“This Claim About Election Fraud is Disputed”:

Mitigating the Spread of Disinformation Through Social Media In American Electoral Contexts

Brendan Leonard

GOVT 595: Policy Writing and Analysis Precapstone

December 18, 2020

Table of Contents

Table of Figures

Abbreviations

CDA	The Communications Decency Act of 1996
FCC	Federal Communications Commission
ICT	Information and Communications Technology
Red Lion	*Red Lion Broadcasting Co. v. FCC* (1969)
Section 230	Section 230 of the Communications Decency Act (47 US Code § 230)

Executive Summary

Growing proportions of the population are turning to social media platforms not only to connect with friends and family, but to discover what is going on in the world at-large. Approximately one-fifth of American adults use social media as their primary source of political news and, consequently, this population stands among the lowest-informed in terms of actual political knowledge.[1] Between coordinated foreign interference in the 2016 election and disinformation spread by domestic leaders for partisan gain in 2020, it is quite clear that the American political system stands susceptible to manipulation spread through online social media channels.

While various social media platforms have begun to ramp up enforcement against disinformation campaigns, federal action aimed at curtailing the spread of such disinformation remains absent. Due to the specific incentives of social media platforms, including profit-maximization, legislators cannot trust these companies to self-moderate to a sufficient degree to protect the common interest of a well-informed electorate without appropriate guardrails. To that end, the federal government should enact legislation aimed at mitigating disinformation spread through social media from both foreign and domestic actors.

Relevant Policy Context

The First Amendment clearly sets up protections for freedom of speech which any future action must take care to avoid infringing upon. While the protections of the First Amendment are broad, they are not without exception. For example, courts may agree that disinformation knowingly spread could be considered unprotected fraudulent speech due to the harm it poses to the electorate.[2] Additionally, past jurisprudence has generally shaped three frameworks through which courts may examine regulation of speech on social media:[3]

- **Company Town:** courts would acknowledge that a platform effectively acts as a government in some capacity and may not be allowed to infringe upon its users' speech.
- **Special Industry:** courts would validate a specific governmental interest in regulating an industry due to the medium of communication and allow weaker protections.

---

[1] Mitchell et al., "Americans Who Mainly Get Their News on Social Media," 3–6.
[2] Killion, "The First Amendment: Categories of Speech," 1–2.
[3] Brannon, "Free Speech and the Regulation of Social Media Content."

- **Newspaper Editor:** courts would assert that platforms utilize protected speech in making decisions about presenting content to a user, including which content to show.

Each of these would have dramatic implications for not only the regulatory relationship between the government and platforms, but also platforms and their users.

Section 230 of the CDA is a federal law which most notably grants wide immunity to ICT companies for the content that they host in two different and complementary ways. The primary provisions of the law state that a platform hosting the speech of another creator cannot be treated as the speaker itself and that platforms have the authority to moderate content.

The fairness doctrine was a policy of the FCC aimed at promoting civic discourse by requiring public broadcasters to devote time to issues of public importance and show multiple sides of an issue.[4] The fairness doctrine is notable due to the Supreme Court upholding its constitutionality and largely creating the basis for the special industries framework.[5] The fairness doctrine could stand as inspiration for future requirements of politically-balanced content.

Recommended Policy Priorities

- **Hole in the Shield:** specifically, introduce legal exceptions to Section 230 immunity for hosting known disinformation to allow for pursuing remedies through judicial channels. Currently any such case would likely be thrown out on Section 230 grounds immediately, rather than examining for potential damages and recourse through the courts.
- **Good-Faith Moderation:** in the spirit of a modern fairness doctrine, create a fiduciary relationship between social media platforms and users. Requirements may include actively moderating content for disinformation and promoting reputable and balanced information to counter echo chambers.
- **Label & Restrict:** in an effort to diminish the power of inauthentic actors and bot accounts, require platforms to label or remove applicable content of hostile actors or non-human accounts spreading disinformation.

[4] "Applicability of the Fairness Doctrine," 10426.
[5] Ruane, "Fairness Doctrine: History and Constitutional Issues," 4.

Introduction

In delivering the opinion of the Supreme Court in *Red Lion Broadcasting Co. v. FCC*, Justice Byron White states, "It is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail…"[6] Since Justice White penned this opinion in 1969, the way information and ideas travel throughout the population has dramatically changed. The advent of the internet, and more specifically social media, has fundamentally changed the landscape of information spreading in the United States. Justice White further goes on to declare, "It is the right of the public to receive suitable access to social, political, esthetic, moral, and other ideas and experiences…"[7] When accessing political or other news an individual relies on for decision making, particularly in the context of an election, do they have a right for this suitable access for that information to not be knowingly and demonstrably false?

Between coordinated disinformation campaigns carried out by Russian operatives in 2016 and the sitting President baselessly calling into question the legitimacy of election results in 2020, recent developments justify asking if this free market of ideas Justice White spoke of still reaches efficiency in the age of social media. A well-informed electorate is crucial for the functioning of liberal democracy. True and validated information not only allows for the electorate to allocate their votes optimally and efficiently, but it simultaneously allows for the building of confidence in institutions. To that end, does the US government have a role in preventing the spread of disinformation in online contexts? Regulation tailored to mitigating the sharing of disinformation on social media platforms can benefit both individuals and society at-large, but it must overcome a variety of regulatory obstacles to do so.

The Problem

Under current circumstances, bad actors are perfectly capable of using freedom of speech protections to infuse dishonest information into the public conversation and shape electoral results. In the final weeks of the 2016 election, 44.3% of US adults "visited an article on an untrustworthy website" making up 6.7% of Americans' hard-news diet in that timeframe.[8] Under current law, there are effectively no provisions to protect American interests from disinformation

[6] Red Lion Broadcasting Co. v. FCC, 395 US at 390.
[7] Red Lion Broadcasting Co. v. FCC, 395 US at 390.
[8] Guess, Nyhan, and Reifler, "Exposure to Untrustworthy Websites in the 2016 US Election," 473.

spread online. Social media companies are left to their own devices to determine content moderation policies and decide what speech is considered acceptable on their own platforms. While the protection of freedom of speech stands as a core tenet of American democracy, policymakers must wrestle with whether the detriment to society and electoral institutions is worth attempts at expanding regulation of speech. Action will likely require passing through all three branches of government: Congress to amend the current laws and provide direction for updated policy, the Executive Branch to carry out policy directives, and the Judicial Branch to uphold constitutionality of laws designed with these problems in mind.

