THE YALE LAW JOURNAL FORUM

NOVEMBER 16, 2021

Platform Realism, Informational Inequality, and Section 230 Reform

Olivier Sylvain

ABSTRACT. Online companies bear few duties under law to tend to the discrimination that they facilitate or the disinformation that they deliver. Consumers and members of historically marginalized groups are accordingly the likeliest to be harmed. These companies should bear the same, if not more, responsibility to guard against such inequalities.

INTRODUCTION

As much as social-media companies have reconnected college roommates and spread awareness about movements like #BLM and #MeToo,[1] they also have contributed to the dysfunction of the current online information environment. They have helped cultivate bigotry,[2] discrimination,[3] and disinformation about highly consequential social facts.[4] Worse still, they have distributed and

1. SARAH J. JACKSON, MOYA BAILEY & BROOKE FOUCAULT WELLES, #HASHTAGACTIVISM: NETWORKS OF RACE AND GENDER JUSTICE, at xxv-xxviii (2020).

2. Joe Tidy, *Twitter Apologises for Letting Ads Target Neo-Nazis and Bigots*, BBC (Jan. 16, 2020), https://www.bbc.com/news/technology-51112238 [https://perma.cc/8FDL-UTD8].

3. Karen Hao, *Facebook's Ad Algorithms Are Still Excluding Women from Seeing Jobs*, MIT TECH. REV. (Apr. 9, 2021), https://www.technologyreview.com/2021/04/09/1022217/facebook-ad-algorithm-sex-discrimination [https://perma.cc/PHJ2-Y9RE]; Corin Faife & Alfred Ng, *Credit Card Ads Were Targeted by Age, Violating Facebook's Anti-Discrimination Policy*, MARKUP (Apr. 29, 2021, 8:00 AM ET), https://themarkup.org/citizen-browser/2021/04/29/credit-card-ads-were-targeted-by-age-violating-facebooks-anti-discrimination-policy [https://perma.cc/P4H5-LXTW].

4. Corin Faife & Dara Kerr, *Facebook Said It Would Stop Recommending Anti-Vaccine Groups. It Didn't*, MARKUP (May 20, 2021, 8:00 ET), https://themarkup.org/citizen-browser/2021/05/20/facebook-said-it-would-stop-recommending-anti-vaccine-groups-it-didnt [https://perma.cc/48DQ-WRJV].

delivered such material without bearing the burden of anticipating or attending to their social harms and costs.

These and other online companies remain free of any such legal obligation because of courts' broad interpretation of Section 230 of the Communications Decency Act (CDA), which immunizes "interactive computer services" that host or remove content posted by third parties.[5] Congress enacted Section 230 in 1996 to limit minors' exposure to pornography, as well as to encourage free expression, self-regulation, and innovation online.[6] Courts have read the provision broadly, generally dismissing complaints (before discovery) in which plaintiffs allege that the defendant service has published unlawful material — or unlawfully removed material from its platform. Courts will only allow a case to proceed when the defendant "contributes materially" to the offending content.[7]

In this way, Section 230 has created a very strong incentive for creative entrepreneurs to build or promote interactive computer services that host user-generated content. Novel social-media companies like Myspace, Foursquare, and Friendster sprang up in the years immediately following its enactment.[8] These services enabled their users to communicate with each other freely, about almost anything. While community sites and online bulletin boards existed before Congress enacted the CDA in 1996, Section 230 gave a distinctive boost to the online applications and services for user-generated content that most people today associate with the internet.[9]

But that was a long time ago. Today, the most popular social-media companies do much more than serve as simple platforms for users' free expression and innovation. The most prominent online services are shrewd enterprises whose main commercial objective is to collect and leverage user engagement for advertisers. Yet, until very recently, courts have allowed these companies to avoid public scrutiny because of the liability shield under Section 230.

It is past time for reform. I have elsewhere argued that courts should more closely scrutinize online intermediaries' designs on user-generated content and

5. Section 230 is codified at 47 U.S.C. § 230.

6. *See* Zeran v. Am. Online, Inc., 129 F.3d 327, 331 (4th Cir. 1997).

7. Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1168 (9th Cir. 2008).

8. Saheli Roy Choudhury, *Foursquare Pioneered the Trend of 'Checking-in' to a Place — Now It Sells Access to Its Data to Companies*, CNBC (Aug. 31, 2017, 10:20 AM EDT), https://www.cnbc.com/2017/08/30/foursquare-pioneered-the-trend-of-checking-in-to-a-place--now-it-sells-your-data-to-companies.html [https://perma.cc/L3KN-BYAL]; Max Chafkin, *How to Kill a Great Idea*, INC. (June 1, 2007), https://www.inc.com/magazine/20070601/features-how-to-kill-a-great-idea.html [https://perma.cc/QZF7-AJNC]..

9. *See generally* JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (2019) (detailing how Section 230 created the framework of the modern internet).

data.[10] And courts have begun to do so. They now are far more attentive to the ways in which internet platforms' prevailing ad-based business models and specific application design features necessarily facilitate harmful content and conduct, online and offline.[11] Courts have also been far more alert to whether defendant intermediaries are acting as publishers or something else, as in the cases in which plaintiffs have successfully alleged that online retailers and homesharing services are not "publishers" within the meaning of Section 230.[12]

But a court-centered approach to reforming Section 230 will not suffice. Courts cannot legislate, and in light of Section 230's plain language, statutory reform is probably the most effective and direct way to update the doctrine.[13] After all, Congress crafted Section 230 to protect interactive computer services that host third-party content. It is therefore up to Congress to adapt its aims to the current state of affairs. Consumers, but especially members of vulnerable and historically marginalized groups, have the most to gain from revamping Section 230 to require all online intermediaries to mitigate the anticipated impacts of their services. Reform is urgently needed because online service designs produce outcomes that conflict with hard-fought but settled consumer-protection and civil-rights laws.

This Essay sets out the reasons why now is the moment for statutory reform. Although this change would not ameliorate all of the social and economic ills for which intermediaries are responsible, it would have the salutary effect of ensuring that companies abide more closely to public-law norms and civic obligations. That is what we expect from actors in other sectors of the economy. Companies that have an outsized influence on public life should at least be held to the same standards—if not stricter ones.

This Essay proceeds in four parts. Part I describes the current social-media market and, in the process, argues that the role of social-media services in hosting and distributing third-party content is incident to their primary objective of holding the attention of their users and collecting their data for advertisers.

---

10. Olivier Sylvain, *Discriminatory Designs on User Data*, KNIGHT FIRST AMEND. INST. COLUM. U. (Apr. 1, 2018), https://knightcolumbia.org/content/discriminatory-designs-user-data [https://perma.cc/E88P-G28M] [hereinafter Sylvain, *Discriminatory Designs*]; Olivier Sylvain, *Intermediary Design Duties*, 50 CONN. L. REV. 203, 214-15 (2018) [hereinafter Sylvain, *Intermediary Duties*].

11. Lemmon v. Snap Inc., 995 F.3d 1085 (9th Cir. 2021); Doe No. 14 v. Internet Brands, Inc., 824 F.3d. 846 (9th Cir. 2016); Force v. Facebook, Inc., 934 F.3d 53 (2d Cir. 2019) (Katzmann, C.J., concurring in part and dissenting in part); Gonzalez v. Google LLC, 335 F. Supp. 3d 1156 (N.D. Cal. 2018), *aff'd*, 2 F.4th 871 (9th Cir. 2021).

12. *See infra* notes 129-131 and accompanying text.

13. *Cf. Gonzalez*, 2 F.4th at 897 ("In light of the demonstrated ability to detect and isolate at least some dangerous content, Congress may well decide that more regulation is needed.").

While this account has been made before, it is important for arguments in the following Parts to illuminate how online companies operate in a manner at odds with the how they and their defenders often describe themselves. Part II outlines the incentives and positive theory for the prevailing laissez-faire approach to content moderation and the legal doctrine that has given rise to the current state of affairs. In consideration of the market imperative to hold consumer attention, the protections under Section 230 doctrine have in fact set out a perverse disincentive to moderate. Part III outlines how courts have started to see online intermediaries, especially the biggest actors, for what they are, impacting the ways in which they make sense of whether a defendant is a "publisher" under Section 230. Finally, in Part IV, I return to a theme about which I have written elsewhere: the ways in which the robust protection under Section 230 has entrenched and, in some cases, deepened inequality in information markets. Free markets might redound to the benefit of consumers, but, as in other legislative fields, patterns of exclusion and subordination proliferate in the absence of legal rules against disparate impacts.

I. PLATFORM REALISM

Not too long ago, many of the most popular social-media companies loudly proclaimed to be champions of authentic voice, free expression, and human connection.[14] Their pronouncements often resembled marketing slogans and branding strategies. But they also reflected an earnest and widely held belief—that internet companies help people discover ideas and acquaintances in ways that legacy media companies in print, radio, television, and cable had not and could not.[15]

---

14. *See, e.g.*, *Mark Zuckerberg Stands for Voice and Free Expression*, FACEBOOK NEWSROOM (Oct. 17, 2019), https://about.fb.com/news/2019/10/mark-zuckerberg-stands-for-voice-and-free-expression [https://perma.cc/BKL8-CKSG]; Nicholas Thompson, *Jack Dorsey on Twitter's Role in Free Speech and Filter Bubbles*, WIRED (Oct. 16, 2018, 6:28 PM), https://www.wired.com/story/jack-dorsey-twitters-role-free-speech-filter-bubbles [https://perma.cc/S7FC-MFQ6].

15. *See* CLAY SHIRKY, HERE COMES EVERYBODY: THE POWER OF ORGANIZING WITHOUT ORGANIZATIONS 56-61 (2008) ("The old bargain of the newspaper—world news lumped in with horoscopes and ads from the pizza parlor—has now ended. The future presented by the internet is the mass amateurization of publishing and a switch from 'Why publish this?' to 'Why not?'"); *id.* at 296-302 (suggesting that social media would lead to "an explosion of new groups pursing new possibilities with new tools" which might be "painful for many existing organizations" and have some "negative effects" but would overall be "beneficial" to society); YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM 278-94 (2006) ("Ubiquitous Internet communications expand something of the freedom of city parks and streets, but also the freedom of cafés and bars—commercial platforms for social interaction—so that it is available everywhere."); John Perry

Today, most people are at best resigned to, and at worst weary of, their experiences with social media.[16] The same internet companies that proclaimed themselves to be champions of free expression just a few years ago have since backtracked. To be sure, they continue to promote themselves as platforms that "give people the power to build community and bring the world closer."[17] But their recent moderation decisions and policies suggest a far more cautious approach. They had no choice. The pressure they received from politicians, advertisers, consumers, and public interest groups has forced them to more aggressively curtail the most corrosive and objectionable material that they host. These efforts have been especially urgent for popular social-media companies, including Facebook and Twitter, that purport to build communities and foster discussion.[18]

But social-media companies generally have little concern for the nature of the communities or discussions that they host. That is because they are not mere platforms for authentic voice, free expression, and human connection. In fact, social media's ability to host and distribute third-party content—and thus to connect people and build communities—is incident to its ravenous ambition to hold the attention of their consumers and collect their data for advertisers.[19]

Barlow, *A Declaration of the Independence of Cyberspace*, ELEC. FRONTIER FOUND. (Feb. 8, 1996), https://www.eff.org/cyberspace-independence [https://perma.cc/283D-LM4D] (purporting that cyberspace will be a space free from existing power structures); Esther Dyson, George Gilder, George Keyworth & Alvin Toffler, *Cyberspace and the American Dream: A Magna Carta for the Knowledge Age*, PROGRESS & FREEDOM FOUND. (Aug. 1994), http://www.pff.org/issues-pubs/futureinsights/fi1.2magnacarta.html [https://perma.cc/4BCZ-K63S] (speculating about the power of cyberspace to reform social and economic structures).

16. *The Future of Tech Policy: American Views*, KNIGHT FOUND. (June 16, 2020), https://knightfoundation.org/reports/the-future-of-tech-policy-american-views [https://perma.cc/8ST7-2E4T] ("[J]ust a few years ago, Americans were overwhelmingly optimistic about the power of new technologies to foster an informed and engaged society. More recently, however, that confidence has been challenged by emerging concerns over the role that internet and technology companies—especially social media—now play in our democracy.").

17. *Our Mission: Give People the Power to Build Community and Bring the World Closer Together*, FACEBOOK, https://about.facebook.com/company-info/?_ga=2.167895777.1341094359.1623844710-1362261438.1613068707 [https://perma.cc/X268-AKKX].

18. Hayley Tsukayama, *Twitter's Asking for Help on How to Be Less Toxic*, WASH. POST (Mar. 1, 2018), https://www.washingtonpost.com/news/the-switch/wp/2018/03/01/twitters-asking-for-help-on-how-to-be-less-toxic [https://perma.cc/QF8J-S8QA]; Barbara Ortutay & Michael Liedtke, *Mark Zuckerberg Wants to Foster Communities, Not Just 'Connections'*, INC. (June 22, 2017), https://www.inc.com/associated-press/mark-zuckerberg-facebook-building-communities.html [https://perma.cc/973Z-MAKD].

19. Karen Hao, *How Facebook Got Addicted to Spreading Misinformation*, MIT TECH. REV. (Mar. 11, 2021), https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation [https://perma.cc/4RFS-7BAZ]. *See generally* TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR HEADS (2016) (discussing how social-media companies aim to hold consumer attention in order to increase ad revenue).

