

Washington University Law Review

Volume 98 | Issue 3

2021

How Content Moderation May Expose Social Media Companies to Greater Defamation Liability

Tanner Bone

Washington University School of Law

Follow this and additional works at: https://openscholarship.wustl.edu/law_lawreview



Part of the [Internet Law Commons](#), and the [Torts Commons](#)

Recommended Citation

Tanner Bone, *How Content Moderation May Expose Social Media Companies to Greater Defamation Liability*, 98 WASH. U. L. REV. 937 (2021).

Available at: https://openscholarship.wustl.edu/law_lawreview/vol98/iss3/10

This Note is brought to you for free and open access by the Law School at Washington University Open Scholarship. It has been accepted for inclusion in Washington University Law Review by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

HOW CONTENT MODERATION MAY EXPOSE SOCIAL MEDIA COMPANIES TO GREATER DEFAMATION LIABILITY

INTRODUCTION

On July 25, 2018, Vice News published an online article discussing allegations from various Twitter users that Twitter was “shadow-banning.”¹ The users accused Twitter of secretly removing or hiding the accounts of several prominent, politically conservative users of the platform.² Twitter has chalked up the perceived shadow-banning to an algorithmic error,³ but these allegations and others like them have led to broader conversations about content moderation on some of the world’s largest social media platforms. Since Vice’s article, accusations of politically motivated content curation have increased in both volume and frequency, including accusations from the President of the United States.⁴ Companies like Twitter, Facebook, and YouTube have evolved into popular sources for breaking news and act as a primary means of communication for millions of people.⁵ The expanding influence of social media has led some to call these platforms the new “town square.”⁶ Business executives are using social media to engage with consumers.⁷ Political leaders are taking to

1. Alex Thompson, *Twitter Appears to Have Fixed “Shadow Ban” of Prominent Republicans Like the RNC Chair and Trump Jr.’s Spokesman*, VICE NEWS (July 25, 2018, 12:37 PM), https://www.vice.com/en_us/article/43paqq/twitter-is-shadow-banning-prominent-republicans-like-the-rnc-chair-and-trump-jrs-spokesman [<https://perma.cc/CB4V-LX38>].

2. *Id.*

3. Vijaya Gadde & Kayvon Beykpour, *Setting the Record Straight on Shadow Banning*, TWITTER BLOG (July 26, 2018), https://blog.twitter.com/en_us/topics/company/2018/Setting-the-record-straight-on-shadow-banning.html [<https://perma.cc/4JMA-BQH7>].

4. Donald J. Trump (@realDonaldTrump), TWITTER (July 26, 2018, 6:46 AM), <https://twitter.com/realDonaldTrump/status/1022447980408983552> [<https://perma.cc/AYM4-TRLF>] (“Twitter ‘SHADOW BANNING’ prominent Republicans. Not good. We will look into this discriminatory and illegal practice at once! Many complaints.”). Just before this Note was published, Twitter permanently suspended President Trump’s personal Twitter account following the storming of the Capitol on January 6, 2021.

5. Elisa Shearer, *Social Media Outpaces Print Newspapers in the U.S. as a News Source*, PEW RSCH. CTR. (Dec. 10, 2018), <https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/> [<https://perma.cc/WQ3K-MH9E>] (explaining “[o]nly five U.S. adults say they often get news via social media, slightly higher than the share who often do so from print newspapers (16%)”).

6. Brian Kane, *Social Media is the New Town Square: The Difficulty in Blocking Access to Public Accounts*, NAT’L ASS’N OF ATT’YS GEN. (Feb. 2018), <https://www.naag.org/publications/nagtri-journal/volume-3-number-1/social-media-is-the-new-town-square-the-difficulty-in-blocking-access-to-public-accounts.php> [<https://perma.cc/3DEQ-RD26>].

7. See Allana Akhtar, *4 Things the Best CEOs Do on Social Media to Make Authentic Connections with Their Followers*, BUS. INSIDER (June 11, 2019, 8:22 AM), <https://www.businessinsider.com>

social media to discuss new initiatives and policies.⁸ Athletes and celebrities are using these platforms to provide social commentary and to galvanize supporters and fans.⁹ In sum, social media has spread rapidly throughout society and is now present in most Americans' lives in innumerable ways.¹⁰

While the evidence of shadow-banning in this instance is speculative and contested,¹¹ these allegations and others have spurred many conversations about social media, content curation, and free speech. More specifically, scholars and politicians are asking how much censorship power is appropriate for these social media companies.¹² In this era of “fake news,” “hot takes,” and political polarization when society seems more and more reliant on social media, we must reevaluate how these companies regulate the content on their platforms. In particular, this means revisiting Section 230 of the Communications Decency Act (CDA) in order to ensure that these powerful companies foster a vibrant and open marketplace of ideas.

This Note will explain the critical distinction between “publishers” and “platforms,” why social media entities are currently considered “platforms,” and why the legal system should reevaluate the liability of social media entities based on how they moderate and regulate content. Part I of this Note will discuss the history of the common-law liability of content providers prior to the invention of the internet. It will also explore the history and rationale for enacting Section 230 of the CDA. Part II of this Note will explain the distinction between “publishers” and “platforms” as it relates to defamation liability. Further, it will discuss the rapid growth of social media

r.com/why-and-how-ceos-should-use-social-media-2019-6 [https://perma.cc/5J9T-25MB] (“Celebrities and influencers aren’t the only ones using social media to bolster their career. Chief Executive Officers . . . are using Facebook, Instagram, and LinkedIn more than ever to build an online following.”). See also James Crowley, *Dave Portnoy Announces Barstool Fund to Aid Small Businesses Impacted by Covid*, NEWSWEEK (Dec. 17, 2020), <https://www.newsweek.com/dave-portnoy-announces-barstool-fund-small-businesses-covid-1555783> [https://perma.cc/9NW3-GHH2] (“In a video he tweeted out, [Barstool Sports CEO Dave] Portnoy said that [Barstool Sports] was setting up the \$500,000 fund to try to help restaurants get to the next month until the coronavirus pandemic ends.”).

8. See JOHN H. PARMALEE & SHANNON L. BICHARD, *POLITICS AND THE TWITTER REVOLUTION: HOW TWEETS INFLUENCE THE RELATIONSHIP BETWEEN POLITICAL LEADERS AND THE PUBLIC* 35–37 (2012).

9. Jiahui Zhuang, *I’m Your Fan – Engaging in Celebrity’s Social Media Page with the Mediation of Parasocial Interaction and Parasocial Relationship*, GRADUATE THESES & DISSERTATIONS (2018), <http://scholarcommons.usf.edu/etd/7252>.

10. As of February 2019, it was estimated that 72% of Americans use at least one type of social media. *Social Media Fact Sheet*, PEW RSCH. CTR. (Feb. 7, 2019), <https://www.pewresearch.org/internet/fact-sheet/social-media/> [https://perma.cc/43A3-82RS].

11. Gadde & Beykpour, *supra* note 3; see also Cal Jeffrey, *USPTO Grants Facebook Patent for Automated Shadow-Banning System*, TECHSPOT (July 16, 2019, 1:32 PM), <https://www.techspot.com/news/80979-uspto-grants-facebook-patent-automated-shadow-banning-system.html> [https://perma.cc/2S87-P2LM] (“[T]here has never been any definitive proof that moderators and social networking platforms execute shadow bans . . .”). Shadow-banning is more thoroughly explained in Part II.b.

12. Cf. Anupam Chander & Vivek Krishnamurthy, *The Myth of Platform Neutrality*, 2 GEO. L. TECH. REV. 400, 404–06 (2018) (arguing that many social media algorithms can create “echo chambers,” spread “fake news,” and ultimately promote relevant information rather than accurate information).

during the internet age and its impact on communication and the spread of information. It will also discuss the cryptic and often vague algorithmic process that social media companies use to decide which content is visible to users. Part III of this Note will analyze the current liability of social media companies as a “platform” and will discuss the argument that social media is the twenty-first century’s “town square.” Part IV will explain three key pieces of recently proposed legislation that may affect Section 230 of the CDA. Part V of this Note will explain specific changes that social media companies must make to avoid the enhanced defamation liability of moving from the “platform” category to the “publisher” category. Part VI will discuss a few legislative and executive solutions to allow Section 230 of the CDA to reflect the current internet landscape by focusing on pushing social media companies toward transparent content-moderation practices.

I. DEFAMATION AND THE CREATION OF THE COMMUNICATIONS DECENCY ACT

To fully grasp the conversations and debate surrounding the legal implications of content-moderation processes such as “shadow-banning,” it is necessary to understand defamation liability as it developed under the common law. It is also critical to understand how this liability developed and changed in order to foster the growth of the internet. This section will discuss defamation liability under the common law through the enactment of the Communications Decency Act.

A. *Defamation Under the Common Law*

Defamation liability arises when three elements are present: (1) there is an unprivileged publication; (2) the publication is false and defamatory; and (3) the publication is actionable irrespective of special harm or is the legal cause of special harm to the other.¹³ Any person or entity that publishes a defamatory statement assumes liability equivalent to having been the initial purveyor of the statement.¹⁴ For entities that provide information or content, the assumed liability is categorized into two principal categories: publishers and distributors.¹⁵

13. RESTATEMENT (FIRST) OF TORTS § 558 (AM. L. INST. 1938).

14. *Immunity for Online Publishers Under the Communications Decency Act*, DIGIT. MEDIA L. PROJECT, <http://www.dmlp.org/legal-guide/immunity-online-publishers-under-communications-decency-act> [https://perma.cc/E6U4-X9WQ] (“Under standard common-law principles, a person who publishes a defamatory statement by another bears the same liability for the statement as if he or she had initially created it.”).