Misinformation vs. Disinformation

When addressing issues regarding the spread of untrue information, an essential understanding is the delineation between misinformation and disinformation. Dictionary.com defines misinformation as "false information that is spread, regardless of intent to mislead."[9] Disinformation, on the other hand, is distinguished by a deliberate intent. Formally, disinformation is considered to originate or spread from a government source. This definition has loosened with colloquial usage to mean more generally "deliberately misleading or biased information; manipulated narrative or facts; propaganda."[10] This distinction becomes important when turning to discussions regarding protections granted by the First Amendment surrounding freedom of speech and the feasibility of a given regulation passing judicial scrutiny.[11] If the federal government is able to implement any regulation targeting the spread of untrue information on social media, it will very likely need to focus on disinformation specifically, rather than misinformation more generally.

Disinformation and Elections

Both the 2016 and 2020 elections faced notable disinformation tactics, though with notable differences in sources of origin and methodologies of the specific campaigns. The 2016 election featured widespread, coordinated, and effective attacks from foreign adversaries, with Russian operatives in particular attaining the most widespread notoriety. While foreign influence still played a role, the 2020 election has been significantly more notable for the role domestic

---

[9] Dictionary.com, "'Misinformation' vs. 'Disinformation.'"
[10] Dictionary.com.
[11] See Types of Speech section.

disinformation from politicians and other influencers had on the contest, both before and after Election Day.

2016 Election

Since 2016, a multitude of evidence has come to light describing the scope and scale of coordinated attempts by Russian operatives at the Internet Research Agency to influence the election.[12] In particular, the foreign agents sought to influence Americans to the benefit of Donald Trump and the Republican Party. The Russians skillfully split Americans into key groups in order to effectively deliver targeted messaging. For example, they sought to stoke reactions among conservatives on gun rights and immigration, while deteriorating the influence of Black voters on the left by undermining faith in electoral processes and spreading disinformation regarding how to vote. As the first truly large-scale and coordinated attack on American elections through disinformation, social media platforms were especially unprepared for fighting back against these efforts. With the lessons learned from 2016, the 2020 election saw much greater levels of collaboration between the government and technology companies in efforts to mitigate disinformation from foreign actors.[13]

2020 Election

While the US response to counter foreign interference through disinformation has strengthened, these strategies were not necessarily effective at reducing the spread of disinformation sowed by US politicians and influencers, particularly President Donald Trump. According to Graham Brookie, director and managing editor of the Atlantic Council's Digital Forensic Research Lab, "Trump is hands down the most significant accelerant and amplifier for disinformation in the election… The scale and scope of domestic disinformation is far greater than anything a foreign adversary could ever do to us."[14]

In the weeks following the 2020 Election, President Donald Trump sent hundreds of tweets to his followers spreading untrue information about the electoral process.[15] The President cast suspicion on routine and expected electoral processes, made claims of unsubstantiated fraud,

---

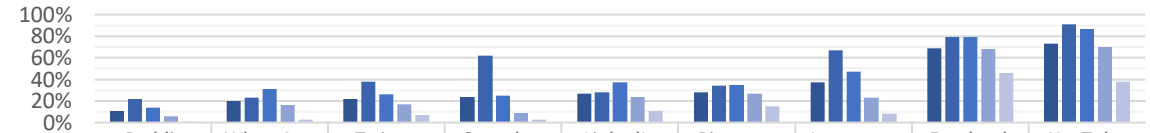
[12] Timberg and Romm, "New Report on Russian Disinformation."
[13] Rosenbaum, "Most Important 2020 Election Misinformation Threat."
[14] Dwoskin and Timberg, "The Unseen Machine Pushing Trump's Social Media Megaphone into Overdrive."
[15] Qiu, "Trump Has Amplified Voting Falsehoods in over 300 Tweets since Election Night."

and outright declared victory by tweeting "I won the Election!"[16] The impact the President had on disinformation spread does not stop with his followers, as right-wing influencers and other followers echo his messages in an incredibly prominent and effective feedback loop.[17] Researchers have found that only 20 right-wing accounts (including @realDonaldTrump) originated 20% of all retweets spreading false information about voting. In addition to information related directly to the election, research has found President Trump to have had the largest impact on COVID-19 misinformation in the first six months of 2020.

The State of Social Media

Social media platforms are websites where users are able to connect to other individuals to share content and communicate with one another. In less than two decades, the usage rate of social media sites by US adults has climbed from 5% in 2005 to 72% in 2019.[18] As of 2019, YouTube (a video sharing platform) and Facebook (a platform targeting person-to-person networking with individuals a user likely knows offline) are by far the most popular social networks in the US, with 73% and 69% of the adult population on these sites, respectively. Other popular social media platforms include Instagram and Snapchat (photo sharing), LinkedIn (professional networking), Twitter (microblogging with short posts or "tweets" of 280 characters or fewer), WhatsApp (messaging), and Reddit (discussion forum). Social media utilization in the US inversely correlates with age, with younger users making up a higher proportion of users than older Americans.

