (2017).

violation of anti sex-trafficking statutes; the law removes immunity for FOSTA-related civil suits brought by victims against online services that knowingly promote or facilitate sex trafficking through such activities as hosting third-party posts, listings and advertisements.<sup>113</sup>

At a July 2018 hearing in the House Judiciary Committee on social media filtering practices, a Congressman suggested to representatives of Facebook, Twitter, and YouTube that further exceptions to CDA immunity could be in the works.<sup>114</sup> At a hearing in December 2018 on Google's business practices, Congress members criticized the broad protections that Section 230 provides to platforms such as Google.<sup>115</sup> One bill has been introduced that would remove Section 230 immunity when social media companies use algorithms to alter and filter content.<sup>116</sup>

---

<sup>113</sup> See Jeffrey Neuburger, *FOSTA Signed into Law, Amends CDA Section 230 to Allow Enforcement Against Online Providers for Knowingly Facilitating Sex Trafficking*, PROSKAUER: NEW MEDIA AND TECH. L. BLOG (Apr. 11, 2018), <https://newmedialaw.proskauer.com/2018/04/11/fosta-signed-into-law-amends-cda-section-230-to-allow-enforcement-against-online-providers-for-knowingly-facilitating-sex-trafficking/>.

<sup>114</sup> See Issie Lapowski, *Lawmakers Don't Grasp the Sacred Tech Law They Want to Gut*, WIRED (July 17, 2018, 5:47 PM), <https://www.wired.com/story/lawmakers-dont-grasp-section-230/>.

<sup>115</sup> See Jeff Kosseff, *Congress Could Still Break the Internet. Here's How (Opinion)*, GOV'T TECH. (Dec. 17, 2018), <http://www.govtech.com/policy/Congress-Could-Still-Break-the-Internet-Heres-How-Opinion.html>.

<sup>116</sup> See Biased Algorithm Deterrence Act of 2019, H.R. 492, 116th Cong. (2019). Based on the press release accompanying the bill's introduction, its primary purpose is to counter alleged social media bias against conservative views. See Press Release, Louie Gohmert, Member,

There are many problems associated with limiting social media sites' immunity. The CDA has been widely credited with allowing the Internet to flourish and changing website immunity from liability risks self-censorship and suppression of lawful speech.<sup>117</sup>

### B. ACTION BY SOCIAL MEDIA SITES

The most obvious solution to misuse of social media is for those sites themselves to crack down on misuse in the form of manipulation and hate speech, although this too comes with risks to free expression.<sup>118</sup> While the sites have community standards and guidelines for use,

---

U.S. House of Representatives, Gohmert Statement on Bill that Removes Liability Protections for Social Media Companies (Dec. 20, 2018), <https://gohmert.house.gov/news/documentsingle.aspx?DocumentID=398676>.

<sup>117</sup> See Jonathan Zittrain, *CDA 230 Then and Now: Does Intermediary Immunity Keep the Rest of Us Healthy?* (Aug. 31, 2018), <https://blogs.harvard.edu/jzwrites/2018/08/31/cda-230-then-and-now/> (The Internet's development over the past twenty years has benefited immeasurably from the immunities conferred by Section 230. We've been lucky to have it. But any honest account must acknowledge the collateral damage it has permitted to be visited upon real people whose reputations, privacy, and dignity have been hurt in ways that defy redress. Especially as that damage becomes more systematized – now part of organized campaigns to shame people into silence online for expressing opinions that don't fit an aggressor's propaganda aims – platforms' failures to moderate become more costly, both to targets of harassment and to everyone else denied exposure to honestly-held ideas.).

<sup>118</sup> While social media sites are not bound by constitutional restrictions on government action, they have been likened to public spaces where restrictions on free speech should be prohibited. See Colby M. Everett, *Free Speech on Privately-Owned Fora: A Discussion on Speech Freedoms and Policy for Social Media*, 28 KAN. J.L. & PUB. POL'Y 113, 127, (2018) (“Social media is a modern-age public forum.”).

and respond to complaints by, in some instances, taking down posts or banning users from their sites, these measures are often not enough to prevent manipulation and violence.<sup>119</sup>