15. Throughout this Note, the terms “platform” and “distributor” will be used interchangeably.

Publishers, including book publishers and newspapers such as the *New York Times*, are liable for any content that appears in their publications.¹⁶ The basis for this liability is that these publishers review and curate their content, so they have full knowledge of the material they were releasing to the public.¹⁷ Thus, common-law publishers are held to a higher legal standard because of their subjective editorialization of content.¹⁸ Conversely, distributors (non-digital “platforms”) such as libraries, bookstores, and newspaper stands are not held liable for the content they disperse.¹⁹ The rationale is that it is not feasible or practical to expect libraries and newspaper stands to know all of the content that they possess and or provide to the public because the content is voluminous and comes from a variety of sources.²⁰ Accordingly, distributors are not subject to defamation liability for the content of the sources they provided. Put in today’s terms with real-world examples:

Typically, publishers are considered to have editorial judgment, while platforms lack it. From this perspective, the *Harvard Business Review*, *The Atlantic*, and *The New York Times* are classic “publishers”—they present highly-curated content, and their editors invest a lot of time in its creation. Google, Facebook, and Twitter are classic “platforms”—they distribute other peoples’ content without as much editorial oversight. But these differences are largely cultural. It’s not technologically difficult for publishers to add platform-like elements, and vice versa.²¹

Under the common law, websites that do not or cannot review all of their content operate as “platforms” or “distributors.”²²

16. *Immunity for Online Publishers*, *supra* note 14.

17. *Id.*

18. *See, e.g.*, *Stratton Oakmont, Inc. v. Prodigy Servs. Co.*, 1995 WL 323710, at *5 (N.Y. Sup. Ct. May 24, 1995) (holding that Prodigy is a publisher because “Prodigy’s conscious choice, to gain the benefits of editorial control, has opened it up to a greater liability than CompuServe and other computer networks that make no such choice.”), *superseded by statute*, 47 U.S.C. § 230.

19. *Immunity for Online Publishers*, *supra* note 14.

20. *Id.* (“The concern is that it would be impossible for distributors to read every publication before they sell or distribute it, and that as a result, distributors would engage in excessive self-censorship.”).

21. Lydia Laurenson, *Don’t Try to Be a Publisher and a Platform at the Same Time*, HARV. BUS. REV. (Jan. 19, 2015), <https://hbr.org/2015/01/dont-try-to-be-a-publisher-and-a-platform-at-the-same-time> [<https://perma.cc/LTF2-KC9A>].

22. *See, e.g.*, *Cubby, Inc. v. CompuServe, Inc.*, 776 F. Supp. 135, 140 (S.D.N.Y. 1991) (“CompuServe has no more editorial control over such a publication than does a public library, book store, or newsstand, and it would be no more feasible for CompuServe to examine every publication it carries for potentially defamatory statements than it would be for any other distributor to do so.”).

B. The Rise of the Internet and the Origin of Section 230 of the CDA

Before 1996, the aforementioned common-law distinctions provided the basis for liability for internet content providers. This common-law standard—as applied to the internet—was called into question after *Stratton Oakmont, Inc. v. Prodigy Services Co.*²³ Prodigy was an internet company that hosted online bulletin boards with over two million subscribers and sixty thousand posts per day.²⁴ Prodigy oversaw the content of the bulletin boards and occasionally removed posts that were “offensive” or “in bad taste.”²⁵ The court held that Prodigy’s practice of moderating some of the content on its platform required that it assume liability for all of the content on its platform.²⁶ In other words, to be a platform and thus avoid liability under the common-law standard of defamation, there could be no content moderation whatsoever.

The *Stratton Oakmont* decision in 1995 initiated a swift legislative response. In an attempt to “regulate obscenity and indecency online,” Congress passed the Communications Decency Act just a year later.²⁷ In response to *Stratton Oakmont*, Congress addressed the question of internet liability through a proposed amendment to the CDA.²⁸

The amendment was intended to encourage free speech online by shielding interactive computer services from most common-law defamation liability.²⁹ The amendment was formalized as Section 230 of the CDA, which effectively grants internet service providers immunity for information provided by a third party, thus treating them differently than publishers in print.³⁰ Specifically, the language of Section 230(c)(1) stipulates that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by

23. 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995), *superseded by statute*, 47 U.S.C. § 230.

24. *Id.*; see also *CDA 230: Legislative History*, ELEC. FRONTIER FOUND., <https://www EFF.ORG/ISSUES/cda230/legislative-history> [<https://perma.cc/JR7E-GFZW>].

25. *CDA 230: Legislative History*, *supra* note 24.

26. *Id.*

27. *Id.*

28. *Id.* (“Worried about the future of free speech online and responding directly to *Stratton Oakmont*, Representatives Chris Cox (R-CA) and Ron Wyden (D-OR) introduced an amendment to the Communications Decency Act that would end up becoming Section 230.”); see also Matt Laslo, *The Fight over Section 230—and the Internet as We Know It*, WIRED (Aug. 13, 2019, 3:18 PM), <https://www.wired.com/story/fight-over-section-230-internet-as-we-know-it/> [<https://perma.cc/UM33-VSDC>] (“Section 230 wasn’t crafted to shield companies for the heck of it. The fear was that if companies could be held responsible for *all* the content its users posted simply because they moderated some of it . . . they wouldn’t moderate anything at all.”).

29. *Immunity for Online Publishers*, *supra* note 14.

30. See Alina Selyukh, *Section 230: A Key Legal Shield for Facebook, Google Is About to Change*, NPR (Mar. 21, 2018, 5:11 AM), <https://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change> [<https://perma.cc/QC36-AHWEJ>].

another information content provider.”³¹ This intent was highlighted in the Fourth Circuit’s conclusion that it was Congress’s intent in Section 230 to treat “distributors” differently than “publishers” under the law.³² Other circuits have reliably agreed with the analysis in *Zeran* by refusing to hold internet entities liable for certain content published by third parties.³³

Congress enacted Section 230 at the dawn of the internet age,³⁴ and the statute has had a profound effect on the development of free speech on the internet over the past quarter of a century.³⁵ Section 230 has been referred to as one of the most critical pieces of legislation impacting the “freedom of expression.”³⁶ This is unsurprising as its primary intent was to foster creativity and a competitive market in the internet space.³⁷ Section 230 defines both interactive computer services³⁸ and information content providers.³⁹ The difference between these terms plays a meaningful role in assessing defamation liability.⁴⁰

An interactive computer service is “any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or

31. 47 U.S.C. § 230(c)(1). *See also* *Malwarebytes, Inc. v. Enigma Software Group USA*, 141 S. Ct. 13, 14 (2020) (J. Thomas denying certiorari) (stating that “§230(c)(1) indicates that an Internet provider does not become the publisher of a piece of third-party content—and thus subjected to strict liability— simply by hosting or distributing that content.”).

32. *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330–31 (4th Cir. 1997).

33. *See, e.g., Ben Ezra, Weinstein, & Co. v. Am. Online Inc.*, 206 F.3d 980, 986 (10th Cir. 2000) (“Congress clearly enacted § 230 to forbid the imposition of publisher liability on a service provider for the exercise of its editorial and self-regulatory functions.”) (citing *Zeran*, 129 F.3d 327); *Klayman v. Zuckerberg*, 753 F.3d 1354, 1359 (D.C. Cir. 2014) (citing *Zeran*, 129 F.3d 327).

34. “An umbrella term for the 21st century, in which information travels around the world in seconds and is made available to people in more countries than ever before. It is also a moniker for high-speed communications, the convergence of computers and consumer electronics (CE) and wireless devices.” *Internet Age*, PCMag, <https://www.pcmag.com/encyclopedia/term/internet-age> [<https://perma.cc/LQ4Z-7WXG>].

35. *See* Selyukh, *supra* note 30 (“This 1996 statute became known as ‘a core pillar of Internet freedom’ and ‘the law that gave us modern Internet’—a critical component of free speech online.”).

36. *See* Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2313 (2014) (“Section 230 . . . [has] been among the most important protections of free expression in the United States in the digital age.”); *see also* *Section 230 of the Communications Decency Act*, ELEC. FRONTIER FOUND., <https://www.eff.org/issues/cda230> [<https://perma.cc/3JRQ-GN89>] (referring to Section 230 as “one of the most valuable tools for protecting freedom of expression and innovation on the Internet”).

37. “Congress enacted the Communications Decency Act (CDA) in part to carve out a sphere of immunity from liability for providers of interactive computer services to preserve that ‘vibrant and competitive free market’ of ideas on the Internet.” Claudia G. Catalano, Annotation, *Validity, Construction, and Application of Immunity Provisions of Communications Decency Act*, 47 U.S.C.A. § 230, 52 A.L.R. Fed. 2d 37 (2011).

38. 47 U.S.C. § 230(f)(2).

39. § 230(f)(3) (“The term ‘information content provider’ means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.”).

40. *See* *Zeran*, 129 F.3d at 330 & n.2 (quoting 47 U.S.C. § 230).

services offered by libraries or educational institutions.”⁴¹ Conversely, information content providers may be, and often are, third-party users.⁴² Section 230 effectively grants interactive computer services immunity from defamation liability⁴³ for third-party content present on their websites.⁴⁴ In practice, this means that if a Twitter user posts a video to Twitter, the user (as the information content provider) may be liable for defamation, but Twitter (as the interactive computer service) would not be liable.⁴⁵ However, Section 230 immunity may be forfeited where the interactive computer service curates or censors the content on its site.⁴⁶

II. SOCIAL MEDIA AND THE INFLUENCE OF SECTION 230 OF THE CDA

A. Overview of the Publisher v. Platform Distinction on the Internet

The early 2000s saw the rapid expansion of the internet, bringing people from around the world online.⁴⁷ Entrepreneurs took advantage of the newfound marketplace, and social networking companies like Facebook, Twitter, and YouTube were born.⁴⁸ Ever since, social media use has boomed—both in terms of quantity of users and influence on users.⁴⁹ Social media is defined as “forms of electronic communication (such as websites

41. § 230(f)(2).