Figure 1: Social Media Usage of US Adults by Platform (February 2019)[19]



	Reddit	WhatsApp	Twitter	Snapchat	LinkedIn	Pinterest	Instagram	Facebook	YouTube
Overall	11%	20%	22%	24%	27%	28%	37%	69%	73%
18–29	22%	23%	38%	62%	28%	34%	67%	79%	91%
30–49	14%	31%	26%	25%	37%	35%	47%	79%	87%
50–64	6%	16%	17%	9%	24%	27%	23%	68%	70%
65+	1%	3%	7%	3%	11%	15%	8%	46%	38%

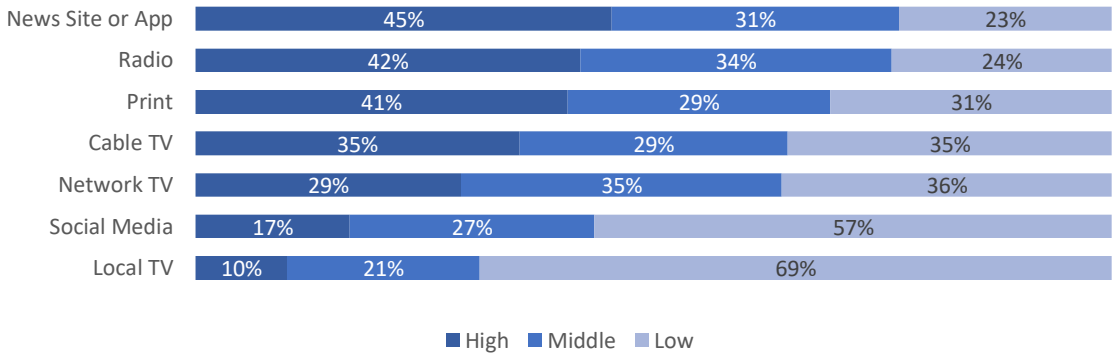
[16] Qiu; Trump, "I Won the Election!"
[17] Dwoskin and Timberg, "The Unseen Machine Pushing Trump's Social Media Megaphone into Overdrive."
[18] Pew Research Center, "Social Media Fact Sheet."
[19] Pew Research Center.

The Democratization of News

As the usage of social media sites in general has grown over time, so has their use as a source for disseminating and retrieving news. In addition to traditional bureaus on social media, these platforms may be appealing to up-and-coming or independent journalists who can build a following independently of ties to any particular organization. Journalism is going through a transformation to accommodate this democratization of reporting where anybody on the site can post information, whether true or not.

Approximately one-in-five American adults reports social media sites as their primary source of accessing political news.[20] Similar to overall trends, using social media as one's primary source of political news inversely trends with age, with 48% of adults under 30 and 40% of those between 30–49 using social media as their primary news source. In comparison, only 9% of those 50–64 and 3% of those over 65 use social media in this fashion. A primary issue with the way people tend to use social media as a news source comes through in their factual knowledge. Adults who report social media as their most common source of political and electoral news are among the least knowledgeable in terms of factual information. In terms of other sources of news, social media users rank only above local TV watchers, with news website or application users coming in as the most well-informed.

Figure 2: Level of Political Knowledge by Most Common Source of Political News[21]



This is a significant proportion of the population getting news from sources with virtually no guardrails. Individual users within social media networks can become isolated in bubbles and

[20] Mitchell et al., "Americans Who Mainly Get Their News on Social Media," 3.
[21] Mitchell et al., 5.

hyper-targeted to the point where it may become impossible to come across balancing or alternative views. The democratization of information has made it much harder to judge both the authenticity and validity of information a user comes across when compared to traditional media outlets who may act as gatekeepers to certain topics by verifying stories before publishing to followings.

Self-Moderation

While policy compelling platforms to monitor and act upon disinformation on their own sites has not come to fruition, several companies have taken steps of various levels of effectiveness to combat the spread of false information. For the 2020 election, these interventions included, but are by no means limited to:

- Twitter and Facebook labeled posts, including those of President Trump, as containing misinformation when appropriate in an attempt to reduce their impact.[22] Labels on Twitter included "This claim about election fraud is disputed" and "Multiple sources called this election differently."[23]
- Twitter delegated the determination of election calls to a number of trusted and independent news bureaus. [24] It additionally appended an appropriate label when tweets made a claim regarding an election that no outlet supported.
- In the months surrounding the election, YouTube removed thousands of channels for "harmful and misleading" content and began to "remove videos that make false claims that widespread fraud or error cost President Trump the election."[25]
- Pinterest took down posts about stolen ballots, prevented search of controversial topics likely to be breeding grounds for misinformation, and cracked down on terms related to conspiracy theories.[26]

Despite fewer personnel and resources, some smaller platforms, such as Pinterest, have gone further to combat false information on their sites when compared to some of the larger players. For these smaller sites, blocking wider swarths of content likely to contain disinformation is not

---

[22] Roose, Isaac, and Frenkel, "Facebook Struggles to Balance Civility and Growth"; Zakrzewski and Lerman, "No Social Media Is Safe."
[23] Trump, "@realDonaldTrump."
[24] Gadde and Beykpour, "Additional Steps We're Taking Ahead of the 2020 US Election."
[25] Telford, "YouTube Removes 8,000 Channels Promoting False Election Claims."
[26] Zakrzewski and Lerman, "No Social Media Is Safe."

only easier, but also more effective in mitigating its spread. Their initiatives are often timelier, as YouTube took more than a month following the election to begin removing videos spreading disinformation about voter fraud in the 2020 election.[27]

Profit Motive

Social media sites face a balancing act between serving as a positive force in society and their own growth and profits. For example, Facebook engineers and scientists conducted a series of experiments related to an algorithm designed to decrease the number of posts users saw that would be considered "bad for the world."[28] The problem for the platform came as the lower amount of objectionable content seen by users correlated to a lower frequency of users logging into the site. Facebook has also developed other technology that could decrease disinformation and other objectionable content that has been rejected by company leadership, such as a notification system to alert users when they have shared untrue information, which executives feared would "disproportionately show notifications to people who shared false news from right-wing websites."[29] Leaders of these platforms do not make decisions in a vacuum and must grapple with both financial and political consequences of each outcome.

Testifying before the Senate Commerce Committee, Facebook CEO Mark Zuckerberg stated, "we don't think tech companies should be making so many important decisions about these important issues alone. I believe we need a more active role for governments and regulators, which is why… I called for regulation on harmful content, privacy, elections, and data portability."[30] While individuals within social media companies may seek altruistic ends of promoting civil debate and the spread of truth, the companies themselves cannot be trusted to carry out this task alone due to the incentives they face as large corporations.