The January 6, 2021 siege on the Capitol building demonstrated that social-media companies can even mobilize seething reactionary mobs. There is little doubt that Twitter and Facebook helped to widely spread former President Trump's spurious claims about the 2020 presidential election, as well as a variety of other assertions and posts that seemed to violate their content guidelines, in the months and years before.[20] Their belated decisions to suspend his accounts made this fact plain as day. But other prominent social-media companies were also complicit, including Parler,[21] Reddit,[22] and YouTube.[23] They facilitated and galvanized the groups that bore down on Washington, D.C. to invade the Capitol. They fostered racist and xenophobic white nationalist online communities, provided forums for coordination among those groups, and helped to distribute information about their plans.[24] What followed was only a matter of time. The

20. Suhauna Hussain & Jeff Bercovici, *How Twitter Made Its Own Rules for Trump to Break*, L.A. TIMES (May 29, 2020), https://www.latimes.com/business/technology/story/2020-05-29/how-twitter-made-its-own-rules-for-trump-to-break [https://perma.cc/C9C2-EFG3]; Casey Newton, *Why Social Networks Keep Tripping over Their Own Content Moderation Policies*, VERGE (Oct. 16, 2019), https://www.theverge.com/interface/2019/10/16/20915658/trump-shooting-video-facebook-ads-twitter-world-leaders-rule [https://perma.cc/72BA-HFD9].

21. Aleszu Bajak, Jessica Guynn & Mitchell Thorson, *When Trump Started His Speech Before the Capitol Riot, Talk on Parler Turned to Civil War*, USA TODAY (Feb. 1, 2021), https://www.usatoday.com/in-depth/news/2021/02/01/civil-war-during-trumps-pre-riot-speech-parler-talk-grew-darker/4297165001 [https://perma.cc/NL7A-6YP2].

22. Brian Heater, *Reddit 'Taking Action' on Site Violations as Rioters Storm US Capitol*, TECHCRUNCH (Jan. 6, 2021), https://techcrunch.com/2021/01/06/reddit-taking-action-on-site-violations-as-rioters-storm-us-capitol [https://perma.cc/PW9G-CGXC].

23. Jennifer Elias, *New Google Union Slams YouTube for 'Lackluster' Response to Trump and Capitol Mob*, CNBC (Jan. 7, 2021), https://www.cnbc.com/2021/01/07/new-google-union-slams-youtube-for-lackluster-response-to-trump.html [https://perma.cc/FT3T-VDMW]. This was probably what instigated the former President's misbegotten effort to regulate social media. *See* Olivier Sylvain, *Solve the Underlying Problem: Treat Social Media as Ad-Driven Companies, Not Speech Platforms*, KNIGHT FOUND. (June 16, 2020), https://knightfoundation.org/articles/solve-the-underlying-problem-treat-social-media-as-ad-driven-companies-not-speech-platforms [https://perma.cc/MVG4-7PQW].

24. Jessica Guynn, *'Burn down DC': Violence that Erupted at Capitol Was Incited by Pro-Trump Mob on Social Media*, USA TODAY (Jan. 6, 2021), https://www.usatoday.com/story/tech/2021/01/06/trump-riot-twitter-parler-proud-boys-boogaloos-antifa-qanon/6570794002 [https://perma.cc/2ZX8-8UZC]; Isobel Asher Hamilton, *Plans to Storm the Capitol Were Circulating on Social Media Sites, Including Facebook, Twitter, and Parler, for Days Before the Siege*, BUS. INSIDER (Jan. 7, 2021), https://www.businessinsider.com/plans-to-storm-the-capitol-circulated-on-social-media-2021-1 [https://perma.cc/C6CB-Z96P]; Logan Jaffe, Lydia DePillis, Isaac Arnsdorf & J. David McSwane, *Capitol Rioters Planned for Weeks in Plain Sight. The Police Weren't Ready*, PROPUBLICA (Jan 7. 2021), https://www.propublica.org/article/capitol-rioters-planned-for-weeks-in-plain-sight-the-police-werent-ready [https://perma.cc/W4DK-QU72].

former President's rhetoric and mendacity lit the match, but online services provided the kindling.

Of course, these companies are not in cahoots with the reactionaries who promoted the attack on the Capitol. Things are more complicated. No matter the platform, demagogues and clever companies exploit a variety of biases for political gain and convince unwitting consumers to do things they might not otherwise do.[25] Still, the big social-media companies' main objective is to hold consumers' attention for advertisers irrespective of these miscreants' aims. The companies' stated goal of fostering community is an incident of their pecuniary imperative to optimize consumer engagement. This objective is not unqualified; social-media companies have long-term incentives to deemphasize content that offends the majority of their consumers, so that consumers will continue using their platforms.[26] But those long-term incentives have not been strong enough to curtail the distribution of hateful, violent, and debasing user content and advertisements. Thus, alongside clips of lawyers inadvertently talking through cat filters and clever dance sequences, newsfeeds and recommendations are also filled with baseless headlines about crackpot conspiracies, bigoted calls for violence, and advertisements for far-right militia merchandise — despite policies that ban "militia content."[27]

All of this has disillusioned many, if not most, consumers and policy officials. Advertisers have noticed. Companies generally do not want to associate their brands with toxic and divisive content.[28] This is why the most popular internet

25. Daniel Kreiss, Regina G. Lawrence & Shannon C. McGregor, *Political Identity Ownership: Symbolic Contests to Represent Members of the Public*, 6 SOC. MEDIA + SOC'Y 2-3 (June 10, 2020); Scott Ikeda, *California Bans Deceptive "Dark Patterns" with Update to State Privacy Law*, CPO MAG. (Mar. 23, 2021), https://www.cpomagazine.com/data-privacy/california-bans-deceptive-dark-patterns-with-update-to-state-privacy-law [https://perma.cc/ES5P-HJGE]; *see also* Jamie Luguri & Lior Jacob Strahilevitz, *Shining a Light on Dark Patterns*, 13 J. LEGAL ANALYSIS 43, 48-58 (2021) (setting out a taxonomy of dark patterns).

26. *See* Nick Clegg, *You and the Algorithm: It Takes Two to Tango*, MEDIUM (Mar. 31, 2021), https://nickclegg.medium.com/you-and-the-algorithm-it-takes-two-to-tango-7722b19aa1c2 [https://perma.cc/ZJP4-B8RX].

27. Jeremy B. Berrill, Tech Firms Profited from Far-Right Militia Content Despite Ban on "Three Percenters," Markup (Jan. 21, 2021, 11:50 AM), https://themarkup.org/news/2021/01/21/tech-firms-profited-from-far-right-militia-content-despite-ban-on-three-percenters [https://perma.cc/RW7U-8X7T].

28. *See, e.g.*, Olivia Solon, *Google's Bad Week: YouTube Loses Millions as Advertising Row Reaches the U.S.*, GUARDIAN (Mar. 25, 2017, 6:00 PM), https://www.theguardian.com/technology/2017/mar/25/google-youtube-advertising-extremist-content-att-verizon [https://perma.cc/2LUU-V85W].

companies today seem to have shifted away from being beacons of free speech.[29] Some internet companies have even called for increased government oversight or regulation.[30]

For several years, legal scholars and social scientists have been recommending creative design tweaks that are more varied than the familiar but unsatisfyingly binary "keep-up versus take-down" framework.[31] Internet companies have

29. Gilad Edelman, *On Social Media, American-Style Free Speech Is Dead*, WIRED (Apr. 27, 2021, 8:00 AM), https://www.wired.com/story/on-social-media-american-style-free-speech-is-dead [https://perma.cc/22AZ-BQ45].

30. *E.g.*, Mark Zuckerberg, *Big Tech Needs More Regulation*, META NEWSROOM (Feb. 18, 2020), https://about.fb.com/news/2020/02/big-tech-needs-more-regulation [https://perma.cc/MB85-JANV]. They have done so even as content moderation is costly and otherwise resource intensive. Jullian 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/MNU4-49AD]; Mike Masnick, *Content Moderation at Scale Is Impossible: Recent Examples of Misunderstanding Context*, TECHDIRT (Feb 26, 2021), https://www.techdirt.com/articles/20210225/10365146316/content-moderation-scale-is-impossible-recent-examples-misunderstanding-context.shtml [https://perma.cc/EVK5-26EC]. *See generally* Tarleton Gillespie, *Content Moderation, AI, and the Question of Scale*, 7 BIG DATA & SOC'Y 1 (2020) (challenging the view that companies should automate content moderation, even assuming they could). But there is good reason to doubt that they are motivated by purely altruistic impulses when they do. After all, the biggest internet companies are likely to benefit from new regulations that, while potentially burdensome, will be easier for them to abide by than it will be for upstarts and smaller rivals. To the extent that an online service moderates content—and most, if not all, do in some way—it is easier for a company with more technological sophistication and absolute resources to do so. *See* Mike Isaac, *Mark Zuckerberg's Call to Regulate Facebook, Explained*, N.Y. TIMES (Mar. 30, 2019), https://www.nytimes.com/2019/03/30/technology/mark-zuckerberg-facebook-regulation-explained.html [https://perma.cc/TW75-R282]. While, hypothetically, smaller companies could, the larger companies are more likely to be in the enviable position of being able to "moderate at scale," as hard as that task is. *See* Bobbie Johnson, *How a Democratic Plan to Reform Section 230 Could Backfire*, MIT TECH. REV. (Feb. 8, 2021), https://www.technologyreview.com/2021/02/08/1017625/safe-tech-section-230-democrat-reform [https://perma.cc/SC9U-SX5A] (quoting Eric Goldman, who said that "Section 230 reform won't stick it to Big Tech. Section 230 reform will deepen the incumbents' competitive moats to make it even harder for new entrants to compete"); Jason Kelley, *Section 230 Is Good, Actually*, ELEC. FRONTIER FOUND. (Dec. 3, 2020), https://www.eff.org/deeplinks/2020/12/section-230-good-actually [https://perma.cc/D8JJ-MKSJ] ("Though calls to reform Section 230 are frequently motivated by disappointment in Big Tech's speech moderation policies, evidence shows that further reforms to Section 230 would make it more difficult for new entrants to compete with Facebook or Twitter—and would likely make censorship worse, not better."); *see also* Matt Perault, *Well-Intentioned Section 230 Reform Could Entrench the Power of Big Tech*, SLATE (June 1, 2021, 9:00 AM), https://slate.com/technology/2021/06/section-230-reform-antitrust-big-tech-consolidation.html [https://perma.cc/9DVA-F9KS] (discussing potential effects of Section 230 on the Big Tech industry).

31. *See, e.g.*, Eric Goldman, *Content Moderation Remedies*, MICH. TECH. L. REV. (forthcoming), https://ssrn.com/abstract=3810580 [https://perma.cc/Z8F7-ZHSQ].

been listening; today, they are taking demonstrable steps to tamp down their most toxic and alarming content through creative adjustments to the designs of their user interfaces.[32] The biggest social-media companies are introducing friction into the ways in which their consumers share and engage content.[33] Twitter and Facebook, for example, started flagging dubious political ads and claims by high-ranking elected officials a couple of years ago.[34] They do this by placing visual and textual "content labels" alongside suspect user-generated posts in order to inform consumers about misleading or harmful content.[35] Other notable design tweaks include "circuit breakers" that limit the amplification or viral spread of toxic content.[36] Twitter, for example, recently started sending users warnings before they post anything that its automated content-review systems anticipate as being potentially harmful or offensive.[37] Research has shown that "frictive prompts" like these may curb people's impulse to post.[38] Twitter also

---

32. *See, e.g.*, *Our Commitments*, YOUTUBE, https://www.youtube.com/howyoutubeworks/our-commitments/curbing-extremist-content [https://perma.cc/47AR-23V9].

33. *See* Ellen P. Goodman, Karen Kornbluh & Eli Weiner, *The Stakes of User Interface Design for Democracy*, GERMAN MARHSALL FUND U.S. (June 30, 2021), https://www.gmfus.org/publications/stakes-user-interface-design-democracy [https://perma.cc/RKL7-GK9Y].

34. Fadel Allassan, *Twitter and Facebook Label Trump Tweet on Absentee Ballot Ruling as Misleading*, AXIOS (Nov. 3, 2020), https://www.axios.com/trump-twitter-facebook-labels-tweet-post-eb34d4cb-eb46-477f-8e1c-a587ad364835.html [https://perma.cc/38LP-8QKY]. Some interventions like these, however, are more effective than others. *See* Tom Dobber, Sanne Kruikemeier, Ellen P. Goodman, Natali Helberger & Sophie Minihold, *Effectiveness of Online Political Ad Disclosure Labels: Empirical Findings*, U. AMSTERDAM INST. INFO. L. (Mar. 8, 2021), https://www.uva-icds.net/wp-content/uploads/2021/03/Summary-transparency-disc-loures-experiment_update.pdf [https://perma.cc/9FCX-W5G8]; Deepa Seetharaman, *Twitter's Labels for Trump Tweets Show Platforms' Split Over Political Speech*, WALL ST. J. (May 28, 2020, 12:57 AM), https://www.wsj.com/articles/twitters-labels-for-trump-tweets-show-platforms-split-over-political-speech-11590621829 [https://perma.cc/F55T-URF3].

35. For a recent discussion of content labeling, see generally Garrett Morrow, Briony Swire-Thompson, Jessica M. Polny, Matthew Kopec & John P. Wihbey, The Emerging Science of Content Labeling: Contextualizing Social Media Content Moderation (Dec. 3, 2020) (unpublished manuscript), https://ssrn.com/abstract=3742120 [https://perma.cc/M6TK-YBY6].

36. Goodman et al., *supra* note 33.

37. Anita Butler & Albeto Parrella, *Tweeting with Consideration*, TWITTER BLOG (May 5, 2021), https://blog.twitter.com/en_us/topics/product/2021/tweeting-with-consideration [https://perma.cc/XM9V-2S8E].