42. See § 230(f)(3); cf. VALERIE C. BRANNON, CONG. RSCH. SERV., LSB10306, LIABILITY FOR CONTENT HOSTS: AN OVERVIEW OF THE COMMUNICATION DECENTRY ACT’S SECTION 230 2 (June 6, 2019), <https://crsreports.congress.gov/product/pdf/LSB/LSB10306>.

43. Section 230 does not provide immunity from intellectual property infringement or violations of federal criminal law. See Mark Sullivan, *The 1996 Law That Made the Web Is in the Crosshairs*, FAST CO. (Nov. 29, 2018), <https://www.fastcompany.com/90273352/maybe-its-time-to-take-away-the-outdated-loophole-that-big-tech-exploits> [<https://perma.cc/8FLU-UWZ6>].

44. “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.” § 230(c)(1). See also *Malwarebytes, Inc. v. Enigma Software Group USA*, 141 S. Ct. 13 (2020) (“In short, the statute suggests that if a company unknowingly leaves up illegal third-party content, it is protected from publisher liability by §230(c)(1); and if it takes down certain third-party content in good faith, it is protected by §230(c)(2)(A).”).

45. This immunity is not without certain exceptions. § 230(e) outlines its effect on other laws, explicitly stating that the provision does not have any effect on criminal law, intellectual property law, state law, communications privacy law, or sex-trafficking law. See § 230(e).

46. VALERIE C. BRANNON, CONG. RSCH. SERV., R45650, FREE SPEECH AND THE REGULATION OF SOCIAL MEDIA CONTENT 12 (2019) (citing *Fair Hous. Council v. Roommates.com, LLC*, 521 F.3d 1157, 1170, 1162 (9th Cir. 2008)), <https://fas.org/sgp/crs/misc/R45650.pdf>.

47. See generally *Internet Growth Statistics*, INTERNET WORLD STATS: USAGE & POPULATION STATS., <https://www.internetworldstats.com/emarketing.htm> [<https://perma.cc/8ZA8-BL5F>].

48. See generally Saqib Shah, *The History of Social Networking*, DIGIT. TRENDS (May 14, 2016), <https://www.digitaltrends.com/features/the-history-of-social-networking/> [<https://perma.cc/6TQT-P3C9>].

49. Andrew Perrin & Monica Anderson, *Share of U.S. Adults Using Social Media, Including Facebook, Is Mostly Unchanged Since 2018*, PEW RSCH. CTR. (Apr. 10, 2019), <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/> [<https://perma.cc/S8AE-2E9D>].

for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos).⁵⁰ Thus, websites such as Twitter and Reddit are essentially open forums, or platforms, providing an arena for candid discussion across the spectrum of ideologies and opinions.⁵¹ In line with the intended protections of Section 230, this characterization states that “[t]he Internet and other interactive computer services offer a forum for a true diversity of political discourse, unique opportunities for cultural development, and myriad avenues for intellectual activity.”⁵²

B. Introduction to Shadow-Banning

Shadow-banning is a method of content moderation that is “used to ban people from forums or message boards without alerting them that they’ve been banned.”⁵³ In other words, it is a means for social media platforms to “hide” or “ban” certain users for posting content they deem to be inappropriate for the platform without informing the user or pointing to a specific guideline or rule that was violated. Shadow-banning is not a new idea; in fact, it has been around nearly as long as interactive computer services have existed.⁵⁴ Moderators of chat rooms and message boards, such as those on Reddit, have long been able to disable comments or terminate accounts of “disruptive participants.”⁵⁵ This was not always effective, however, because the users could simply create a new account and continue to participate in the online community.⁵⁶ To solve this problem, moderators

50. *Social Media*, MERRIAM WEBSTER, <https://www.merriam-webster.com/dictionary/social%20media> [<https://perma.cc/DJF2-LAMS>].

51. Distinguishing between “publishers” and “platforms” is not always easy. *See* Malwarebytes, Inc. v. Enigma Software Group USA, 141 S. Ct. 13 (2020) (“To be sure, recognizing some overlap between publishers and distributors is not unheard of. Sources sometimes use language that arguably blurs the distinction between publishers and distributors.”).

52. 47 U.S.C. § 230(a)(3).

53. *See* Samantha Cole, *Where Did the Concept of ‘Shadow Banning’ Come From?*, VICE (July 31, 2018), https://www.vice.com/en_us/article/a3q744/where-did-shadow-banning-come-from-trump-r-publicans-shadowbanned [<https://perma.cc/F92K-696D>]; *see also* G.F., *The Economist Explains: What is ‘Shadowbanning’?*, THE ECONOMIST (Aug. 1, 2018), <https://www.economist.com/the-economist-explains/2018/08/01/what-is-shadowbanning> [<https://perma.cc/2GR6-FEU3>]. The article explains:

Shadowbanned users are not told that they have been affected. They can continue to post messages, add new followers and comment on or reply to other posts. But their messages may not appear in the feed, their replies may be suppressed and they may not show up in searches for their usernames. The only hint that such a thing is happening would be a dip in likes, favourites or retweets—or an ally alerting them to their disappearance.

Id.

54. *See* Barbara Ortutay, *AP Explains: What Is Shadow Banning?*, AP NEWS (Sept. 5, 2018), <https://www.apnews.com/8ee05a6abfe54131874428b0671b1e15> [<https://perma.cc/5XYP-7GUA>].

55. *Id.* (“At Reddit, shadow banning was long the only tool available to moderators.”).

56. *Id.*

began to disable these users secretly, allowing the user to believe he or she was still an active participant, albeit with no engagement from other users.⁵⁷

Politically conservative users have recently alleged they have been shadow-banned and have argued that prominent social media sites like Twitter and YouTube have used the process to suppress conservative talking points in favor of liberal points of view.⁵⁸ Specifically, the allegations are that when a user searches for specific, notable conservative political figures in the search box, the names do not appear.⁵⁹ Of course, this is not necessarily evidence of political censorship. But the belief that companies are engaging in such censorship is prevalent and not solely held by purported victims—a majority of Americans believe that social media companies engage in political censorship.⁶⁰

Despite the pervasive belief that political censorship is occurring, evidence of explicit shadow-banning is scarce. While this may be because reports of shadow-banning are merely rumors, it may also be because it is nearly impossible to collect non-circumstantial evidence without more information from the social media companies.⁶¹ Those who make the argument that shadow-banning is a problem point to changes in the algorithms used by the platforms, which control both the visibility of content and the prioritization of the content shown to all users.⁶² In fact, many that deny the existence of shadow-banning believe that shadow-banning is simply a conspiracy theory born out of the lack of transparency reflected in the algorithms used to determine which content appears on a user's feed.⁶³ Social media companies like Twitter that use algorithms to curate their content acknowledged that the algorithms may result in certain tweets not appearing as frequently as others, which may lead to certain

57. *Id.*

58. See Harper Neidig, *PragerU Sues Google, YouTube for 'Censoring' Conservative Videos*, THE HILL (Oct. 24, 2017, 5:22 PM), <https://thehill.com/policy/technology/356966-prageru-sues-google-youtube-for-censoring-conservative-videos> [<https://perma.cc/MK6Y-ASVU>].

59. G.F., *supra* note 53. See also Philip Bump, *Trump 'Shadow Ban' Tweet: A F.A.Q.*, WASH. POST (July 26, 2018, 7:58 AM), <https://www.washingtonpost.com/news/politics/wp/2018/07/26/trump-shadow-ban-tweet-a-f-a-q/> [<https://perma.cc/RQ7J-T2J8>] (“Normally, if you go to the search bar on Twitter.com and type someone's name, you'll get a list of suggested accounts. Type 'Trump,' for example and, as of writing, you get the president, the first lady and Ivanka Trump, among others. I get other options, too, like @trumphop, a bot I created that retweets old Trump tweets on the same day and time that he originally tweeted them. Which is to say: The results are customized.”).

60. *Roughly Seven-in-Ten Americans Think It Likely That Social Media Platforms Censor Political Viewpoints*, PEW RSCH. CTR. (June 26, 2018), https://www.pewresearch.org/internet/2018/06/28/public-attitudes-toward-technology-companies/pi_2018-06-28_tech-companies_0-01/ [<https://perma.cc/WAUS-2YAZ>].

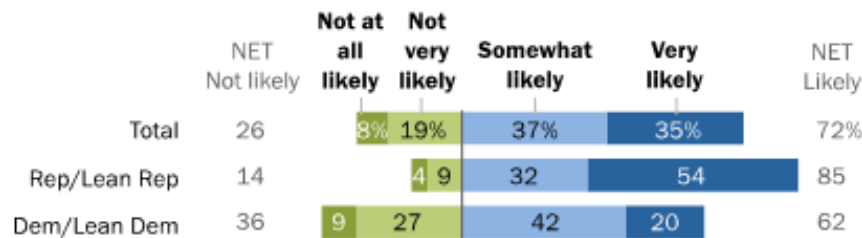
61. *Cf. Personalization Based on Where You See Twitter Content Across the Web*, TWITTER BLOG, <https://help.twitter.com/en/using-twitter/tailored-suggestions> [<https://perma.cc/S7FJ-QSH9>].