John Bowers and Jonathan Zittrain of Harvard Law School go further to assert that today's predominant models of content governance, described as "public relations-oriented," do not work well "because they assume that sensitive interests of individuals and society – like, say, the preservation of democratic norms in the face of disinformation – can be given reasonable treatment under corporate 'customer service' processes architected to defuse PR pressure and

---

[27] Telford, "YouTube Removes 8,000 Channels Promoting False Election Claims."
[28] Roose, Isaac, and Frenkel, "Facebook Struggles to Balance Civility and Growth."
[29] Roose, Isaac, and Frenkel.
[30] Zuckerberg, Hearing Before the United States Senate Committee on Commerce, Science, and Transportation, 2.

protect profitability."[31] While ICT firms at-large may have incentives to reduce disinformation within their control, these come largely from public pressure, rather than any inherent motive. Social media companies will not act upon their own accord unless the incentives align in such a way to encourage behavior, and too often it seems that the profit motive comes at odds with moderation policies that go beyond public appeasement. Government oversight of content governance, therefore, should play a more active role in the clear interest of safeguarding the spread of true information and promoting respectful and civil debate in online forums, "even at the cost of the shareholder."[32]

Relevant Policy History & Application

Three pieces of American policy stand out as key historical baselines to provide context for any future regulation of disinformation spread through social media. The first of these is the First Amendment to the Constitution, specifically the Free Speech clause. Second is Section 230 of the Communications Decency Act of 1996 which protects internet platforms from issues arising from the content created by their users. Finally, the third is the former policy of the FCC known as the Fairness Doctrine which sought to balance national conversations through the regulation of broadcast television content.

Free Speech Protections

The First Amendment of the Constitution lays out that, "Congress shall make no law… abridging the freedom of speech…" While this may seem straightforward enough, the reality of these protections is far more nuanced and dependent on the particular context. This constitutional level is of critical importance when designing policy that must pass judicial muster informed by past jurisprudence and arguments of academics for future interpretation. When considering regulation of social media, a court will consider both the type of speech regulated as well as a judicial framework shaped by past precedent to decide the constitutionality of a provision. These frameworks can be described as treating social media platforms equivalently to company towns, special industries, or news editors.[33]

---

[31] Bowers and Zittrain, "Answering Impossible Questions," 5.
[32] Bowers and Zittrain, 5.
[33] Brannon, "Free Speech and the Regulation of Social Media Content," 22–23.

Types of Speech

The Supreme Court has historically distinguished between protected speech and a limited number of unprotected categories.[34] While most speech is generally considered protected, previously enumerated categories of unprotected speech include obscenity, defamation, fraud, incitement, fighting words, true threats, speech integral to criminal conduct, and child pornography. While a court may consider disinformation to fall under the category of fraud, it is critical to recognize:

> '[S]ome false statements are inevitable if there is to be an open and vigorous expression of views in public and private conversation,' the Supreme Court has rejected a categorical First Amendment exception for false statements. Nevertheless, the Court has stated that false statements can form the basis for other 'legally cognizable harm[s]' such as defamation or fraud. In general, the government may regulate fraudulent speech in order to prevent public or consumer deception.[35]

Thus, the constitutionality of a regulation that moderates disinformation seems to be much more plausible than one seeking to squash misinformation at-large, as disinformation inherently includes an intent to defraud a targeted population. It is possible that courts may be willing to consider disinformation to be unprotected speech under the category of fraud as the harm done to the electorate through bad-faith actors causes public deception.

Company Town Framework

Under the framework of a company town, the Supreme Court has ruled that under certain circumstances, a private entity should be effectively treated as the government itself and held to the same constitutional standards.[36] While courts have so far held that social media platforms do not meet this standard, various scholars have argued that the public function of social media sites primarily providing a public forum may be enough to apply the company town framework. If courts were to adopt this view of social media companies with regard to freedom of speech, it would largely kill any ability, including that currently provided by Section 230, of social media platforms to regulate speech through their own content moderation policies. Further, the government would also be extremely limited in their ability to require the removal of anything

---

[34] Killion, "The First Amendment: Categories of Speech," 1–2.
[35] Killion, 2.
[36] Brannon, "Free Speech and the Regulation of Social Media Content," 23.

considered protected speech. Were courts to adopt this framework, it would be extremely detrimental to any efforts to mitigate disinformation spread in virtual public squares.

Special Industries Framework

The next framework is that of special industries, including broadcasters, cable providers, or common carriers.[37] Under this view, the Supreme Court has traditionally allowed the government greater regulation over specific industries due to nuanced interests, such as an inherent particularity of a medium. When considering the regulation of broadcast television in *Red Lion Broadcasting Co. v. FCC*, the Supreme Court upheld that the scarcity of radio frequencies and the interest of the public to receive content consistent with the First Amendment allowed for greater regulation than other forms of communication may receive.[38] The 1997 Supreme Court case *Reno v. ACLU* found that the internet at-large did not justify the greater regulation afforded to other media due to its historical lack of government regulation, the fact that the internet is not a scarce commodity, and the less-invasive role the internet played in the lives of Americans compared television, radio, and other forms of communication.[39]

The landscape of cyber has changed dramatically since 1997 and no major case has touched on these issues in the years since. The argument that the internet is not an invasive part of American life in particular seems ripe for reevaluating. A total of 37% of Americans prefer to get their local news via online sources including social media (15%) and news websites or applications (23%).[40] This is in comparison to 41% who prefer to access local news via television, 13% who prefer print media, and 8% who prefer radio. Additionally, 89% of Americans ever get news from websites or social media. This part of the justification of denying the government regulatory capacity over the internet seems to have changed enough to warrant re-evaluation by the courts.

The inherent nature of social media platforms and specifically the way they curate content may indeed render them more functionally similar to broadcasters than initially apparent. Social media sites tend to determine most content ordering and presentation independently of the user. While an individual may choose to view a particular page with specific content (e.g., the

---

[37] Brannon, 27.
[38] Ruane, "Fairness Doctrine: History and Constitutional Issues," 5.
[39] Brannon, "Free Speech and the Regulation of Social Media Content," 20, 30.
[40] Pew Research Center, "For Local News, Americans Embrace Digital," 13.

profile page of a friend or news source), most home pages or feeds are determined by an algorithm based upon who a user follows and what interests the platform has decided the user may have. The fact that a user has relatively limited say in the way content is presented to them, especially as platforms move away from simple reverse-chronological timelines to custom algorithms, seems to position social media companies as broadcasters who decide what relatively limited content is available for a user to come across and discover when accessing the platform. Additionally, social media platforms tend to prioritize large, commercial entities over individuals or small organizations which has created forms of scarcity in certain types of speech.[41] Were the courts to adopt this updated thinking and overturn precedent rejecting the special interests of regulating internet and social media activities, it would allow for a potential path forward on mitigating disinformation through public policy action.