38. *See generally* Ellen Goodman & Karen Kornbluh, *The Stakes of User Interface Design for Democracy*, GERMAN MARSHALL FUND U.S. (June 2021), https://www.gmfus.org/publications/stakes-user-interface-design-democracy [https://perma.cc/V6GN-AH9Q] (delineating how neutral design principles can be used to empower, instead of exploit, users); Ellen Goodman, *Digital Information Fidelity and Friction*, KNIGHT FIRST AMEND. INST. COLUM. UNIV. (Feb. 26, 2020), https://knightcolumbia.org/content/digital-fidelity-and-friction [https://perma.cc/FG2D-FBNC].

recently made changes to the ways in which it crops the photographs that its users post.[39] Before Twitter made those changes, research suggested that its "saliency algorithm" featured images of white people more than those of Black people and focused on women's chests or legs rather than their other physical attributes.[40]

These tweaks are a sign that social-media companies are alert to consumers' distaste for certain kinds of content and content-distribution methods. But these changes elide far more pressing problems. As I explain below in Section II.B and Part III, under current law, companies may still distribute harmful or illegal content even if they control the ways in which they deliver that material.[41] Under Section 230, Congress created a safe harbor to encourage online companies to host and moderate third-party content, "unfettered" by government regulation.[42] This, according to the courts, was the drafters' unrestrictive approach to encouraging innovation and content regulation, both at once.[43]

In this way, through Section 230, Congress promulgated a court-administered innovation policy that aimed to promote a certain kind of online business design—platforms for user-generated content. But legislators in 1996 could not anticipate how shielding this form of "interactive computer service" would beget companies whose main objective would be to optimize consumer engagement for advertisers, unencumbered by the social costs and harms that online content

---

39. Rumman Chowdhury, *Sharing Learnings About Our Image Cropping Algorithm*, TWITTER ENG'G (May 19, 2021), https://blog.twitter.com/engineering/en_us/topics/insights/2021/sharing-learnings-about-our-image-cropping-algorithm.html [https://perma.cc/T6AW-RWLG].

40. *Id.* Also consider the company's reported plan for a tiered subscription service. The premium level of service would cost about three dollars per month and feature an "undo tweets" function and a folder for bookmarked content. Kim Lyons, *Twitter May Be Working on Twitter Blue, a Subscription Service that Would Cost $2.99 Per Month*, VERGE (May 15, 2021, 11:40 AM EDT), https://www.theverge.com/2021/5/15/22437690/twitter-blue-subscription-service-299-undo-tweets [https://perma.cc/X2DV-MTB8]; Chaim Gartenberg, *Twitter Is Surveying Users on What Features They'd Want from a Subscription*, VERGE (July 31, 2020, 12:46 PM EDT), https://www.theverge.com/2020/7/31/21349644/twitter-subscription-service-survey-undo-sent-analytics-feature-ideas [https://perma.cc/NYL8-97FP]. Other creative recommendations focus on processes for "accountability and repair" for the victims of cyberharassment, as well as their deplatformed attackers. *See, e.g.*, Sarita Schoenebeck, Carol F. Scott, Emma Hurley, Tammy Chang & Ellen Selkie, *Youth Trust in Social Media Companies and Expectations of Justice: Accountability and Repair After Online Harassment*, 5 HUM.-COMPUT. INTERACTION, Apr. 2021, at 14, https://yardi.people.si.umich.edu/pubs/Schoenebeck_AccountabilityRepair2021.pdf [https://perma.cc/GY46-4VJ9].

41. *See, e.g.*, Vanessa Barbara, Opinion, *Miracle Cures and Magnetic People. Brazil's Fake News Is Utterly Bizarre.*, N.Y. TIMES (July 5, 2021), https://www.nytimes.com/2021/07/05/opinion/brazil-fake-news-bolsonaro.html [https://perma.cc/PR35-FC9M].

42. 47 U.S.C. § 230(b)(2) (2018).

43. *See, e.g.*, Zeran v. Am. Online, 129 F.3d 327, 331 (4th Cir. 1997).

and expressive conduct imposes on others.[44] The design tweaks I enumerate above may arise from the market imperative to keep consumer demand, but it also bumps up against a far more compelling market incentive to hold and quantify consumer attention for advertisers. Social-media companies, wedded to the extraordinary amounts of ad revenue that they generate, are today in no position to redress this incentive structure. The current information environment is proof enough.

II. MARKET FOR MODERATION

Only Congress, the architect of Section 230's regulatory scheme, is capable of reforming the prevailing incentive structure motivating social-media companies. In Section II.A below, I sketch out the argument for the current laissez-faire approach, before turning to reasons for legislative or regulatory reform in Section II.B.

A. The Laissez-Faire Logic for Online Platforms

Market pressures evidently affect the ways in which social-media companies choose to distribute content. This presents a challenge for people who believe that legislation, regulation, civil litigation, or criminal enforcement (and the threat of their occurrence) affect internet companies.[45] After all, due to the expansive protection courts afford platforms under Section 230,[46] few articulated legal rules prefigure how intermediaries may distribute or moderate content. To the extent any exist, they influence (but do not resolve) how intermediaries may distribute online content that violates criminal law,[47] intellectual-property law,[48]

---

44. *The Impact of the Law That Helped Create the Internet and an Examination of Proposed Reforms for Today's Online World: Hearing on the PACT Act and Section 230 Before the Subcomm. on Commc'ns, Tech., Innovation & the Internet*, 116th Cong. 6 (2020) (statement of Rep. Cox).

45. *Cf.* Book Note, *In the Shadow of the Law by Kermit Roosevelt*, 120 HARV. L. REV. 1367 (2007); Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950 (1979).

46. *See infra* Section II.B and Part iii.

47. *See* 47 U.S.C. § 230(e)(1) (2018).

48. *See id.* § 230(e)(2); *see also* Digital Millennium Copyright Act, 17 U.S.C. § 512(c) (2018) (limiting liability of online service providers for copyright infringement).

telephone consumer-privacy law,[49] and sex-trafficking laws.[50] Yet, despite the absence of such constraints, companies appear to be taking the initiative by, among other things, deplatforming demagogues and fashioning new ways to tamp down the distribution of disinformation.[51]

This is why advocates of the current regulatory regime, as spare as it is, stand on good ground when they defend the status quo. They tend to subscribe to a beguiling classical conception of free markets.[52] For them, the unregulated market for "interactive computer services" has promoted experimentation, innovation, learning, and discovery in ways that could never be possible were the law more heavy-handed about restricting content. Freedom may have its costs, they allow, but those are the incidents of progress and learning. Even the most quotidian and frivolous of online exchanges could be valuable.[53] Law should not chill free authentic democratic deliberation, such as it is.

---

49. *See* 47 U.S.C. § 230(e)(4) (2018); *see also* 18 U.S.C. § 2511 (2018) (prohibiting "[i]nterception and disclosure of wire, oral, or electronic communications" with certain exemptions for communications service providers).

50. *See* 47 U.S.C. § 230(e)(5) (2018); *see also* 18 U.S.C. §§ 1591, 1595 & 2421(a) (2018) (criminalizing sex trafficking).

51. *See Facebook Suspends Trump Accounts for Two Years*, BBC (June 5, 2021), https://www.bbc.com/news/world-us-canada-57365628 [https://perma.cc/VG94-37MD]; Kari Paul, *Twitter Targets Covid Vaccine Misinformation with Labels and 'Strike' System*, GUARDIAN (Mar. 1, 2020, 6:05 PM EST), https://www.theguardian.com/technology/2021/mar/01/twitter-coronavirus-vaccine-misinformation-labels [https://perma.cc/E2MG-7XKA].

52. *See generally* Olivier Sylvain, *'AOL v. Zeran': The Cyberlibertarian Hack of §230 Has Run Its Course*, LAW.COM (Nov. 10, 2017, 1:15 AM), https://www.law.com/therecorder/sites/therecorder/2017/11/10/aol-v-zeran-the-cyberlibertarian-hack-of-%C2%A7230-has-run-its-course [https://perma.cc/WPN9-4N9L] (identifying the claims that the most celebrated advocates of broad protection for online speech in the mid-1990s made).

53. Some exchanges and online conversations may seem inconsequential but may nevertheless be affecting and productive, as when one Twitter user decided to question the talent of a celebrity musician. "I don't think pretty privilege has carried anyone better than it carried Alicia Keys," the user wrote. @seyi-baby, TWITTER (May 8, 2021, 12:21 PM), https://twitter.com/seyi_baby/status/1391065839026196482 [https://perma.cc/HV22-3PMU]. This was playful if poorly thought-through banter. As innocuous as it may have seemed, however, the tweet became a momentary flashpoint for discussions about musicianship, race, and contested conceptions of beauty. It set off a torrent of reactions, the vast majority of which disagreed. *See, e.g.*, @Nerdiac, TWITTER (May 8, 2021, 2:20 PM), https://twitter.com/Nerdiac/status/1391095639744856067 [https://perma.cc/G5PA-A83D] ("Alicia Keys is one of the few times the industry got it right. Total package: writer, producer, musician, beautiful, vocals, and young. They debuted her quick because they'd be fools not to. Her being pretty is a given, ALL mainstream music stars are good looking. You clownin."). The original tweet hit a nerve and Twitter's moderation algorithms shuttled it from user to user until it started "trending." Thousands of similar episodes likely take place across the internet every day. They create opportunities for engagement—explicit and sublimated—on the challenging and not-so-challenging preoccupations of the day. This is presumably what Twitter founder Jack Dorsey had

Other proponents of the status quo laissez-faire approach might recognize that regulation would be necessary if the market for online services was not filled with variety. But, as there are none of the same trappings of scarcity online as there are in other industries, they might argue, regulation is not necessary. Users enjoy an abundance of options for content and services. The means of production are different, too. Indeed, the barriers to entry for content creators are low: almost anyone can publish anything on some platform. Or they can create their own substack, Medium account, or webpage. And the market for online content moderation responds to consumer demand. Users who want a heavily curated and moderated online experience can patronize a variety of familiar content producers. On the one hand, prominent producers of online content like Amazon, Netflix, Peacock, and Disney Plus offer heavily curated entertainment that features their content and excludes much content of other companies. Consumers can also find or subscribe to matching and recommendation services in specialized areas: everything from music sharing sites like SoundCloud or BandCamp, to health technology sharing sites for doctors like Doximity, to sites that facilitate the buying and selling of unregistered firearms like Armslist. Consumers may also find services that restrict harmful content, such as services that forbid sexually explicit material like Instagram. Other consumers will prefer services that are ostensibly far more permissive like Parler. Still others will look for services that host the most alarming and offensive content, including websites like Gab that distribute material that is racist, misogynist, white-nationalist, and antisemitic.[54]

This is the unfettered market for online content and services. Proponents of the laissez-faire approach contend that, as vibrant as the information ecosystem is, neither legislatures nor regulators should intervene; the free market for content moderation and recommendation is robust in ways that, for them, is normatively desirable and consistent with prevailing First Amendment doctrine.[55]

---

in mind when, in explaining the social-media company's decision to deplatform former President Trump, he asserted that social media was a platform for democratic deliberation. *See* Jack Dorsey (@jack), TWITTER (Jan. 13, 2021, 7:16 PM ET), https://twitter.com/jack/status/1349510784620003330?lang=en [https://perma.cc/JY9K-5MCX] ("I believe the internet and global public conversation is our best and most relevant method of achieving this. I also recognize it does not feel that way today. Everything we learn in this moment will better our effort, and push us to be what we are: one humanity working together.").

54. *See* Micah Lee, *Inside Gab, the Online Safe Space for Far-Right Extremists*, INTERCEPT (Mar. 15, 2021, 6:00 AM), https://theintercept.com/2021/03/15/gab-hack-donald-trump-parler-extremists [https://perma.cc/WAJ9-895R]; Tanya Basu, *The "Manosphere" Is Getting More Toxic as Angry Men Join the Incels*, TECH. REV. (Feb. 7, 2020), https://www.technologyreview.com/2020/02/07/349052/the-manosphere-is-getting-more-toxic-as-angry-men-join-the-incels [https://perma.cc/SB7D-WVC2].

55. *See, e.g.*, Daphne Keller, *Amplification and Its Discontents*, KNIGHT FIRST AMEND. INST. COLUM. UNIV. (June 8, 2021), https://knightcolumbia.org/content/amplification-and-its-

The positive case for the laissez-faire approach resonates with an emerging view that companies, especially internet companies, have a constitutional right to decide which ideas to distribute or promote and which ideas to demote or block.[56] The strongest version of this view conceives of almost all information flows, even overtly commercial ones, as presumptively protected communicative acts under the First Amendment—notwithstanding the fact that in other contexts, overtly commercial speech is afforded less protection than other expressive acts.[57] Scholars have labeled this emergent view the "New *Lochner*" because of the ways in which courts have applied the strong constitutional interest in free speech to shield commercial activities that historically have not been protected.[58] These writers invoke *Lochner v. New York*,[59] a Supreme Court case notorious for its grotesquely expansive view of freedom to contract that overrode Congress's interest in minimum-wage and maximum-hours labor legislation.[60] Some of the more recognizable contemporary artifacts of the New *Lochner* are the Supreme Court's decisions on campaign-finance regulation (such as *Citizens United v. FEC*,[61] in which the Court invalidated limits on contributions to issue advertising) and targeted marketing (such as *Sorrell v. IMS Health Inc.*,[62] in which the Court struck down limits on promotional campaigns by pharmaceutical companies).[63]

---

discontents [https://perma.cc/EB92-N9R4] ("[P]latforms' own algorithmic ranking and recommendation have been held to constitute protected speech. A law explicitly prohibiting such speech—or requiring platforms to replace their own preferred algorithm with the state's preferred algorithm, as a chronological ranking mandate would do—is likely to face real constitutional problems."); *see also* Packingham v. North Carolina, 137 S. Ct. 1730, 1735-36 (2017) ("[S]ocial media users employ these websites to engage in a wide array of protected First Amendment activity on topics 'as diverse as human thought.'" (quoting Reno v. ACLU, 521 U.S. 844, 870 (1997))).