62. G.F., *supra* note 53.

63. *Id.*

Roughly seven-in-ten Americans think it likely that social media platforms censor political viewpoints

% of U.S. adults who think it is ___ that social media sites intentionally censor political viewpoints they find objectionable



Note: Respondents who did not give an answer are not shown.

Source: Survey conducted May 29-June 11, 2018.

"Public Attitudes Toward Technology Companies"

PEW RESEARCH CENTER

political commentators and users being unintentionally prioritized over others.⁶⁴

The apprehension over the algorithms results from potential bias in how these entities present search results. Twitter ranks tweets higher if it believes they are relevant to the user or tweets that are popular at the time of search.⁶⁵ Conversely, tweets from "bad-faith actors who intend to manipulate or divide the conversation are ranked lower."⁶⁶ Facebook generates search results based on factors such as pages a user follows, groups the user has joined, events a user has "liked," a user's previous searches, and information available on a user's profile.⁶⁷

Twitter (via founder Jack Dorsey) has recognized that its search-result algorithm has previously biased search results, admitting that the "algorithms were unfairly filtering 600,000 accounts, including some members of Congress, from [their] search auto-complete and latest

64. Cf. Selyukh, *supra* note 30.

65. *Search Result FAQ's*, TWITTER BLOG, <https://help.twitter.com/en/using-twitter/top-search-results-faqs> [<https://perma.cc/R8UK-YHZD>] (the other factors are not discussed in detail).

66. Matt Southern, *Twitter Reveals How It Ranks Tweets in Search Results*, SEARCH ENGINE J. (July 31, 2018), <https://www.searchenginejournal.com/twitter-reveals-how-it-ranks-tweets-in-search-results/263869/> [<https://perma.cc/THT9-W89J>].

67. Matt Southern, *Facebook Explains How Its Search Results Work*, SEARCH ENGINE J. (Nov. 30, 2018), <https://www.searchenginejournal.com/facebook-explains-how-its-search-results-work/281003/#close> [<https://perma.cc/76SD-URG3>].

results.”⁶⁸ Although Mr. Dorsey went on to assure the public that Twitter had taken steps to correct for the partiality in the algorithm, there is currently no legal requirement to do so under Section 230 of the CDA.⁶⁹

III. SOCIAL MEDIA: THE TWENTY-FIRST CENTURY’S “TOWN SQUARE”?

As it stands, social media companies are considered platforms, meaning they are not liable for third-party content. Rather, they provide the opportunity for large numbers of third-party users to exchange information.⁷⁰ Because they do not screen user-generated content before it is posted on the platform, all content moderation is *ex post*.⁷¹ Generally, social media sites limit their content moderation to subjects that clearly fall within the specific categories listed in Section 230,⁷² including posts encouraging violence or terrorism, sexual exploitation, adult content, or illegal activities.⁷³ However, some have begun to argue that given the role that social media plays in society today, it should be subject to the First Amendment’s Free Speech clause.⁷⁴ Despite the First Amendment only limiting governmental actors, advocates of this position focus their argument on the public forum doctrine.⁷⁵

68. Jack Dorsey (@jack), TWITTER (Sept. 5, 2018, 12:56 PM), https://twitter.com/jack/status/1037399156460081152?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwtterm%5E1037399156460081152&ref_url=https%3A%2F%2Fwww.engadget.com%2F2018%2F09%2F05%2Ftwitter-shadow-banning-bug-accounts-jack-dorsey%2F [<https://perma.cc/JQK6-3SL6>].

69. See 47 U.S.C. § 230.

70. John Samples, *Why the Government Should Not Regulate Content Moderation of Social Media*, CATO INST. (Apr. 9, 2019), <https://www.cato.org/publications/policy-analysis/why-government-should-not-regulate-content-moderation-social-media> [<https://perma.cc/QJY9-H5BF>].

71. *Id.* at 3. Social media platforms’ use of *ex post* content moderation is a significant difference from how traditional publishers moderate content. Typical publishers, such as newspapers, conduct all of their content moderation *ex ante*.

72. § 230(c)(2)(A) (stating that no interactive computer service may be held liable to restrict material that the provider or users consider “obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.”).

73. See, e.g., *The Twitter Rules: Safety*, TWITTER, <https://help.twitter.com/en/rules-and-policies/twitter-rules> [<https://perma.cc/78JE-77MP>]; *Community Standards*, FACEBOOK, <https://www.facebook.com/communitystandards/introduction> [<https://perma.cc/KW2J-8E EC>].

74. See generally David L. Hudson Jr., *In the Age of Social Media, Expand the Reach of the First Amendment*, 43 A.B.A. HUM. RTS. MAG. 2, 4 (2019) (“The Court should interpret the First Amendment to limit the ‘unreasonably restrictive and oppressive conduct’ by certain powerful, private entities—such as social media entities—that flagrantly censor freedom of expression.”).

75. *Id.* at 3. However, the Supreme Court has been unwilling to go so far, holding that “[p]roviding some kind of forum for speech is not an activity that only governmental entities have traditionally performed. Therefore, a private entity who provides a forum for speech is not transformed by that fact alone into a state actor.” *Manhattan Cmty Access Corp. v. Halleck*, 139 S. Ct. 1921, 1930 (2019).

A. *The Public Forum Doctrine*

The Supreme Court created the public forum doctrine in 1939 as a way to preserve freedom of expression amongst citizens in public spaces.⁷⁶ The public forum doctrine “prescribes rules limiting the government’s ability to regulate speech in areas created for the purpose of speech”⁷⁷ Traditional public forums have consisted of “places which by long tradition or by government fiat have been devoted to assembly and debate” such as streets, parks, and town squares.⁷⁸ Proponents of applying the public forum doctrine suggest that, given the significant control and power that companies like Twitter and Facebook exhibit over communication today, these companies “are analogous to a governmental actor” in spreading public information.⁷⁹ Proponents conclude, therefore, that the First Amendment requirement to not infringe on free speech ought to apply to such entities because they have become a modern day town square.⁸⁰

In *Packingham v. North Carolina*,⁸¹ the Supreme Court looked favorably on such a comparison. In *Packingham*, the Court appeared receptive to social media platforms being akin to a town square, recognizing social media’s widespread use and its function as a forum to exchange ideas and viewpoints.⁸² The Court found that social media platforms enable individuals to engage in protected First Amendment activity, touching a wide variety of topics and interests.⁸³ When deciding whether a state may

76. Many trace the origins of the public forum doctrine to *Hague v. Comm. for Indus. Org.*, 307 U.S. 496 (1939). In the opinion, Justice Owen Roberts held:

Wherever the title of streets and parks may rest, they have immemorially been held in trust for the use of the public and, time out of mind, have been used for purposes of assembly, communicating thoughts between citizens, and discussing public questions. Such use of the streets and public places has, from ancient times, been a part of the privileges, immunities, rights, and liberties of citizens. The privilege of a citizen of the United States to use the streets and parks for communication of views on national questions may be regulated in the interest of all; it is not absolute, but relative, and must be exercised in subordination to the general comfort and convenience, and in consonance with peace and good order; but it must not, in the guise of regulation, be abridged or denied.

Id. at 515–16.

77. Micah Telegen, *You Can’t Say That!: Public Forum Doctrine and Viewpoint Discrimination in the Social Media Era*, 52 U. MICH J.L. REFORM 235, 238 (2018).

78. *Perry Educ. Ass’n v. Perry Local Educators’ Ass’n*, 460 U.S. 37, 45 (1983).

79. Hudson Jr., *supra* note 74, at 3; *see id.* (“This societal development and change in communications capacities require that the antiquated state action doctrine be modified lest the law become ossified. The time has come to recognize that the reach of the First Amendment be expanded.”).

80. *Id.*

81. 137 S. Ct. 1730 (2017).

82. “While in the past there may have been difficulty in identifying the most important places (in a spatial sense) for the exchange of views, today the answer is clear. It is cyberspace—the ‘vast democratic forums of the Internet’ in general . . . and social media particular. Seven in ten American adults use at least one Internet social networking service.” *Id.* at 1735 (quoting *Reno v. Am. Civ. Liberties Union*, 521 U.S. 844, 868 (1997)).

83. *Packingham*, 137 S. Ct. at 1735.

prevent an individual from accessing social media, the Court held that a statute preventing social media access infringes on an individual's First Amendment rights because "[t]hese websites can provide perhaps the most powerful mechanisms available to a private citizen to make his or her voice heard."⁸⁴

The *Packingham* decision illustrates that access to social media is not a luxury, but is critical to furthering public discourse—thereby supporting the idea that social media is closer than ever before to a “public forum.”⁸⁵ This idea is continuing to gain legal support. For example, the Second Circuit recently held that President Trump could not block Twitter users from reading his tweets because he uses Twitter⁸⁶ for government business.⁸⁷ The court suggested that debate between government officials and political opposites is a necessary and vital part of open online discussion—further fueling the notion that social media is a public forum, serving as the new town square.⁸⁸

B. The Future of Section 230

Washington has turned its focus to Section 230 as social media has grown more prevalent, both politically and more generally. The recent debate regarding the future of social media regulation has led to many legislators proposing new regulations for “Big Tech.”⁸⁹ Proponents of

84. *Id.* at 1737; *see id.* (further explaining “to foreclose access to social media altogether is to prevent the user from engaging in the legitimate exercise of First Amendment rights. . . . Even convicted criminals . . . might receive legitimate benefits from these means for access to the world of ideas, in particular if they seek to reform and to pursue lawful and rewarding lives.”).