Newspaper Editor Framework

The final framework that a court may choose to view government regulation of speech through is that of a newspaper editor.[42] Newspapers, and other similar industries which engage in content moderation, act as more than conduit for speech of others, but rather partake in protected speech of their own. Under this framework, were the government to proscribe that an organization must or must not publish something, it would be taking away this editorial control and infringing on the organization's own freedom of speech. Several federal trial courts have previously ruled that search engines can fall into this category as they determine whether and how to present relevant information to a user. Of particular relevance to the discussion of disinformation, the Supreme Court has ruled, "the creation and dissemination of information are speech within the meaning of the First Amendment. Facts, after all, are the beginning point for much of the speech that is most essential to advance human knowledge and conduct human affairs."[43] Whether or not the opposite of this, the curation of deliberately untrue facts, would also be considered speech seems to be an open question for the Court. Nevertheless, were a court to proscribe the newspaper editor framework, any regulation would therefore undergo strict scrutiny when examining implications for speech. However, even with this, the government may

[41] Brannon, "Free Speech and the Regulation of Social Media Content," 31.
[42] Brannon, 33.
[43] Sorrell v. IMS Health Inc., 564 US at 570.

pass scrutiny and be able to regulate speech to combat extraordinary issues with specifically-tailored interventions without too much burden on speech.[44]

Section 230 Shield

Then-Representatives Chris Cox (R-CA) and Ron Wyden (D-OR) collaborated to introduce Section 230 into the controversial Communications Decency Act of 1996.[45] The measure received very little attention or opposition, as few could have anticipated the impact the relatively short and straightforward law would have and attention was focused on more immediately impactful issues in the bill. In his book detailing the history of Section 230, Jeff Kosseff describes the legacy of the statute as, "those twenty-six words have fundamentally changed American life."[46] Further, internet law scholar David Post describes the statute with, "[n]o other sentence in the U.S. Code…has been responsible for the creation of more value than that one."[47] Finally, In testimony before the Senate Commerce Committee, Mark Zuckerberg described Section 230 as a "foundational law" that both "encourages freedom of expression" and "allows platforms to moderate content."[48] This seemingly simple law has created much of the environment needed for social media and the internet as a whole to reach the point it has to this day.

As Zuckerberg alluded to, the key provisions of the law appear as the two halves of Section 230(c) which governs the relationship between "interactive computer services" (e.g., social media platforms) and "information content providers" (i.e., entities creating content online).[49] In application to social media sites, Section 230(c)(1) grants immunity to platforms hosting the speech of users as that speech cannot be reflected upon the platform as its own. However, when a platform is both an interactive computer service and an information content provider, this provision does not take effect.[50] This particularly comes into play on questions of whether a platform had a hand in creating speech when it edits, augments, or otherwise changes the speech of its users. The second provision of Section 230(c) grants the ability for online

---

[44] Brannon, "Free Speech and the Regulation of Social Media Content," 39.
[45] Kosseff, *The Twenty-Six Words That Created the Internet*, 2–3.
[46] Kosseff, 3.
[47] Post, "A Bit of Internet History."
[48] Zuckerberg, Hearing Before the United States Senate Committee on Commerce, Science, and Transportation, 1.
[49] See Appendix: 47 US Code § 230(c) for full text.
[50] Brannon, "Free Speech and the Regulation of Social Media Content," 12.

services to, in good faith, restrict or remove "obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable" content on their platforms.

When combined, these two provisions do create much of the environment needed for social media to have reached the point it has. Section 230(c)(1) allows platforms to choose not to censor their sites as they have immunity from what is posted on them.[51] Section 230(c)(2) instead applies when a platform does choose to moderate offensive material. These aspects of the law both ensure a vested interest in social media sites allowing relatively wide amounts of freedom of speech, while simultaneously allowing for explicit ability to moderate content, including the removal of constitutionally protected speech by the private parties. Additionally, both provisions shield online organizations from lawsuits through immunity regarding both what they allow and disallow to be hosted on their platforms.

Fairness Doctrine

In looking to shape modern federal policy aimed at shaping communication in the public sphere, it is critical to understand past attempts at similar aims. For three decades between the 1940s and the 1980s, the FCC implemented a policy that was come to be known as the "fairness doctrine."[52] This policy sought to require public broadcasters who were licensed by the government to "devote a reasonable portion of broadcast time to the discussion and consideration of controversial issues of public importance" and "affirmatively endeavor to make… facilities available for the expression of contrasting viewpoints held by responsible elements with respect to the controversial issues presented."[53] Essentially, broadcasters were charged not only with allocating time to issues of importance, but also to use their editorial discretion in determining a responsible way to fairly show various opposing viewpoints surrounding these issues.[54] As previously discussed in considering First Amendment protections, the Supreme Court upheld FCC regulatory authority over broadcasters in *Red Lion* due to the scarce nature of frequencies and the regulation doing so in the public interest.

The federal government could seek to replicate aspects of the fairness doctrine in a modernized fashion in an effort to once again promote civic discourse. In doing so, they would

---

[51] Brannon, 14.
[52] Ruane, "Fairness Doctrine: History and Constitutional Issues," 1.
[53] "Applicability of the Fairness Doctrine," 10426.
[54] Ruane, "Fairness Doctrine: History and Constitutional Issues," 2.

be pursuing regulation of social media companies through a special industries framework. Possible policy inspiration drawn from the fairness doctrine may include pushing social media companies to display trustworthy and balanced information to users, rather than allowing for the formation of highly-polarized information bubbles within networks.

