NetChoice, Regulatory Competition, and the Real Battle Behind Social Media Regulations

Tao Huang[*]

ABSTRACT

Regulating online social media platforms has been a fiercely debated issue for years. The NetChoice *case decided by the US Supreme Court 2023 Term arises from a circuit split regarding Florida and Texas laws that regulate social media content moderation. Most discussions on regulation of social media content moderation have focused on whether social media platforms are speakers (editors) or carriers (conduits), and whether their moderation practice constitutes editorial judgment (i.e., protected speech). This Article argues that this framing of the debate ignores another more important role of social media platforms—as regulators. Platforms, when they enact and enforce their content rules, are regulating the speech of users. When the government prescribes how platforms should moderate content, it is using its public regulatory power to preempt the platforms' private regulatory power.*

In the social media context, regulatory competition and preemption among state or national governments are common. Different powers are trying to shape the platforms according to their own normative visions of free speech. However, there exist multiple and competing visions, the two most prominent of which are the US vision and the European Union vision. Hence, the real concern of the Texas and Florida laws is not that they are content-based, but that they have imposed one particular vision of free speech values upon the global "public square," a place that is characterized by legal and cultural pluralism. Accommodating such value heterogeneity and conflict is not easy. This Article proposes three possible ways forward: (1) a judicial approach that embraces open-ended balancing instead of strict categoricalism; (2) an administrative approach that nudges procedural governance by platforms and democratic participation by users; and (3)

[*] Assistant Professor of Law, City University of Hong Kong. S.J.D., Duke University School of Law. LL.M., Harvard Law School.

a technological approach that aims to decentralize the structure of social media.

TABLE OF CONTENTS

I. INTRODUCTION

In 2021, Florida and Texas each passed legislation regulating the content moderation practices of social media platforms.[1] The Florida law bans platforms from willfully deplatforming political candidates, requires consistently applied content moderation, and mandates that platforms provide notice and explanation to the users whose content has been removed.[2] The Texas law demands viewpoint neutral content moderation that provides users with notifications about any modifications to their posts.[3] In addition, the law requires

1. S. 7072, 2021 Leg., Reg. Sess. (Fla. 2021); H.B. 20, 87th Leg., 2d Sess. (Tex. 2021).
2. Fla. S. 7072.
3. Tex. H.B. 20.

platforms to provide users with appeal mechanisms to challenge the moderation of their posts.[4]

Such legislative endeavors reflect the growing popular call to restrain the power of social media giants who exert tremendous and nearly unbridled power over the modern public discourse.[5] However, these laws have also generated controversy.[6] Social media companies responded quickly by initiating lawsuits against these laws in federal courts, arguing that the laws unconstitutionally abridge the platforms' free speech rights and unduly preempt the immunity conferred by Section 230 of the Communications Decency Act (Section 230).[7] On appeal of the Florida law, the United States Court of Appeals for the Eleventh Circuit found that the platforms' content moderation constitutes editorial judgment protected by the First Amendment of the US Constitution.[8] Thus, the Eleventh Circuit held that the Florida law was an impermissible content-based regulation of speech.[9] By contrast, on appeal of the Texas law, the United States Court of Appeals for the Fifth Circuit reasoned that a platform's content moderation is not speech but censorship.[10] Accordingly, the court found that the Texas law was content-neutral and survived intermediate scrutiny.[11]

In response to this circuit split and the issue's growing importance, the Supreme Court of the United States granted certiorari.[12] The case was set to shape the substantive rules of how social media platforms can be regulated and develop a uniform

---

4. *Id.*

5. *See* Jennifer Huddleston & Liam Fulling, *Examining State Tech Policy Actions in 2021*, AM. ACTION F. (July 21, 2021), https://tinyurl.com/2vhftt42 [perma.cc/WC2T-Q4TK]; Zoe Bedell & John Major, *What's Next for Section 230? A Roundup of Proposals*, LAWFARE (July 29, 2020, 9:01 AM), https://www.lawfaremedia.org/article/whats-next-section-230-roundup-proposals [perma.cc/P6SA-2ASY].

6. *See* discussion *infra* Section II.A.

7. NetChoice, LLC v. Moody, 546 F. Supp. 3d 1082, 1085 (N.D. Fla. 2021), *aff'd in part, vacated in part, remanded sub nom.* NetChoice, LLC v. Att'y Gen. 34 F.4th 1196 (11th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024); NetChoice, LLC v. Paxton, 573 F. Supp. 3d 1092, 1101 (W.D. Tex. 2021), *vacated and remanded sub nom.* NetChoice, LLC v. Paxton, 49 F.4th 439 (5th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

8. NetChoice, LLC v. Att'y Gen., 34 F.4th 1196, 1203 (11th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

9. *Id.*

10. NetChoice, LLC v. Paxton, 49 F.4th 439, 448 (5th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

11. *Id.* at 480.

12. *See Granted & Noted List October Term 2023 Cases for Argument*, SUP. CT. OF THE U.S., https://www.supremecourt.gov/orders/23grantednotedlist.pdf [https://perma.cc/QCQ5-2H6W] (last visited Nov. 27, 2024).

standard among courts.[13] Importantly, it offered a chance to define the boundary between government and platform power in shaping the landscape of online communications.[14] The high stakes make the case a trending issue, in both media and scholarship.[15]

Despite these high stakes and high hopes, the Court punted the issue back to the lower courts. In its 2023 term, the Court remanded the case to the lower courts in *Moody v. NetChoice* for further consideration of the scope of the laws and offered several binding principles on how regulations on social media may implicate the First Amendment.[16]

Current debates over social media regulation in general, and the Florida and Texas laws in particular, have focused primarily on two issues: whether content moderation by social media platforms constitutes editorial judgment—and thus subject to First Amendment protection—and, if so, whether the government regulation targeting the moderation practices are content-based—and thus subject to the mostly fatal strict scrutiny.[17]

This Article argues, however, that the current debate is out of focus. By analogizing social media to forums and scrutinizing the regulations according to rigid categorical rules, the current analysis fails to recognize the multiple roles of social media platforms and the real battle behind their regulations. The current discussion over whether platforms are editors (speakers) or conduits (common carriers) captures only two roles of those intermediaries, while ignoring another more important role: platforms as regulators. Platforms, when they enact and enforce their content moderation rules, are regulating the

---

13. *See* Opening Brief of Appellants at *1, *Att'y Gen.*, 34 F.4th 1196 (No. 21-12355), ("Whether and how such power ought to be regulated is among the most consequential and controversial policy issues in America today.").

14. *See* Jameel Jaffer, *What Matters Most in the Supreme Court's Upcoming Social Media Cases*, KNIGHT FIRST AMEND. INST. (Oct. 31, 2023), https://knightcolumbia.org/blog/what-matters-most-in-the-supreme-courts-upcoming-social-media-cases [https://perma.cc/5N5P-EQVU].

15. *See* Joe Schneider & Alex Barinka, *Texas Social-Media Law on Web Censorship Upheld by Federal Appeals Court*, BLOOMBERG (Sept. 16, 2022, 5:52 PM), https://www.bloomberg.com/news/articles/2022-09-16/texas-social-media-law-upheld-by-federal-appeals-court [https://perma.cc/9UAC-YE5S]; Adam Liptak, *Supreme Court to Hear Challenges to State Laws on Social Media*, N.Y. TIMES (Sept. 29, 2023), https://www.nytimes.com/2023/09/29/us/supreme-court-social-media-first-amendment.html [https://perma.cc/538A-S7QM]; Adam Candeub, *Editorial Decision-Making and the First Amendment*, 2 J. FREE SPEECH L. 157, 201–02 (2022); Dawn Carla Nunziato, *Protecting Free Speech and Due Process Values on Dominant Social Media Platforms*, 73 HASTINGS L.J. 1255, 1256 (2022).

16. 603 U.S. 707, 716–17 (2024).

17. *See infra* Part II.

speech posted by users. Those private powers regulate and impact speech in the same way government regulation would.[18]

Platforms simultaneously assume three roles—speaker, common carrier, and regulator—but with varying degree and importance. Sometimes, platforms explicitly express themselves as speakers, like through the public statements of their top executives.[19] This is most evident when they explain the rationale for a content moderation decision in response to public pressure.[20] Platforms also serve as carriers or conduits for others' speech and have primarily labeled themselves as such.[21] But despite the carrying function they assume, they are not *common* carriers.[22] On the contrary, platforms continuously and frequently deny users' speech on the basis of content.[23] Moreover, a platform's role as a speaker is minor on most occasions; the massive scale of moderation decisions—which are dispersed, largely unrelated, and sometimes automated—cannot be described as acts of speaking.[24] Instead, in most cases, content moderation is an act of regulation, the aim of which is to manage and preserve the online speech environment.

Part II of this Article introduces the tripartite roles of social media platforms, describing how the existing discussion of social media platforms as content moderators has emphasized the roles of carriers and speakers while overlooking the role of regulators.

Part III examines the implications of this paradigm shift. Focusing on the regulatory role of platforms reminds us of the triangular scenario of free speech raised by legal scholar Jack Balkin.[25] But here, the government, as exemplified by the Texas and Florida

---

18. *See* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1599 (2018) [hereinafter Klonick, *New Governors*].

19. *See, e.g.*, Mark Zuckerberg, *Bringing the World Closer Together*, FACEBOOK (Mar. 15, 2021), https://www.facebook.com/notes/393134628500376/ [https://perma.cc/8P66-KCP4].

20. For example, Facebook had made an open statement explaining its moderation decision of first deleting, then restoring, an iconic child nudity image. *See* Mark Scott & Mike Issac, *Facebook Restores Iconic Vietnam War Photo It Censored for Nudity*, N.Y. TIMES (Sept. 9, 2016), https://www.nytimes.com/2016/09/10/technology/facebook-vietnam-war-photo-nudity.html [https://perma.cc/AU62-CFQX].

21. *See* David Cohen, *Mark Zuckerberg Q&A: Dislike Button, Ferguson, Graph Search, News Feed Study Controversy*, ADWEEK (Dec. 12, 2014), http://www.adweek.com/digital/mark-zuckerberg-qa-121114 [https://perma.cc/2HMX-K3R5].

22. *See* discussion *infra* Section II.B.

23. *See* NetChoice, LLC v. Att'y Gen., 34 F.4th 1196, 1214 (11th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

24. *See infra* Section II.C.

25. Jack M. Balkin, Essay, *Free Speech Is a Triangle*, 118 COLUM. L. REV. 2011, 2020 (2018) [hereinafter Balkin, *Triangle*]; *see* discussion *infra* Part III.

laws, is not trying to suppress the ideas expressed by the platforms as speakers nor impede their function as conduits of speech.[26] Rather, the government aims to preempt the private regulation by platforms with a regulatory scheme endorsed by the government.[27]

Under this perspective, the central question for judicial review is not whether the government's regulation is content-based or content-neutral, but whether the regulatory preemption truly advances the speech rights of the users in regulating the platforms. Current debates concentrate upon government intention and platform rights, without paying due respect to the users and their free speech rights.[28] Thus, the real concern over social media regulations and regulatory competition should be determining which kind of regulation truly advances the free speech rights of the users. Different regulatory efforts reflect different managerial plans that shape the background structure of the communicative environment on the platforms.[29]

However, determining which background rules or norms better facilitate users' speech rights is not easy. As Part IV demonstrates, there are multiple and competing visions of free speech, all of which are based upon different normative conceptions of the relationships among state, society, and individuals.[30] For example, Americans and Europeans endorse distinct views regarding human agency, vulnerability, and the comparative threat of public and private powers.[31] Determining how to accommodate these cultural conflicts in the global public square is the biggest challenge that must be addressed by social media regulations and the stakeholders behind them.

Value conflict is the thorniest issue to be solved. To provide some reference points and invite more exploration, Part V of this Article offers proposals for accommodating the competing visions of free speech on platforms. Courts, governments, and social media companies can actively engage in this enterprise. Courts can turn away from formulaic categorical review and embrace a more open-ended balancing that takes more interests into account. Governments can nudge platforms to improve their internal governance, such as by facilitating user participation. Lastly, social media companies can develop technical ways to decentralize platforms and enable the coexistence of competing free speech norms.

---

26. *See* S. 7072, 2021 Leg., Reg. Sess. (Fla. 2021); H.B. 20, 87th Leg., 2nd Sess. (Tex. 2021).
27. *See, e.g.*, Fla. S. 7072.
28. Bedell & Major, *supra* note 5.
29. *See id.*
30. *See infra* Part IV.
31. *See infra* Part IV.

The aim of this Article is not to argue that we should stop discussing issues of editorial discretion and content neutrality, issues that lie at the core of the current judicial doctrine.[32] The problem is not that they are unimportant, but that they have captured too much of our attention and agenda.[33] Such capture limits our willingness and capacity to think beyond the existing framework and look for new ways of framing the issues behind social media regulations, as well as new approaches to solving them. By dissecting the various roles assumed by social media platforms in moderating content, we acquire a clearer view of the competition between different regulatory powers, each endorsed by a normative vision of free speech.[34] Future content moderation discussions should address the conflicts between those competing visions through legal, social, or technical means.

II. SPEAKER, CARRIER, AND REGULATOR

Social media platforms assume various roles in their moderation practices: speaker, carrier, and regulator.[35] They carry the speech of others, but unlike common carriers, they do not host indiscriminately.[36] Platforms sometimes express messages directly, but only in limited circumstances.[37] Most importantly, platforms are regulators of speech.[38]

A. The Speaker-Carrier Debate that Dominates Current Discussion

The platforms challenging the Florida and Texas laws rely primarily on the argument that content moderation constitutes editorial judgment, a constitutionally protected category of speech.[39]

---

32. *See* discussion *infra* Section IV.B.

33. *See infra* Part IV.

34. *See infra* Part VI.

35. *See* Sofia Grafanaki, *Platforms, the First Amendment and Online Speech: Regulating the Filters*, 39 PACE L. REV. 111, 122–23 (2018).

36. John Villasenor, Commentary, *Social Media Companies and Common Carrier Status: A Primer*, BROOKINGS INST. (Oct. 27, 2022), https://www.brookings.edu/articles/social-media-companies-and-common-carrier-status-a-primer/ [https://perma.cc/22MN-HKUT].

37. *See, e.g.*, Zuckerberg, *supra* note 19.