85. “To determine whether a public forum has been created, courts look ‘to the policy and practice of the government’ as well as ‘the nature of the property and its compatibility with expressive activity to discern the government’s intent.’” *Knight First Amend. Inst. at Columbia Univ. v. Trump*, 928 F.3d 226, 237 (2d Cir. 2019) (quoting *Cornelius v. NAACP Legal Def. & Educ. Fund, Inc.*, 473 U.S. 788, 802 (1985); *see also id.* (“Opening an instrumentality of communication ‘for indiscriminate use by the general public’ creates a public forum.”) (quoting *Perry Educ. Ass’n v. Perry Local Educators’ Ass’n*, 460 U.S. 37, 47 (1983)).

86. *Id.* at 237 (explaining President Trump’s Twitter account “was intentionally opened for public discussion when the President, upon assuming office, repeatedly used the Account as an official vehicle for governance and made its interactive features accessible to the public without limitation” and holding that “this conduct created a public forum”).

87. Charlie Savage, *Trump Can’t Block Critics from His Twitter Account, Appeals Court Rules*, N.Y. TIMES (July 9, 2019), <https://www.nytimes.com/2019/07/09/us/politics/trump-twitter-first-amendment.html> [<https://perma.cc/GM33-R3QE>].

88. *See generally Knight*, 928 F.3d at 226.

89. VALERIE C. BRANNON, CONG. RSCH. SERV., LSB10309, REGULATING BIG TECH: LEGAL IMPLICATIONS 1 (Sept. 11, 2019), <https://fas.org/sgp/crs/misc/LSB10309.pdf> [<https://perma.cc/5W9R-SDVC>]. “Big Tech” often refers to five tech companies: Alphabet, Amazon, Apple, Facebook, and Microsoft. *Id.* However, many of the pieces of pending legislation seem to define “big tech” in terms of monthly users, such as companies “with more than 30 million active monthly users in the U.S., more than 300 million active monthly users worldwide, or who have more than \$500 million in global annual

amending Section 230 correctly point out that at the time of its enactment, the internet was nothing like the internet we have today.⁹⁰ While the complaints concerning shadow-banning have predominantly come from Republicans, amending Section 230 has drawn attention from both sides of the aisle.⁹¹ Conservatives have argued that social media companies violate “the spirit of the law” by censoring conservative users.⁹² Liberals have argued that social media companies have done a poor job taking down “problematic content” and “tackling harassment” due to Section 230 protections.⁹³ As it stands, there are three prominent pieces of proposed legislation focused on amending Section 230: (1) The “Ending Support for Internet Censorship Act,”⁹⁴ (2) the “Biased Algorithm Deterrence Act,”⁹⁵ and (3) the “Algorithmic Accountability Act.”⁹⁶ The first two suggest that social media companies ought to be required to relax their content-moderation tactics; the third suggests these companies ought to be able to moderate further.

IV. PROPOSED LEGISLATION IMPACTING SOCIAL MEDIA PLATFORMS

The “Ending Support for Internet Censorship Act,” proposed by Senator Josh Hawley (R-MO) in June of 2019, would allow social media platforms to be sued “unless companies submit to an external audit by the Federal Trade Commission [FTC] that proves by ‘clear and convincing evidence’ that they do not moderate content ‘in a politically biased manner.’”⁹⁷ The bill has three primary objectives: (1) remove automatic immunity from “big

revenue.” See *Senator Hawley Introduces Legislation to Amend Section 230 Immunity for Big Tech Companies*, JOSH HAWLEY, U.S. SENATOR FOR MISSOURI (June 19, 2019), <https://www.hawley.senate.gov/senator-hawley-introduces-legislation-amend-section-230-immunity-big-tech-companies> [<https://perma.cc/28JQ-5ZSR>].

90. See Laslo, *supra* note 28 (quoting Senator Hawley, “The world has changed. The Internet has changed, and I think we need to keep pace with change. The dominant, monopoly-sized platforms that exist today didn’t exist then. The business model they employ today wasn’t employed then.”).

91. *Id.* (“In Congress, both parties have singled Section 230 out for attack, with some Democrats saying it allows tech companies to get away with not moderating content enough, while some Republicans say it enables them to moderate too much.”). Importantly, Section 230 is facing scrutiny for other reasons as well, including its application to fake news. See, e.g., Andrea Butler, Note, *Protecting the Democratic Role of the Press: A Legal Solution to Fake News*, 96 WASH. U. L. REV. 419 (2018).

92. Daisuke Wakabayashi, *Legal Shield for Social Media Is Targeted by Trump*, N.Y. TIMES (May 28, 2020), <https://www.nytimes.com/2020/05/28/business/section-230-internet-speech.html> [<https://perma.cc/W5BA-6DZL>].

93. *Id.*

94. Ending Support for Internet Censorship Act, S. 1914, 116th Cong. (2019).

95. Biased Algorithm Deterrence Act of 2019, H.R. 492, 116th Cong. (2019).

96. Algorithmic Accountability Act, S. 1108, 116th Cong. (2019).

97. James Pethokoukis, *Josh Hawley’s New Tech Bill Would Blow Up the Internet*, THE WEEK (June 20, 2019), <https://theweek.com/articles/848015/josh-hawleys-new-tech-bill-blow-internet> [<https://perma.cc/KL9N-9A7D>].

tech companies”; (2) allow these big tech companies the ability to regain that immunity by submitting to an FTC audit; and (3) preserve the currently existing immunity for small and mid-size tech companies.⁹⁸

The “Biased Algorithm Deterrent Act of 2019,” proposed by Rep. Louie Gohmert (R-TX), “would provide that social media services ‘shall be treated as a publisher or speaker’ of certain user-generated content if they display it ‘in an order other than chronological order.’”⁹⁹ Further, the bill would hold any social media company liable as a publisher of user-generated content if a company or its algorithm delays the display of some content and not others, or if it “hinders the display of such content” for reasons other than to carry out the user’s direction or to restrict material for reasons not explicitly listed in Section 230 of the CDA.¹⁰⁰

Each of these two pieces of proposed legislation seeks to curtail the immunity offered by Section 230 drastically. The “Ending Support for Internet Censorship Act” essentially requires that big tech companies prove to the government that they should be afforded Section 230 immunity by requiring them to demonstrate that they are politically neutral platforms.¹⁰¹ Many have expressed concerns with the bill, specifically citing its failure to narrowly define “political viewpoint,” as well as its implication that all political positions and groups are “equally legitimate.”¹⁰² The “Biased Algorithm Deterrent Act of 2019” goes even further, stipulating that platforms that seek to be afforded Section 230 immunity may only organize and present their content in one way.¹⁰³ In either case, the implications would be broad, as they may disincentivize entrepreneurs from entering the social media space altogether.

The “Algorithmic Accountability Act,” proposed by Senators Cory Booker (D-NJ) and Ron Wyden (D-OR), would have the FTC enact rules which analyze “‘highly sensitive’ automated systems,” primarily evaluating algorithmic tools for discrimination, bias, and privacy risks.¹⁰⁴ This bill principally focuses on increasing transparency in the decision-making process of interactive computer service’s algorithms and how they present

98. See Senator Hawley Introduces Legislation to Amend Section 230 Immunity for Big Tech Companies, *supra* note 89.

99. BRANNON, *supra* note 89, at 3.

100. H.R. 492.

101. See Ending Support for Internet Censorship Act, S. 1914, 116th Cong. (2019).

102. Eric Goldman, *Comments on Sen. Hawley’s “[Ending] Support for Internet Censorship Act”*, TECH. & MARKETING L. BLOG (July 10, 2019), <https://blog.ericgoldman.org/archives/2019/07/comment-s-on-sen-hawleys-ending-support-for-internet-censorship-act.htm> [<https://perma.cc/BJ3A-JURG>].

103. See H.R. 492.

104. Adi Robertson, *A New Bill Would Force Companies to Check Their Algorithms for Bias*, THE VERGE (Apr. 10, 2019, 3:52 PM), <https://www.theverge.com/2019/4/10/18304960/congress-algorithmic-accountability-act-wyden-clarke-booker-bill-introduced-house-senate>. [<https://perma.cc/493M-NHGY>]. The bill is “aimed at major companies with access to large amounts of information. It would apply to companies that . . . primarily act as data brokers that buy and sell consumer data.” *Id.*

information regarding “race, color, national origin, political opinions, religion” and more.¹⁰⁵

Unlike the previously discussed bills, the “Algorithmic Accountability Act” focuses primarily on the unintentional discriminatory decisions made by a social media company’s algorithm.¹⁰⁶ This legislation focuses less on perceived political biases and more on “tech practices that lead to ‘houses that you never know are for sale, job opportunities that never present themselves, and financing that you never become aware of.’”¹⁰⁷ While the focus is not the same, the underlying premise is similar—that social media companies must be held accountable for failing to maintain neutrality with respect to how they decide to present user-generated content.¹⁰⁸ Each of these pieces of legislation would add regulation to internet services already subject to Section 230, and if enacted, would vastly change the internet as we know it.

V. HOW SOCIAL MEDIA COMPANIES CAN REFORM CONTENT MODERATION PRACTICES AND MAINTAIN “PLATFORM” LIABILITY

Neither extending the First Amendment protections to the realm of social media¹⁰⁹ nor enacting any of the aforementioned pieces of legislation¹¹⁰ is necessary (nor preferred) in order to reach the outcome that many seek. While it is clear that there may be a time where Section 230 will need to be revisited,¹¹¹ there are some interim solutions that social media companies may put in place to balance their Section 230 immunity with the growing concerns of content-moderation practices. It is worth noting that courts have

105. Algorithmic Accountability Act of 2019, S. 1108, 116th Cong. (2019); *see also* Jerry Barbanel, *A Look at the Proposed Algorithmic Accountability Act of 2019*, IAPP (Apr. 29, 2019), <https://iapp.org/news/a/a-look-at-the-proposed-algorithmic-accountability-act-of-2019/> [<https://perma.cc/SD58-2V8X>].