Policy Recommendations

A number of policy mechanisms exist as feasible options for addressing the problem of disinformation spread online. The recommendations outlined here represent a number of legislative options that should be enacted in concert to disrupt disinformation spread on a number of fronts. Additionally, Bowers and Zittrain assert that future regulation "must consider both solutions that assign platforms new duties with respect to content governance, and those that would delegate important aspects of the content governance process outside of the platforms themselves."[55] These proposals further seek to adhere to this idea and place burden on both ICT companies as well as additional avenues of resolution, including expanded ability of the courts to intervene on such matters.

Hole in the Shield

A variety of litigants have attempted to hold platforms liable for perceived harms of their content moderation only to have legal challenges thrown out on the grounds of the broad immunity shield granted by Section 230.[56] Amending Section 230 or implementing additional legislation creating express exemptions to this immunity would allow for plaintiffs to bring cases to court and be heard without immediate dismissal. Currently, most perceived harms imposed by ICT companies acting as platforms have no judicial avenue for recourse. Previously, Congress has created specific exemptions from Section 230 protections, such as the Allow States and Victims to Fight Online Sex Trafficking Act of 2017.[57] It could thus choose to once again create a hole in the broad immunity shield, this time to specifically target disinformation.

Jeff Kosseff notes, "Congress giveth Section 230, and Congress can taketh it away. Technology companies—the most direct beneficiaries of Section 230—have a responsibility to

[55] Bowers and Zittrain, "Answering Impossible Questions," 5.
[56] Brannon, "Free Speech and the Regulation of Social Media Content," 9–10.
[57] Brannon, "Regulating Big Tech," 3.

prove that they are using this extraordinary immunity to the benefit of society."[58] Immunity for hosting disinformation could be revoked in whole, or in tandem with additional legislation such as that discussed in the following section. Congress could stipulate that platforms not implementing satisfactory good-faith efforts to moderate disinformation and other misleading content that defrauds the public would lose Section 230 immunity for that content.

In a 2018 white paper discussing social media and technology regulation, Senator Mark Warner raises a possible initiative to allow for the courts to find ICT firms liable for failure to remove content that state courts have ruled violate various laws (including defamation, false light, and public disclosure of private facts).[59] Under such process, victims with a favorable ruling from a court would notify a platform of court action. This process then removes Section 230 immunity for hosting the content once notified and if the company fails to remove or prevent future re-uploading of relevant materials. This arrangement addresses various issues at stake as it provides a judicial path for remedying harms caused by some types of disinformation and places such burden on hosting platforms without overextending new holes in Section 230 immunity.

While any specific modification to Section 230 may seem straightforward, it would not be without consequence. Kosseff further warns regarding rolling back or repealing Section 230:

> The modern Internet in the United States is built on the foundation of Section 230. To eliminate Section 230 would require radical changes to the Internet…The internet without Section 230 would be an Internet in which litigation threats could silence the truth. If websites were liable for third-party content after receiving complaints, they probably would quickly remove the content. Some of this content has little social value. But much of it does.[60]

Opening up a wider path for the removal of content through judicial means is then a double-edged sword. While it may successfully result in the removal of disinformation, weakened protections will likely also create collateral damage as platforms remove legitimate content as the easier route when faced with potential lawsuits. Platforms would almost certainly retain the protections of Section 230(c)(2) granting them the ability to remove content as they see fit, and potentially have these powers expanded were disinformation to be added as an explicit category of approved moderation. Thus, any weakening of Section 230 would seemingly tend to result in ICT companies removing more content than under status quo protections. As Congress shapes

---

[58] Kosseff, *The Twenty-Six Words That Created the Internet*, 277.
[59] Warner, "Potential Policy Proposals for Regulation of Social Media and Technology Firms," 8–10.
[60] Kosseff, *The Twenty-Six Words That Created the Internet*, 278–79.

legislation aimed at targeting disinformation through Section 230 reform, it must balance the benefit and detriment to the public good and put forth a well-tailored intervention.

Good-Faith Moderation

Due to the inherent nature of social media, regulatory action must also occur at the level of the ICT companies themselves. In this day and age, information travels and replicates with incredible speed when compared to traditional media. Solutions must exist to both prevent and take down content without the need for much more time consuming judicial intervention or it often will be too little too late. Government regulators must prescribe platforms a duty to moderate content in good-faith for their users. This act could, in some regards, be seen as a modern-day implementation of the fairness doctrine. In that same spirit of furthering public discourse, regulation may expand beyond the scope of simply requiring increased moderation of disinformation to also include promotion of well-balanced and reputable media consumption at-large. This would fight disinformation by effectively providing a counterbalance to falsehoods through verified journalism, increasing access to factual information, and bursting media bubbles and echo chambers.

Such action would leave a good deal of room for interpretation of exact implementation, as did the fairness doctrine. Instead of outlining specific actions broadcasters must have adopted, the FCC instead relied on assessing the reasonability of actions and whether they were taken in good faith in support of the doctrine.[61] Bowers and Zittrain offer a similar proposal for future solutions in addressing the content governance of disinformation through the creation of a fiduciary relationship between platforms and users.[62] They argue that in many relationships where knowledge asymmetries exist (e.g., between a doctor and patient), the law recognizes and protects the disadvantaged party by requiring the fiduciary to act in the client's best interests, even at the expense of their own.

Congress can use this framework to regulate a fiduciary relationship between platforms and users. Considering examples previously discussed when examining the profit motive of corporations, these platforms seem to ultimately answer solely to shareholders despite potential benefits to society at-large. Such a relationship would require these platforms to instead put their

---

[61] Ruane, "Fairness Doctrine: History and Constitutional Issues," 3.
[62] Bowers and Zittrain, "Answering Impossible Questions," 6.

users' interests first. This relationship should then in turn benefit liberal democratic function through the promotion of a well-informed electorate. Bowers and Zittrain cite an example that as a fiduciary, a platform would show users potentially objectionable content (e.g., conspiracy theories) when actively seeking them out, but not when casually perusing their feed at risk of "exposing a potentially unwitting individual to manipulation."[63] In general, this model of addressing the asymmetrical relationships takes precedent from a number of other professional fields and would address a number of issues with the current status quo of largely unregulated interaction of social media platforms and their users.