38. NetChoice, LLC v. Paxton, 49 F.4th 439, 448 (5th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

39. *See, e.g.*, Complaint for Declaratory & Injunctive Relief ¶ 80, NetChoice, LLC v. Moody, 546 F. Supp. 3d 1082 (N.D. Fla. 2021) (No. 4:21-cv-00220-RH-MAF) [hereinafter Complaint for Declaratory & Injunctive Relief in *Moody*] ("Online businesses that make editorial decisions regarding what content to publish, including content created or posted by third parties, engage in speech that is fully protected by the First Amendment."); Complaint for Declaratory & Injunctive Relief ¶ 6, NetChoice, LLC v. Paxton, 573 F. Supp. 3d 1092 (W.D. Tex. 2021) (No. 1:21-cv-00840)

Meanwhile, the state governments counter that the social media platforms are more akin to common carriers and conduits that host and distribute others' speech, rather than their own.[40] Both parties found support from existing case law.[41] The platforms offered a series of Supreme Court precedent to support their argument: *Miami Herald Publishing Co. v. Tornillo*,[42] *Pacific Gas & Electric Company v. Public Utilities Commission of California*,[43] and *Hurley v. Irish-American Gay, Lesbian & Bisexual Group of Boston*.[44] Meanwhile, the states pointed to two different cases to support their view: *Rumsfeld v. Forum for Academic & Institutional Rights, Inc.* (*FAIR)*[45] and *PruneYard Shopping Center v. Robins.*[46]

From the trial to the appeal, the parties, amici and courts relied on these precedents, endeavoring to analogize and differentiate social media platforms with the entities discussed in the cases cited by the

---

[hereinafter Complaint for Declaratory & Injunctive Relief in *Paxton*] ("Though the laws differ in their specifics, Florida's law and H.B. 20 here both infringe on the editorial discretion that the First Amendment protects, and the Texas Attorney General himself has called the two laws 'similar.'").

40. *See, e.g.*, Defendants' Response in Opposition to Plaintiffs' Motion for Preliminary Injunction at 27–28, *Moody*, 546 F. Supp. 3d 1082 (No. 4:21-cv-00220-RH-MAF); Defendant's Response to Motion for Preliminary Injunction at 3, *Paxton*, 573 F. Supp. 3d 1092 (No. 1:21-cv-00840-RP).

41. *See, e.g.*, Defendants' Response in Opposition to Plaintiffs' Motion for Preliminary Injunction, *supra* note 40, at 15; *Moody*, 546 F. Supp. 3d at 1091–93 ("The plaintiffs say, in effect, that they should be treated like any other speaker. The State says, in contrast, that social-media providers are more like common carriers, transporting information from one person to another . . . . More generally, the plaintiffs draw support from three Supreme Court decisions in which a state mandate for a private entity to allow unwanted speech was held unconstitutional. On the State's side are two Supreme Court decisions in which a state or federal mandate for a private entity to allow unwanted speech was held constitutional. Each side claims the cases on its side are dispositive . . . .").

42. Miami Herald Publ'g Co. v. Tornillo, 418 U.S. 241, 258 (1974) (ruling that Florida's "right to reply" law unduly infringed upon the editorial function of the newspapers, and invalidating the law).

43. Pac. Gas & Elec. Co. v. Pub. Utils. Comm'n, 475 U.S. 1, 8–9 (1986) (plurality opinion) (invalidating a California regulation that required a private utility company to include third-party speech in its newsletters within the billing envelops, holding that the newsletter was similar to a small newspaper).

44. Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Bos., 515 U.S. 557, 569–70 (1995) (holding that the parade organizer's decision to include other groups was protected speech).

45. Rumsfeld v. F. for Acad. & Institutional Rts., Inc., 547 U.S. 47, 69–70 (2006) (holding that a law requiring law school campuses to provide equal access to military recruitment did not violate the First Amendment because the law schools' speech would not be affected by hosting others' speech).

46. PruneYard Shopping Ctr. v. Robins, 447 U.S. 74, 88 (1980) (holding that the public generally will not connect an individual's speech to the shopping center, the shopping center could not prohibit individual speech when a private shopping center open to the public prohibited individuals from engaging in expressive activity on the property).

parties.[47] Once the cases reached the Supreme Court in *Moody*, the Court also based its ruling upon analogies with *Tornillo*, *Pacific Gas & Electrics*, *Hurley, FAIR, and PruneYard*.[48] Such a reason-by-analogy approach is not surprising—the approach also finds support in academic scholarship. For example, some scholars argue that "[t]he Supreme Court has treated regulation of each of these forms of media differently under the First Amendment, and thus courts' analys[es] of social media regulations may depend on which older media they find most closely analogous to social media."[49]

The five forums discussed by the Supreme Court precedents in *Moody*—newspapers, newsletters, parades, shopping centers, and school recruitment spaces—are properties or spaces with different and varying degrees of connection between the property owners and the speech they host. [50] A newspaper, for example, is a First Amendment institution (the press) that is itself expressive.[51] In contrast, a shopping center is a business entity that is only remotely related to speech.[52] Indeed, it is nearly impossible to group these five forums into one or two categories. Any analogy of social media platforms to one of these forums will necessarily capture some features of the properties while downplaying others. Each of the five forums are similar to social media in some aspects while vastly different in others.[53] Thus, borrowing doctrines from existing, analogous categories is not only inattentive,[54] but also misleading, for "analogies are always imperfect, and there is no universal agreement on which of the

47. *See, e.g.*, NetChoice, LLC v. Paxton, 49 F.4th 439, 456–63 (5th Cir. 2022), *vacated and remanded sub nom*. Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

48. *Miami Herald Publ'g Co.*, 418 U.S. at 258; *Pac. Gas & Elec. Co.*, 475 U.S. at 8–9; *Hurley*, 515 U.S. at 569–70; *Rumsfeld*, 547 U.S. at 69–70; *PruneYard Shopping Ctr.*, 447 U.S. at 88; Moody v. NetChoice, LLC, 603 U.S. 707, 733 (2024).

49. Andrew J. Ceresney, Jeffrey P. Cunard, Courtney M. Dankworth & David A. O'Neil, *Regulating Harmful Speech on Social Media: The Current Legal Landscape and Policy Proposals*, *in* SOCIAL MEDIA, FREEDOM OF SPEECH AND THE FUTURE OF OUR DEMOCRACY XXIII, XXX–XXXI (Lee C. Bollinger & Geoffrey R. Stone eds., 2022).

50. *Moody*, 603 U.S. at 707.

51. Michael J. Burstein, *Towards a New Standard for First Amendment Review of Structural Media Regulation*, 79 N.Y.U. L. REV. 1030, 1045 (2004). A First Amendment institution refers to an institution that is explicitly mentioned and protected by the First Amendment, such as the press. *See id.*

52. *Moody*, 603 U.S. at 730.

53. Nathaniel Persily, *Platform Power, Online Speech, and the Search for New Constitutional Categories*, *in* SOCIAL MEDIA, FREEDOM OF SPEECH AND THE FUTURE OF OUR DEMOCRACY 193, 195, 211 (Lee C. Bollinger & Geoffrey R. Stone eds., 2022).

54. Monroe E. Price, *The Newness of New Technology*, 22 CARDOZO L. REV. 1885, 1892–93 (2001).

analogies best suits platforms."[55] Indeed, the five court decisions with five different forums cannot be said to generate a consistent case law.[56] For example, newspapers have editorial rights, but with regard to other entities such as shopping centers, these rights are much more ambiguous and complex.[57]

Several approaches have been proposed to distinguish these forum precedents in the social media realm. For example, some approaches ask whether the "forced" carry of others' speech adversely affects the property owner's right of expression, whether listeners may confuse the user's speech with the platform's speech, and whether there exists a common theme underlying the owner's behavior of hosting.[58] Those distinctions, however, have received contrasting treatment by different courts: while the Fifth Circuit partly relied on these three distinctions to uphold the Texas regulation,[59] the Eleventh Circuit found the distinctions irrelevant to the Florida regulation.[60]

The various interpretations of the five forum precedents reflect different understandings of the functions of social media.[61] Thus, we must first grasp the functions of social media platforms before scrutinizing the regulations of them. In the Texas and Florida cases, the states and platforms debated whether social media platforms are editors like newspapers or common carriers like telephone companies.[62] Both are partially right, since the newspaper metaphor captures the moderation function of platforms while the common carrier analogy emphasizes the hosting function. However, both parties ignored important points of the different roles played by platforms. As a complex and powerful entity, a social media platform plays multiple, rather than singular, roles.[63] The remainder of Part II of this Article and Part III will address the multiple roles of a platform.[64]

---

55. Grafanaki, *supra* note 35, at 135.

56. Villasenor, *supra* note 36.

57. Ashutosh Bhagwat, *Do Platforms Have Editorial Rights?*, 1 J. FREE SPEECH L. 97, 100–01 (2021).

58. *See, e.g.*, Defendants' Response in Opposition to Plaintiffs' Motion for Preliminary Injunction, *supra* note 40, at 16–23.

59. NetChoice, LLC v. Paxton, 49 F.4th 439, 460–65 (5th Cir. 2022), *vacated and remanded sub nom*. Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

60. NetChoice, LLC v. Att'y Gen., 34 F.4th 1196, 1219–20 (11th Cir. 2022), *vacated and remanded*, 603 U.S. 707 (2024).

61. *See id.* at 1212; *Paxton*, 49 F.4th at 459.

62. *See Att'y Gen.*, 34 F.4th at 1212; *Paxton*, 49 F.4th at 459.

63. *See* Eugene Volokh, *Treating Social Media Platforms Like Common Carriers?*, 1 J. FREE SPEECH L. 377, 440 (2021).

64. *See infra* Part III. This Article's analysis is limited within the First Amendment's role. Other roles assumed by those companies, such as contractor, competitor, or employer, are covered

B. Why Social Media Platforms Are Not Common Carriers

Common carriers generally refer to entities or corporations that hold themselves out as providing services to the public in a non-discriminatory fashion.[65] Social media platforms are carriers, but they are not common carriers.[66] Acting as carriers, hosting and distributing speech posted by users constitutes a major function of those platforms.[67] That function and its facilitation of speech has made platforms the "modern public square."[68] However, platforms are not *common* carriers for two reasons.

First, social media platforms do not indiscriminately carry others' speech—they selectively host speech according to their rules and standards.[69] The defining feature of a common carrier is that the service is offered to the public in an indiscriminate way.[70] Typical examples include telephone lines, railroads, and broadband providers.[71] Social media platforms do not fit into this category.[72] Just take a look at the community rules of three major platforms, Facebook,[73] X (formerly known as Twitter),[74] and YouTube.[75] As evidenced by the community rules maintained by each of these platforms, each platform has rules that permits certain types of content while prohibiting others.[76] The existence of these community rules indicates that these platforms' services are not offered indiscriminately.[77] Because the platforms' services are not offered indiscriminately, the platforms as carriers are

---

by other fields of law. In addition, among all the activities conducted by the platforms, only content moderation will be discussed in this Article.

65. Villasenor, *supra* note 36.

66. *Id.*

67. *See id.*

68. Packingham v. North Carolina, 582 U.S. 98, 107 (2017).

69. NetChoice, LLC v. Att'y Gen., 34 F.4th 1196, 1221–22 (11th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

70. Villasenor, *supra* note 36.

71. *See id.*

72. NetChoice, LLC v. Paxton, 573 F. Supp. 3d 1092, 1107 (W.D. Tex. 2021), *vacated and remanded sub nom.* NetChoice, LLC v. Paxton, 49 F.4th 439 (5th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

73. *See Facebook Community Standards*, META, https://transparency.fb.com/policies/community-standards/ [https://perma.cc/7TAT-A6PS] (last visited Sept. 8, 2024).

74. *See The X Rules*, X, https://help.twitter.com/en/rules-and-policies/x-rules [https://perma.cc/YZ89-N9TU] (last visited Sept. 8, 2025).

75. *See Community Guidelines*, YOUTUBE, https://www.youtube.com/howyoutubeworks/policies/community-guidelines/ [https://perma.cc/RB4G-KHV6] (last visited Sept. 8, 2024).

76. *See Facebook Community Standards*, *supra* note 73; *The X Rules*, *supra* note 74; *Community Guidelines*, *supra* note 75.

77. *See Facebook Community Standards*, *supra* note 73; *The X Rules*, *supra* note 74; *Community Guidelines*, *supra* note 75.

not "common," but rather selective and restrictive; they "bear essentially no similarities to" the historically recognized common carriers.[78]

Second, legislative mandates that require platforms to become common carriers would destroy the whole scheme of content moderation and jeopardize the companies' business model.[79] The rationale for selectively hosting speech is to make the forum appropriate and attractive to users: if platforms cannot be free to limit and prevent spam, hatred, and abusive content, both users and advertisers would choose to leave these hostile platforms.[80] Indeed, requiring platforms to become indiscriminate common carriers would be an unbearable regulatory effort ostensibly fatal to the existence of platforms, and users might lose such spaces to speak as a result.[81]

Platforms do not always moderate in good faith to attract users' attention—hateful, fake, and offensive content is often easier to make viral.[82] The conflict between the platforms' commercial interest and the public interest is a legitimate ground for regulation. However, this regulatory need should not support a radical approach that makes platforms common carriers. Transforming platforms into common carriers would compromise rather than promote the public interest, because destroying such important forums of speech would deprive users of many chances to express themselves in the digital age.[83]

---

78. Ashutosh Bhagwat, *Why Social Media Platforms Are Not Common Carriers*, 2 J. FREE SPEECH L. 127, 143 (2022).

79. *See id.*

80. *See* Complaint for Declaratory & Injunctive Relief in *Moody*, *supra* note 39, at ¶ 46. The plaintiff has expressed similar concern in the case about Texas law, *see* Complaint for Declaratory & Injunctive Relief in *Paxton*, *supra* note 39, at ¶ 23 ("[A]dvertisers will not permit their products and services to be displayed in an editorial context of harmful or offensive content. And the proliferation of such objectionable content will cause many users to use the platforms less, or stop using them entirely.").

81. *See* Eric Goldman & Jess Miers, *Online Account Terminations/Content Removals and the Benefits of Internet Services Enforcing Their House Rules*, 1 J. FREE SPEECH L. 191, 193–96, 211 (2021); Karanjot Gill, *Regulating Platforms' Invisible Hand: Content Moderation Policies and Processes*, 21 WAKE FOREST J. BUS. & INTELL. PROP. L. 171, 175 (2021).