56. *See, e.g.*, Search King, Inc. v. Google Tech., Inc., No. CIV-02-1457-M, 2003 WL 21464568, at *3-4 (W.D. Okla. May 27, 2003); Zhang v. Baidu.com Inc., 10 F. Supp. 3d 433, 438 (S.D.N.Y. 2014); Prager Univ. v. Google LLC, 951 F.3d 991, 996-97 (9th Cir. 2020).

57. *See, e.g.*, Sorrell v. IMS Health Inc., 564 U.S. 552 (2011).

58. *See, e.g.*, Amanda Shanor, *The New* Lochner, 2016 WIS. L. REV. 133, 136 (tracing the ways in which companies and commercial interests have successfully invoked the First Amendment as a powerful deregulatory tool across legislative fields over the past few decades).

59. 198 U.S. 45 (1905).

60. *See generally* Jamal Greene, *The Anticanon*, 125 HARV. L. REV. 379, 380 (2011) (explaining the ways in which scholars, teachers, lawyers, and writers invoke "anticanonical cases" like *Lochner* to "dispel dissensus about or sanitize the Constitution" in ways that may obfuscate the evolution and contested nature of the founding document).

61. 558 U.S. 310 (2010).

62. 564 U.S. 552 (2011).

63. *See also* Manhattan Cmty. Access Corp. v. Halleck, 139 S. Ct. 1921, 1928 (2019) (noting that the Free Speech Clause "prohibits only *governmental* abridgment of speech," not "*private*

Against this backdrop, proponents of the current regime warn about the unintended consequences of regulation. Social-media companies have been demonstrably creative, innovative, and prolific; they have transformed the internet into a vast bazaar of goods for everyone. They fear that legal oversight will have a chilling effect on innovation and expression.[64] Developers anxious about attracting legal trouble will be less creative and adventurous about pursuing untested business models and novel content, effectively entrenching the power of the biggest companies.[65] They also posit that it might backfire against vulnerable groups and minorities that espouse unpopular views.[66] This could have the affect of silencing social movements for reform, including, for example, #BlackLivesMatter and #MeToo. Some even worry that entrepreneurs in this country would lose their competitive edge if the United States imposes legal constraints on what internet companies can develop or sell.[67]

---

abridgment of speech"). Justice Kavanaugh, who was not on the bench when the Court announced *Citizens United* and *Sorrell*, is now the clearest proponent of this muscular conception of speech on the Court. *See, e.g.*, U.S. Telecom Ass'n v. FCC, 855 F.3d 381, 426-31 (D.C. Cir. 2017) (Kavanaugh, J., dissenting) (arguing that broadband-service providers have a First Amendment right to be free from open-internet regulation).

64. *See, e.g.*, Christopher Mims, *How Congress Might Upend Section 230, the Law Big Tech Is Built on*, WALL ST. J. (Feb. 13, 2021, 12:00 AM ET), https://www.wsj.com/articles/how-congress-might-upend-section-230-the-internet-law-big-tech-is-built-on-11613172368 [https://perma.cc/45RL-YNPA] (explaining that, if social-media companies have to return to their pre-1996 legal status, they would likely "have to drastically narrow the scope and volume of what's permitted on their platforms"); *see also* Daphne Keller, *Amplification and Its Discontents*, KNIGHT FIRST AMEND. INST. COLUM. UNIV. 9-17 (June 8, 2021), https://knightcolumbia.org/content/amplification-and-its-discontents [https://perma.cc/5HMG-EL4Z] (arguing that laws that regulate content amplification would encourage companies to "over-enforce and suppress lawful speech").

65. *See* Mark Weinstein, *Small Sites Need Section 230 to Compete*, WALL ST. J. (Jan. 25, 2021, 2:16 PM ET), https://www.wsj.com/articles/small-sites-need-section-230-to-compete-11611602173 [https://perma.cc/EDU8-826K] (arguing that revoking Section 230 "would significantly harm smaller companies" and "new startups that compete with the tech giants," as only the biggest companies would have the resources to "hire the massive modernization and legal teams that would be necessary to defend themselves" against liability); Matt Perault, *Well-Intentioned Section 230 Reform Could Entrench the Power of Big Tech*, SLATE (June 1, 2021, 9:00 AM), https://slate.com/technology/2021/06/section-230-reform-antitrust-big-tech-consolidation.html [https://perma.cc/4CCD-92AZ] (making a similar argument).

66. Billey Easley, *Revising the Law that Lets Platforms Moderate Content Will Silence Marginalized Voices*, SLATE (Oct. 29, 2020, 5:43 PM), https://slate.com/technology/2020/10/section-230-marignalized-groups-speech.html [https://perma.cc/98J9-3D2L] ("There are likely marginalized groups today who have not yet come to more mainstream acceptance that will be denied this same opportunity [to speak freely] if they do not have open internet platforms that allow users to generate their own content.").

67. *Comments of Consumer Technology Association, in the Matter of Section 230 of the Communications Act of 1934*, CONSUMER TECH. ASS'N (Sept. 2, 2020) https://cdn.cta.tech/cta/media/media/advocacy/issues/comments-of-cta-rm-11862.pdf [https://perma.cc/M2HJ-AVAS] ("The

B. The Disincentive to Moderate

But none of this means that social-media companies are unaffected by law. Even though there is no positive law regulating content moderation, internet companies have been free to develop moderation standards because of the protection under Section 230. For over two decades, the courts have concluded that, pursuant to Section 230(c)(1),[68] online intermediaries are not liable for the unlawful material that their users create or develop.[69] Nor, under Section 230(c)(2), are companies legally responsible for their decisions to remove or block third-party content or make filtering technology available.[70] Congress concluded that if such companies were held liable for any of these activities, the free flow of ideas and information would slow and stall.[71] Courts accordingly have shielded intermediaries from liability to the extent those companies provide platforms for third-party content, no matter how heinous the material is.[72]

This protection goes beyond the protections that the First Amendment provides. That is, Section 230 shields defendants from liability for third-party content that falls outside the scope of First Amendment protections, including

unique balance of protections afforded by Section 230 has enabled the United States to be the global leader in internet innovation.").

68. The term Section 230 is colloquial as much as a simple indication of its codification. *See* 47 U.S.C. § 230 (2018). Congress enacted the provision as part of the Communications Decency Act of 1996 (CDA), which amended Title 47 in several places. *See, e.g.*, *id.* § 223 (setting forth prohibitions, which the CDA amended, on sending obscene or harassing comments by means of a telecommunications device). The Supreme Court invalidated those other provisions, *see* Reno v. ACLU, 521 U.S. 844 (1997) (holding that two CDA provisions violated the freedom of speech), so that, now, Section 230 is all that remains.

69. *See, e.g.*, Zeran v. Am. Online, Inc., 129 F.3d 327, 328 (4th Cir. 1997); Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1120-21 (9th Cir. 2003); Chi. Laws.' Comm. for C.R. Under L., Inc. v. Craigslist, Inc., 519 F.3d 666, 671 (7th Cir. 2008); Doe v. MySpace, Inc., 528 F.3d 413, 418 (5th Cir. 2008); Jones v. Dirty World Ent. Recordings LCC, 755 F.3d 398, 402 (6th Cir. 2014); Doe v. Backpage.com, LLC, 817 F.3d 12, 15 (1st Cir. 2016); Herrick v. Grindr Holding Co., 765 F. App'x 586, 591 (2d Cir. 2019); Dyroff v. Ultimate Software Grp., 934 F.3d 1093, 1094 (9th Cir. 2019).

70. *See* 47 U.S.C. § 230(c)(2)(A) (2018). Section 230(c)(2)(B) provides a third protection for entities that provide content filtering — that is, for "interactive computer service[s]" that "enable or make available to information content providers or others the technical means to restrict access" to content. This provision is rarely, if ever, litigated.

71. *See Zeran*, 129 F.3d at 331 ("The specter of tort liability in an area of such prolific speech would have an obvious chilling effect. It would be impossible for service providers to screen each of their millions of postings for possible problems. Faced with potential liability for each message republished by their services, interactive computer service providers might choose to severely restrict the number and type of messages posted.").

72. *See* Gonzalez v. Google LLC, 2 F.4th 871, 897-99 (9th Cir. 2021).

defamation or commercial speech.[73] Evidently, Congress's aim was to ensure that all views and ideas, even abhorrent ones, are exposed and debated in the marketplace of ideas.

Courts have read Section 230 to bar third-party liability suits against intermediaries, even when the companies know that their services will distribute unlawful or illicit content.[74] This is a departure from traditional publisher liability rules in common-law tort, as well as the general principles of third-party liability across legislative fields.[75] Under the first of the protections — Section 230(c)(1) — plaintiffs must successfully establish that an intermediary has "contribute[d] materially" to the development of the content in order for such suits to proceed to discovery, let alone succeed on the merits.[76] Pursuant to the second protection — Section 230(c)(2)(A)[77] — an online service's decision to takedown or block third-party content is not actionable if that decision is voluntary and in "good faith."[78] The third, far less litigated protection shields intermediaries who make filtering technology available.[79]

In its foundational interpretation of the first of those provisions almost a quarter century ago, a Fourth Circuit panel concluded in *Zeran v. America Online, Inc.* that "[t]he specter of tort liability in an area of such prolific speech would

---

73. This is to say nothing of its other procedural advantages. *See generally* Eric Goldman, *Why Section 230 Is Better than the First Amendment*, 95 NOTRE DAME L. REV. REFLECTION 33 (2019) (explaining that Section 230 provides defendants with more substantive and procedural benefits than the First Amendment does).

74. *See, e.g.*, *Dirty World Ent. Recordings LLC*, 755 F.3d at 414; *Backpage.com, LLC*, 817 F.3d at 21.

75. *Zeran*, 129 F.3d at 333-34.

76. Fair Hous. Council v. Roommates.com, LLC , 521 F.3d 1157, 1168 (9th Cir. 2008) (interpreting 47 U.S.C. § 230(c)(1) (2018) to impose a material-contribution requirement for liability).

77. 47 U.S.C. § 230(c)(2)(A) (2018).

78. Enigma Software Grp. U.S.A., L.L.C. v. Malwarebytes, Inc., 946 F.3d 1040, 1051 (9th Cir. 2019) (holding that § 230(c)(2)(A) does not apply to claims that the defendant blocked access to the plaintiff's services for anticompetitive reasons). For what it is worth, courts have also dismissed plaintiffs' claims that "deplatforming" decisions or takedowns violate the First Amendment because they are biased against certain viewpoints. *See, e.g.*, Prager Univ. v. Google, 951 F.3d 991, 997-98 (9th Cir. 2020) (holding that YouTube's decision to demonetize plaintiff's YouTube channel did not violate the First Amendment because YouTube is not a state actor and does not perform a "quintessential public function" that implicates that constitutional provision). When rejecting such claims, courts have explained that plaintiffs have it backwards. The prevailing doctrine, the Ninth Circuit recently clarified, presumes that private companies, like individuals, have the constitutional right to choose what content to carry, no matter how big they are; consumers do not have a corollary right to speak freely on those services. *Id.*; *see also* Manhattan Cmty. Access Corp. v. Halleck, 139 S. Ct. 1921, 1928 (2019) (holding that the First Amendment "prohibits only governmental abridgment of speech," not "*private* abridgment of speech").

79. *See* 47 U.S.C. § 230(c)(2)(B) (2018).

have an obvious chilling effect" because of the "staggering" amount of third-party content that flows through their servers, some of which is surely unlawful or harmful.[80] The panel accordingly decided to read Section 230's protections broadly, but without specifying which of the three protections among Section 230(c)(1), 230(c)(2)(A), or 230(c)(2)(B) it was interpreting.[81] Otherwise, the Fourth Circuit reasoned, online companies would likely choose "to severely restrict the number and type of messages posted."[82] Under that logic, the doctrine should be broadly protective and generous "to avoid any such restrictive effect."[83] Moreover, the court inferred a congressional belief that intermediaries would have it in their commercial self-interest to regulate content to keep their consumers happy; consumer demand would be regulation enough and, in any case, a far better judge of which content ought to be allowed.[84]

Federal and state courts across the country have since adopted this reasoning.[85] Almost two decades later, in a case involving an online service that notoriously facilitated sex trafficking of minors, the First Circuit elaborated that this "hands-off approach is fully consistent with Congress's avowed desire to permit the continued development of the internet with minimal regulatory interference."[86]

But as understood by the courts, Congress did more than simply set out a legal protection for "interactive computer service[s]."[87] It privileged services that host and distribute user-generated content in particular by removing all

80. 129 F.3d at 331.

81. *Id.* The panel did not distinguish among the protections under §§ (c)(1), (c)(2)(A), or (c)(2)(B). It also altogether omitted the fact that Congress appeared to have been most interested in the operative Good Samaritan aims under the second of the provisions (§ 230(c)(2)), and not § 230(c)(1). *See* Force v. Facebook, Inc., 934 F.3d 53, 79-80 (2d Cir. 2019) (Katzmann, C.J., concurring in part and dissenting in part) ("Congress emphasized the narrow civil liability shield that became § 230(c)(2), rather than the broad rule of construction laid out in § 230(c)(1)."); *see also* Doe v. G.T.E. Corp., 347 F.3d 655, 659-60 (7th Cir. 2003) (speculating that such a broad protection is a disincentive to moderate tortious user-generated content).