A primary consideration that policymakers must decide on this action is the ability of the federal government to compel social media companies to regulate content. If legal challenges, in the view of Congress, would be successful challenging this power, lawmakers may instead opt to design a system to incentivize companies to willingly adopt enhanced moderation policies. Compelling private entities to either not publish disinformation or to publish reliable journalism may impede upon the protected speech rights of the platforms in the eyes of the court. If Congress thinks that this line of argument may be successful in overturning legislation, it could then opt to create a series of incentives for platforms, such as tax breaks or grants. These incentives would be in exchange for adopting stricter moderation of content in order to combat the negative perverse incentives introduced by their profit motives.

As with the fairness doctrine, the implementation of such policy will likely once again rely on the FCC. Previously, the FCC itself made decisions on how broadcasters carried out the directives to present fair and balanced perspectives of issues.[64] In developing policy, Congress should delegate enforcement once again to this agency and allow for suitable mechanisms for enforcement as determined by lawmakers. The FCC would then be charged with evaluating how individual platforms carry out policy directives and whether they are done to a satisfactory degree. As the goal of this action is to drive action on the platform level, oversight to ensure adhesion to regulations is of critical importance.

Label & Restrict

[63] Bowers and Zittrain, 6.
[64] Ruane, "Fairness Doctrine: History and Constitutional Issues," 3.

The final recommendation would be to require either labeling or removal of content posted by bots, foreign actors, or inauthentic accounts. Under this proposal, platforms would be charged with a duty to identify content created by each category of user and deal with it appropriately. Such interventions could include conspicuously labeling posts with sources of origin or removing them altogether. The intent behind this act would be to dampen the power of disinformation spread through such accounts by contextualizing media as potentially misleading. These reforms could also be seen as a further extension of the fiduciary duty discussed in the previous section as social media platforms lessen the magnitude of the information asymmetry through such disclosures or actions.

The 2018 Warner white paper considering policy options for regulating ICT firms also navigates each of the three options for labeling or restricting different categories of actors.[65] Of these, the research finds the least issue with "clearly and conspicuously" labeling bot accounts.[66] As technology has been developing, bots are becoming harder and harder to distinguish from humans, and ICT firms could protect the public interest by making known what content is ultimately coming from non-human accounts, including those provided by platforms themselves. California has previously adopted a 2018 law restricting using "an automated online account" to make certain communications.[67] Federal action could now follow.

The second set of content that the government could take action on limiting would be posts of foreign origin or from inauthentic accounts. Besides automation, the anonymous nature of the internet leaves room for bad-faith actors to commandeer aspects of social media for nefarious purposes.[68] Users can claim to be located in any part of the world while situated in another entirely and introduce disinformation aimed at undermining the target locale. Further, individuals or groups can create accounts based on false identities with little impediment. These accounts can also be used to introduce and spread disinformation. These two situations, however, introduce additional issues than bot tracking due to the human component involved. Specifically, tracing and publicizing information about a user or their whereabouts introduces complex privacy concerns. As pointed out in the Warner white paper, such restrictions could harm marginalized groups (e.g., real name identify verification policies affecting LGBT, Native

---

[65] Warner, "Potential Policy Proposals for Regulation of Social Media and Technology Firms," 6–8.
[66] Warner, 6.
[67] Brannon, "Regulating Big Tech," 3.
[68] Warner, "Potential Policy Proposals for Regulation of Social Media and Technology Firms," 6.

American groups), increase user tracking, and potentially allow attacks on freedom of speech and privacy in oppressive regimes. Additionally, determining the ultimate location of a user is made more difficult due to VPN and IP masking technology.

As with previous proposals, disclosure requirements or removal of speech could come at odds with constitutional protections. Depending on the framework of First Amendment rights adopted by a court, judicial review may find that compelling additional speech (e.g., labeling content with disclaimers) or requiring a platform to not publish certain speech (e.g., removing posts) infringes upon the rights of the site to edit their own content.[69] However, as with *Red Lion*, courts may instead find a clear enough interest, along with narrowly tailored regulation, to uphold such action as a special industry. Additionally, lower courts have previously ruled that computer-generated speech may be liable for protections as long as it communicates information to humans, as it is ultimately the speech of the human author of the code.[70] This further complicates matters as issues of whether mandating the labeling or removal of bot-generated content would likely need to be reviewed by the courts.

Ultimately, lawmakers must balance these challenges with the public interest when shaping future policy aimed at disinformation. Such legislation must show clear intent in purpose and must not be widely written enough to overstep any clear and defined goal. If such regulation is deemed to not pass scrutiny, this also could be included as a part of legislation with incentives for social media companies to adopt reforms in an attempt to counter profit motives without direct implications on government-imposed restrictions on speech.

Conclusion

Should the US government decide to attempt to regulate the spread of disinformation shared through social media sites, it will face numerous challenges both politically and logistically. Most prominently, policymakers will need to navigate the variety of freedom of speech protections that may apply in this case. While it is not inconceivable to be able to draft law which passes constitutional muster and appropriately handles the traditionally broad immunities granted through Section 230, it will need to be done carefully and intentionally. Such interventions may include allowing a wider judicial channel for perceived harms by creating

---

[69] Brannon, "Regulating Big Tech," 4.
[70] Brannon, "Free Speech and the Regulation of Social Media Content," 17.

exceptions in Section 230 immunity, compelling or incentivizing good-faith moderation through creation of fiduciary relationships, or hampering the efficacy of inauthentic accounts through labeling or removal. Each of these ultimately aims to limit the amount of disinformation in the marketplace that Justice White spoke of in the *Red Lion* decision.

Hostile foreign operatives or the President of the United States themselves may each have incentives to introduce false information into the public consciousness to sway electoral or political institutions. This fact stood true decades ago and remains true to this day. What has changed since Justice White's analysis is the effectiveness in technology in spreading such disinformation rapidly and effectively. The President can now speak directly to more than 80 million people through a single tweet and Russian agents can pretend to be red-blooded Americans directly from Moscow. The marketplace of ideas has changed dramatically and now requires renewed protections to safeguard the interests of both the electorate and democratic institutions. Options for reform and regulation can aim to do just that.

Bibliography

"Applicability of the Fairness Doctrine in the Handling of Controversial Issues of Public Importance." In *Federal Register 29*, 145:10416–27, 1964.