82. *See* Manoel Horta Ribeiro, Virgílio A. F. Almeida, Raphael Ottoni, Wagner Meira Jr. & Robert West, *Auditing Radicalization Pathways on YouTube*, *in* FAT* '20: PROCEEDINGS OF THE 2020 CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 131, 131 (2020); *see also* Gill, *supra* note 81, at 180.

83. Goldman & Miers, *supra* note 81, at 213–15.

C. Why Content Moderation Is Mostly Not Speech

Apart from host or carriers, social media platforms can also act as speakers.[84] The platforms challenging the Texas and Florida statutes argued that platforms' decisions regarding which content should be allowed and which should not is similar to editorial decisions made by newspapers.[85] Therefore, the content moderation decisions made by platforms should also be protected.[86] This Section argues that the case for classifying content moderation as speech is, on most occasions, weak.

The first case against classifying platforms as speakers is that they are corporate entities, and not treated the same as individuals under the First Amendment law.[87] Although the Court has extended First Amendment protections to corporate speakers,[88] the non-human identity of platforms matters. On one hand, autonomy interests are unique to human speakers.[89] On the other hand, the corporate speaker is a privileged powerholder that must be checked.[90]

As a fictitious legal entity, a social media company cannot speak on behalf of its shareholders.[91] The corporate entity is "not an association of its members; it is a distinct legal entity separate from its stockholders, managers, creditors, and the like."[92] Indeed, the corporate

---

84. Moody v. NetChoice, LLC, 603 U.S. 707, 716 (2024) ("[S]ome platforms, in at least some functions, are indeed engaged in expression.").

85. Reply Brief for Petitioners at 8, Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

86. *See* NetChoice, LLC v. Moody, 546 F. Supp. 3d 1082, 1090–91 (N.D. Fla. 2021), *aff'd in part, vacated in part, remanded sub nom.* NetChoice, LLC v. Att'y Gen. 34 F.4th 1196 (11th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024); *see also* NetChoice, LLC v. Paxton, 573 F. Supp. 3d 1092, 1106 (W.D. Tex. 2021), *vacated and remanded sub nom.* NetChoice, LLC v. Paxton, 49 F.4th 439 (5th Cir. 2022), *vacated and remanded sub nom.* Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

87. *E.g.*, *Paxton*, 573 F. Supp. 3d at 1107.

88. *See* Citizens United v. Fed. Election Comm'n, 558 U.S. 310, 343 (2010) ("The Court has thus rejected the argument that political speech of corporations or other associations should be treated differently under the First Amendment simply because such associations are not 'natural persons.'"); s*ee also* Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495, 1502 (2013).

89. The autonomy interest of free speech values refers to the argument that freedom of speech can promote individual autonomy of humans. *See generally* Martin H. Redish, *Value of Free Speech*, 130 U. PA. L. REV. 591 (1982) (elaborating free speech's function in promoting individual autonomy).

90. *See* Nathan Cortez & William Sage, *The Disembodied First Amendment*, 100 WASH. U. L. REV. 707, 716 (2023); Morgan N. Weiland, *Expanding the Periphery and Threatening the Core: The Ascendant Libertarian Speech Tradition*, 69 STAN. L. REV. 1389, 1437 (2017).

91. *See* Alexander Tsesis, *Balancing Free Speech*, 96 B.U. L. REV. 1, 43 (2016) ("Shareholders of for-profit companies, especially those that are publicly traded, ordinarily have hardly any say in corporate affairs apart from voting for directors and on issues like mergers and dissolutions.").

92. *See generally* Cortez & Sage, *supra* note 90.

entity contains no intrinsic value but only instrumentally facilitates the flourishing of human individuals.[93] Without a speaker's interest at stake, the only ground for protecting the corporation's expressive right is listener-based.[94] But on many occasions, corporations' commercial interests are in tension with the listeners' interests, and platforms can easily influence, control, and manipulate their listeners.[95]

If the argument of content moderation as speech is listener-based, then the expression must be understood by potential listeners to be constitutionally protected.[96] That is also required by the *Spence* test, which prescribes that in order to have First Amendment protection, a speaker must have the intent to convey a message, and it must be likely that the message can be understood by listeners.[97] However, in reality, the message conveyed by platforms' moderation decisions, which are numerous, amorphous, and complex, can hardly been meaningfully captured by listeners.[98]

Content moderation decisions are made by thousands of moderators, artificial intelligence (AI) tools, or a combination of both.[99] Thus, with so many players involved, moderation decisions are made without a consistent or coherent theme, which is a key distinction between social media and newspapers.[100] For social media platforms, outsourced moderators and AI robots are just mechanically moderating content, rather than expressing ideas.[101] Algorithms used to moderate content are written before the user content has been generated, and the developers or coders have no idea how their codes can be used in future occasions.[102] Thus, without a consistent decision-making process, we cannot identify which message a platform's moderation conveys.

---

93. C. Edwin Baker, *Turner Broadcasting: Content-Based Regulation of Persons and Presses*, 1994 SUP. CT. REV. 57, 81 (1994). For example, social media, and thus the companies that control social media platformsit, promote an individual's message by amplifying it to others around the world. *Id.* It allows individuals to attract an attention to their ideas that they would have never gotten without the platform that social media companies provide. *Id.*

94. *Id.* at 83–84.

95. Weiland, *supra* note 88, at 1396 ("It is deeply ambiguous whether the Court's deregulatory holdings actually benefit listeners, though corporate interests are always served.").

96. *See* Candeub, *supra* note 15, at 180.

97. *See* Spence v. Washington, 418 U.S. 405, 410–11 (1974).

98. Candeub, *supra* note 15, at 187–88.

99. *See How Review Teams Work*, META (Jan. 19, 2022), https://transparency.meta.com/enforcement/detecting-violations/how-review-teams-work/ [https://perma.cc/R43C-NWQU].

100. *See* Bhagwat, *supra* note 57, at 112–13.

101. *See* Pauline Trouillard, *Social Media Platforms Are Not Speakers: Why Are Facebook and Twitter Devoid of First Amendment Rights?*, 19 OHIO ST. TECH. L.J. 257, 293 (2023).

102. *Id.*

Platforms argue that their enactment and enforcement of community standards reflect a message concerning the community they seek to build.[103] This argument cannot stand for three reasons. First, while enactment of community rules surely expresses the norms and ethos of platforms, rule enforcement is either too dispersed or automated.[104] Thus, the distance between a platform's community standards and its community-building goal is remote at best.[105] Second, the community-building-as-expression argument can be used by nearly every corporation.[106] For example, a grocery store decides which products will be in what location or level of the shelves in the store. In that scenario, does the layout of the grocery store convey an expressive message about what kind of shopping community the store wants to offer its customers?[107] Should that message receive First Amendment protections? Answering "yes" to these questions would make the First Amendment boundless.[108] And a boundless First Amendment would invalidate most government regulations for their impact on expressive messages.[109]

A platform may argue that content moderation decisions at least convey the platform's approval or disapproval of the message.[110] Yet, platforms have explicitly denied that the content they host reflects the platforms' opinions.[111] Normatively, if platforms truly believe that

---

103. Complaint for Declaratory & Injunctive Relief in *Paxton*, *supra* note 39, at ¶ 56.

104. Sarah Roberts has documented that the work of content moderators is fractured, low-paid, and outsourced; few moderators know their work before being recruited, and they have nearly no connection with the employer platforms as well as other moderators. *See* SARAH T. ROBERTS, BEHIND THE SCREEN: CONTENT MODERATION IN THE SHADOWS OF SOCIAL MEDIA 39–71 (2019). In this context, we can hardly say that the moderators, most of whom are subcontracted and offshored, are acting consciously in line with the company's goal of community building. Rather, they are merely enforcing the rules to the content, without any conception of the holistic picture of their work of enforcement as well as that of their peer workers.

105. *Id.*

106. *See* Wu, *supra* note 88, at 1529 ("Google hopes to convey ideas like 'quality' or 'usefulness,' but then so too did the designers of my coffeemaker.").

107. Heather Whitney, *Search Engines, Social Media, and the Editorial Analogy*, KNIGHT FIRST AMEND. INST. (Feb. 27, 2018), https://knightcolumbia.org/content/search-engines-social-media-and-editorial-analogy [https://perma.cc/Q2U5-DU3X].

108. Candeub, *supra* note 15, at 198.

109. *See id.*

110. *See* NetChoice, LLC v. Att'y Gen., 34 F.4th 1196, 1211, 1217 (11th Cir. 2022), *cert. granted in part sub nom.* Moody v. NetChoice, LLC, 144 S. Ct. 477 (2023), *and cert. denied*, 144 S. Ct. 69 (2023), *and vacated and remanded*, 144 S. Ct. 2382 (2024).

111. *See Twitter Terms of Service*, X, https://twitter.com/en/tos/previous/version_13#us [https://perma.cc/9J5P-HN4R] (last visited Sept. 8, 2024) ("We do not endorse, support, represent or guarantee the completeness, truthfulness, accuracy, or reliability of any Content or communications posted via the Services or endorse any opinions expressed via the Services."); s*ee also* Ravi Somaiya, *How Facebook Is Changing the Way Its Users Consume Journalism*, N.Y. TIMES

moderation decisions communicate approval or disapproval of content, the platforms should be liable for the content of posts, like traditional publishers.[112] Descriptively, the mere scale of content on platforms demonstrates that platforms hardly know—let alone endorse—individual posts generated by users.[113] Indeed, social conventions imply that both users and platforms do not attribute user-generated content to the platforms.[114] The platforms' refusal to disclose their moderation decisions reveals that there hardly exists any intention to express ideas through content moderation decisions.[115] If platforms truly want to use moderation as a means of expression, their moderation decisions should not be kept in a black box.

Third, even if we accept that moderators (human or AI) express themselves through reviewing content, that expression is not the target of government regulation.[116] In other words, the government's regulations are not targeting what the platforms (or the moderators working for them) want to express in reviewing the content, but rather what the moderation will achieve in setting up the background architecture of the speech platforms.[117] For example, the target of the Florida and Texas laws is the expressive experience of the users, not the platforms' speech, if any, cumulated by the individual moderation decisions.[118]

The target of the government regulation is important because it is a necessary limitation to the expanding scope of the First Amendment and a useful tweak to the *Spence* test.[119] If a mere communicative element is sufficient to earn First Amendment protection, then almost all conduct is communicative and most government regulations will be in jeopardy.[120] Moreover, the *Spence* test is not a fully reliable tool to

---

(Oct. 26, 2014), https://www.nytimes.com/2014/10/27/business/media/how-facebook-is-changing-the-way-its-users-consume-journalism.html [https://perma.cc/ZLK4-J9M8] (quoting a Facebook engineer stating, "We try to explicitly view ourselves as not editors . . . . We don't want to have editorial judgment over the content that's in your feed").

112. *See* Bhagwat, *supra* note 57, at 105–06; Volokh, *supra* note 63, at 454–57.

113. *See* Trouillard, *supra* note 101, at 275; Ceresney et al., *supra* note 49.

114. Trouillard, *supra* note 101, at 276.

115. Candeub, *supra* note 15, at 202.

116. *Id.* at 158.

117. *See* Amanda Shanor, *First Amendment Coverag*e, 93 N.Y.U. L. REV. 318, 366 (2018).

118. *See* S. 7072, 2021 Leg., Reg. Sess. (Fla. 2021); H.B. 20, 87th Leg., 2nd Sess. (Tex. 2021).

119. *See* Spence v. Washington, 418 U.S. 405, 410–11 (1974).

120. *See* Robert Post, *Recuperating First Amendment Doctrine*, 47 STAN. L. REV. 1249, 1252 (1995) [hereinafter Post, *Recuperating*] (arguing that "any action can at any time be made communicative in a manner that satisfies the *Spence* test"); *see also* Nat'l Inst. of Fam. & Life Advocs. v. Becerra, 585 U.S. 755, 782 (2018) (Breyer, J., dissenting).

determine expressive conduct,[121] for it ignores social context and the government's purpose behind speech regulations.[122] Consider a Jackson Pollock painting, which can have a meaning difficult for viewers to understand but up for interpretation and his personal diary written only for self-expression without any intent to communicate to his painting's viewers.[123] Government target of regulation is key here. If the government is not targeting the communicative element of Pollock's painting, but other elements, such as the copyright of the painting or the rule of its transaction, then the regulation will not trigger First Amendment scrutiny.[124] If, by contrast, the government imposes regulations because it fears the communicative impact of the painting or because it disagrees with the content of the private diary, then the regulation would undergo First Amendment scrutiny.[125]

Similarly, even if we assume that content moderation does contain communicative elements, regulations like the Texas and Florida laws should not trigger First Amendment scrutiny because the government regulation is not targeting what the platforms want to *express*, but how the platforms *regulate* users' speech.[126] In other words, the regulatory purposes of the Florida and Texas laws are unrelated to the platforms' communication.[127]

III. REGULATING THE REGULATOR

A. Revisiting the Free Speech Triangle

The foregoing analysis reveals that social media platforms play multiple roles: their hosting function makes them similar to carriers or conduits and their recommendation function makes them analogous to editors or speakers.[128] But these two roles are not the primary ones targeted by government regulations,[129] nor are they what matters most

121. Shanor, *supra* note 117, at 323 (arguing that "the Supreme Court's test for First Amendment coverage, articulated in *Spence v. Washington*, has been analytically undone, even if not fully rejected by the courts" (citing *Spence*, 418 U.S. 405)).

122. *See* Post, *Recuperating*, *supra* note 120, at 1255; *see also* Jed Rubenfeld, *The First Amendment's Purpose*, 53 STAN. L. REV. 767, 773 (2001) (stating that "the *Spence* test merely states sufficient, not necessary, criteria for determining if conduct is expressive").

123. *See* Rubenfeld, *supra* note 122, at 823.

124. *See id.* at 777–78; *see also* Wu, *supra* note 88, at 1514, 1516–17; John Fee, *The Freedom of Speech-Conduct*, 109 KY. L.J. 81, 85 (2020).

125. *See* Fee, *supra* note 124, at 85–86.

126. *See id.* at 120.

127. *See id.* at 116–17.

128. *See* Volokh, *supra* note 63, at 409–10.