82. *Zeran*, 129 F.3d at 331.

83. *Id.*

84. *Id.*

85. *See* Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 157, 1179-80 (9th Cir. 2008) (en banc).

86. Doe No. 1 v. Backpage.com, LLC, 817 F.3d 12, 19 (1st Cir. 2016).

87. 47 U.S.C. § 230(f)(2) (2018) ("The term 'interactive computer service' means 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 services offered by libraries or educational institutions.").

affirmative duties under law to monitor, moderate, or block illegal content.[88] Legislators, of course, did not invent online forums like these. Electronic bulletin boards, newsgroups, and similar online communities proliferated in the years before legislators enacted Section 230. Indeed, Congress intervened in 1996 because a trial-level state court decision in New York assigned secondary liability to an online service that distributed defamatory statements by one of their users.[89] Section 230 overturned that decision.[90] Through its legislation, Congress signaled a policy preference for a certain kind of user-focused service design.[91]

Silicon Valley responded almost immediately. Investors and internet entrepreneurs eagerly started developing services that feature "user-generated content" with the knowledge that they would not be held legally responsible for any of it.[92] Put differently, the protection under Section 230(c) and the *Zeran* rule that soon followed established a new disincentive for companies to create, develop, or showcase their own content. It is no surprise, then, that emergent companies at this early stage shied away from content production and instead created services for "user-generated content" without fear of legal exposure,[93] even if they knew or could reasonably anticipate that their consumers would use the new services to do harm. Congress and the courts, in short, have created the statutory equivalent of an invisibility cloak for services that feature (but do not contribute to) third-party content.[94]

---

88. *See Roommates.com*, 521 F.3d at 1171-72 ("The claim against the website was, in effect, that it failed to review each user-created profile to ensure that it wasn't defamatory. That is precisely the kind of activity for which Congress intended to grant absolution with the passage of Section 230. With respect to the defamatory content, the website operator was merely a passive conduit and thus could not be held liable for failing to detect and remove it."); Fed. Trade Comm'n v. LeadClick, 838 F.3d 158, 174 (2d Cir. 2016) (stating that Section 230 "bars 'lawsuits seeking to hold a service provider liable for its exercise of a publisher's traditional editorial functions—such as deciding whether to publish, withdraw, postpone or alter content'" provided by another for publication (quoting Jones v. Dirty World Ent. Recordings LLC, 755 F.3d 398, 407 (6th Cir. 2014))).

89. *See* 141 CONG. REC. 22,044-45 (statement of Rep. John Cox discussing *Stratton Oakmont v. Prodigy*, No. 031063/94, 1995 WL 323710 (N.Y. Sup. Ct. May 24, 1995)).

90. Force v. Facebook, 934 F.3d 53, 79 (2d Cir. 2019).

91. Zeran v. Am. Online, Inc., 129 F.3d 327, 330-31 (observing Section 230 reflects Congress's decision "not to deter harmful online speech through the separate route of imposing tort liability on companies that serve as intermediaries for other parties' potentially injurious messages").

92. KOSSEFF, *supra* note 9, at 120-22.

93. Content producers today like Amazon, Netflix, and Disney aggressively develop and control access to content on the internet. But they do not enjoy Section 230 protection generally.

94. The Supreme Court has recognized the "technology-forcing" aspect of law. *See* Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 49 (1983). *State Farm*, for example, is less known for its significant place in technology law than in administrative-law doctrine,

Ever since, a ready-made populist ideology has supplied more normative heft to this legal fiction, going beyond the laissez-faire justifications set out above. It posits that, with the internet, consumers no longer have to abide by the unilateral designs and service terms of powerful legacy media companies and retailers. The internet will empower consumers to be the architects of their own online experiences.[95] Social-media companies are the apotheosis in this framing. They are the necessary outgrowth of the legal protection that Congress created.

So, in spite of its title, the CDA discourages companies to act with decency. Under prevailing judicial interpretations of Section 230, companies are free to leave up or take down unlawful or harmful content as they please. This is a policy that, in practice, disincentivizes moderation and incentivizes the distribution of third-party material. It explains, at least in part, Silicon Valley companies' desire to hold consumer attention and collect consumer data. In this way, the laissez-faire policy approach set out by Congress and elaborated by the courts has become a perversion of the statute's titular objective. It is to this to which I turn next in Part III below.

III. COMMERCIAL DESIGNS AND PERVERSE INCENTIVES

Today, social-media companies do much more than simply host or distribute user-generated content. They solicit, sort, deliver, and amplify content that holds consumer attention for advertisers.[96]

Most companies' targeted-content delivery systems are not as sophisticated as those of large and powerful companies like Facebook and YouTube.[97] But they, too, fashion their sites with advertisers in mind.[98] The Experience Project, for example, was a website that innocuously aimed to make connections between anonymous users based on the information that users entered into a straightforward query box.[99] As with most popular online services today, automated decision-making systems were essential to building community groups on the site. Typing something as simple as "I like dogs" or "I believe in the paranormal"

but it provides an important lesson on the impact of law in prefiguring how companies develop and deploy technologies.

95. This is the conception that animates the idea of the internet "user."

96. *See, e.g.*, Hao, *supra* note 19.

97. *See* Matt Perault, *Well-Intentioned Section 230 Reform Could Entrench the Power of Big Tech*, SLATE (June 1, 2021), https://slate.com/technology/2021/06/section-230-reform-antitrust-big-tech-consolidation.html [https://perma.cc/2LPN-5LMH].

98. *See, e.g.*, Dryoff v. Ultimate Software Grp., 934 F.3d 1093, 1094-95 (9th Cir. 2019).

99. *Id.* at 1094.

would be enough for the service to make a connection.[100] It would also send users an email notification whenever other users on the site responded to related inquiries.[101] The company generated income through advertisements, donations, and the sale of tokens that users could spend to communicate with others in their groups.[102]

As the Experience Project's userbase grew, so too did the range of community groups that emerged. The company shuttered its service in 2016[103] because, as commentators in this field like to observe, "moderation at scale" is difficult if not impossible.[104] The Experience Project claimed it had to go offline because of the "bad apples" that were flocking to the site.[105] And by "bad apples," it was referring to a sexual predator who had used the site to entrap underage victims and a murderer who killed a woman he met through its services.[106]

Before it closed, the Experience Project's automated system sent a notification to Wesley Greer, who was using the site to meet people and find heroin.[107] Greer found what he was looking for—and more, as it turns out. He died of fentanyl poisoning after unknowingly purchasing heroin laced with fentanyl from another user.[108] Greer's mother sued the Experience Project for wrongful death.

---

100. Liz Spikol, *The Experience Project Is the Worst Social Networking Website*, PHILADELPHIA (May 8, 2013, 10:58 AM), https://www.phillymag.com/news/2013/05/08/social-network-experience-project-terrible [https://perma.cc/3TJP-DY9Q].

101. *Dryoff*, 934 F.3d at 1095.

102. *See* Answering Brief of Appellee Ultimate Software Grp., Inc. at 15, *Dryoff*, 934 F.3d 1093 (No. 3:17-cv-05359-LB).

103. *See Dryoff*, 934 F.3d at 1095.

104. *See* Mike Masnick, *Content Moderation at Scale Is Impossible: Recent Examples of Misunderstanding Context*, TECHDIRT (Feb. 26, 2021, 9:37 AM), https://www.techdirt.com/articles/20210225/10365146316/content-moderation-scale-is-impossible-recent-examples-misunderstanding-context.shtml [https://perma.cc/6UTM-8BRK] (arguing that large-scale moderation is impossible because of the difficulty of understanding context).

105. *Until We Meet Again*, EXPERIENCE PROJECT, http://www.experienceproject.com/until-we-meet-again [https://web.archive.org/web/20160409092804/http:/www.experienceproject.com/until-we-meet-again].

106. *See* Teri Knight, *Trial Too Traumatic for Victim's Family; Chilling Blaze Destroys Dairy Barn; Arts Guild Receives Large Grant*, KYMN RADIO (Jan. 13, 2017), https://kymnradio.net/2017/01/13/trial-traumatic-victims-family-chilling-blaze-kills-animals-arts-guild-receives-large-grant [https://perma.cc/37AG-BSE7]; Rebecca Roberts, *Sting Busts Troy Man Trying to Have Sex with 13-Year-Old Girl and Her Mom*, KTVI FOX 2 (Apr. 10, 2015), https://fox2now.com/news/sting-busts-troy-man-trying-to-have-sex-with-13-year-old-girl-and-her-mom [https://perma.cc/PTK7-Z6XD].

107. Bob Egelko, *Defunct Website Not Culpable in Death of Man from Fentanyl, Court Rules*, S.F. CHRON. (Aug. 21, 2019), https://www.sfchronicle.com/nation/article/Defunct-website-not-culpable-in-death-of-man-from-14368906.php [https://perma.cc/33LJ-YCV2].

108. *Id.*

She argued that he would not have obtained the drugs that killed him without its services. But she could not prevail in court. Her claim failed because the Experience Project was immune under Section 230 from legal liability. According to the court, the service merely connected users to likeminded people and communities they sought out.[109] Greer's mother could not even proceed to an initial hearing on the question of whether the Experience Project's service design somehow contributed to the death of her son.

As complicit as social-media sites may seem, courts have barred cases where plaintiffs allege legal fault for design features like anonymity,[110] notifications,[111] recommendations,[112] and location tracking.[113] They have held that those functions only help to deliver the user-generated content that the intermediary receives and, as such, do not rise to the level of "material contribution."[114] They have also repeatedly refused to impose duties on an "interactive computers service" to monitor for malicious use of their service or implement safety measures to protect against known informational harms.[115]

In many regards, this legal regime is upside down. By way of comparison, Greer's experience with the Experience Project resembles that of the young adults who have jumped off "the Vessel," a "spiraling staircase" in Manhattan's Hudson Yards with waist-high guardrails that whimsically climb sixteen stories into the air.[116] The developers closed access to visitors after a third young person committed suicide by leaping from the structure.[117] There can be little doubt that, like the Vessel, certain design features for distribution, delivery, and

---

109. Dryoff v. Ultimate Software Grp., 934 F.3d 1093, 1099 (9th Cir. 2019) ("The recommendation and notification functions helped facilitate this user-to-user communication, but it did not materially contribute, as Plaintiff argues, to the alleged unlawfulness of the content.").

110. *Id.* at 1100.

111. *Id.* at 1101.

112. *See* Force v. Facebook, Inc., 934 F.3d 53, 70 (2d Cir. 2019); Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1124 (9th Cir. 2003).

113. *See* Herrick v. Grindr LLC, 765 F. App'x. 586, 590-91 (2d Cir. 2019).

114. *Dryoff*, 934 F.3d at 1000.

115. *Herrick*, 765 F. App'x. at 589-90.

116. Ed Shanahan & Kimiko de Freytas-Tamura, *150-Foot Vessel Sculpture at Hudson Yards Closes After 3rd Suicide*, N.Y. TIMES (Jan. 12, 2021), https://www.nytimes.com/2021/01/12/nyregion/hudson-yards-suicide-vessel.html [https://perma.cc/X9N3-4FEX]; *accord* Snejana Farberov, *Vessel at NYC's Hudson Yards Is CLOSED Indefinitely Just Days After Murder Suspect Jumped to His Death From 16-Story Sculpture—The Third Suicide Since It Opened Less Than Two Years Ago*, DAILY MAIL (Jan. 14, 2021), https://www.dailymail.co.uk/news/article-9148201/Vessel-NYCs-Hudson-Yards-closed-indefinitely-suicide.html [https://perma.cc/R4ZD-LUDT].

117. Farberov, *supra* note 116.

amplification of information can sometimes be predictably dangerous, even if ostensibly innocuous.

Over the past couple of years, courts have become more scrutinizing of the ways in which interactive computer-service design impacts online behavior.[118] Consider the *Force v. Facebook, Inc.* case, decided in the Second Circuit.[119] There, plaintiffs had argued that Facebook materially supported terrorism by making friend recommendations and supporting online groups.[120] The panel rejected that argument, holding that Facebook could not be sued for enabling foreign terrorists to meet and collaborate in violation of federal antiterrorism laws.[121] The case is notable, however, because of then-Chief Judge Katzmann's separate concurring and dissenting opinion in which he argued that "the CDA does not protect Facebook's friend- and content-suggestion algorithms."[122] In Chief Judge Katzmann's reading, neither the statutory text nor the stated purposes of the statute supported the view that an intermediary gets immunity when it showcases user-generated data or content.[123] He would have held that Facebook's recommendations should not count as "publishing" under Section 230(c)(1) because Facebook is, first, communicating its own views about who among its users should be friends and, second, creating "real-world (if digital) connections" with demonstrably real-world consequences.[124]

Chief Judge Katzmann's separate opinion in *Force* marks an important inflection point in the evolution of the doctrine. Over the past couple of years in particular, courts have started to look far more carefully at the ways in which the designs of interactive computer services cause informational harm.[125] Chief

---

118. I have elsewhere argued that courts should be far more alert than they have been to the nuances of internet services' designs and targeted delivery of information to consumers, especially when they differentially impact racial minorities and other historically marginalized and vulnerable groups. *See* Sylvain, *Discriminatory Designs*, *supra* note 10; Sylvain, *Intermediary Duties*, *supra* note 10.

119. 934 F.3d 53 (2d Cir. 2019).

120. *Id*. at 59.

121. Force v. Facebook, Inc., 934 F.3d 53, 57 (2d Cir. 2019).

122. *Id.* at 82 (Katzmann, C.J., concurring in part and dissenting in part).

123. *Id.*

124. *Id.* at 82-83.