Bowers, John, and Jonathan Zittrain. "Answering Impossible Questions: Content Governance in an Age of Disinformation." *Harvard Kennedy School Misinformation Review* 1, no. 1 (January 14, 2020). https://doi.org/10.37016/mr-2020-005.

Brannon, Valerie C. "Free Speech and the Regulation of Social Media Content." Washington, DC: Congressional Research Service, March 27, 2019. https://crsreports.congress.gov/product/details?prodcode=R45650.

———. "Regulating Big Tech: Legal Implications." Washington, DC: Congressional Research Service, September 11, 2019.

Dictionary.com. "'Misinformation' vs. 'Disinformation': Get Informed On The Difference." *Dictionary.Com* (blog), May 15, 2020. https://www.dictionary.com/e/misinformation-vs-disinformation-get-informed-on-the-difference/.

Dwoskin, Elizabeth, and Craig Timberg. "The Unseen Machine Pushing Trump's Social Media Megaphone into Overdrive." *Washington Post*, October 30, 2020. https://www.washingtonpost.com/technology/2020/10/30/trump-twitter-domestic-disinformation/.

Gadde, Vijaya, and Kayvon Beykpour. "Additional Steps We're Taking Ahead of the 2020 US Election." *Twitter Blog* (blog), November 2, 2020. https://blog.twitter.com/en_us/topics/company/2020/2020-election-changes.html.

Guess, Andrew M., Brendan Nyhan, and Jason Reifler. "Exposure to Untrustworthy Websites in the 2016 US Election." *Nature Human Behaviour* 4, no. 5 (May 2020): 472–80. https://doi.org/10.1038/s41562-020-0833-x.

Killion, Victoria L. "The First Amendment: Categories of Speech." Washington, DC: Congressional Research Service, January 16, 2019. https://crsreports.congress.gov/product/pdf/IF/IF11072.

Kosseff, Jeff. *The Twenty-Six Words That Created the Internet*. Ithaca: Cornell University Press, 2019.

Mitchell, Amy, Mark Jurkowitz, J. Baxter Oliphant, and Elisa Shearer. "Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable." Pew Research Center, July 30, 2020. https://www.journalism.org/wp-content/uploads/sites/8/2020/07/PJ_2020.07.30_social-media-news_REPORT.pdf.

Pew Research Center. "For Local News, Americans Embrace Digital but Still Want Strong Community Connection," March 26, 2019. https://www.journalism.org/wp-content/uploads/sites/8/2019/03/PJ_2019.03.26_Local-News_FINAL.pdf.

———. "Social Media Fact Sheet." Pew Research Center, June 12, 2019. https://www.pewresearch.org/internet/fact-sheet/social-media/.

Post, David. "A Bit of Internet History, or How Two Members of Congress Helped Create a Trillion or so Dollars of Value." *Washington Post*, August 27, 2015. https://www.washingtonpost.com/news/volokh-conspiracy/wp/2015/08/27/a-bit-of-internet-history-or-how-two-members-of-congress-helped-create-a-trillion-or-so-dollars-of-value/.

Qiu, Linda. "Trump Has Amplified Voting Falsehoods in over 300 Tweets since Election Night." *The New York Times*, November 16, 2020. https://www.nytimes.com/2020/11/16/technology/trump-has-amplified-voting-falsehoods-in-over-300-tweets-since-election-night.html.

Red Lion Broadcasting Co. v. FCC, 395 US 367 (1969).

Roose, Kevin, Mike Isaac, and Sheera Frenkel. "Facebook Struggles to Balance Civility and Growth." *The New York Times*, November 24, 2020, sec. Technology. https://www.nytimes.com/2020/11/24/technology/facebook-election-misinformation.html.

Rosenbaum, Eric. "Most Important 2020 Election Misinformation Threat Is Not Coming from Overseas: Facebook Former Security Chief Alex Stamos." CNBC, October 30, 2020. https://www.cnbc.com/2020/10/30/most-important-2020-election-misinformation-threat-not-from-overseas.html.

Ruane, Kathleen Ann. "Fairness Doctrine: History and Constitutional Issues." Washington, DC: Congressional Research Service, July 13, 2011.

Sorrell v. IMS Health Inc., 564 US 552 (2011).

Telford, Taylor. "YouTube Removes 8,000 Channels Promoting False Election Claims." *Washington Post*, December 9, 2020. https://www.washingtonpost.com/business/2020/12/09/youtube-false-2020-election-claims/.

Timberg, Craig, and Tony Romm. "New Report on Russian Disinformation, Prepared for the Senate, Shows the Operation's Scale and Sweep." *Washington Post*, December 17, 2018. https://www.washingtonpost.com/technology/2018/12/16/new-report-russian-disinformation-prepared-senate-shows-operations-scale-sweep/.

Trump, Donald J. "I Won the Election!" Tweet. *@realDonaldTrump* (blog), November 16, 2020. https://twitter.com/realDonaldTrump/status/1328334945148952576.

———. "@realDonaldTrump." Twitter. Accessed November 25, 2020. https://twitter.com/realDonaldTrump.

Warner, Mark R. "Potential Policy Proposals for Regulation of Social Media and Technology Firms," July 23, 2018. https://www.warner.senate.gov/public/_cache/files/d/3/d32c2f17-cc76-4e11-8aa9-897eb3c90d16/65A7C5D983F899DAAE5AA21F57BAD944.social-media-regulation-proposals.pdf.

Zakrzewski, Cat, and Rachel Lerman. "No Social Media Is Safe: How Election Misinformation Spread on LinkedIn, Pinterest and Nextdoor." *Washington Post*. Accessed November 25, 2020. https://www.washingtonpost.com/technology/2020/11/07/pinterest-linkedin-election-disinfo/.

Zuckerberg, Mark. Hearing Before the United States Senate Committee on Commerce, Science, and Transportation (2020). https://www.commerce.senate.gov/services/files/E017B34E-F87F-4127-88A7-2C32B6BC3810.

Appendix: 47 US Code § 230(c)

Protection for "Good Samaritan" blocking and screening of offensive material

(1) Treatment of Publisher or Speaker

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

(2) Civil Liability

No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

(B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).