129. *See generally id.* at 408.

for the public discourse. Most importantly, platforms are regulators that define the background rules and boundaries of online expression.[130]

The view that online platforms are regulators of speech is not new to legal scholars.[131] Lawrence Lessig has argued that computer code is law and algorithmic design is as powerful as legal rules in shaping expressive freedom.[132] Jack Balkin has proposed the triangular model of free speech in which private regulators play a no less prominent role than state regulators.[133] But surprisingly, the role of platforms as regulators has been rarely mentioned in the current debate of social media regulations, a debate overwhelmed by the expressive right of platforms as speakers.[134] This echoes Lessing and Balkin's diagnosis that the First Amendment has been weaponized or Lochnerized to entrench corporate interests and foil regulatory intervention.[135] Whatever its cause, focusing on platforms as speakers distracts our attention from what is truly at stake.

First and foremost, platforms are regulators.[136] Platforms use a complex set of rules, personnel, and procedures to regulate the content they host.[137] The regulatory toolkit they have includes general principles,[138] concrete rules,[139] and enforcement procedures.[140]

The role of platforms as speakers is marginal and only secondary to their role as regulators.[141] Platforms do not edit and issue coherent content product like newspapers.[142] Nor do they act like other ordinary

---

130. *See id.*

131. *See, e.g.*, *id.* at 454–57.

132. *See* Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 507–10 (1999).

133. *See* Balkin, *Triangle*, *supra* note 25, at 2014. In Balkin's triangle, governments and private corporations are two regulators, regulating the third node of the triangle (individual speakers). *Id.*

134. *See id.* at 2015.

135. *See* Weiland, *supra* note 90, at 1393 ("[C]orporate litigants . . . increasingly use the First Amendment to prioritize new applications of the freedom of speech over regulations designed to protect consumers and citizens.").

136. *See id.* at 1399.

137. *See, e.g.*, *Community Principles*, TIKTOK (Apr. 17, 2024), https://www.tiktok.com/community-guidelines/en/community-principles/ [https://perma.cc/ZE9A-73V3].

138. *See, e.g.*, *id.*

139. *See, e.g.*, *Violent or Graphic Content Policies*, YOUTUBE, https://support.google.com/youtube/answer/2802008?hl=en&ref_topic=9282436 [https://perma.cc/2FWQ-MSW3] (last visited Sept. 8, 2024).

140. *See, e.g.*, *Counting Strikes*, META (Oct. 4, 2022), https://transparency.fb.com/enforcement/taking-action/counting-strikes/ [https://perma.cc/GG6C-C65R].

141. *See infra* Section VI.

142. *See* discussion *supra* Section II.C.

participants in the marketplace of ideas.[143] Instead, platforms shape the structure of the marketplace by determining who can participate and in what manner.[144] Platforms are managers of the marketplace. Primarily viewing and protecting platforms as speakers harms First Amendment values; it renders individual speakers much more vulnerable to the will of platforms and entrenches the current distribution of power.[145] As powerful regulators, the platforms already enjoy tremendous power and leverage in shaping the marketplace of ideas. That power grants them a role as speakers under the First Amendment, which makes it significantly harder to regulate their market behavior and limit their power.[146] This Article's framing of social platforms' roles is different from that of the Supreme Court. The Supreme Court defines the regulatory role of platforms as expression; it argues that when the platforms curate user-generated content and choose which content to include or exclude, those platforms engage in expressive activity.[147] Such selective curation constitutes editorial discretion protected by the First Amendment and is exactly why the Court deemed the Texas law unlikely to succeed.[148]

This Article argues that the Court's framing ignores another, more important role of platforms—their role as regulators. It also highlights the cultural conflict underlying the regulatory endeavor. Indeed, more important than the platforms' expression of their own views is the platforms' selective curation or moderation that regulates the platforms' speech environment according to political, cultural, and ideological values of free speech.[149] Framing such value choices as expression rather than regulation detracts from the thorny cultural tension underlying regulatory endeavors.

In contrast to governments, which are public regulators, platforms are private regulators.[150] However, both public and private regulators regulate the free speech of individuals—a scenario described by Balkin as a triangle.[151] When the government regulates such freedom not by directly passing rules on people's expression but by ordering how the platforms should regulate, the triangle takes a new

143. *See* discussion *infra* Section V.C.

144. *See, e.g.*, *Violent or Graphic Content Policies*, *supra* note 139.

145. Asaf Wiener, *A Speaker-Based Approach to Speech Moderation and First Amendment Analysis*, 31 STAN. L. & POL'Y REV. 187, 226 (2020).

146. *See id.* at 230–31.

147. Moody v. NetChoice, LLC, 603 U.S. 707, 728 (2024).

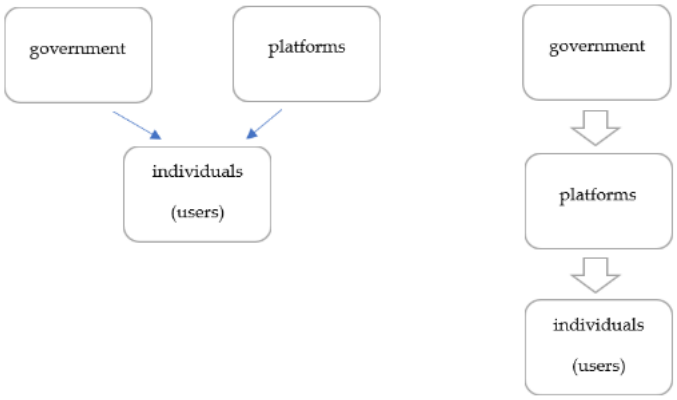
148. *See id.* at 2402, 2405.

149. *See generally* Klonick, *New Governors*, *supra* note 18, at 1635.

150. *See generally id.* at 1601–02.

151. Balkin, *Triangle*, *supra* note 25, at 2015.

form (see the graph below). Now, what is to be noted, but often ignored, is that the government is not prescribing the rules of public discourse directly, but indirectly "commandeering" another powerful regulator, the platforms.[152] In this context, the true object of government regulation is not how these platforms speak, but how users speak under the platforms' private regulation.



The above graphic illustrates a scenario of regulatory competition, in which government regulators are channeling platforms for shaping users' expression online. On the one hand, the multiplicity of regulators poses greater challenges to users' expressive rights.[153] Distrust of government is no longer the only concern—platforms exert tremendous power over communications that is more transnational, invisible, and less accountable than state power.[154] Of course, the danger of the state should not be ignored, given its monopoly over the use of force and its ubiquitous exercise of power as a sovereign.[155] On the other hand, regulatory competition may be beneficial to users' freedom since it may function similarly to checks and balances.[156] This Madisonian idea reminds us that "having two powerful regulators, rather than only one, can sometimes strengthen individuals' freedom,

152. *See* S. 7072, 2021 Leg., Reg. Sess. (Fla. 2021); H.B. 20, 87th Leg., 2nd Sess. (Tex. 2021).

153. *See* Balkin, *Triangle*, *supra* note 25, at 2055.

154. Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2308–09 (2014); Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1149, 1153 (2018); Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. PA. L. REV. 11, 13–14 (2006); *see* Marvin Ammori, *Another Worthy Tradition: How the Free Speech Curriculum Ignores Electronic Media and Distorts Free Speech Doctrine*, 70 MO. L. REV. 59, 64 (2005) [hereinafter Ammori, *Worthy Tradition*].

155. Ammori, *Worthy Tradition*, *supra* note 154, at 73.

156. *See id.* at 76.

liberty, and security because often it takes a powerful regulator to challenge and check another powerful regulator."[157]

B. Putting Users' Rights at the Core

Judicial review ensures that both the public regulatory power of governments and the private regulatory power of platforms are kept within boundaries that do not harm the interests of platform users.[158] That is, a focus on the interests and rights of users should be at the center of judicial scrutiny. Yet, the current debate on content moderation largely ignores users' rights.[159]

Once courts recognize the multiplicity of regulators and the centrality of users' rights, the key issue for judicial scrutiny should not be about the content-based infringement of a platform's editorial rights. Instead, judicial scrutiny should focus on whether the regulatory preemption promotes or impairs the expressive interests of users.[160] In this context, the government uses its own regulatory scheme to override and preempt the content moderation schemes of private platforms.

Recognizing the centrality of users' rights is not a novel idea. Rather, it echoes the First Amendment traditions that predate the entrenchment of the present Lochnerized, libertarian tradition.[161] In *Red Lion Broadcasting Co. v. FCC*, the Court stressed the importance of users' rights, stating "[i]t is the right of the viewers and listeners, not the right of the broadcasters, which is paramount."[162] Yet, today, First Amendment jurisprudence has lost sight of its focus on the individual user; it has been captured by corporate rights and deregulatory agendas.[163] But a return to the interests of individual speakers and listeners, rather than corporate regulators, is still possible. Such reframing invites consequential and normative considerations into our currently formalistic thinking.[164]

However, setting boundaries around users' rights is not an easy task. Indeed, determining what is good for free speech unavoidably

---

157. Kristen Eichensehr, *Digital Switzerlands*, 167 U. PA. L. REV. 665, 715 (2019).

158. Moody v. NetChoice, LLC, 603 U.S. 707, 716 (2024).

159. *See supra* Section I.

160. Persily, *supra* note 53, at 196 ("Government regulation of social media might implicate the constitutional rights of users, but the platforms themselves . . . should not be viewed as having rights that stand in the way of regulation of content moderation or other practices.").

161. *See* Genevieve Lakier, *The First Amendment's Real* Lochner *Problem*, 87 U. CHI. L. REV. 1241, 1247–48 (2020).

162. Red Lion Broad. Co. v. FCC, 395 U.S. 367, 390 (1969).

163. Lakier, *supra* note 161, at 1298–99.

164. *See id.* at 1248.

carries value judgments that defy easy solutions.[165] Sometimes what is good for the short-term interests of users may harm their long-term interests.[166] But the complications do not stop there. The several competing and equally convincing visions of free speech that dictate how the online public square should be regulated complicate the issue even further.[167]

Regulatory competition occurs when different regulators regulate the same speech forum.[168] In the context of social media, there are state regulators and private regulators.[169] State regulators alone constitute a multitude of different governments with equally differing views and regulatory schemes.[170] Moreover, regulatory competition is even further complicated by the fact that the multiplicity of regulators occurs not only at a horizontal level, but also at a vertical level.[171] The polycentric governance of the internet stems from global, regional, national, and subnational actors.[172] Motivated by different normative concerns, these regulators impose competing visions of free speech on the platforms.[173] These background norms, and the competition among them, constitute the biggest challenge for governing online platforms.[174]

---

165. *See id.*

166. *See* Balkin, *Triangle*, *supra* note 25, at 2047–49. Take the current business model of social media as an example. Social media platforms provide free services to users in exchange for the attention and personal information extracted from the users. *See id.* at 247. This model drove the popularity of social media and expanded the chances of users to freely communicate with others. But in the long term, the fast-growing industry out of this model may bring a lot of problems that harm users' rights, such as surveillance, manipulation, and echo chambers. *See* Jack M. Balkin, Lecture, *The First Amendment in the Second Gilded Age*, 66 BUFF. L. REV. 979, 986–96 (2018) [hereinafter Balkin, *Second Gilded Age*].

167. *See* Balkin, *Triangle*, *supra* note 25, at 2015.

168. *See, e.g.*, Denver Area Educ. Telecomms. Consortium, Inc. v. FCC, 518 U.S. 727, 732 (1996).

169. *See* Balkin, *Triangle*, *supra* note 25, at 2055.

170. Not only different nation states, but also different governments within a nation state can issue different laws regulating social media. Texas and Florida laws are examples of this kind. *See* S. 7072, 2021 Leg., Reg. Sess. (Fla. 2021); H.B. 20, 87th Leg., 2nd Sess. (Tex. 2021).

171. Jan Aart Scholte, *Polycentrism and Democracy in Internet Governance*, *in* THE NET AND THE NATION STATE 165, 168 (Uta Kohl ed., 2017).

172. *Id.* at 168–69.

173. *Id.* at 169.

174. *See id.* at 173.

IV. COMPETING VISIONS OF FREE SPEECH

A. Why Background Norms Are Indispensable for Free Speech

There is no such thing as an unmediated speech environment.[175] From traditional public forums, like streets and parks, to the online world, background norms govern expression.[176] Because free speech is value-oriented (i.e., speech is not protected for its own sake but for realizing values or ideals), mere non-censorship is far from enough for realizing the goal of free speech. [177] Rather, there must be institutions and norms that intermediate the public discourse,[178] so that discourse is not only "uninhibited," but also "robust, and wide-open."[179]

The background norms of free speech, also described as "communications order"[180] or civility rules,[181] refer to the rules that govern the behavior of communications in a given forum.[182] These norms aim to preserve the forum's vibrancy and promote the values associated with free speech while also delineating the boundaries of dialogic behaviors.[183] These norms do so by encouraging behaviors facilitating free speech values and prohibiting or punishing others that hinder their realization.[184] They are not only rules that regulate the manner, scope, and medium of speech, but also norms about content.[185]

Without background rules, public discourse would become chaos.[186] One notable experiment is the Wikitorial project launched by

---

175. *See* Jack M. Balkin, *To Reform Social Media, Reform Informational Capitalism*, *in* SOCIAL MEDIA, FREEDOM OF SPEECH, AND THE FUTURE OF OUR DEMOCRACY 233, 240 (Lee C. Bollinger & Geoffrey R. Stone eds., 2022) [hereinafter Balkin, *Reform Informational Capitalism*].

176. *See id.*

177. *See* Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) ("[T]he power of the thought to get itself accepted in the competition of the market . . . ."); ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT 27 (1948) (justifying the protection of free speech upon its role in facilitating democratic self-government); Redish, *supra* note 89, at 593 (arguing that free speech primarily serves the value of self-realization and self-development).

178. Jack M. Balkin, *How to Regulate (and Not Regulate) Social Media*, 1 J. FREE SPEECH L. 71, 78 (2021) [hereinafter Balkin, *Regulate Social Media*].

179. *See* N.Y. Times Co. v. Sullivan, 376 U.S. 254, 270 (1964).

180. Baker, *supra* note 93, at 85.

181. Robert C. Post, *The Constitutional Concept of Public Discourse: Outrageous Opinion, Democratic Deliberation, and* Hustler Magazine v. Falwell, 103 HARV. L. REV. 601, 629–30 (1990).

182. *Id.* at 639.

183. *See generally* Baker, *supra* note 93, at 91.

184. *See id.* at 62.

185. *See id.* at 104.

186. TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA 5 (2018).

the Los Angeles Times.[187] The Wikitorial project is a libertarian and anarchical version of Wikipedia, free from background norms and moderation.[188] The result, unsurprisingly, was total failure; within several days, the site was filled with offensive and obscene content.[189] The L.A. Times quickly terminated the experiment.[190] As exemplified by the Wikitorial project, a totally hands-off approach without any moderation creates an online community full of hatred, spam, and extremism.[191]

The need for mediation or moderation according to background norms is greater in the online environment. Before the internet era, communications were not anonymous and the reach of a message was limited by physical space.[192] In cyberspace, speed, anonymity, and low cost of communication make it much easier to spread hateful and damaging content.[193] Without active moderation, the online public sphere will collapse into a landfill of information.[194] Information overload on the internet creates a scenario in which more information is not better, but merely adds noise that decreases the efficiency of communication.[195]

The key difference between social media and twentieth century mass media is not that the latter is curated while the former is not, but that the latter is curated under a set of professional ethics and journalistic norms, while the former is curated under a totally different logic.[196] The logic of the information economy mandates that platforms draw users' attention and collect their data.[197] For business purposes, social media companies need to moderate their platforms in a way that is appealing to users and advertisers.[198] For public interest purposes,

187. Dan Glaister, *LA Times 'Wikitorial' Gives Editors Red Faces*, GUARDIAN (June 22, 2005, 7:23 AM), https://www.theguardian.com/technology/2005/jun/22/media.pressandpublishing [https://perma.cc/D7LY-RLBR].

188. *See id.*

189. *See id.*

190. *See id.*

191. *See* James Grimmelmann, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 44–45 (2015).

192. Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353, 1361 (2018).

193. *Id.* at 1358–59.

194. Justin "Gus" Hurwitz, *Noisy Speech Externalities*, 3 J. FREE SPEECH L. 159, 180 (2023).

195. *Id.* at 164–70.

196. Larry Kramer, *A Deliberate Leap in the Opposite Direction: The Need to Rethink Free Speech*, *in* SOCIAL MEDIA, FREEDOM OF SPEECH AND THE FUTURE OF OUR DEMOCRACY 17, 23 (Lee C. Bollinger & Geoffrey R. Stone eds., 2022).

197. *Id.* at 34.

198. *See id.* at 38.

the platforms should be moderated in a manner that effectively hosts public discourse.[199] These two interests do not always align. Without any government intervention, platform moderation would likely sacrifice the public interest in pursuit of private profits.[200] Of course, public pressure and reputational concern can occasionally drive platforms to look after the public interest.[201] But public pressure is not enough.[202] Just take a glimpse at the scandals or controversies caused by platforms.[203] Indeed, relying on news headlines to exert public pressure on the platforms is not a sustainable strategy to hold them accountable. Thus, government regulation is necessary to ensure that the content moderation does not substantially deviate from the public interest in free speech.[204]

Therefore, adherence to neutrality in content moderation is a suicide pact.[205] Platforms must choose at least one set of norms to manage content.[206] Determining the boundaries of public discourse involves many factors, such as "balancing offense and importance; reconciling competing value systems; mediating when people harm one another, intentionally or otherwise; honoring the contours of political discourse and cultural taste; grappling with inequities of gender, sexuality, race, and class; extending ethical obligations across national, cultural, and linguistic boundaries."[207] However, norms governing online speech are not easily agreed upon and are in constant competition with each other.[208] The two most

199. *See id.*

200. *See* Margaret Sullivan, *Facebook Is Harming Our Society. Here's a Radical Solution for Reining It In.*, WASH. POST (Oct. 5, 2021, 6:00 AM), https://www.washingtonpost.com/lifestyle/media/media-sullivan-facebook-whistleblower-haugen/2021/10/04/3461c62e-2535-11ec-8831-a31e7b3de188_story.html [https://perma.cc/GDQ4-2XFH] ("'Facebook has realized that if they change the algorithm to be safer, people will spend less time on the site, they'll click on less ads, they'll make less money' . . . .") (quoting Frances Haugen, former member of Facebook's civic integrity team).

201. *See* Balkin, *Regulate Social Media*, *supra* note 178, at 88.

202. *See id.* at 87–88.

203. *See Protecting Kids Online: Testimony from a Facebook Whistleblower: Hearing Before the Subcomm. on Consumer Prot., Prod. Safety, & Data Sec. of the S. Comm. on Com., Sci., & Transp.*, 117th Cong. 17–19 (2021) (statement of Frances Haugen, Facebook Whistleblower).

204. Evelyn Douek, *The Siren Call of Content Moderation Formalism*, *in* SOCIAL MEDIA, FREEDOM OF SPEECH AND THE FUTURE OF OUR DEMOCRACY 139, 141 (Lee C. Bollinger & Geoffrey R. Stone eds., 2022) ("And while platforms' interests will not align with the public's, there is room (albeit limited) for stakeholders like government, civil society, advertisers and users to push platforms to reform their rules to respond to their preferences.").

205. *See generally* Anupam Chander & Vivek Krishnamurthy, *The Myth of Platform Neutrality*, 2 GEO. L. TECH. REV. 400, 415–16 (2018).

206. *See id.* at 216.

207. GILLESPIE, *supra* note 186, at 10.

208. *See id.*

influential competing normative visions are the US vision and the European Union (EU) vision.[209]

B. US Vision vs. EU Vision

Before the emergence of the internet, the norms governing free speech were either national or local.[210] They came from historical traditions, shared social and moral values, and positive law,[211] but they largely remained within state borders.[212] Today's social media platforms create an entirely different landscape. Although Facebook, X, and YouTube are US companies,[213] the majority of their users are from other countries and their non-US user base outside the US is growing faster than that within the US.[214] These international features make value pluralism on these platforms inevitable.

Because the biggest platforms are based in the US, the US vision of free speech surely affects them most.[215] The US vision is characterized by its adherence to neutrality, distrust of government, and the preferred status of free speech in the constitutional rights system.[216] But platforms are transnational corporations with users in different parts of the world.[217] With so many users active on these platforms outside of the US, "the First Amendment is merely a local ordinance."[218]

In such global platforms, we see another competing regulatory vision that is playing an increasingly important role: the EU vision.[219] The distinctions between the two visions are stark. First, Europeans do

---

209. *See* Gill, *supra* note 81, at 182.

210. *See* Jeremy Waldron, *Free Speech Apart from Law*, 2 J. FREE SPEECH L. 107, 108 (2022).

211. *See* Damian Tambini, *Reconceptualizing Media Freedom*, *in* REGULATING BIG TECH: POLICY RESPONSES TO DIGITAL DOMINANCE 299, 299 (Martin Moore & Damian Tambini eds., 2022).

212. *See id.*

213. *See* Alexis C. Madrigal, *Are Facebook, Twitter, and Google American Companies?*, ATLANTIC (Nov. 1, 2017), https://www.theatlantic.com/technology/archive/2017/11/are-facebook-twitter-and-google-american-companies/544670/ [https://perma.cc/5KF7-LAEN].

214. *See* Eichensehr, *supra* note 157, at 697.

215. *See* Thomas E. Kadri & Kate Klonick, Facebook v. Sullivan*: Public Figures and Newsworthiness in Online Speech*, 93 S. CALIF. L. REV. 37, 71 (2019).

216. *See* Post, *Recuperating*, *supra* note 120, at 1276 (stating that the American commitment of free speech "immunizes public discourse from the legal imposition of community norms of decency and respect").

217. *See* Eichensehr, *supra* note 157, at 697.

218. Marvin Ammori, *The "New"* New York Times*: Free Speech Lawyering in the Age of Google and Twitter*, 127 HARV. L. REV. 2259, 2278 (2014).

219. *See* Gill, *supra* note 81, at 182.

not have the same level of animosity toward content-based speech regulations.[220] The EU vision views content-based regulation as "benign."[221] Second, unlike the US jurisprudence that views free speech as a preferred liberty that enjoys special treatment,[222] the EU vision takes a more balanced approach, under which free speech is put on the same level with other values such as privacy, dignity, and equality.[223] Third, the EU vision does not share the same ethos of distrust of government.[224] This is exemplified by Europe's pervasive prohibition of hate speech, a category of speech generally protected in the US.[225]

In sum, the EU vision trusts governments more than individuals.[226] It allows governments to play a more active role in shaping public debate and emphasizes the vulnerability of individuals.[227] The EU vision seems to view human capabilities as weak, myopic, and credulous when it comes to speech considerations.[228] Yet, it also takes a more holistic view toward power dynamics by recognizing that private power can also threaten liberties.[229] In contrast, the US vision places greater trust in individuals than governments[230] and touts that human autonomy and agency is best achieved without government interference.[231] It places more confidence in individuals to counter public and private powers and views governments as more dangerous than private actors.[232]

Both visions have exerted influence on global platforms' governance and, as a result, regulatory competition has emerged as a

---

220. *See* ANU BRADFORD, THE BRUSSELS EFFECT: HOW THE EUROPEAN UNION RULES THE WORLD 39 (2020) [hereinafter BRADFORD, BRUSSELS EFFECT]. For American courts' fatal approach against content-based laws of speech, *see* Barry P. McDonald, *Speech and Distrust: Rethinking the Content Approach to Protecting the Freedom of Expression*, 81 NOTRE DAME L. REV. 1347, 1351–52 (2006).

221. Guy E. Carmi, *Dignity Versus Liberty: The Two Western Cultures of Free Speech*, 26 B.U. INT'L L.J. 277, 347 (2008).

222. Richard H. Fallon, Jr., *Strict Judicial Scrutiny*, 54 UCLA L. REV. 1267, 1289 (2007).

223. *See* Fredrick Schauer, *The Exceptional First Amendment*, *in* AMERICAN EXCEPTIONALISM AND HUMAN RIGHTS 29, 42 (Michael Ignatieff ed., 2005); *see also* Gehan Gunatilleke, *Justifying Limitations on the Freedom of Expression*, 22 HUM. RTS. REV. 91, 96 (2021).

224. *See* Waldron, *supra* note 210, at 115–16.

225. Michel Rosenfeld, *Hate Speech in Constitutional Jurisprudence: A Comparative Analysis*, 24 CARDOZO L. REV. 1523, 1523 (2003).

226. *See* BRADFORD, BRUSSELS EFFECT, *supra* note 220, at 39.

227. *See* Carmi, *supra* note 221.

228. *See id.* at 367.

229. *See id.* at 365.

230. *See* Tambini, *supra* note 211, at 299, 301.

231. *See* BRADFORD, BRUSSELS EFFECT, *supra* note 220, at 39.

232. *See* Tambini, *supra* note 211, at 301.

salient phenomenon.[233] Two examples illustrate how Facebook's rules of governance have been shaped by the competing visions from the two sides of the Atlantic.

The first example involves Facebook's content rules. A quick glimpse over Facebook's Community Standards reveals its similarities to the EU vision.[234] Like the EU vision, the platform prohibits a broad range of speech, including violence, incitement, deception, nudity, harassment, hate speech, and misinformation,[235] many of which are protected categories of speech under the First Amendment.[236] The Community Standards deviate from First Amendment jurisprudence, which takes a more libertarian approach to speech,[237] and embraces the EU vision. Under the EU vision, speech should be more actively regulated and balanced with other rights or interests to improve democratic discourse.[238]

The influence of the US tradition of promoting free expression is still apparent in Facebook's Community Standards. Unlike the EU vision, which places free speech on par with other values, Facebook explicitly states that "[o]ur commitment to expression is paramount," thus placing higher value on expressive freedom than other interests.[239] In short, even though Facebook's approach is generally more balanced and therefore analogous to the EU vision, it places free speech at a preferred status over other competing values, exemplifying the influence of US tradition.[240]

The second example involves Facebook's reconciliatory treatment of hate speech. Hate speech regulations sharply distinguish the EU vision from the US vision. Generally, European countries regulate hate speech in a far more comprehensive way than in the US.[241] In Europe, it is constitutionally permitted to ban hate speech per

---

233. *See* Kadri & Klonick, *supra* note 215; Gill, *supra* note 81, at 182.

234. *Facebook Community Standards*, *supra* note 73; Paolo Cavaliere, *Digital Platforms and the Rise of Global Regulation of Hate Speech*, 8 CAMBRIDGE INT'L L.J. 282, 295–96 (2019).

235. *See Facebook Community Standards*, *supra* note 73.

236. *See* Matal v. Tam, 582 U.S. 218, 246 (2017) ("Speech that demeans on the basis of race, ethnicity, gender, religion, age, disability, or any other similar ground is hateful; but the proudest boast of our free speech jurisprudence is that we protect the freedom to express 'the thought that we hate.'" (quoting United States v. Schwimmer, 279 U.S. 644, 655 (1929))); United States v. Alvarez, 567 U.S. 709, 710 (2012) ("[T]he Court has instructed that falsity alone may not suffice to bring the speech outside the First Amendment.").

237. *See Facebook Community Standards*, *supra* note 73; *Matal*, 582 U.S. at 246.

238. *See Facebook Community Standards*, *supra* note 73.

239. *See id*.

240. *See id.*

241. *See* Cavaliere, *supra* note 234.

se, while such a categorical prohibition is only constitutional in the US when the speech is sufficiently likely to cause imminent harm.[242]

Facebook and other social media platforms largely follow the EU approach and extensively ban hate content in their platforms.[243] The Facebook Community Standards explicitly prohibit hate speech, which is an endorsement of the EU vision.[244] Facebook and other platforms' endorsement of the EU vision can also be seen from the Code of Conduct on Countering Illegal Hate Speech Online, which several major social media platforms voluntarily signed with the European Commission.[245]

More complicated, however, are the decisions of Facebook's quasi-judicial Oversight Board, which adjudicates appeals concerning Facebook's moderation decisions.[246] The Board has been hesitant to clarify whether a likelihood of imminent harm, the very condition that distinguishes the US vision from the EU vision, is strictly required for censoring hate speech.[247]

For example, upon reviewing Facebook's decision to remove a demeaning slur toward an ethnic group, Facebook's Board seemed to endorse the EU vision of prohibiting hate speech.[248] It determined that, even without the likelihood of an imminent threat, hate speech can be legitimately banned from the platform because it has the effect of "creating a discriminatory environment that undermines the freedom of others to express themselves."[249]

Yet, in another case concerning a similarly pejorative slur, the Board expressed a different view, causing a split among Board members.[250] There, the Board overturned Facebook's removal decision after a majority of its members voted to do so.[251] According to the majority, even though the post referred to a sword, it was a permitted

242. *See id.*; Rosenfeld, *supra* note 225, at 1529.

243. *See Facebook Community Standards*, *supra* note 73.