106. *See* Taylor Hatmaker, *Democrats Draw Up Bill That Would Require Tech Platforms to Assess Algorithmic Bias*, TECHCRUNCH (Apr. 10, 2019, 5:26 PM), <https://techcrunch.com/2019/04/10/algorithmic-accountability-act/> [<https://perma.cc/L2PL-64N9>] (“By requiring large companies to not turn a blind eye toward[] unintended impacts of their automated systems, the Algorithmic Accountability Act ensures 21st Century technologies are tools of empowerment, rather than marginalization, while also bolstering the security and privacy of all consumers,” Rep. Clarke said.”).

107. *Id.* (quoting Sen. Cory Booker).

108. *See* S. 1108.

109. *Supra* Part III.

110. *Supra* Part IV.

111. Some argue that the “political will currently exists to make changes to the ways in which online platforms are held accountable for the content posted on their sites” Frank Ready, *Timing Is Ripe for Section 230 Amendments—But the ‘How’ Is Missing*, LAW.COM (Oct. 24, 2019, 11:00 AM), <https://www.law.com/legaltechnews/2019/10/24/timing-is-ripe-for-section-230-amendments-but-the-how-is-missing/> [<https://perma.cc/UW9P-5YSR>]. With rising concern over fake news, censorship, and hate speech on online platforms, some think we are more likely to see changes to Section 230 sooner rather than later. *See id.*

“invok[ed] a broad interpretation of immunity”¹¹² and have generally not made much of a distinction between social media companies which exercise editorial discretion and those that do not.¹¹³ However, there are signs that this distinction could eventually be made. Notably, the Seventh Circuit has long questioned the practical effects of failing to hold websites and ISPs which exercise editorial control to a different standard than those that do not.¹¹⁴ If courts do seek to make a distinction, social media companies should take three key steps to maintain their immunity from defamation liability.

A. Establish Clear Guidelines for Content That Is Appropriate for Their Platform

Social media companies must establish consistent and clear guidelines for appropriate (and inappropriate) content to minimize user confusion and outrage. Most of these companies have vague and broad guidelines setting forth what content is allowed,¹¹⁵ but the guidelines may be applied subjectively by a small group of employees.¹¹⁶ Additionally, content moderation relies predominately on a concept called “community policing,” where users report other users for violations—or perceived violations—of the content guidelines.¹¹⁷ Concerns have been raised that Twitter’s content guidelines, for example, may have a disparate impact on some viewpoints

112. See *Malwarebytes, Inc. v. Enigma Software Group USA*, 141 S. Ct. 13 (2020) (explaining that, with respect to §230 “many courts have construed the law broadly to confer sweeping immunity on some of the largest companies in the world.”); see also *Wakabayashi*, *supra* note 92; see *id.* (further explaining that “Section 230 has also provided legal cover for the complicated decisions regarding content moderation. Facebook and Twitter have recently cited it to defend themselves in court when users have sued after being barred from the platforms.”).

113. Butler, *supra* note 91, at 434 (“[A]n interpretation of § 230 that treats websites and ISPs exercising editorial control the same as those that do not defeats the original policy goals of the ‘Good Samaritan’ law and likely serves few of the purposes Congress intended.”).

114. *Doe v. GTE Corp.*, 347 F.3d 655, 660 (7th Cir. 2003). The *Doe* court discussed that the principal purpose of § 230 was to afford immunity to websites and ISPs that exercised editorial discretion to remove “offensive material,” highlighting the section entitled “Protection for ‘Good Samaritan’ blocking and screening of offensive material.” *Id.* In other words, the broad immunity afforded to these companies was to encourage the removal of offensive content, not simply whatever content they wanted to remove.

115. See, e.g., Dieter Bohn, *One of Twitter’s New Anti-Abuse Measures Is the Oldest Trick in the Forum Moderation Book*, THE VERGE (Feb. 16, 2017, 9:07 AM), <https://www.theverge.com/2017/2/16/14635030/twitter-shadow-ban-moderation> (“Twitter itself has been intentionally vague about what its precise policies and tools are . . .”).

116. Jillian C. York & Corynne McSherry, *Content Moderation Is Broken. Let Us Count the Ways*, ELEC. FRONTIER FOUND. (Apr. 29, 2019), <https://www.eff.org/deeplinks/2019/04/content-moderation-broken-let-us-count-ways> [<https://perma.cc/B4HA-VH7M>] (explaining that throughout the history of social media, guidelines have been applied subjectively by employees).

117. *Id.*; see also *The Joe Rogan Experience, #1258 – Jack Dorsey, Vijaya Gadde & Tim Pool*, (March 5, 2019) (downloaded using iTunes) (explaining that a certain potentially offensive and violent video may not have been removed because it may not have been reported).

as opposed to others based on the likelihood of certain people to report material as offensive.¹¹⁸ Ultimately, it is likely that social media companies use information gathered from community policing in different ways, though in the past companies like Facebook have kept this process close to the chest, revealing little information to the public about this procedure.¹¹⁹

Companies have attempted to make the moderation process more consistent and less subjective by replacing much of the human component of content moderation with the aforementioned algorithms and other artificial intelligence.¹²⁰ While this may eliminate the human biases present when a person reviews content, concerns of bias and inconsistent results largely persist, as previously discussed.¹²¹ Ultimately, social media companies must strive to develop specific guidelines that may be applied consistently, irrespective of political preference or cultural view.¹²² One suggested solution is to tie content guidelines to the UN Guiding Principles on Business and Human Rights.¹²³ In doing so, companies could rely on a generally agreed upon set of guidelines which “provide a standard . . . for companies to prevent and address the risk of adverse impacts on human rights.”¹²⁴ This idea would establish clear guidelines, but private businesses may reject the idea of tying themselves to such a standard by arguing it to be unworkable and overbroad. However, the idea to build off an existing set of guidelines would certainly help ensure that content-moderation decisions are consistent.¹²⁵ A consistent process will push the content-moderation practices closer to the “platform” category rather than the “publisher” category. Adopting a clear and agreed-upon set of rules that may be applied across all content regardless of viewpoint or user mitigates opportunity for editorial decision-making—a critical part of being a “publisher.”¹²⁶

118. See *The Joe Rogan Experience*, *supra* note 117 (explaining that some users have been banned for misgendering a transgender person in an attempt to assert their opinion on biology and gender).

119. Jillian C. York, *Policing Content in the Quasi-Public Sphere*, OPENNET INITIATIVE (Sept. 2010), <https://opennet.net/policing-content-quasi-public-sphere> [<https://perma.cc/8SKM-L7GJ>] (“Facebook does, however, offer a function through which users may report one another with the simple click of a button. The company has not spoken publicly about how this process works, but one hypothesis is that when a critical mass of users reports a profile, that profile is automatically disabled, possibly for later review by a staff member.”).

120. Kalev Leetaru, *The Problem with AI-Powered Content Moderation Is Incentives Not Technology*, FORBES (Mar. 19, 2019, 2:34 PM), <https://www.forbes.com/sites/kalevleetaru/2019/03/19/the-problem-with-ai-powered-content-moderation-is-incentives-not-technology/#639da14955b7> [<https://perma.cc/96RH-2HJW>].

121. See also Gadde & Beykpour, *supra* note 3.

122. See Jillian C. York, *UN Report Sets Forth Strong Recommendations for Companies to Protect Free Expression*, ELEC. FRONTIER FOUND. (June 27, 2018), <https://www EFF.org/deeplinks/2018/06/un-report-sets-forth-strong-recommendations-companies-protect-free-expression> [<https://perma.cc/99FM-HJK8>].

123. *Id.*

124. *Id.*

125. *Id.*

126. See generally *Immunity for Online Publishers*, *supra* note 14. See also *supra* Part II.a.

B. Provide Initial Warnings and Clear Explanations When a Post Is Removed, or a User Is Banned

In a perfect world, there would be no need for content moderation. But we do not live in a perfect world; to truly foster open discussion, specific rules are required. Since rules are bound to be broken, social media companies should have detailed procedures to handle such violations. In conjunction with a detailed and defined set of rules for appropriate content, these companies should have a precise order of operations to inform users when a post or a profile has been suspended, removed, or terminated. YouTube's copyright strike-system is one example of a relatively transparent process that informs users of their violations, explains the violations, and offers a warning for their first mistake.¹²⁷

This process should inform users immediately when a post or a profile is flagged or when visibility parameters are changed. Currently, this is not always the case, and the lack of transparency in this regard is a significant reason for the recent outcry against companies like Twitter.¹²⁸ In late 2018, the Associated Press detailed Twitter's process to maintain the "health" of the platform as follows:

Twitter outlined a new approach¹²⁹ intended to reduce the impact of disruptive users, or trolls, by reading "behavioral signals" that tend to indicate when users are more interested in blowing up conversations than in contributing. . . . While accounts flagged this way don't technically violate Twitter policy, the company now wants to protect the "health" of users' online conversations. (That word is now a staple in the company's lexicon; CEO Jack Dorsey used "health," "healthy" or "unhealthy" 31 times in prepared congressional testimony Tuesday.)¹³⁰ So Twitter will reduce their

127. *Copyright Strike Basics*, YOUTUBE HELP, <https://support.google.com/youtube/answer/2814000?hl=en> [<https://perma.cc/SF2P-TVURU>].