Some sites have acknowledged as much and made recent efforts to do better, particularly in response to the violence that has resulted in Myanmar and India. In July 2018, Facebook pledged that it would begin removing information in Myanmar that could lead to people being physically harmed.<sup>120</sup> In August of that year, Facebook banned Myanmar's military chief and nineteen other individuals and organizations from the network in order to prevent the hate speech and propaganda.<sup>121</sup> And three months later, Facebook admitted it did not do enough to prevent the violence and hate speech in Myanmar, vowing to "do more" to protect human rights.<sup>122</sup> In India, WhatsApp attempted to crack down on the viral spread of misinformation through its service by limiting the ability to forward messages to only twenty people.<sup>123</sup>

In the U.S., Facebook removed the account of suspect Cesar Sayoc, but only after the story emerged of

---

<sup>119</sup> See Everett, at 120–21 ("Social media websites utilize community guidelines and moderation to retain users and protect business interests. These websites outsource a vast majority of moderation to armies of overseas contractors who screen flagged information and make judgment calls based on guideline compliance. The rest is left to algorithms.").

<sup>120</sup> See Frenkel, *supra* note 24.

<sup>121</sup> See Ortutay, *supra* note 15.

<sup>122</sup> See *id.*

<sup>123</sup> See Hern, *supra* note 85.

Sayoc's mailing of pipe bombs.<sup>124</sup> A Facebook representative said that several of Sayoc's previous posts violated Facebook's community standards and had been removed before his arrest but that none of his posts, reported to or discovered by Facebook, contained violations of its rules severe enough to remove the account entirely.<sup>125</sup> And while a political commentator had complained to Twitter about Sayoc's threats to her prior to his sending the pipe bombs, Twitter had said he had not violated their rules against abuse.<sup>126</sup>

Criticism prompted by the livestreamed New Zealand shooting led to Facebook's institution of measures to limit livestreaming<sup>127</sup> and to combat use of its site by white nationalists.<sup>128</sup> On March 27, 2019, Facebook announced that it was imposing "a ban on

---

<sup>124</sup> See Jason Hanna, Evan Perez, Scott Glover, Steve Almasy & Ray Sanchez, *Bomb Suspect Arrest: What We Know About Cesar Sayoc*, CNN Politics (Oct. 26, 2018, 10:56 PM), <https://www.cnn.com/2018/10/26/politics/suspicious-packages-arrest/index.html>.

<sup>125</sup> See *id.*

<sup>126</sup> See Donie O'Sullivan, *Bomb Suspect Threatened People on Twitter, and Twitter Didn't Act*, CNN (Oct. 27, 2018, 10:12 AM), <https://www.cnn.com/2018/10/26/tech/cesar-sayoc-twitter-response/index.html>.

<sup>127</sup> See Heather Kelly, *Facebook Changes Livestream Rules After New Zealand Shooting*, CNN (May 15, 2019), <https://www.cnn.com/2019/05/14/tech/facebook-livestream-changes/index.html>.

<sup>128</sup> Facebook Newsroom, Mar. 27, 2019, <https://newsroom.fb.com/news/2019/03/standing-against-hate/>.

praise, support and representation of white nationalism and white separatism on Facebook and Instagram.”<sup>129</sup>

After criticism that it was becoming a hub for “nut-case” conspiracy theories, YouTube began appending Wikipedia blurbs to videos espousing such theories and began giving priority to reliable news sources over partisans in search results for breaking news stories.<sup>130</sup> Next, YouTube announced it was changing its recommendations algorithm to reduce the spread of “borderline content and content that could misinform users in harmful ways.”<sup>131</sup> Most recently, YouTube announced plans to remove videos and channels “that advocate neo-Nazism, white supremacy and other bigoted ideologies in an attempt to clean up extremism and hate speech” on its service.<sup>132</sup>

In an extreme case, a social media site that does not control hate speech may find itself closed off from the Internet altogether. After the Pittsburg synagogue shooting and revelations that the suspected gunman used the social network site Gab to threaten Jews, numerous

---

<sup>129</sup> *Id.*

<sup>130</sup> See Kevin Roose, *YouTube Unleashed a Conspiracy Theory*, N.Y. TIMES (Mar. 26, 2018, 6:30 PM), <https://www.nytimes.com/2019/02/19/technology/youtube-conspiracy-stars.html>.

<sup>131</sup> *See id.*

<sup>132</sup> Kevin Roose & Kate Conger, *YouTube to Remove Thousands of Videos Pushing Extreme Views*, N.Y. TIMES (June 5, 2019), <https://www.nytimes.com/2019/06/05/business/youtube-remove-extremist-videos.html>.