244. *See Facebook Community Standards–Hate Speech*, META TRANSPARENCY CTR., https://transparency.fb.com/policies/community-standards/hate-speech/ [https://perma.cc/HXT2-AQAH] (last visited Sept. 29, 2024).

245. *Code of Conduct on Countering Illegal Hate Speech Online* (2016), https://ec.europa.eu/newsroom/just/document.cfm?doc_id=42985 [https://perma.cc/LJK5-D7GR].

246. *See generally* Kate Klonick, *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*, 129 YALE L.J. 2418, 2476–77 (2020).

247. *Compare Case Decision 2020-003-FB-UA*, META OVERSIGHT BOARD, § 8.3 (2020), https://www.oversightboard.com/decision/FB-QBJDASCV/ [https://perma.cc/66RU-VPYW], *with Case Decision, 2020-007-FB-FBR*, META OVERSIGHT BOARD, § 8.3 (2020), https://www.oversightboard.com/decision/FB-R9K87402/ [https://perma.cc/8G4V-N2SH].

248. *See Case Decision 2020-003-FB-UA*, *supra* note 247.

249. *Id.*

250. *See Case Decision, 2020-007-FB-FBR*, *supra* note 247.

251. *See id.*

criticism rather than a violent threat because "imminent . . . physical harm was unlikely to result from this post."[252] The minority of the Board dissented from the majority's insistence on an imminence standard, arguing that "Facebook should not wait for violence to be imminent before removing content that threatens or intimidates those exercising their right to freedom of expression."[253] Notably, the majority's decision did not mark a shift of the Board's view. The Board later switched its stance by upholding a removal decision even when no imminent harm exists.[254]

Such inconsistency and sway reveal the platform's struggle between the two visions of free speech. Although the EU vision is the mainstream for the Western world and appeals to many users,[255] the major social media platforms are US corporations that are conditioned to accept First Amendment traditions.[256]

Business incentives also help explain Facebook's reluctance to strictly adhere to one vision of free speech. On one hand, a highly permissive attitude toward speech, like that of the US, offers the platform a philosophy that endorses diversity and openness, which is beneficial to its business interests—more speech means more profit.[257] On the other hand, a laissez faire marketplace of ideas can be unappealing to users and advertisers seeking to avoid offensive or controversial messaging.[258] In those cases, the EU vision provides the best reference point for the platform to police its communicative spaces.[259]

Because Europe takes a more proactive approach to regulating internet platforms than the US, current platform rules are primarily shaped by the EU vision.[260] This phenomenon echoes what law professor Anu Bradford calls the "Brussels Effect," a term used to describe the dominance of EU regulations upon the world market.[261]

In the context of online speech regulation, the competitive advantage of European countries derives from their early participation

252. *Id.*

253. *Id.*

254. *Case Decision 2022-001-FB-UA*, META OVERSIGHT BOARD, § 8.3 (2022), https://www.oversightboard.com/decision/FB-JRQ1XP2M/ [https://perma.cc/PCD8-DXQV].

255. Carmi, *supra* note 221, at 362.

256. *See* Kadri & Klonick, *supra* note 215.

257. *See* Klonick, *New Governors*, *supra* note 18, at 1627.

258. *See id.*

259. *See* Gill, *supra* note 81, at 182.

260. *See id.*

261. *See generally* BRADFORD, BRUSSELS EFFECT, *supra* note 220, at xiv.

in the regulatory race regarding the internet platforms.[262] Two salient examples include Germany's Network Enforcement Act (NetzDG)[263] and the EU's Digital Services Act.[264] The former requires platforms to remove illegal content within twenty-four hours, and the latter mandates strict transparency requirement upon the platforms' content moderation.[265] These rules are shaping the regulatory landscape of social media.[266]

The Florida and Texas laws at issue in *Moody* seem to be an attempt by US state governments to exert influence on global speech governance.[267] The real concern behind the Florida and Texas laws should not be their interference with the platforms' editorial rights, but their imposition of one vision of free speech upon the global and plural public square.

In an attempt to mandate neutrality within political speech, the Florida law requires platforms to host all speech from all political candidates.[268] Similarly, the Texas law imposes viewpoint neutrality upon the entire platform.[269] Both statutes hold neutrality as a central value and take a libertarian stance toward the marketplace of ideas.[270] The concern raised by these statutes lies not only in their imposition of the US vision into a space of value pluralism, but also in that the US vision is itself a minority view of free speech in the world.[271] Such imposition is "American First Amendment imperialism over the world's speech marketplace."[272]

Even though both the Florida and Texas laws limit their force within the respective jurisdictions of the two states,[273] it would be technically and economically impractical for platforms to devise a

---

262. Gill, *supra* note 81, at 182.

263. *See Netzwerkdurchsetzungsgesetz - NetzDG*, BUNDESAMT FÜR JUSTIZ (Jan. 9, 2017), https://www.gesetze-im-internet.de/netzdg/BJNR335210017.html [https://perma.cc/V72F-ZCMQ].

264. *See The Digital Services Act*, EUR. COMM'N, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en [https://perma.cc/4692-XRRJ] (last visited Nov. 27, 2024).

265. *See Netzwerkdurchsetzungsgesetz – NetzDG*, *supra* note 262; *The Digital Services Act*, *supra* note 263.

266. *See* Gill, *supra* note 81, at 182.

267. Anu Bradford has raised a similar idea when she borrowed the term "preemptive federalism" to describe the US effort of preempting the European regulatory hegemony. *See* BRADFORD, THE BRUSSELS EFFECT, *supra* note 220, at 261.

268. *See* S. 7072, 2021 Leg., Reg. Sess. (Fla. 2021).

269. H.B. 20, 87th Leg., 2d Sess. (Tex. 2021).

270. *See id.*; Fla. S. 7072.

271. Frederick Schauer, *The Exceptional First Amendment, in* AMERICAN EXCEPTIONALISM AND HUMAN RIGHTS, 29, 30 (Princeton Univ. Press ed., 2005).

272. Persily, *supra* note 53, at 205.

273. *See* Fla. S. 7072; Tex. H.B. 20.

particular set of community standards that apply only to Florida and Texas users. Technically, users can easily circumvent the IP-based geo-blocking that would identify Florida and Texas users by using a VPN service or encryption technology.[274] Economically, imposing content moderation that satisfies jurisdiction-specific regulations is costly.[275] Not only does it require tailored training for AI moderators, but it also requires retraining human moderators.[276]

The scale of content moderation further tips in favor of a uniform regulation of speech across platforms. Platforms have reasons to achieve consistency in their enforcement of content rules, since moderating in a country- and culture-specific way is much more costly, if not technically infeasible.[277]

Besides, as global public squares, the social media platforms are attractive precisely because of their global nature.[278] The key variable that drives the platforms' prosperity is the network effects, which became open and borderless upon the introduction of the internet.[279] If social media platforms were fragmented and became national rather than global, free speech values would be compromised because dialogues on these platforms would also be fragmented.

Therefore, the extraterritorial effect of the Florida and Texas laws cannot be ignored. Such effect will either balkanize the global speech forum if platforms tailor their rules according to state laws, or force a race to the bottom, as the most restrictive rule will become common.[280] Both results would hardly be acceptable.

Apart from the state regulations' substantive value conflicts, another concern is procedural in nature. Imposing one country's legal rules onto users from other countries causes a legitimacy concern because the lawmakers are not accountable to the people of other countries.[281] Preempting a platform's existing content moderation rules with one state's law may be censured as regulatory imperialism, for it

274. *See* BRADFORD, THE BRUSSELS EFFECT, *supra* note 220, at 165.

275. Langvardt, *supra* note 192, at 1376.

276. *Id.*

277. BRADFORD, THE BRUSSELS EFFECT, *supra* note 220, at 166.

278. *See* Guy Berger, *The Universal Norm of Freedom of Expression – Towards an Unfragmented Internet: Interview with Guy Berger*, in THE NET AND THE NATION STATE: MULTIDISCIPLINARY PERSPECTIVES ON INTERNET GOVERNANCE 28 (Uta Kohl ed., 2017).

279. *Id.*

280. Graham Smith, *Cyberborders and the Right to Travel in Cyberspace*, in THE NET AND THE NATION STATE: MULTIDISCIPLINARY PERSPECTIVES ON INTERNET GOVERNANCE 126, 136 (Uta Kohl ed., 2017).

281. BRADFORD, THE BRUSSELS EFFECT, *supra* note 220, at 250.

"undermines the ability of foreign governments to serve their citizens in accordance with their democratically established preferences."[282]

The US and EU visions play predominant roles in shaping social media regulatory norms.[283] However, the juxtaposition of the two visions does not indicate a lack of internal variations or that the norms and principles are unchangeable.[284] It also does not mean that they are the only visions that govern the global public square. Recently, China has increased its salience in shaping the rules of the digital economy.[285] Indeed, the three superpowers (EU, US, and China), all deploying their own platform regulations, are competing to set a global standard.[286]

Thus, the inherent tension between free speech universalism and national sovereignty has been on the rise, and, with it, the tension between global values and local traditions.[287] Determining how to move forward in the face of such regulatory competition and value conflict is fraught with difficulty. Platforms are struggling with how to maintain the global speech forum while simultaneously accommodating local norms.[288] For example, TikTok lists "respect for local context" as one principle of its community rules, stating that it will recognize local cultures on its transborder platform.[289] However, TikTok's enforcement of public interest exceptions to the content rules has emphasized consistency and uniformity: "[o]ur approach to content moderation uses the same criteria, no matter who creates it."[290] In a recent enforcement quarterly report, TikTok described how it has accommodated EU and US data protection laws in an effort to fulfill its "localised approach to regulatory compliance".[291] However, thus far, TikTok has only addressed its "respect for local context" policy with respect to data

282. *Id.*

283. ANU BRADFORD, DIGITAL EMPIRES: THE GLOBAL BATTLE TO REGULATE TECHNOLOGY 6 (2023) [hereinafter BRADFORD, DIGITAL EMPIRES].

284. *See, e.g.*, Jud Campbell, *The Emergence of Neutrality*, 131 YALE. L.J. 861, 865–69 (2022) (documenting that the notion of neutrality only entrenched itself in American constitutional jurisprudence in recent decades).

285. *See, e.g.*, Matthew Erie & Thomas Streinz, *The Beijing Effect: China's Digital Silk Road as Transnational Data Governance*, 54 N.Y.U. J. INT'L L. & POL. 2, 2 (2021).

286. *See generally* BRADFORD, DIGITAL EMPIRES, *supra* note 283, at 6.

287. *See id.* at 9.

288. *See id.* at 15.

289. *Community Guidelines Enforcement Report*, TIKTOK (June 30, 2023), https://www.tiktok.com/transparency/en/community-guidelines-enforcement-2023-1/ [https://perma.cc/Q8HG-9SK4].

290. *Enforcement*, TIKTOK (Apr. 17, 2024), https://www.tiktok.com/community-guidelines/en/enforcement/ [https://perma.cc/2E3D-FZP4].

291. *Community Guidelines Enforcement Report*, *supra* note 289.

protection; it has not yet detailed incidences involving local cultural norms and speech moderation.[292]

Reframing the central issue of social media moderation from "editorial discretion as expression" to "regulatory competition derived from conflicting values" has merits.[293] Such reframing liberates us from the current vocabulary and provides us with a chance of looking for new ways forward. It also forces us to take cultural conflicts seriously, rather than hiding them. The Supreme Court emphasized in *Moody* that "a State may not interfere with private actors' speech to advance its own vision of ideological balance."[294] But private actors must choose from the ideologies (or free speech visions) in regulating their platforms. Recognizing the inevitability inherent in a private actor's choice is the first step to approaching the issue of social media regulation.

V. THREE PROPOSALS FOR MOVING FORWARD

If the real battles behind social media regulation are value conflicts between different visions of free speech, then how should different stakeholders proceed in front of this moral and legal quagmire? This Part offers three proposals: (1) what the judiciary can do, (2) what the government can do, and (3) what the platforms can do. The proposals do not purport to solve the issue of social media regulation completely—we may never completely solve it. However, these proposals provide a strong basis for the future of social media regulation and online speech.

A. Embracing More Open-Ended Balancing in Judicial Review

When reviewing social media regulations, courts should bear in mind that it is not the speaker's expressive right being regulated, but two regulators competing for dominance in the online communicative space.[295] Putting users' rights at the core of their analysis and recognizing the competing visions of background speech norms requires the judiciary to move from categorical scrutiny to a more holistic and open-ended balancing approach.

The categorical approach's effectiveness is diminishing for three reasons. First, the rigid distinction between content-based and content-neutral regulations is inapposite, because content rules are

292. *Id.*; *see* Streinz, *supra* note 285, at 47.

293. *See* BRADFORD, DIGITAL EMPIRES, *supra* note 283, at 12; *see also* Candeub, *supra* note 15, at 158–59.

294. Moody v. NetChoice, 603 U.S. 707, 741 (2024).

295. BRADFORD, DIGITAL EMPIRES, *supra* note 283, at 183.

necessary to ensure the integrity and civility of online discourse.[296] The rigidity of the content neutrality principle ignores the value orientation of the First Amendment, for it is how the content fits with free speech values, rather than the inquiry of neutrality, that should really matter for judicial scrutiny.[297]

Second, the categorical approach deals with new scenarios—such as the emergence of social media platforms—by analogizing and differentiating them from existing categories of speakers, such as newspapers, broadcasting systems, and shopping centers.[298] However, the new technologies and features of the speech ecosystem may be too novel to fit into these existing categories.[299] Formulaic categorical thinking is not capable of keeping up with the rapidly evolving digital landscape.[300] As Justice Breyer observed, categorical approaches "import law developed in very different contexts into a new and changing environment, and they lack the flexibility necessary to allow government to respond to very serious practical problems."[301] In view of the complexity and novelty of contemporary First Amendment issues, as defined by rapid technological development and the growing scope of its coverage, adherence to the old days of formalism is inappropriate.[302]

Third, by focusing on the identification and application of traditional categories, the categorical approach is too narrow to adequately consider all relevant interests in speech cases, especially those of non-parties.[303] The current debate on social media regulations clearly reflects this point. Dominated by the judiciary's categorical approach, the debate has focused on the editorial rights of the platforms and interests behind the government regulations, while ignoring users' expressive rights.[304] The categorical and analogical reasoning is

296. Wilson R. Huhn, *Assessing the Constitutionality of Laws That Are Both Content-Based and Content-Neutral: The Emerging Constitutional Calculus*, 79 IND. L.J. 801, 802 (2004).

297. *Id.* at 803, 821.