128. See *Ortutay*, *supra* note 54. But see Sara E. Needleman, *Facebook Says It Is Removing All Content Mentioning 'Stop the Steal'*, WALL STREET J. (Jan. 11, 2021), <https://www.wsj.com/articles/facebook-says-it-is-removing-all-content-mentioning-stop-the-steal-11610401305> [<https://perma.cc/3ASA-XXWG>] (Facebook decided to remove "all content mentioning 'stop the steal,' a phrase popular among supporters of President Trump's claims about the election, as part of a raft of emergency measures to stem misinformation and incitements to violence on its platform. . . .").

129. Del Harvey & David Gasca, *Serving Healthy Conversation*, TWITTER BLOG (May 15, 2018), https://blog.twitter.com/official/en_us/topics/product/2018/Serving_Healthy_Conversation.html [<https://perma.cc/YQL4-KMB6>].

130. *Twitter: Transparency and Accountability: Hearing Before the H. Comm. on Energy & Com.*, 115th Cong. (2018) (statement of Jack Dorsey, Chief Executive Officer of Twitter, Inc.), <https://docs.house.gov/meetings/IF/IF00/20180905/108642/HHRG-115-IF00-Wstate-DorseyJ-20180905.pdf> [<https://perma.cc/895M-EPDS>].

visibility in certain ways, by displaying them less prominently in search results or conversation threads.¹³¹

Importantly, when rolling out this policy, Twitter did not inform the impacted users—or at least some of the impacted users—of the changes made to their profile visibility.¹³² Restricting content and user visibility based on grounds outside content guidelines is not technically shadow-banning.¹³³ However, the increased editorialization may push social media companies closer to the “publisher” category than they were before.¹³⁴ Instead, companies should seek to clearly define categories such as “healthy” and “unhealthy” and include them in their content guidelines. Defined content guidelines will help social media companies make consistent content-moderation decisions, while also providing users more insight into content that may be flagged or removed.

Next, social media companies should consider providing clear explanations to affected users regarding the changes to their posts or accounts, including the specific policy that was violated. This policy will reinforce objective content-moderation decisions, as it suggests that companies are applying rules mechanically rather than editorializing. Proactively citing a violated rule when issuing a warning or taking stronger action suggests the decision was decided objectively rather than subjectively—an important distinction in the “publisher v. platform” discussion.¹³⁵

C. Err on the Side of Caution When Deciding to Remove a User or a Post

Platforms do not face defamation liability under the common law or Section 230 because they use little to no editorial discretion.¹³⁶ It is true that Section 230 has granted some editorial discretion to companies such as Twitter and Facebook, though its grant is limited and targeted in nature, specific to offensive material.¹³⁷ To maintain a clear distinction from publishers, these companies should sparingly ban or hide users. Whether content is flagged as a result of a few unhappy users or as a result of an

131. Ortutay, *supra* note 54.

132. Bump, *supra* note 59.

133. As previously mentioned, a user is shadow-banned when the shadow-banned user is able to post normally, but no other user on the platform is able to see their account or their posts. *See supra* Part II.b.

134. *See supra* Part I.a.

135. *Id.* As mentioned, the one key distinction between a publisher and a platform is that platforms do not subjectively curate their content, whereas publishers do.

136. *See* Laurenson, *supra* note 21; *see also supra* Part I.a.

137. 47 U.S.C. § 230(c) (entitled “Protection for ‘Good Samaritan’ blocking and screening of offensive material”).

outrage mob,¹³⁸ social media companies should err on the side of the user when possible to ensure they maintain their treatment as platforms, relying on and communicating specific evidence when making final decisions about removing content or banning users.

Evidence-based policymaking—making content-moderation decisions based on factual data rather than general user unrest or discontent—is a step in the right direction because it demonstrates an objective, rather than subjective, approach to content moderation.¹³⁹ The less subjective content-moderation decisions are, the less an entity looks like a publisher. Employing evidence-based policymaking for content moderation ensures that social media companies will wait to make decisions on a post or a user until they have determined all the necessary facts (through the help of researchers and experts), preventing the rash (and often incorrect) decisions that frequently arise during a period of social media outrage.¹⁴⁰ An additional step in prioritizing objective content-moderation decisions is to implement an appeals process. York and McSherry propose such a process, advocating that when content is flagged for violating content community standards, “absent exigent circumstances companies must notify the user and give them an opportunity to appeal before the content is taken down. If they choose to appeal, the content should stay up until the question is resolved.”¹⁴¹ The more precautions taken by Twitter, Facebook, and others before deciding to suspend or remove a user, the less they appear to be a publisher using editorial discretion, and the more they appear to be a platform merely enforcing rules objectively.

138. “Outrage mobs” occur when a large number of people become angry at an individual, a group, or an idea and aggressively target those individuals, groups, or ideas on social media. See Erik Kain, *Internet Mob Justice Isn't Justice at All*, FORBES (May 14, 2015, 4:51 PM), <https://www.forbes.com/sites/erikkain/2015/05/14/internet-mob-justice-isnt-justice-at-all/#2037e0eb616b> [<https://perma.cc/U5GY-BMRS>]; Brooke A. Rogers, *We've Got to Stop Letting 'Outrage Mobs' Get People Fired*, N.Y. POST (Aug. 5, 2018, 6:16 PM), <https://nypost.com/2018/08/05/weve-got-to-stop-letting-outrage-mobs-get-people-fired/> [<https://perma.cc/Q4TY-RYMD>]; Mark Molloy, *Online Shaming: The Dangerous Rise of the Internet Pitchfork Mob*, TEL. (June 25, 2018, 4:39 PM), <https://www.telegraph.co.uk/news/2018/06/25/online-shaming-dangerous-rise-internet-pitchfork-mob/>. “Outrage mobs” are often considered a part of the broader “cancel culture” as well. See Brooke Kato, *What is Cancel Culture? Everything To Know About the Toxic Online Trend*, N.Y. POST (July 10, 2020, 4:44 PM), <https://nypost.com/article/wh-at-is-cancel-culture-breaking-down-the-toxic-online-trend/> [<https://perma.cc/CX6K-KAQJ>] (describing “cancel culture” as “the phenomenon of promoting the ‘canceling’ of people, brands and even shows and movies due to what some consider to be offensive or problematic remarks or ideologies”).

139. See York & McSherry, *supra* note 116 (encouraging social media entities to utilize experts and researchers when making content-moderation decisions).

140. *See id.*

141. *Id.*

VI. LEGAL SOLUTIONS TO ENSURE TRANSPARENCY IN CONTENT MODERATION

The internet in 2020 is vastly different from the internet in 1996.¹⁴² When Section 230 was initially passed nearly twenty-five years ago, the internet was in its infancy, and many websites were simply “digital versions of real-world things.”¹⁴³ As the internet expanded, business models changed to create a more personalized experience for each user.¹⁴⁴ To foster further engagement with users, “internet companies began ‘personalizing’ their sites so that each user would have a different and unique experience” that was crafted to specifically engage that user.¹⁴⁵ “Websites became highly curated experiences served up by algorithms” that used the user’s browsing history and personal data to create this custom experience.¹⁴⁶

Section 230 protects internet providers from defamation liability for “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”¹⁴⁷ This provision is critical because without it, social media companies would be treated like platforms in the truest sense, liable for any curation of their content.¹⁴⁸ If not for this provision, social media companies would be incentivized to allow any and all content on their platform so as to avoid defamation liability.¹⁴⁹ With the evolution of internet platforms, it is time to revisit Section 230 to ensure we are honoring the intent of the legislation and to confirm that social media companies are not capitalizing on what has been called an “outdated loophole.”¹⁵⁰

It is difficult to say that lawmakers could have predicted the creation of social media when passing this legislation—much less its explosive growth

142. Reuben Fischer-Baum, *What ‘Tech World’ Did You Grow Up In?*, WASH. POST (Nov. 26, 2017), <https://www.washingtonpost.com/graphics/2017/entertainment/tech-generations/> [<https://perma.cc/FL3G-EATS>] (comparing year-over-year statistics on internet usage and consumption of digital media).

143. Sullivan, *supra* note 43.

144. *Id.*

145. *Id.*

146. *Id.*

147. 47 U.S.C. § 230(c)(2)(A) (emphasis added).

148. See Elizabeth Nolan Brown, *Section 230 Is the Internet’s First Amendment. Now Both Republicans and Democrats Want to Take It Away*, REASON (July 29, 2018, 8:01 AM), <https://reason.com/2019/07/29/section-230-is-the-internets-first-amendment-now-both-republicans-and-democrats-want-to-take-it-away/> [<https://perma.cc/96TH-5DZC>].

149. *Id.*

150. Sullivan, *supra* note 43 (quoting former White House technology staffer and long-time Google employee Mikey Dickerson).

in recent years. However, we do have an indication of their intent for Section 230 and the CDA. In its preamble, Section 230 states:

The Congress finds the following:

- (1) The rapidly developing array of Internet and other interactive computer services available to individual Americans represent an extraordinary advance in the availability of educational and informational resources to our citizens.
- (2) These services offer users a great degree of control over the information that they receive, as well as the potential for even greater control in the future as technology develops.
- (3) The Internet and other interactive computer services offer a forum for a *true diversity of political discourse, unique opportunities for cultural development, and myriad avenues for intellectual activity.*
- (4) The Internet and other interactive computer services have flourished, to the benefit of all Americans, with a minimum of government regulation.
- (5) Increasingly Americans are relying on interactive media for a variety of *political, educational, cultural, and entertainment services.*¹⁵¹

It is clear from the preamble that the lawmakers anticipated, to some degree, the significant role that the internet would play in the American discourse. It also appears clear that they sought to allow the internet to grow naturally, with minimal government intrusion.¹⁵² What may not have been clear is just how pivotal social media would become; today, social media executives wield power that potentially shapes public discourse. Though some of the solutions may be internal to the companies themselves,¹⁵³ the immense importance of social media for political and social discourse in the United States requires legal solutions as well.