125. *See, e.g.*, Gonzalez v. Google LLC, 2 F.4th 871, 913-18 (9th Cir. 2021) (Berzon, J., concurring) (quoting and discussing Chief Judge Katzmann's opinion); *id.* at 919 (Gould, J., concurring in part and dissenting in part) (same); Lemmon v. Snap, Inc., 995 F.3d 1085, (9th Cir. 2021) (holding that the social-media company was not immune under Section 230 in a claim for negligent design). The Federal Trade Commission (FTC) has been at the forefront of this work. It successfully argued, for example, that an online company that promoted false stories about nutritional supplements across a network of affiliated websites was not a mere "publisher" within the meaning of Section 230. FTC v. LeadClick Media, LLC, 838 F.3d 158 (2d Cir.

Judge Katzmann's opinion cites *HomeAway.com v. City of Santa Monica*,[126] one of a handful of cases concerning municipal ordinances that impose nonpublishing duties on online homesharing services to report or register short-term rentals with public officials. In cases from Boston to New York to San Francisco, federal courts have rejected the Section 230 defense on the view that Section 230 "does not mandate 'a but-for test that would provide immunity . . . solely because a cause of action would not otherwise have accrued but for the third-party content.'"[127] An interactive computer service may, at once, distribute third-party content without fear of liability, but also be subject to legal duties surrounding how it designs and provides its systems to consumers. Thus, the courts have determined that homesharing sites, for example, are not free to ignore whether their guests and hosts are lawfully registered under local housing or hotelier laws.[128]

Amazon, too, has been on the losing end in federal- and state-court litigation in which plaintiffs have alleged that the retail behemoth is a seller (subject to products liability for product defects) rather than a mere publisher of information from third-party manufacturers.[129] Federal and state courts across the country have been taken by the way Amazon controls the marketing, pricing, delivery logistics, and general political economy of online consumer retail purchasing. Even if, in any given case, Amazon may not have a duty to warn or monitor third-party products, the courts have generally concluded that the work that it does behind the scenes is not "publishing."[130] It is, rather, a seller—for the purposes of product-liability law, at least. This is to say that, in the eyes of most

2016). The Tenth Circuit has also rejected the Section 230 defense in a federal enforcement action the FTC brought against a company that sold personal data about people to all comers, including telephone records, in violation the Telecommunications Act. FTC v. Accusearch Inc., 570 F.3d 1187, 1197-1201 (10th Cir. 2009). That company operated a website that sold confidential information it paid third-party researchers to acquire upon its request. This, for the panel, was not "information provided by another information content provider," within the meaning of Section 230. *Id.* at 1197. The company was "responsible for the development of the specific content that was the source of the alleged liability." *Id.* at 1198.

126. 918 F.3d 676 (9th Cir. 2019).

127. *Force*, 934 F.3d at 82 (Katzmann, C.J., concurring in part and dissenting in part) (quoting *HomeAway.com, Inc.*, 918 F.3d at 682) (internal quotation marks omitted); *see also HomeAway.com, Inc.*, 918 F.3d at 682 (rejecting the but-for test); Airbnb v. City of Boston, 386 F. Supp. 3d 113, 120-22 (D. Mass. 2019) (same).

128. *See, e.g.*, *HomeAway.com, Inc.*, 918 F.3d 676; Airbnb, Inc. v. City of San Francisco, 217 F. Supp. 3d 1066 (N.D. Cal. 2016).

129. *See, e.g.*, Oberdorf v. Amazon.com Inc., 930 F.3d 136, 153 (3d Cir. 2019), *vacated*, 936 F.3d 182 (3d Cir. 2019); Erie Ins. Co. v. Amazon.com, Inc., 925 F.3d 135, 137 (4th Cir. 2019); Bolger v. Amazon.com, LLC, 267 Cal. Rptr. 3d 601, 604 (Ct. App. 2020); State Farm Fire & Cas. Co. v. Amazon.com Servs., Inc., 137 N.Y.S.3d 884, 887-89 (Sup. Ct. 2020).

130. *See, e.g.*, Bolger v. Amazon.com, LLC, 267 Cal. Rptr. 3d 601, 626 (Ct. App. 2020).

courts, Amazon is a market actor deeply embedded in the marketing and pricing of consumer products, even if not their manufacturing. An exception, however, arises in Texas, where the court of last resort held that Amazon is not a "seller" under state law if the originating manufacturer "do[es] not relinquish title to [its] products."[131]

Finally, in *Lemmon v. Snap, Inc.*, the Ninth Circuit rejected a Section 230 defense to a claim alleging wrongful death. In that case, the plaintiffs argued that Snapchat, a popular social-media app through which users share disappearing photos and videos with annotations, negligently contributed to their teenage sons' fatal car crash.[132] The Snapchat feature at issue allows users to track their land speed and share that information with friends through a "Speed Filter," which superimposed a real-time speedometer over another image.[133] One plaintiff's son was allegedly using this filter while driving, shortly before running off the road at 113 miles per hour and ramming into a tree.[134] The other plaintiff was in the passenger seat. According to the plaintiffs, Snap (the owner of Snapchat) knew that users went faster than 100 miles per hour on the mistaken belief that they would be rewarded with in-app "trophies,"[135] but it did not do anything that effectively dissuaded them from doing so.[136] The young occupants' parents sued, alleging that Snap's "Speed Filter" negligently caused their son's death.[137] Snapchat moved to dismiss, arguing that the parents' suit sought to impose liability for publishing user content—here, the driving speed.

The Ninth Circuit rejected Snap's Section 230 defense, reversing the lower court's decision. The panel concluded that the plaintiffs' claims targeted the

---

131. *See* McMillan v. Amazon.com, Inc., 2 F.4th 525, 525 (5th Cir. 2021) (mem.) (reversing trial court after seeking and receiving answer from the Supreme Court of Texas in *Amazon.com, Inc. v. McMillan*, 625 S.W.3d 101 (Tex. 2021), on the certified question of whether Amazon could be a seller under state law).

132. Lemmon v. Snap, Inc., 995 F.3d 1085, 1088-90 (9th Cir. 2021).

133. *Id.* at 1088. Snap discontinued the Speed Filter in 2021 following criticism that it encouraged reckless driving. *See* Bobby Allyn, *Snapchat Ends 'Speed Filter' That Critics Say Encouraged Reckless Driving*, NPR (June 17, 2021, 11:58 AM ET), https://www.npr.org/2021/06/17/1007385955/snapchat-ends-speed-filter-that-critics-say-encouraged-reckless-driving [https://perma.cc/N7CX-HRS5].

134. 995 F.3d at 1088.

135. *Id.* at 1088-89 (describing the ways in which Snapchat rewards users "with trophies, streaks, and social recognitions based on the snaps they send" and some users' suspicions that these rewards are unlocked for using the Speed Filter at more than 100 mph).

136. *Id.* at 1089-90.

137. *Id.* at 1090. The court laid out as a requirement for the plaintiffs' negligence claim that "a reasonable person would conclude that 'the reasonably foreseeable harm'" of Snap's Speed Filter outweighs its utility. *Id.* at 1092 (quoting Merrill v. Navegar, Inc., 28 P.3d 116, 125 (Cal. 2001)).

design of the application, not Snap as a publisher.[138] According to the Ninth Circuit, plaintiffs' negligent-design claim faulted Snap solely for Snapchat's design; the parents contended that the application's "Speed Filter and reward system worked together to encourage users to drive at dangerous speeds."[139] It does not matter, the panel explained, that the company provides "neutral tools," as long as plaintiffs' allegations are not addressed to the content that users generate with those tools.[140] This conclusion, echoing Chief Judge Katzmann's partial concurrence in *Force*, may very well instigate creative new lawsuits that could better tailor Section 230 doctrine to our times.[141]

This emerging view among judges is refreshing because it suggests that courts are no longer so immediately taken by the pretense that internet companies are mere publishers or distributors of user-generated content — or that they are mere platforms that do little more than facilitate free expression and human connection. Sometimes they are. But often they are not. The sooner that policy makers dispense with the romantic story about beneficent online platforms for user-generated content and recognize their uncontestable pecuniary aims, the better.[142]

IV. INFORMATIONAL INEQUALITY

The problem with prevailing Section 230 doctrine today is not only that it protects online services that amplify and deliver misleading or dangerous information by design. To be sure, this is bad enough because the entities that are most responsible for distributing and delivering illicit or dangerous content are the least likely to be held responsible for it. But the principal problems that this Part highlights are the ways in which powerful online application and service designs harm people for whom hard-fought public-law consumer protections (such as civil-rights laws or rules against unfair or deceptive trade practices) are essential. And yet, under the prevailing Section 230 doctrine, it appears that

138. *Id.* at 1093-94.

139. *Id.* at 1093.

140. *Id.* at 1094.

141. *See, e.g.*, Sam Dean, *A Teen Who Was Bullied on Snapchat Died. His Mom Is Suing to Hold Social Media Liable*, L.A. TIMES (May 10, 2021), https://www.latimes.com/business/story/2021-05-10/lawsuit-snap-teen-suicide-yolo-lmk [https://perma.cc/RTW2-DWQE].

142. *Cf.* Danielle Keats Citron & Mary Anne Franks, *The Internet as a Speech Machine and Other Myths Confounding Section 230 Reform*, 2020 U. CHI. LEGAL F. 45, 47 (noting that while "their assessments of the problem differ, lawmakers agree that Section 230 needs fixing . . . [i]n a few short years, Section 230 reform efforts have evolved from academic fantasy to legislative reality").

interactive computer services can never be held liable for any of the harms that they contribute to or cause.

What is worse, most courts resolve Section 230 disputes at the motion to dismiss stage, before discovery can unfold.[143] They do this even as the text of the statute does not explicitly advert that intention. Courts have chosen to limit their analysis to the question of whether plaintiffs have alleged that the respective defendant interactive computer service is acting as a publisher or distributor of third-party content or is materially contributing to unlawful content.[144] Courts rarely, if ever, scrutinize how deeply involved the defendant service is in creating or developing the offending content.[145] In practice, then, the doctrine has effectively foreclosed any opportunity for the vast majority of plaintiffs (or the public generally) to scrutinize internet companies' role in the alleged harm.[146] This presents a substantial hurdle to holding intermediaries accountable, even if, in the end, they are not liable pursuant to plaintiffs' legal theory of the case. This is especially true given the ways in which the most powerful internet companies today zealously resist public scrutiny of the systems that animate user experience.[147]

My reform proposal is simple: online intermediaries should not be immune from liability to the extent that their service designs produce outcomes that conflict with hard-won but settled legal protections for consumers—including consumer-protection and civil-rights laws and regulations. At a minimum, law in this area should be far more skeptical of online intermediaries to the extent that they know that their systems are causing informational harms and that they have the capacity to stop or prevent. And that burden should be substantial when the offending content or online conduct harms consumers and members of historically marginalized groups.[148]

---

143. *See, e.g.*, Force v. Facebook, Inc., 934 F.3d 53 (2d. Cir. 2019); Jane Doe No. 1 v. Backpage.com, LLC, 817 F.3d 12 (1st Cir. 2016).

144. *See, e.g.*, Zeran v. Am. Online, Inc., 129 F.3d 327, 331 (4th Cir. 1997); Fair Hous. Council v. Roommates.com, LLC, 521 F.3d 1157, 1168 (9th Cir. 2008).

145. *See Zeran*, 129 F.3d at 331; *Roomates.com*, 521 F.3d at 1168. *But see* FTC v. Accusearch, 570 F.3d 1187, 1190 (10th Cir. 2009) (holding Accusearch, an internet-service provider, accountable for published content on the grounds that "Accusearch's actions were not 'neutral' with respect to generating offensive content . . . its actions were intended to generate such content").

146. *See* Goldman, *supra* note 73, at 39–42.

147. *Cf.* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 8 (2015) ("Without knowing what Google actually *does* when it ranks sites, we cannot assess when it is acting in good faith to help users . . . [a]ll these [algorithms] are protected by laws of secrecy and technologies of obfuscation.").

148. I do not here take up the question of which agency should do this, though it seems obvious that the Federal Trade Commission would be the best fit.

A. Knowingly Entrenching Inequality

Informational harms spread unevenly across politically or culturally salient groups. Owing to the way in which structural inequality permeates all aspects of society, historically marginalized groups and consumers are likely to be the worst off in the absence of regulatory checks.[149] Safiya Noble identified this problem a few years ago in the context of online search.[150] She explained that ostensibly innocuous terms return and, with each query, entrench prevailing racist and misogynist meanings. In other words, before Google rectified this problem, when someone searched for "black girls," the top results were likely to include sexualized or debasing terms, while the top search results for "white girls" tended to be less degrading.

These same problems continue, even as Noble's writing (and that of others) has raised awareness about the ways in which putatively neutral technologies perpetuate or entrench extant inequalities.[151] These consequences could be high-stakes — even life-or-death. Consider, for instance, that some social-media companies distributed information about COVID-19 safety, treatment, and vaccinations to Black people far less than to other groups during the height of the pandemic.[152] Indeed, those belonging to all other racial categories saw significantly more public-health announcements from the Department of Health and Human Services and other public-health bodies.[153] Consider also the ways in

---

149. *See generally* DARIA ROTHMAYR, REPRODUCING RACISM: HOW EVERYDAY CHOICES LOCK IN WHITE ADVANTAGE (2014) (arguing that white socioeconomic advantage is self-perpetuating); William Julius Wilson, *Toward a Framework for Understanding Forces That Contribute to or Reinforce Racial Inequality*, 1 RACE & SOC. PROBS. 3 (2009) (proposing a framework for understanding the formation and perpetuation of racial inequality).