298. *See, e.g.*, NetChoice, LLC v. Paxton, 573 F. Supp. 3d 1092, 1113 (W.D. Tex. 2021), *vacated and remanded*, NetChoice, LLC v. Paxton, 49 F.4th 439 (5th Cir. 2022), *vacated and remanded sub nom*. Moody v. NetChoice, LLC, 603 U.S. 707 (2024).

299. *See* Bradford, Digital Empires, *supra* note 283, at 183.

300. David S. Han, *Constitutional Rights and Technological Change*, 54 U.C. DAVIS L. REV. 71, 113 (2020) ("[T]he benefits rooted in the predictability, administrability, and judicial constraint associated with a rule-like regime are outweighed by the increasing lack of fit between the rigid rules in question and the rapidly evolving world in which they are applied.").

301. Denv. Area Educ. Telecomms. Consortium v. FCC, 518 U.S. 727, 741 (1996).

302. Han, *supra* note 300, at 121.

303. *See id.* at 119.

304. Pierre J. Schlag, *An Attack on Categorical Approaches to Freedom of Speech*, 30 UCLA L. REV. 671, 676 (1983).

self-reinforcement, which deprives judges of the ability to analyze detailed facts and values in specific cases.[305] Instead of examining value questions, the categorical approach merely fits fact patterns into existing categories and applies pre-ordained rules despite its incompatibility with social media.[306]

After the issues of editorial right and content neutrality are settled, the outcome of categorical review is mostly fixed, leaving little room for judges to consider the users' rights.[307] Under the categorical approach, "all the important work in litigation is done at the outset. Once the relevant right and mode of infringement have been described, the outcome follows, without any explicit judicial balancing of the claimed right against the government's justification for the infringement."[308]

Recognizing the limits of categorical strictness, courts should instead embrace a more open-ended balancing that considers multiple rights and interests, especially those of the users. The aim of regulation is not neutrality per se, but to shape the incentives of platforms to align with the public ends.[309] A new, "Breyerian" balancing approach would be more inclusive, accounting for the relevant interests of all sides.[310] This approach not only considers the expressive and proprietary interests of platforms and the regulatory interests of the government, but also (and perhaps more importantly) gives due weight to the expressive interests of the general public.[311]

Open-ended balancing considers a variety of rights and interests not in an arbitrary way, but according to the normative values of free speech.[312] The appropriate question for judges to ask is not whether platforms enjoy editorial discretion in choosing what content to host, but whether their discretion could promote the discovery of knowledge, the democratic discourse, and the personal autonomy of their users.[313]

Such holistic scrutiny may give the government and courts more latitude to rein in the platforms, since some content-based regulations

---

305. *Id.* at 735.

306. *Id.* at 733–34.

307. Kathleen M. Sullivan, *Post-Liberal Judging: The Roles of Categorization and Balancing*, 63 U. COLO. L. REV. 293, 306 (1992).

308. *Id.* at 293.

309. Balkin, *Regulate Social Media, supra* note 178, at 71, 90.

310. Jerome A. Barron, *The Electronic Media and the Flight from First Amendment Doctrine: Justice Breyer's New Balancing Approach*, 31 U. MICH. J.L. REFORM 817, 883 (1998).

311. *Id.*

312. *See* Alan Z. Rozenshtein, *Silicon Valley's Speech: Technology Giants and the Deregulatory First Amendment*, 1 J. FREE SPEECH L. 337, 371 (2021).

313. *See id.* at 371.

would be upheld for their beneficial effect on the public discourse.[314] However, the judiciary should not defer to any legitimate reason the government may offer to justify a regulation. The court must closely scrutinize the intention and rationale of the regulation, as well as its design. For example, although many provisions of the Texas and Florida laws may serve important goals, such as the public's right to know the internal workings of moderation,[315] they may still be viewed as poorly drafted and fail to achieve their goals in a proportionate way.

B. Nudging Private Governance

The second proposal is nudging private governance. This Article uses "nudge" rather than "regulate" to refer to measures that do not directly impose substantive content rules to the platforms, but that require the platforms to enact, revise, and enforce their content rules in a better way. Here, a "better way" can carry several meanings; it can mean more transparency, more democracy, and more accountability. Nudging uses various mechanisms—hard and soft, legal and nonlegal—to ensure that platform governance conforms to constitutional and democratic values.[316]

By refraining from regulating the substantive rules of content moderation directly, nudging avoids thorny First Amendment issues and the risks of taking sides on value conflicts. [317] Taking sides on value conflicts is objectionable because it would entrench government-endorsed values while suppressing future dialogue on values.[318] Focusing on procedural issues may be the best short-term solution.[319] Rather than imposing one set of values on global platforms, governments can make an effort to ensure that platforms are governing themselves through better procedures.

Nudging aims to incentivize.[320] Because of the misalignment between platforms' business interests and the public interest, platforms have the motivation, at least on certain occasions, to abuse their power and sacrifice the users' interests.[321] Nudging re-aligns the two types of

314. *See id.* at 361–62.

315. *Id.* at 361.

316. GIOVANNI DE GREGORIO, DIGITAL CONSTITUTIONALISM IN EUROPE: REFRAMING RIGHTS AND POWERS IN THE ALGORITHMIC SOCIETY 295 (2022).

317. Persily, *supra* note 53, at 208.

318. *See id.*

319. PAUL GOWDER, THE NETWORKED LEVIATHAN: FOR DEMOCRATIC PLATFORMS 53 (2023).

320. Balkin, *Regulate Social Media, supra* note 178, at 72.

321. *Id.* at 90.

interests and motivates the platforms to adopt public law principles[322] as their guidance in designing their private governance.[323]

Section 230 is one important tool for nudging.[324] Many argue for changing the nearly absolute immunity Section 230 imposes on social media platforms to conditional immunity by, for example, "condition[ing] the immunity on social media companies' agreeing to accept a new set of public interest obligations."[325] Another avenue for improving platform governance is through industry initiatives, such as the Santa Clara Principles[326] and the UN Guiding Principles on Business and Human Rights (UNGP).[327] These voluntary agreements reflect social media platforms' commitment to observing human rights norms.[328] Governments should encourage and nudge platforms by elaborating and enforcing these norms.

Through conditioning Section 230 immunity and channeling industry initiatives, governments can nudge platforms to fulfill "due process" obligations in their content moderation schemes.[329] These obligations can include disclosing more information about content moderation, reporting serious abuses to the government and public, and coordinating with various stakeholders to collectively improve moderation. To be sure, these obligations can be directly imposed on platforms through legislation, as exemplified by the Texas and Florida laws.[330] Indeed, the disclosure and transparency requirements of the Texas and Florida laws are the most likely provisions to be upheld by the Supreme Court.[331]

Users should also contribute to the platforms' governance structure. So far, governments, platforms, industrial associations, and other international human rights bodies facilitate the accountability

322. Such as accountability, democracy, and transparency. *Id.* at 89–90.

323. *Id.* at 90.

324. 47 U.S.C. § 230 (2018).

325. Balkin, *Reform Informational Capitalism, supra* note 175, at 253.

326. *The Santa Clara Principles on Transparency and Accountability in Content Moderation*, SANTA CLARA PRINCIPLES (2018), https://santaclaraprinciples.org/ [https://perma.cc/NL2Z-HS4N].

327. *UN Guiding Principles on Business and Human Rights*, UNITED NATIONS (2011), https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf [https://perma.cc/2E47-33PH].

328. *The Santa Clara Principles on Transparency and Accountability in Content Moderation*, *supra* note 325; *UN Guiding Principles on Business and Human Rights*, *supra* note 326.

329. Balkin, *Reform Informational Capitalism, supra* note 175, at 253.

330. PACT Act, S. 797, 117th Cong. (2021).

331. *See* Rozenshtein, *supra* note 312.

and transparency of platform governance.[332] Noticeably absent from this list is users.[333] One step toward resolving value conflicts in global platforms is empowering users to enact civility norms. In other words, users are not only participants but also lawmakers of the public sphere. Ideally, they should not only be consumers but also self-governing citizens.[334]

The laws and norms of a global public sphere should be enacted through a global electorate. Currently, platform rules are enacted by their internal teams under the leadership of top executives.[335] However, there are no institutional arrangements for users' participation in this "lawmaking" process,[336] and "[r]evisions of the [community] guidelines often come only in response to outcries and public controversies."[337] Lack of democratic input makes the legitimacy of the rulemaking deficient from a normative perspective.[338]

To address the legitimacy deficit and promote quality governance, platforms should be required to provide more avenues for user participation. To facilitate industry-wide cooperation and decrease the regulatory discrepancies across different platforms, governments should encourage platforms to establish a social media self-regulatory council[339] or organization.[340] Such an industry-wide self-governing entity would facilitate more coordination across different sectors, solicit more feedback from both experts and the public, and promote more transparency regarding platform moderation.[341]

For example, law professor Paul Gowder proposed establishing a "multilevel participatory governance organization."[342] According to Gowder, such organization can be an industry-wide council composed of several participatory institutions that operates at local, regional, and

---

332. Balkin, *Reform Informational Capitalism, supra* note 175, at 89; Bradford, DIGITAL EMPIRES, *supra* note 283, at 6.

333. GILLESPIE, *supra* note 186, at 212.

334. *Id.* at 212.

335. *See id.* at 197.

336. *Id.* at 209.

337. *Id.* at 67.

338. *Id.*

339. Jennifer Grygiel & Nina Brown, *Are Social Media Companies Motivated to be Good Corporate Citizens? Examination of the Connection Between Corporate Social Responsibility and Social Media Safety*, 43 TELECOMMS. POL'Y 445, 457 (2019).

340. Katherine Adams, *Martin Baron, Lee C. Bollinger, Hillary Clinton, Jelani Cobb, Russ Feingold, Christina Paxson & Geoffrey R. Stone*, *Report of the Commission*, *in* Social Media, Freedom of Speech and the Future of Our Democracy 320 (Lee C. Bollinger & Geoffrey R. Stone eds., 2022).

341. *Id.*

342. GOWDER, *supra* note 319, at 3.

global levels.[343] Through representation, coordination, and interaction, the "system would give ordinary people, including ordinary people from currently unrepresented groups, a substantial amount of power to affect platform decisions."[344]

Proposals like Gowder's are meaningful experiments to conduct in future efforts to reform social media governance. In essence, these proposals reflect and enforce the core tenet that "multi-stakeholderism remains the overriding organizing principle for internet governance."[345] The democratic participation of users in the form of consultation, public comments, and voting will improve the legitimacy of platforms' rules and standards.[346]

Several issues must be addressed to implement this approach. First, in the absence of government involvement, democratic participation on private and transnational platforms must be sufficient to address their legitimacy deficit.[347] Second, even under a functional concept of democracy that does not limit itself to governments,[348] the problem of the clash between democratic legitimacy of the state and that of the platforms may remain.[349] The solution must also account for the potential conflicting visions of free speech held by a clash of a global public (the basis of democratic legitimacy for platforms) with a national public of a democratic state. For instance, if the latter wants to preempt the democratically enacted laws of the platform with a national law, whose vision of democracy should prevail?

Third, questions remain as to how to incentivize multiple stakeholders, especially ordinary users, to participate in the governance of platforms. Indeed, designing the mechanism to generate consensus or at least compromises in the rulemaking process remains an issue. Although platform-wide consensus or compromise can emerge, ensuring that the output of the process can be duly respected by the platform companies is paramount.[350]

---

343. *Id.* at 183–84.

344. *Id.* at 187.

345. David Kaye, *Legitimacy, Collective Authority and Internet Governance: A Reflection on David Caron's Study of the UN Security Council*, 46 ECOLOGY L.Q. 135, 145 (2019).

346. Mike Masnick, *Protocols, Not Platforms: A Technological Approach to Free Speech*, KNIGHT FIRST AMEND. INST. (Aug. 21, 2019), https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-to-free-speech [https://perma.cc/Z4T2-CG9N].

347. GOWDER, *supra* note 319, at 168 ("[I]t may be that the only morally valuable forms of democracy go through nation-states. But the question is at least open.").

348. *Id.*

349. *Introduction: Internet Governance and the Resilience of the Nation State*, *in* UTA KOHL & CARRIE FOX, THE NET AND NATION STATE: MULTIDISCIPLINARY PERSPECTIVES ON INTERNET GOVERNANCE 1, 14 (2017).

350. GOWDER, *supra* note 319, at 69.

Fourth, although popular participation may provide legitimacy dividends, there are also limitations inherent to popular opinion, like "censorship by popular demand."[351] Mechanisms must also be in place to address when democratically endorsed content rules of platforms disproportionally suppress the marginal and unpopular speakers.

C. Decentralizing the Platform

Currently, most major social media platforms, including X, Facebook, and YouTube, are moderating and managing content in a centralized way.[352] That is, each platform applies a uniform set of platform-wide content rules enacted by a central authority within the platform.[353] They also rely on their top-down bureaucratic structure to update and enforce these rules.[354]

Centralized governance poses two dilemmas to the effort of improving content moderation. The first dilemma derives from the tension between the uniformity of norms and the heterogeneity of cultures, communities, and regulatory visions in a platform.[355] As this Article explains, in a global public sphere, multiple visions of free speech exist and are shaped by different cultural, legal, and social traditions.[356] Thus, content moderation constantly faces tension between the consistency of rules across platforms and sensitivity to local contexts.[357] Platforms have already confronted the pressures from EU and US regulators who impose different visions of free speech.[358] On one hand, platforms have incentives to preserve consistency in their moderation.[359] On the other hand, platforms are incentivized to adapt to local context and respond to different regulatory calls.[360]

The second dilemma comes from the business model behind the centralized structure. Every system of free speech depends on a business model—after all, someone must pay for maintaining the public

---

351. Langvardt, *supra* note 192, at 1385.

352. Alan Z. Rozenshtein, *Moderating the Fediverse: Content Moderation on Distributed Social Media*, 3 J. FREE SPEECH L. 217, 221 (2023).

353. *Id.* at 228.

354. Robyn Caplan, *Content or Context Moderation? Artisanal, Community-Reliant, and Industrial Approaches*, 6 DATA & SOC'Y 1, 24 (Nov. 14, 2018), https://datasociety.net/library/content-or-context-moderation/ [https://perma.cc/Z3QD-CA8Q].

355. *See* Rozenshtein, *Moderating the Fediverse*, *supra* note 352, at 222.

356. *Id.* at 231.

357. Caplan, *supra* note 354.

358. *See* discussion *supra* Section IV.B.