A. A Non-Partisan Government Watchdog Should Be Employed to Evaluate Potential Disparate Impact in Content Moderation

Senator Hawley's proposed legislation requires that social media companies submit themselves to an FTC audit and clearly demonstrate that they are politically neutral in order to earn immunity.¹⁵⁴ However, complete

151. § 230(a) (emphasis added).

152. See § 230(b)(2) (explaining "[i]t is the policy of the United States . . . to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation"); cf. BRANNON, *supra* note 42, at 4.

153. See *supra* Part V.

154. *Senator Hawley Introduces Legislation to Amend Section 230 Immunity for Big Tech Companies*, *supra* note 89.

neutrality is not a reasonable expectation given the volume of users and posts,¹⁵⁵ and the current inability to effectively moderate content completely using algorithms.¹⁵⁶ Furthermore, political leaders on both sides of the aisle may be wary of enacting any sort of restriction or limitation on the private companies operating these social media sites.¹⁵⁷ However, an oversight committee charged with reviewing and rating algorithms and content-moderation practices of social media companies may be a useful measure. As mentioned, a majority of Americans believe that social media companies engage in political censorship.¹⁵⁸ Requiring online platforms with a particular user volume (like the “Big Tech”¹⁵⁹ social media companies, for example) to submit to a review and rating by a non-partisan oversight group may help both to encourage social media entities –potentially including web-hosting services such Amazon Web Services (AWS)¹⁶⁰ –to minimize biased editorial discretion and to make transparent the facts about whether a bias exists.¹⁶¹ The goal of this proposal is not to require neutrality, but to provide consumers with information on how each company handles politically or culturally sensitive issues. As mentioned, the factual basis for intentional “shadow-banning” is sparse,¹⁶² though public-opinion data

155. “Facebook, the largest social media platform in the world, has 2.4 billion users. Other social media platforms including Youtube and Whatsapp also have more than one billion users each.” Esteban Ortiz-Ospina, *The Rise of Social Media*, OUR WORLD IN DATA (September 18, 2019), <https://ourworldindata.org/rise-of-social-media> [<https://perma.cc/2GZT-G75C>]. With this volume of users, it is impractical to expect complete neutrality without an upheaval of the current business model, though it is reasonable to expect social media entities to apply content moderation rules and technologies in a consistent and generally fair manner.

156. See Megan Rose Dickey, *Unbiased Algorithms Can Still Be Problematic*, TECHCRUNCH (Sept. 30, 2018, 3:45 PM), <https://techcrunch.com/2018/09/30/unbiased-algorithms-can-still-be-problematic/> [<https://perma.cc/7LC5-5DWK>].

157. John Morris, who handled internet policy issues at the Commerce Department’s National Telecommunications and Information Administration, said “[a]lthough the government may be able to support and assist online platforms’ efforts to reduce hate and violence online, the government should not try to impose speech regulations on private platforms. As politicians from both sides of the political spectrum have historically urged, the government should not be in the business of regulating speech.” Margaret Harding McGill & Daniel Lippman, *White House Drafting Executive Order to Tackle Silicon Valley’s Alleged Anti-Conservative Bias*, POLITICO (Aug. 7, 2019, 3:07 PM), <https://www.politico.com/story/2019/08/07/white-house-tech-censorship-1639051> [<https://perma.cc/N7W9-GVPK>].

158. *Roughly Seven-in-Ten Americans Think It Likely That Social Media Platforms Censor Political Viewpoints*, *supra* note 60.

159. BRANNON, *supra* note 46, at 3.

160. Following the attack on the Capitol on January 6, 2021, Parler, a social-media platform largely viewed as politically conservative, was removed on both the Apple and Google app stores. Furthermore, Amazon banned Parler from its web-hosting service AWS because Amazon “ wasn’t confident in [Parler’s] ability to sufficiently police content on its platform that incites violence.” See Keach Hagey, *Parler Sues Amazon After Tech Giant Kicks Site Off Its Servers*, WALL STREET J. (Jan. 11, 2021), <https://www.wsj.com/articles/parler-sues-amazon-kicks-site-off-its-servers-11610363052> [<https://perma.cc/8H6P-BYZF>].

161. This proposal combines elements from both the “Ending Support for Internet Censorship Act” and the “Algorithmic Accountability Act.” See *supra* Part IV.

162. See *supra* note 11 and accompanying text.

suggests the broader population seems to believe it is happening.¹⁶³ Amending Section 230 to include a required review by a non-partisan government agency provides a minimally intrusive solution that may incentivize consistent content-moderation practices and also provide clarity about content-moderation processes for users.

B. An Executive Order Requiring Social Media Companies to Limit Content Moderation to the Precise Language in Section 230

An alternative to formally amending Section 230 is to issue an executive order reinforcing the original language of Section 230. The idea of issuing an executive order to reaffirm an existing law is not new—President Trump issued an executive order that reinforced universities’ requirement to promote free speech on their campuses in order to receive federal research funding.¹⁶⁴ President Obama similarly used an executive order expand upon an executive order¹⁶⁵ enacted by President Lyndon Johnson to prohibit “federal contractors from discriminating against workers on the basis of sexual orientation or gender identity.”¹⁶⁶ While an executive order would not fully address the issue, it would serve as an effective intermediary while legislators, regulators, and company executives continue to determine the future of Section 230 and the liability of social media companies for user content.¹⁶⁷ The White House began to discuss this idea after the initial reports of shadow-banning.¹⁶⁸ In a proposed executive order entitled “Protecting Americans from Online Censorship,” the White House seeks to empower the Federal Communications Commission (FCC) to monitor the content-moderation policies of social media entities, similar to Senator Hawley’s proposed legislation.¹⁶⁹ The proposed order primarily takes aim at the good-faith provision of Section 230,¹⁷⁰ enabling the FCC to take a more restrictive view of the provision, which has given broad immunity to social media platforms and allowed them to curate content mostly

163. Cf. *Roughly Seven-in-Ten Americans Think It Likely That Social Media Platforms Censor Political Viewpoints*, *supra* note 60.

164. Exec. Order No. 13,864, 84 Fed. Reg. 11,401 (2019). See also McGill & Lippman, *supra* note 157.

165. Exec. Order No. 11,246 (1965), 3 C.F.R. 339 (1964–1965).

166. Exec. Order No. 13,672, 41 C.F.R. 60 (2014); see also McGill & Lippman, *supra* note 157.

167. See generally Cristiano Lima, *How a Widening Political Rift over Online Liability Is Splitting Washington*, POLITICO (July 9, 2019, 2:07 PM), <https://www.politico.com/story/2019/07/09/online-industry-immunity-section-230-1552241> [<https://perma.cc/Z5CH-JCV9>].

168. Brian Fung, *White House Proposal Would Have FCC and FTC Police Alleged Social Media Censorship*, CNN BUS. (Aug. 10, 2019, 8:15 AM), <https://www.cnn.com/2019/08/09/tech/white-house-social-media-executive-order-fcc-ftc/index.html> [<https://perma.cc/CK96-DHP6>].

169. *Id.*

170. 47 U.S.C. § 230(c)(2)(A).

unquestioned.¹⁷¹ Most recently, at the direction of the White House, the Commerce Department filed a petition with the FCC encouraging the FCC to “use its authorities to clarify ambiguities in Section 230 so as to make its interpretation appropriate to the current internet marketplace and provide clearer guidance to courts, platforms, and users.”¹⁷² Ultimately, this approach will serve merely as a stop-gap until the legislature determines a further course of action.

CONCLUSION

Section 230 of the Communications Decency Act laid a foundation for the internet to develop by offering defamation liability protections to interactive computer services for the actions of third-party users. This protection, provided exclusively for online platforms, allowed the internet to flourish and opened the door for social media companies like Twitter and Facebook. Concerns over the content-moderation practices—such as shadow-banning—of such companies have given rise to new discussions about whether such liability protection should continue to apply given the power that social media companies have over information. To ensure that they maintain their protections, social media companies should follow a three-step process to restructure their content-moderation practices by articulating clear rules and guidelines, communicating all potential violations to users, and considering all other avenues before deciding to remove a post or ban a user. This will produce an objective content-moderation system, in line with the spirit of Section 230. Furthermore, Congress should revisit and update Section 230 to reflect the current internet age and ensure that social media companies are not exploiting the wide latitude they have been granted. Solutions, such as employing a non-partisan oversight group to analyze and publicize content-moderation practices of social media companies, will incentivize social media companies to act objectively in their content-moderation practices without requiring complete neutrality. These solutions will also inform users about how social

171. See Fung, *supra* note 167 (“Under [President Trump’s] draft proposal, the FCC will be asked to find that social media sites do not qualify for the good-faith immunity if they remove or suppress content without notifying the user who posted the material, or if the decision is proven to be evidence of anticompetitive, unfair or deceptive practices.”).

172. John Hendel, *Administration Asks FCC to Carry Out Trump’s Tech Crackdown*, POLITICO (July 27, 2020 7:04 PM), <https://www.politico.com/news/2020/07/27/administration-fcc-trump-crackdo-wn-383330> [<https://perma.cc/3VQ3-BZ7L>].

media companies determine what content appears and what content does not.

*Tanner Bone**

* J.D. (2021) Washington University School of Law, B.S.I.B. (2016) University of Arkansas. I would like to thank all the members of the Washington University Law Faculty who guided writing process and the editors of the *Washington University Law Review* for their invaluable contributions, suggestions, and questions regarding this Note. Finally, a special thanks to my wife Blakely for acting as a sounding board throughout the writing process.