150. *See generally* SAFIYA NOBLE, ALGORITHMS OF OPPRESSION: HOW SEARCH ENGINES REINFORCE RACISM (2018); Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROC. MACH. LEARNING RSCH. 77 (2018), (showing that automated facial-recognition software has lower error rates for lighter-skinned men than darker-skinned women).

151. *See* Michelle Ruiz, *Safiya Noble Knew the Algorithm Was Oppressive*, VOGUE (Oct. 21, 2021), https://www.vogue.com/article/safiya-noble [https://perma.cc/UKZ6-3ZZQ] (discussing ongoing problems); Chris Stokel-Walker, *Is Social Media Racist by Design?*, ESQUIRE (Nov. 20, 2020), https://www.esquire.com/uk/culture/a34532613/social-media-racism [https://perma.cc/GAJ8-YLXJ] (same).

152. Corin Faife & Dara Kerr, *Official Information About COVID-19 Is Reaching Fewer Black People on Facebook*, MARKUP (Mar. 4, 2021, 8:00 AM), https://themarkup.org/citizen-browser/2021/03/04/official-information-about-covid-19-is-reaching-fewer-black-people-on-facebook [https://perma.cc/XK8B-MPT9].

153. *Id.* Perversely, reports suggest that these communities were also more likely to suffer from biased risk-prediction algorithms in medicine. *See* Donna M. Christensen, Jim Manley & Jason Resendez, *Medical Algorithms Are Failing Communities of Color*, HEALTH AFFS. (Sept. 9,

which political operatives sought to deflate faith in the administration of elections among Black and Brown people with misleading and false information, effectively disenfranchising those groups.[154]

These are direct informational harms that social media can stop and prevent. There can be little doubt about this last point. In the lead up to the 2020 election, for example, Facebook proudly announced that it would ratchet down content from putative news sources that are notorious for distributing disinformation, only to revert back to amplifying that material in the month or so after electors registered their votes.[155]

There are few legal remedies for the distribution of high-stakes falsehoods and informational harms like these. Of course, there are rules that forbid the distribution of false election or nutritional or ingredient information—that is, information about the time and place of an election or information about food and drugs.[156] But these laws probably do not prohibit falsity in campaign material or dangerous medical advice from laypeople.[157] In which case, online companies would owe no legal obligation to take these down under a reformed Section 230 doctrine.[158]

But there are certain kinds of informational harms for which the stakes are so high that their distribution is or should be unlawful. Current laws and judicial

---

2021), https://www.healthaffairs.org/do/10.1377/hblog20210903.976632/full [https://perma.cc/C4KK-HZDS].

154. Shannon Bond, *Black and Latino Voters Flooded with Disinformation in Election's Final Days*, NPR (Oct. 30, 2020, 7:49 AM), https://www.npr.org/2020/10/30/929248146/black-and-latino-voters-flooded-with-disinformation-in-elections-final-days [https://perma.cc/3GWF-TGRA].

155. Nick Statt, *Facebook Tweaked the News Feed to Highlight More Mainstream News Sources After the Election*, VERGE (Nov. 24, 2020, 11:35 AM), https://www.theverge.com/2020/11/24/21612728/facebook-news-feed-us-election-change-mainstream-news-misinformation [https://perma.cc/FW32-VARP].

156. *See* Minn. Voters All. v. Mansky, 138 S. Ct. 1876, 1889 n.4 (2018) ("We do not doubt that the State may prohibit messages intended to mislead voters about voting requirements and procedures."); Bolger v. Youngs Drugs Prods. Corp., 463 U.S. 60, 68 (1983) ("Advertisers should not be permitted to immunize false or misleading product information from government regulation simply by including references to public issues.").

157. *See generally* Richard Hasen, *Drawing the Line Between False Election Speech and False Campaign Speech*, KNIGHT FIRST AMEND. INST. COLUM. UNIV. (Oct. 12, 2021), https://knightcolumbia.org/blog/drawing-the-line-between-false-election-speech-and-false-campaign-speech [https://perma.cc/Q2TL-WPLJ].

158. This is say nothing of the First Amendment. *See* Prager Univ. v. Google, 951 F.3d 991 (9th Cir. 2020) (holding that the state-action doctrine under the First Amendment bars constitutional scrutiny of an internet company's decisions to remove or takedown third-party content); *cf.* United States v. Alvarez, 567 U.S. 7090 (2012) (holding that a federal statute that criminalized false claims about military service is unconstitutional because it "targets falsity and nothing more").

doctrine, for example, strictly forbid advertising that discriminates against people on the basis of protected classifications like race, gender, or age in high stakes areas like hiring, consumer finance, and housing.[159] These are rules that apply with equal force to advertisers and third-party intermediaries (that are not online) as much as the companies whose services and products are being marketed.[160] And yet, due to the protection for interactive computer services under Section 230, restrictive discriminatory advertising has proliferated across these sectors.[161] Under the broad and prevailing interpretation of Section 230(c)(1), nothing in the law obligates these companies to take down or prevent its distribution even when they know it exists.

Online companies' capacity to control the delivery of these unlawful kinds of content, particularly when the disparities ostensibly leave historically marginalized groups less well off, would go unaddressed were it not for intrepid journalism. More to the point, the prevailing interpretation of Section 230 permits companies like these to rest easy and proceed "unfettered,"[162] even when they know that their services facilitate informational disparities. Reddit, like Backpage before it, knowingly hosts child pornography.[163] Dating sites like Tinder and OKCupid do not screen for sexual predators.[164] The law likely does not require

---

159. *See, e.g.*, 42 U.S.C. § 2000e-2(k) (2018) (rules under Title VII of the Civil Rights Act); 12 C.F.R. pt. 1002, supp. I, ¶ 6(a)-2 (2021) (rules under the Equal Credit Opportunity Act); 24 C.F.R. § 100.500(c)(1) (2021) (rules under the Fair Housing Act); *see also* Tex. Dep't of Hous. & Cmty. Affs. v. Inclusive Cmtys., 135 S. Ct. 2507, 2521 (2015) (recognizing that disparate-impact claims are "consistent with the [Fair Housing Act's] central purpose"). *See generally* Stephen Hayes & Kali Schellenberg, *Discrimination Is "Unfair": Interpreting UDA(A)P to Prohibit Discrimination*, STUDENT BORROWER PROT. CTR. (Apr. 2021), https://protectborrowers.org/wp-content/uploads/2021/04/Discrimination_is_Unfair.pdf [https://perma.cc/JSV7-AKNK] (proposing a theory of unfairness under Section 5 of the FTC Act, codified at 15 U.S.C. § 45, that would help to fill gaps in current antidiscrimination law).

160. *See, e.g.*, 42 U.S.C. § 3604(c) (2018) (Fair Housing Act); 12 C.F.R. pt. 1002, supp. I, ¶ 6(a)-2 (2021) (rules implementing the Fair Housing Act).

161. *See, e.g.*, Aaron Rieke & Corrine Yu, *Discrimination's Digital Frontier*, ATLANTIC (Apr. 15, 2019), https://www.theatlantic.com/ideas/archive/2019/04/facebook-targeted-marketing-perpetuates-discrimination/587059 [https://perma.cc/2PJ7-7B8C]; Rory Cellan-Jones, *Facebook Accused of Allowing Sexist Job Advertising*, BBC (Sept. 9, 2019), https://www.bbc.com/news/technology-58487026 [https://perma.cc/7BZT-VYC4].

162. 47 U.S.C. § 230(b)(2) (2018).

163. Bob Van Voris, *Reddit Provides 'Safe Haven' for Child Porn, Lawsuit Claims*, BLOOMBERG L. (Apr. 23, 2021), https://www.bloomberglaw.com/bloomberglawnews/tech-and-telecom-law/X45M1264000000 [https://perma.cc/2QN9-RSQF]; *see also* Doe No. 1 v. Backpage.com, LLC, 817 F.3d 12 (1st Cir. 2016).

164. Hillary Flynn, Keith Cousins & Elizabeth Naismith Picciani, *Tinder Lets Known Sex Offenders Use the App. It's Not the Only One*, BUZZFEED NEWS (Dec. 2, 2019, 10:30 AM),

them to do anything in these settings — even when they have knowledge that illegal conduct is afoot.[165]

This prevailing interpretation contorts Section 230's purposes in at least two ways. First, it flips the ostensible Good Samaritan purpose of the statute upside down by removing the burden to abide by consumer-protection laws. The doctrine in this way creates a disincentive to care rather than an incentive to self-regulate, as the *Zeran* panel presumed.[166] Second, it leaves consumers without any effective legal mechanism to mitigate or redress online informational harms. Danielle Citron, Ben Wittes, and Mary Anne Franks have proposed that intermediaries enjoy the benefit of the Good Samaritan safe harbor if they take "reasonable steps to prevent or address unlawful uses of" their service.[167] This reform aims to operationalize the stated objectives of the Good Samaritan purposes of the statute by creating a functional safe harbor for companies that actually regulate third-party content. Holding companies responsible for the harmful material to take reasonable steps to takedown or block such content would help to redress some of the power imbalances at work, particularly because those companies control content distribution. Such a duty would also shift some of the costs of unlawful or harmful content on to the entities best equipped to prevent them.

Legislators could also impose a burden on the companies that know that their services cause harm. There is nothing earth shattering in this idea. After all, it has been taken as an article of faith among torts professors for over five decades now.[168] As to internet companies in particular, this reform idea only echoes insights set out over a decade ago by Rebecca Tushnet in a law-review essay on the

---

https://www.buzzfeednews.com/article/hillaryflynn/tinder-lets-known-sex-offenders-use-the-app-its-not-the [https://perma.cc/RHH6-XY8H].

165. Herrick v. Grindr LLC, 765 F. App'x 586, 591 (2d Cir. 2019); Doe No. 14 v. Internet Brands, Inc., 824 F.3d 846, 853 (9th Cir. 2016). Under the prevailing doctrine, companies only owe a duty to warn their users about third-party malfeasance on their site when they obtain actual knowledge of such activity "from an outside source" (i.e., not through any content or material that consumers directly post to their services). *Internet Brands*, 824 F.3d at 853. In that circumstance, they are no longer acting as an online publisher within the meaning of the statute. *Id.*; *see also* HomeAway.com, Inc. v. City of Santa Monica, 918 F.3d 676, 682 (9th Cir. 2019) (finding that, in *Internet Brands*, the "underlying legal duty at issue did not seek to hold the defendant liable as a 'publisher or speaker' of third-party content" (quoting *Internet Brands*, 824 F.3d at 853)).

166. Zeran v. Am. Online, Inc., 129 F.3d 327, 331 (4th Cir. 1997).

167. Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 FORDHAM L. REV. 401, 419 (2017); Citron & Franks, *supra* note 142, at 71.

168. *See* GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS 312 (1970) (arguing that costs should be borne by those best-equipped to prevent them). *See generally* RICHARD A. POSNER, TORT LAW: CASES AND ECONOMIC ANALYSIS (1982) (same).

subject.[169] There, just twelve years after Congress passed Section 230, she presciently argued that legislators and policy makers should be alert to the relative social costs of granting broad speech rights or immunity to intermediaries for third-party material.[170] It is not at all obvious, she explained, that those protections would engender a sense of responsibility to moderate because the pecuniary aims of those companies, even then, did not necessarily align with those of users.[171] Thus, she explained, courts have adjusted "the procedure, rather than the substance, of speech torts in order to balance the costs of harmful speech with the benefits of speech that is useful but vulnerable to chilling effects."[172]

The law of defamation in particular offers helpful insights into how Section 230 reformers might tinker with knowledge requirements in order to serve other important public-regarding norms. Among other considerations, defamation law conditions the size and nature of penalties on intentionality as well as the respective target of the content.[173] Specifically, journalistic norms of truth-seeking and verification have been important to the courts' adjudication of defamation claims that require courts to determine whether a defendant journalist acted maliciously.[174] The qualified reporter's privilege also offers some helpful

---

169. *See* Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 76 GEO. WASH. L. REV. 986 (2008).

170. *Id*. at 1002-04.

171. *Id.* at 1010-11 ("Ironically – given § 230's title – immunity alone has not generally been sufficient to convince ISPs to monitor content.").

172. *Id*. at 1013.

173. *Id*. at 1013-14. I mention here without discussing in any detail that Justices Thomas and Gorsuch recently mused about reforming the law of defamation for the internet age in their separate dissents to the Court's decision to deny certiorari in *Berisha v. Lawson*. 141 S. Ct. 2424, 2424 (2021) (mem.) (Thomas, J., dissenting from the denial of certiorari); *id.* at 2425 (Gorsuch, J., dissenting from the denial of certiorari). There, Thomas characteristically observed that there is nothing in the constitutional text to support the public-figure rule set out in *New York Times v. Sullivan*, the seminal case in the area of defamation law. *Id.* at 2425 (Thomas, J., dissenting). Gorsuch meanwhile doubts that the technological changes to the "media landscape" support such a rule. *Id.* at 2427 (Gorsuch, J., dissenting). They are not the only Supreme Court Justices to have expressed reservations about the current doctrine. When she was an academic, Justice Elena Kagan penned two articles on the topic, expressing skepticism toward the doctrine if not the same level of hostility as Thomas or Gorsuch. *See* Elena Kagan, *A Libel Story:* Sullivan *Then and Now*, 18 L. & SOC. INQUIRY 197, 202 (1993) (reviewing ANTHONY LEWIS, MAKE NO LAW: THE SULLIVAN CASE AND THE FIRST AMENDMENT (1991)); Elena Kagan, *Libel and the First Amendment*, *in* ENCYCLOPEDIA OF THE AMERICAN CONSTITUTION 1608-09 (Leonard W. Levy & Kenneth L. Karst eds., 2d ed. 2000).