359. *See, e.g.*, Nina I. Brown, *Regulatory Goldilocks: Finding the Just and Right Fit for Content Moderation on Social Platforms*, 8 TEX. A&M L. REV. 451, 473 (2021).

360. *Id.* at 475.

sphere.[361] Currently, the model is top-down, relying on massive collection of user data by the central authority.[362] Some scholars have called it "surveillance capitalism"[363] or "informational capitalism."[364] Because a platform's private interests of profit earning often collide with the public interest that users expect the platform to serve, platforms are not trusted to manage content.[365] The devil is in their business models, which incentivize them "to act irresponsibly and amplify false and harmful content."[366] All major platforms rely on maximizing user engagement to sell more ads, which incentivizes them to promote content that is emotional, sensational, as well as those that appeal to the dark side of human desires.[367] And because platforms' business models have built upon excessive collection of data and strategic moderation of content, they have little incentive to make their internal processes transparent and subject to public oversight.[368]

This Article argues that it is technically and economically unfeasible to customize speech rules according to the user's location.[369] Instead, encouraging government nudges on procedural mandates and imposing a more flexible and holistic judicial balancing incentivizes platforms to act responsibly.[370] However, this argument has a condition: the content moderation of platforms must remain centralized, and the business model must remain unchanged.

But the possibility of a future in which platforms are decentralized should not be ignored. Decentralizing social media platforms is a more radical way of responding to the current controversy behind moderation. The aim of decentralization is to split platforms' uniform management of a transnational community into numerous small communities that can create, enforce, and maintain norms.[371] This is similar to a vertical separation of powers; it delegates some powers from the central to the local or peripheral units, preserves local

361. Balkin, *Second Gilded Age*, *supra* note 166, at 986.

362. Sarita Schoenebeck & Lindsay Blackwell, *Reimagining Social Media Governance: Harm, Accountability, and Repair*, 23 YALE J.L. & TECH. 113, 150 (2021).

363. *See generally* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (2019).

364. *See generally* JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM (2019).

365. Brown, *Regulatory Goldilocks, supra* note 359, at 451.

366. Balkin, *Reform Informational Capitalism*, *supra* note 175, at 234.

367. *See id.* at 243.

368. *See id.* at 244.

369. *See supra* discussion Section IV.B.

370. Evelyn Douek, *Content Moderation as Systems Thinking*, 136 HARV. L. REV. 526, 591 (2022).

371. *See, e.g.*, Grimmelmann, *supra* note 191, at 51.

values, and increases legitimacy.[372] This new model would increase user agency by replacing the "one-size-fits-all" approach with a kind of governance that incorporates local control.[373]

Decentralization may be achieved in varying degrees.[374] Several platforms, such as Reddit and Mastodon, have successfully established decentralization in moderation.[375] Reddit allows its users to build, moderate, and manage communities (known as sub-Reddits).[376] However, those user-managed communities are subject to the platform rules imposed by the central authority of Reddit.[377] That is, users are free to enact and enforce content rules of their choice in the sub-Reddits, so long as they are in line with the central policies.[378] Mastodon also provides more user capacity to moderate locally.[379] For example, users can not only edit the filters of content,[380] but can also create their own servers and customize the speech environment they abide by.[381]

The most recent attempt at decentralizing social media is the BlueSky project.[382] BlueSky, a new social network, calls for the return to the structure of the early internet, when connections between people were built not upon centrally controlled platforms but end-to-end protocols.[383] BlueSky board member Mike Masnick has envisioned a decentralized model that does not rely on a single and centralized governance structure but rather grants any user the ability to create their own rules and enforce them.[384] However, BlueSky has not gone that far. Instead, it has combined elements from both platform-based

372. Caplan, *supra* note 354 at 21.

373. Schoenebeck & Blackwell, *supra* note 362, at 151.

374. *See* Langvardt, *supra* note 192 at 1357.

375. *See Reddit Content Policy*, REDDIT, https://www.redditinc.com/policies/content-policy [https://perma.cc/4KJN-KTRJ] (last visited Oct. 1, 2024).

376. *Id.*

377. *Id.*

378. *See Moderator Code of Conduct*, REDDIT (Sept. 25, 2023), https://www.redditinc.com/policies/moderator-code-of-conduct [https://perma.cc/K9XK-2A9H].

379. *See Dealing With Unwanted Content*, MASTODON, https://docs.joinmastodon.org/user/moderating/ [https://perma.cc/U3VA-MU9E] (last visited Oct. 1, 2024).

380. *See id.*

381. *See Running Your Own Server*, MASTODON, https://docs.joinmastodon.org/user/run-your-own/ [https://perma.cc/UM9Q-24JW] (last visited Oct. 1, 2024).

382. *See* Martin Kleppmann, Paul Frazee, Devin Ivy, Jake Gold, Jeromy Johnson, Jay Graber, Bryan Newbold, Daniel Holmgren & Jaz Volpert, *Bluesky and the AT Protocol: Usable Decentralized Social Media* (Feb. 5, 2024), https://arxiv.org/abs/2402.03239 [https://perma.cc/UTZ8-DYQY].

383. *About Bluesky*, BLUESKY, https://bsky.social/about/faq [https://perma.cc/S5ZA-T9XM] (last visited Nov. 27, 2024).

384. Masnick, *Protocols, Not Platforms, supra* note 346.

moderation and protocol-based moderation.[385] It has added community labeling into the existing two layers of moderation—centralized moderation by company admins and localized moderation by server admins.[386] Such a hybrid system aims to strike a balance between minimum consistency of community norms and platform safety on one hand, and flexibility and user agency on the other.[387]

Decentralization reform aims to devolve the power of prescribing speech rules from the monolithic host company (such as Meta, X, and Google) to a marketplace in which various participants—media organizations, nongovernmental organizations, and interested individuals—can design different sets of speech rules, or filters, that can be selected and adopted by different platform users.[388] Competition in such a marketplace of filters can not only realize individual users' personalized preferences, but also drive moderation practices toward a path that is more aligned with users' rights through market forces.[389] Moderation at the local, rather than global, level enables a "federation" of speech norms that is divided not by national borders or cultural traditions, but by individual choices and market dynamics.[390]

Although promising, the decentralized vision is far from a silver bullet. Several concerns should be further explored and addressed before the new model is fully embraced.

First, one of the biggest concerns for decentralized social media is that it may exacerbate the effects of filter bubbles and echo chambers.[391] Indeed, customizing content rules and user experiences may reinforce each user's previously held preferences, beliefs, and biases.[392] Such customization harms public discourse and democratic

385. Mike Masnick, *Why Bluesky Remains The Most Interesting Experiment In Social Media, By Far*, TECHDIRT (Mar. 27, 2024, 12:08 PM), https://www.techdirt.com/2024/03/27/why-bluesky-remains-the-most-interesting-experiment-in-social-media-by-far/ [https://perma.cc/ZK36-K6TG].

386. Jay Graber, *Composable Moderation*, BLUESKY (Apr. 13, 2023), https://bsky.social/about/blog/4-13-2023-moderation [https://perma.cc/JW9G-394G].

387. *See* Masnick, *supra* note 385.

388. Masnick, *Protocols, Not Platforms, supra* note 346.

389. *See* The Hill: *Government regulation of social media would kill the internet – and free speech*, PAC. LEGAL FOUND. (Aug. 12, 2019), https://pacificlegal.org/the-hill-government-regulation-of-social-media-would-kill-the-internet-and-free-speech/ [https://perma.cc/2Y34-S736].

390. Masnick, *Protocols, Not Platforms, supra* note 346.

391. Rozenshtein, *Moderating the Fediverse*, *supra* note 352, at 231.

392. *Combatting Online Harms Through Innovation,* FED. TRADE COMM'N 1, 48 (June 16, 2022), https://www.ftc.gov/system/files/ftc_gov/pdf/Combatting%20Online%20Harms%20Through%20Innovation%3B%20Federal%20Trade%20Commission%20Report%20to%20Congress.pdf[https://perma.cc/FG47-A63X] (reporting to Congress).

self-governance, which rely on cultivating tolerance and engagement with different views.[393]

The second concern lies in a sneaky devil in the details—the default settings in the decentralized networks. Many, if not most, people may not bother to change and customize the content rules and simply defer to a platform's default community rules.[394]

Third, one promise of the decentralized vision is that it ceases to rely on the current business model fueled by an endless thirst for user attention and data.[395] But a new business model must be identified before bidding farewell to the old one.[396] One possible mechanism is to develop token-based cryptocurrency,[397] but how that vision can be implemented remains to be seen.

Fourth, not everyone has the incentive (let alone expertise) to develop the content filters.[398] Governments may issue their content filters and organizations like the American Civil Liberties Union or New York Times may develop their own versions.[399] Some independent users with experience may also create content filters.[400] However, ordinary users do not have the capability to develop the filters or moderation tools.[401] Additionally, moderation operations requires huge costs, such as human resources, which may be unbearable to the decentralized communities.[402] If the only feasible future is a market of filters for users to choose, like all other markets, it is likely subject to market failures.[403] For instance, the market may be dominated by a few powerful entities offering the most popular filters, which may reproduce the power inequalities in the current centralized social media.

---

393. *See* Langvardt, *supra* note 192, at 1357.

394. *But see id.* at 1381.

395. Balkin, *Second Gilded Age*, *supra* note 166, at 999.

396. *See* Masnick, *Protocols, Not Platforms, supra* note 346.

397. *See id.*

398. *See generally How to Filter, Block, and Report Harmful Content on Social Media,* RAINN, https://rainn.org/articles/how-filter-block-and-report-harmful-content-social-media [https://perma.cc/C324-M2J3] (last visited Oct. 1, 2024).

399. *See* Masnick, *Protocols, Not Platforms, supra* note 346.

400. *See id.*

401. *See* Nicholas P. Dickerson, *What Makes the Internet So Special? And Why, Where, How, and By Whom Should its Content be Regulated?*, 46 HOUS. L. REV. 61, 77 (2009).

402. Rozenshtein, *Moderating the Fediverse*, *supra* note 352, at 229.

403. John Rothchild, *Protecting the Digital Consumer: The Limits of Cyberspace Utopianism*, 74 IND. L.J. 893, 952 (1999).

VI. CONCLUSION

Regulating social media platforms is regulating our public sphere.[404] And the goal of regulating social media platforms is to shape and define the norms of online discourse.[405] This Article argues that reviewing such regulations requires a holistic scrutiny of the value tradeoffs and cultural conflicts behind them. It is not only an issue of determining whether content moderation is editorial discretion and whether platforms are speakers. Rather, platforms are regulators that enact and enforce rules about the online public square—which communications are allowed, and which are not.[406]

Platforms are in quagmires of dilemmas when assuming their roles as regulators. They must make a choice amidst the cultural division between the US vision of free speech and the EU vision.[407] They also have to pay due respect to many other traditions in the world; each may be endorsing a particular understanding of what is taboo in public communications.[408] For Americans, banning the display of Nazi Swastika may seem a little over restrictive on free expression, but still understandable.[409] But what about the ban on any visual depiction of the Prophet Muhammad in Muslim countries?[410] What about the ban on any negative evaluation of national leaders and revolutionary martyrs in China?[411]

These issues are not novel—the conflict of free speech norms is a defining feature of a plural and heterogeneous society.[412] Although social media platforms did not create this problem, they have made it more salient.[413] For the first time in the history of human communications, people from across the globe can communicate and

404. Langvardt, *supra* note 192, at 1378.

405. *See* Brown, *Regulatory Goldilocks, supra* note 359, at 493.

406. Masnick, *Protocols, Not Platforms, supra* note 346.

407. *See* Balkin, *Second Gilded Age*, *supra* note 166, at 999.

408. *See* Rozenshtein, *Moderating the Fediverse*, *supra* note 351, at 222.

409. See Neil Schoenherr, *WashU Expert: The First Amendment and the Nazi Flag*, WashU (Aug. 16, 2017), https://source.washu.edu/2017/08/washu-expert-first-amendment-nazi-flag/ [https://perma.cc/W8AA-7L5X].

410. *See* Chris Johnston, *Facebook blocks Turkish page that 'insults prophet Muhammad*, GUARDIAN (Jan. 27, 2015, 11:24 AM), https://www.theguardian.com/technology/2015/jan/27/face-book-blocks-turkish-page-insults-prophet-muhammad [https://perma.cc/PFK5-2FQA].

411. *See* John Ruwitch, *China Makes First Use of Law Banning Defamation of National Heroes*, REUTERS (May 22, 2018, 4:24 AM), https://www.reuters.com/article/us-china-martyr-law-suit-idUSKCN1IN0J0/ [https://perma.cc/5NPF-DAHE].

412. *See* Rozenshtein, *Moderating the Fediverse*, *supra* note 352, at 231.

413. Langvardt, *supra* note 192, at 1358.

deliver their message to almost every corner of the world.[414] To maintain and manage a public square with such scope and scale, public and private regulators should confront the difficult conflicts among different free speech traditions and their underlying values. Several options are available to regulators: impose one ideal onto the entirety of the global platform, make a best effort in achieving compromises among different traditions, or decentralize platforms and delegate powers to local communities and users.

It is a difficult choice, to be sure, but one that must be made. This Article does not aim to make that choice, but to highlight the stakes and meaning of making it. In making this choice, stakeholders should step beyond First Amendment jurisprudence, especially the jurisprudence from recent decades, which takes content neutrality as a golden rule and treats corporate editorial rights as a powerful shield against government regulation.[415] This Article proposes three ways of reform: (1) a judicial approach that embraces open-ended balancing instead of strict categoricalism; (2) an administrative approach that nudges procedural governance by platforms and democratic participation by users; and (3) a technological approach that aims to decentralize the structure of social media. These proposals can serve as starting points that invite further discussion and experimentation. Adequately addressing this issue requires imagination and insight, as well as moral empathy and doctrinal flexibility.

414. *See* Peter Suciu, *Social Media Offers Plenty of Ways to Connect – But Are There Too Many Options?*, FORBES (June 1, 2023, 9:20 AM), https://www.forbes.com/sites/peter-suciu/2023/06/01/social-media-offers-plenty-of-ways-to-connect—but-are-there-too-many-options/ [https://perma.cc/SV7Y-JKAG].

415. *See generally* Morgan N. Weiland, *Expanding the Periphery and Threatening the Core: The Ascendant Libertarian Speech Tradition*, 69 STAN. L. REV. 1389, 1453 (2017).