Internet infrastructure providers, including payment processing firm Paypal, hosting service Joyent, domain register Godaddy, and numerous app stores, banned the site.<sup>133</sup> Paypal explained its decision by stating that it would not affiliate with a site that “explicitly allow[s] the perpetuation of hate, violence or discriminatory intolerance.”<sup>134</sup> A similar fate befell the neo-Nazi website Daily Stormer after the violence during a white supremacist rally in Charlottesville, Virginia.<sup>135</sup>

### C. FURTHER ACTION

There are enormous difficulties in policing speech on social media sites like Facebook, because the site must respond to users who have customs that are unfamiliar to Facebook employees and who speak in varying dialects.<sup>136</sup> But social media sites should continue to “do more.” Facebook needs to ensure that it has the manpower and knowledge to do the reviewing of content that it promises in its community standards.<sup>137</sup> Social media sites should continue to tweak algorithms and recommendations to favor verifiable news over false news and conspiracy theories, so as not to add to the

---

<sup>133</sup> See Ivana Kottasova & Sara Ashley O’Brien, Gab, the Social Network Used by the Pittsburgh Suspect, Has Been Taken Offline, Cnn: Business (Oct. 29, 2018, 12:00 PM), <https://www.cnn.com/2018/10/29/tech/gab-offline-pittsburgh/index.html>.

<sup>134</sup> *Id.*

<sup>135</sup> See Everett, *supra* note 118 at 114. (2018).

<sup>136</sup> See McKirdy, *supra* note 12 (noting the difficulty Facebook faces in monitoring the rise of hate speech in Myanmar because of language difficulties).

<sup>137</sup> See *Facebook Community Standards*, FACEBOOK, <https://www.facebook.com/communitystandards/> (last visited Apr. 2, 2019).

problem when the false stories trend. Adding disclaimers next to false stories is a good step.

Second, government should continue to step in when social media manipulation equates to foreign interference in elections. The U.S. should make election interference a priority of foreign policy diplomacy, attempt to negotiate cross-border prosecution of offenders, or enter treaties where members promise not to use cyber election interference.

Third, hand in hand with its engagement abroad, government must cooperate with the technology industry and social media websites themselves in finding solutions. Congress can work with Facebook and other sites to make progress on these issues without stifling Internet speech or enacting technological solutions that will be quickly obsolete. New Zealand's Prime Minister has issued a "Christchurch Call," an effort to enlist companies like Facebook, Twitter and YouTube to do more to curb violent and extremist content.<sup>138</sup> Seventeen countries, the European Commission, and eight large technology companies have signed onto the call.<sup>139</sup>

Finally, one solution will be self-enacting. With time, social media users should become more Internet-

---

<sup>138</sup> See Jamie Tarabay, *As New Zealand Fights Online Hate, the Internet's Darkest Corners Resist*, N.Y. TIMES (July 5, 2019), [https://www.nytimes.com/2019/07/05/world/asia/new-zealand-internet.html?nl=todaysheadlines&emc=edit\\_th\\_190706?campaign\\_id=2&instance\\_id=10726&segment\\_id=14968&user\\_id=4022f9f0b977acdc8bf6eb8211796809&regi\\_id=555121730706](https://www.nytimes.com/2019/07/05/world/asia/new-zealand-internet.html?nl=todaysheadlines&emc=edit_th_190706?campaign_id=2&instance_id=10726&segment_id=14968&user_id=4022f9f0b977acdc8bf6eb8211796809&regi_id=555121730706). Prime Minister Jacinda Ardern also responded to the Christchurch shootings by refusing to utter the shooter's name, banning the sharing or viewing of the shooter's manifesto, and tightening gun laws.

<sup>139</sup> *Id.*

savvy, particularly in areas where Internet use is still developing. Additionally, monopolies on users' social media attention should break down, as we already see users abandoning Facebook in favor of sites like Snapchat. And with time, more content can be developed to drown out untruthful or hateful content.

#### IV. CONCLUSION

Social media is here to stay, and with it come true dangers as well as benefits. The 2016 presidential election has shown us that our social media sites can be used to taint the very basis of our democracy – our election system. Violence in India and Myanmar has erupted in direct response to false stories that could only be spread with such speed and breadth via social media sites. And other violence, like mass shootings, can be tied directly to social media sites where troubled individuals find like-minded peers and fan the flames of hate and derangement. These problems must be addressed by social media sites and by the government, but with care not to destroy the freedom of speech that social media epitomizes. As the Supreme Court noted in 2017, new technologies will be exploited for criminal means. The solution includes improving the technology, not destroying it.