174. *See, e.g.*, St. Amant v. Thompson, 390 U.S. 727, 733 (1968); Dodds v. Am. Broad. Co., 145 F.3d 1053, 1062 (9th Cir. 1998); Tavoulareas v. Piro, 817 F.2d 762, 791 (D.C. Cir. 1987). The professional journalistic norms of truth-seeking and verification do not fit well with today's attention-driven business models for online companies. This might be because these norms are inapposite, having emerged from the particular nineteenth-century political economy of the

cues. There, courts explicitly consider whether, during the course of grand jury proceedings or even in a criminal trial, a subpoenaed reporter who bears the indicia of a professional journalist may decline a government request to testify.[175] In the context of both defamation and the reporter's privilege, the question of whether the defendant or witness engages in the process of truth seeking is not dispositive, but important. These doctrines protect companies' good-faith efforts to attend to the quality or veracity of the content they publish.

It would be fully consistent with these longstanding principles to obligate companies with the knowledge of and demonstrable capacity to control the (automated) distribution, amplification, and delivery of harmful content. Two reform proposals in the current session of the House of Representatives do precisely that.[176] One, for example, would exempt from Section 230 protection services that amplify online material that violates civil rights in particular.[177] Another proposal would carve out civil-rights violations, as well as other informational harms to consumers and historically marginalized groups.[178] This reform, moreover, would not have to require that companies affirmatively monitor for illegal or illicit content that they publish or amplify. One of the more notable bipartisan recent reform proposals in the Senate, for example, would require interactive computer services to remove any material that a court has adjudged to be unlawful within four days of receiving notice of that order.[179]

---

mass media, print news business. Midcentury regulatory interventions like the Federal Communications Commission's (FCC's) fairness doctrine and equal-time rule—policy reforms to which a handful of contemporary activists and writers have sought to retrofit to the internet—do not reveal much about this because objectivity and verification had emerged decades before the FCC promulgated those rules. This is to say nothing of whether such rules are administrable in an ecosystem as complex and varied as the current networked information market.

175. *See, e.g.*, Farr v. Pitchess, 522 F.2d 464 (9th Cir. 1975); *In re* Brown & Williamson Tobacco Corp. v. Wigand, 228 A.D. 187 (N.Y. App. Div. 1996). *See generally* Branzburg v. Hayes, 408 U.S. 632 (1972) (Powell, J., concurring) (recognizing a qualified privilege for journalists to maintain the confidentiality of a newsperson's sources and information).

176. *See* Protecting Americans from Dangerous Algorithms Act, H.R. 8636, 116th Cong. (2020); Justice Against Malicious Algorithms Act, H.R. 5596, 117th Cong. (2021).

177. *See* H.R. 8636. This provision would do better to cover all civil-rights laws rather than carve out amplification that violates any specific statute—let alone one that is not as commonly invoked by aggrieved plaintiffs.

178. *See* Press Release, Sen. Mark R. Warner, Warner, Hirono, Klobuchar Announce the Safe Tech Act to Reform Section 230 (Feb. 5, 2021), https://www.warner.senate.gov/public/index.cfm/2021/2/warner-hirono-klobuchar-announce-the-safe-tech-act-to-reform-section-230 [https://perma.cc/P2BD-DLKQ].

179. *See* Platform Accountability and Consumer Transparency Act, S. 797, 117th Cong. § 5 (c)(1)(A)(i) (2021).

B. Discriminatory Restrictive Targeting

Internet companies also facilitate discrimination on their platforms when they deliver targeted advertising that is personal to each consumer. These advertisements present a different kind of harm because, most of the time, their scope is difficult for outsiders to discern. Worse, however, are the ways in which intermediaries openly enable advertisers to target audiences on the basis of protected categories like race, gender, and age (among hundreds, if not thousands, of other characteristics) in commercial campaigns for housing, employment, and consumer finance, where hard fought civil-rights laws forbid advertisements that explicitly or intentionally solicit or exclude audiences on the basis of those dimensions.[180] Google reportedly allowed employers and landlords to discriminate against nonbinary and some transgender people, pledging to crack down on the practice only after being alerted to it by journalists.[181] Facebook's Ad Manager openly enables advertisers to target audiences by including or excluding people on the basis of thousands of demographic categories or proxies for those categories.[182]

Again, restrictive targeting is especially pernicious because victims are never the worse for knowing that they have been excluded. That is, even as many social-media companies offer individual users relatively particularized explanations for why they see any given advertisement, consumers do not learn about the content that they do not see. It takes resource-intensive analyses by intrepid researchers from outside of the company to uncover these practices. Unrelenting "data journalism" by *Politico* and *The Markup* in particular has uncovered patterns of discrimination in housing, employment, and consumer credit on

---

180. Julia Angwin & Terry Parris Jr., *Facebook Lets Advertisers Exclude Users by Race*, PROPUBLICA (Oct. 28, 2016, 1:00 PM), https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race [https://perma.cc/852Y-7RZS]. *See generally* Sylvain, *Discriminatory Designs*, *supra* note 10 (discussing how Facebook Ads enables the exclusion of racial minorities from the audience for housing advertisements and old people from the audience for job advertisements); Sylvain, *Intermediary Duties*, *supra* note 10 (discussing the ability to target different demographic groups with social-media advertisements); Gillian B. White, *When Algorithms Don't Account for Civil Rights*, ATLANTIC (Mar. 7, 2017), https://www.theatlantic.com/business/archive/2017/03/facebook-ad-discrimination/518718 [https://perma.cc/S9UC-PU83] (discussing discriminatory advertising on social-media platforms and other websites).

181. Jeremy B. Merrill, *Google Has Been Allowing Advertisers to Exclude Nonbinary People from Seeing Job Ads*, MARKUP (Feb. 11, 2021, 8:00 AM), https://themarkup.org/google-the-giant/2021/02/11/google-has-been-allowing-advertisers-to-exclude-nonbinary-people-from-seeing-job-ads [https://perma.cc/P2PU-JH6W].

182. Alexia Fernández Campbell, *Job Ads on Facebook Discriminated Against Women and Older Workers, EEOC Says*, VOX (Sept. 25, 2019, 2:20 PM), https://www.vox.com/identities/2019/9/25/20883446/facebook-job-ads-discrimination [https://perma.cc/L5A7-4FN4].

Facebook's Ad Manager.[183] For example, they have revealed that the Ad Manager allows credit-card companies, housing brokers, and employers to exclude, respectively, young people, racial minorities, and women from advertisements about their services and programs.[184] Facebook eventually promised to stop the practice, agreeing with civil-rights groups and plaintiffs in 2019 to bar the use of protected categories in those areas.[185] But in spite of these promises, patterns of age and sex discrimination persist.[186]

The solutions here are straightforward. Commercial content of any kind, whether online or in the physical world, that has the direct effect of discriminating against consumers on the basis of legislatively protected characteristics (e.g., race, sex, and age) in markets where civil-rights laws forbid the practice (e.g., housing, education, employment, and consumer finance) is (or should be) forbidden.[187] This reform would attend to outcomes rather than the input variables or decision-making processes on which companies rely to deliver content. This is because automated decision-making systems discover salient patterns in the combination of the most innocuous consumer variables (like food tastes and thousands of others), even when developers exclude them; together, they apparently act as a virtual proxy for protected categories.[188] Recall that, in spite of the

183. The principal person to credit for this line of reporting is Julia Angwin who, before starting *The Markup*, led teams that reported on the Ad Manager at *Politico*.

184. *See* Angwin & Parris, *supra* note 180; *see also supra* note 181 (discussing Google allowing advertisers to exclude nonbinary people from job advertisements); *infra* note 186 (discussing Facebook's advertisement algorithms excluding women from job advertisements).

185. Pema Levy, *Facebook Settles Civil Rights Lawsuits Over Ad Discrimination*, MOTHER JONES (Mar. 19, 2019), https://www.motherjones.com/politics/2019/03/facebook-settles-civil-rights-lawsuits-over-ad-discrimination [https://perma.cc/F76Z-G7WN]; *Fair Housing Groups Settle Lawsuit with Facebook: Transforms Facebook's Ad Platform Impacting Millions of Users*, NAT'L FAIR HOUS. ALL. (Mar. 18, 2019), https://nationalfairhousing.org/2019/03/18/national-fair-housing-alliance-settles-lawsuit-with-facebook-transforms-facebooks-ad-platform-impacting-millions-of-users [https://perma.cc/2G5S-Z5PE]. I was a consultant in one of the five civil-rights cases that Facebook settled in March 2019. *See* Olivier Sylvain, *A Watchful Eye on Facebook's Advertising Practices*, N.Y. TIMES (Mar. 28, 2019), https://www.nytimes.com/2019/03/28/opinion/facebook-ad-discrimination-race.html [https://perma.cc/RBA9-S4GK].

186. Hao, *supra* note 3; Faife & Ng, *supra* note 3.

187. *See, e.g.*, 42 U.S.C. § 3604(c) (2018); 12 C.F.R. § 1002.4 (2021).

188. *See* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 677-87 (2016). *See generally* Talia Gillis, *The Input Fallacy*, MINN. L. REV. (forthcoming 2022), https://ssrn.com/abstract=3571266 [https://perma.cc/L58J-2AYA] (showing that laws that forbid the use of race, gender, and other protected categories as input variables in automated decision-making systems in lending do not protect against discriminatory outcomes on those dimensions). Such outcomes may also be the result of defects in the problem definition in the early stages of development, the nature and source of the data on which the system is "trained" and "tested," or a mismatch with the specific context in which they are deployed. *See* Barocas & Selbst, *supra*.

March 2019 settlement with plaintiffs in which Facebook agreed to forbid the use of protected categories in ads for housing, employment, and consumer finance, patterns of discrimination on the basis of age and gender persisted over two years after the settlement.[189] Companies in these circumstances should not be able to invoke the Section 230 defense — at least not before discovery uncovers whether and how they serve advertisements to their users in ways that violate civil rights and other consumer protections.

Proponents of the current regime may protest that such reforms would unduly burden companies and chill the distribution of too much lawful third-party content. But that is a burden that the drafters and advocates of these hard-fought laws imposed in consideration of the social costs of discrimination and disparate consumer harm. The challenge is finding the right balance between innovation and speech on the one hand and equality on the other.

Today, Section 230 protections have set this balance exceptionally out of whack. The burdens that practically all other companies in practically all other legislative fields must carry should certainly be applicable to social-media companies, too, especially considering their enormous influence on public life. Of course, the reforms I propose would not eradicate racism and other consumer harms from internet platforms. Patterns of subordination pervade all public life. But if civil-rights law and other consumer protections are to be effective in today's networked information economy, internet companies ought to be held accountable for the ways in which their services entrench inequality. Current Section 230 doctrine makes that terrifically difficult, if not impossible.

CONCLUSION

Powerful internet companies design their services in ways that facilitate illegal discrimination and other consumer harms. This is reason enough for reform.

This Essay's lessons are twofold. First, most internet companies have the formidable capacity to redress these practices, but do not do so until they are called to task pursuant to a court order or an explosive news report.[190] Second, even

189. Jeff Horwitz, *Facebook Algorithm Shows Gender Bias in Job Ads, Study Finds*, WALL ST. J. (Apr. 9, 2021, 1:12 PM ET), https://www.wsj.com/articles/facebook-shows-men-and-women-different-job-ads-study-finds-11617969600 [https://perma.cc/TN22-UK84]; Faife & Ng, *supra* note 3.

190. Consider by way of ironic context that Google deliberately blocks advertisers from using terms associated with Black Lives Matter and related progressive social-justice movements to post ads alongside YouTube videos and channels. Leon Yin & Aaron Sankin, *Google Blocks Advertisers from Targeting Black Lives Matter YouTube Videos*, MARKUP (Apr. 9, 2021), https://themarkup.org/google-the-giant/2021/04/09/google-blocks-advertisers-from-targeting-black-lives-matter-youtube-videos [https://perma.cc/Y9XJ-PMEM].

when called to task, these companies invoke Section 230 immunity before discovery can reveal whether or how they are implicated in the unlawful content or conduct. It is no surprise that the companies would proceed in this way. Courts have been relatively solicitous of online companies' claims to the protection. It is only recently that some prominent courts and judges have evinced skepticism, in recognition of the ways that service designs necessarily create the harms that ensue.[191]

While I have elsewhere argued that courts should do more,[192] it is time for Congress to reform the statute to comport to the current state of affairs. It can do this by amending the statute so that companies bear the legal duty to block or prevent the amplification or delivery of online content that they know to be unlawful. This is especially important for content that harms consumers and members of historically marginalized groups—that is, people who have few if any other legal avenues for redress.

Today, Section 230's broad protections empower social-media companies to think that they are above the fray, inoculated from bearing responsibility for the services they develop. This turn away from public obligation is corrosive. The time for course correction is overdue.

Olivier Sylvain, Professor of Law, Fordham University Law School; Senior Advisor to the Chair, Federal Trade Commission (FTC). The author wrote the bulk of this piece before his appointment to the FTC. Nothing here represents the FTC's position on any of the matters discussed. *The author is grateful for excellent student research assistance from Nicholas Loh, Laura Reed, and Ryan Wilkins, as well as incredibly resourceful support from Fordham Law School librarian Janet Kearney.*

191. *See, e.g.*, Lemmon v. Snap Inc., 995 F.3d 1085 (9th Cir. 2021); Force v. Facebook, Inc., 934 F.3d 53 (2d Cir. 2019) (Katzmann, J., concurring).

192. *See generally* Sylvain, *Intermediary Duties*, *supra* note 10; Sylvain, *Discriminatory Designs*, *supra* note 10.