

# Social Media as Contractual Networks: A Bottom Up Check on Content Moderation

*Niva Elkin-Koren, \* Giovanni De Gregorio\*\* & Maayan Perel\*\*\**

*ABSTRACT: Is there any limit to social media platforms' ostensibly unfettered discretion to indefinitely suspend users' accounts or remove content?*

*The fierce debate over the exercise of discretionary power by platforms to terminate users' accounts and remove content has primarily focused on free speech ramifications and the constitutional restraints on top-down legal interventions. While suspension and removal decisions by platforms often trigger questions situated in public law, they also raise important challenges to private law. Cutting the livelihood of small businesses, independent creators and political activists, termination and removal decisions may carry irreparable financial and reputational harms.*

*When moderating content, digital platforms exercise discretionary powers conferred under boilerplate contracts defining their Terms of Service. So far, platforms have successfully invoked contractual provisions as a shield against lawsuits of users claiming that unjustified suspension was a breach of contract. This Article argues that courts have often erroneously dismissed users' claims because they have misinterpreted the agreement between platforms and users as dyadic, namely involving two contracting parties. The interpretation of such contracts as establishing bilateral/vertical obligations only, undermines the true intention of the contracting parties and overlooks the plethora of rights and obligations created by such contracts to multiple*

---

\* Professor, Tel-Aviv University Faculty of Law; and a Faculty Associate, Berkman Klein Center at Harvard University.

\*\* Postdoctoral Researcher, Programme in Comparative Media Law and Policy, Centre for Socio-Legal Studies, University of Oxford.

\*\*\* Assistant Professor, Netanya Academic College, Faculty of Law. This Research was supported by THE ISRAEL SCIENCE FOUNDATION (grant No. 1820/17).

Special thanks are due to Michael Birnhack, Catalina Goanta, Hanoach Dagan, Daphne Keller, Hans-W, Micklitz, Lilla Montagnani, Neil Netanel, Rory van Loo, Moran Yemini, and the participants in the S. Horowitz Institute for Intellectual Property Work-in-Progress Workshop, and the TAU Law & Information Technology Workshop for their helpful comments.

stakeholders.

*The purpose of this Article is to highlight this blind spot in current contractual analysis and offer courts an interpretive framework for addressing contractual claims involving digital platforms. We argue that platforms’ contracts should be interpreted as contractual networks. This analytical framework is based on a growing body of literature which focuses on interrelated contractual obligations among independent agents who share a common goal. Users in social media platforms, we argue, collaborate in creating the shared economic and social value generated by social media. By framing the contractual relationship between platforms and users as a contractual network, courts are called to consider this complexity and the extent to which removals or terminations meet the contractual expectations of the networks’ members and advance their common goal.*

*This approach to contract interpretation may facilitate a bottom-up check on content moderation via private ordering, thus increasing platforms’ accountability. Specifically, if users could effectively raise contractual claims against platforms and hold them accountable for capricious, biased, or unfair removal decisions, they could pressure platforms to align content moderation policies with the shared interests of the community of users. To that end, contract law could empower users by offering a decentralized and diversified check over the platforms’ content moderation practices. Holding platforms accountable for content moderation practices via private ordering could also facilitate more diversity and exploration, enabling the emergence of different models for moderating digital content and promoting a more pluralist public discourse.*

- I. INTRODUCTION.....989
- II. UNFOLDING THE PRIVATE DIMENSIONS OF CONTENT MODERATION..... 995
  - A. CONTENT MODERATION BY SOCIAL MEDIA PLATFORMS ..... 995
  - B. PRIVATE IMPLICATIONS OF CONTENT MODERATION..... 997
  - C. THE FORCES THAT SHAPE CONTENT MODERATION..... 1000
    - 1. Bottom-up Pressures ..... 1000
    - 2. Top-down Pressures ..... 1003
  - D. A BOTTOM-UP CHECK ON CONTENT MODERATION BY PLATFORMS..... 1007
- III. THE LEGAL BARRIERS TO CONTESTING CONTENT REMOVALS BY PLATFORMS..... 1010
  - A. CONSTITUTIONAL LAW BARRIERS..... 1011
  - B. TORT IMMUNITY AND EXEMPTIONS FROM LIABILITY ..... 1014
  - C. TERMS OF SERVICE AND BOILERPLATE CONTRACTS ..... 1019
  - D. THE BLIND SPOT OF CONTRACT LAW ..... 1024

IV. CONTRACTUAL NETWORK THEORY.....	1025
A. <i>BEYOND BINARY CONTRACT CLASSIFICATION</i> .....	1026
B. <i>THE LEGAL IMPLICATIONS OF CONTRACTUAL NETWORK THEORY</i> .....	1030
V. RECONFIGURING PLATFORM CONTRACTS.....	1032
A. <i>PLATFORM CONTRACTS: BEYOND BILATERAL CONTRACTS</i> .....	1032
B. <i>SOCIAL MEDIA CONTRACTS AS NETWORK CONTRACTS</i> .....	1035
C. <i>NETWORK ANALYSIS OF PLATFORM CONTRACTS</i> .....	1037
D. <i>ADDITIONAL LEGAL IMPLICATIONS</i> .....	1041
1. Limiting the Discretionary Power of Platforms.....	1041
2. Predictability of Content Moderation Rules .....	1042
3. Due Process and the Right to an Explanation .....	1044
4. Fairness and Equal Treatment.....	1045
5. Damages .....	1047
VI. CONCLUSIONS.....	1049

## I. INTRODUCTION

“I feel like I’m being silenced,” wrote Nyome Nicholas-Williams, a plus-size black model, when half-naked photographs of her were removed by Instagram.<sup>1</sup> Her followers rallied to share the censored photos under the hashtag “#IwanttoseeNyome.”<sup>2</sup> The controversy stimulated a community of fans for Nyome and led to a change in Instagram’s official policy on depictions of female nudity. In a different case, the YouTuber DJ Short-E (real name Erik Mishiyev) had over 110 million views and a quarter of a million subscribers before his channels were terminated by YouTube based on alleged copyright infringements, which Mishiyev denied. Mishiyev sued Google (YouTube’s owner) for loss of earnings in 2019; as of this writing, his accounts remain closed.<sup>3</sup> Similar claims of unfair treatment causing substantial economic loss were recently raised in a class action suit brought by Grammy award-winning artist Maria Schneider, which seeks to force YouTube to provide independent creators with access to Content ID, its copyright enforcement tool. Currently this tool is available only to the major music labels, meaning that—according

---

1. Nosheen Iqbal, *Instagram ‘Censorship’ of Black Model’s Photo Reignites Claims of Race Bias*, GUARDIAN (Aug. 9, 2020, 2:13 PM), <https://www.theguardian.com/technology/2020/aug/09/instagrams-censorship-of-black-models-photo-shoot-reignites-claims-of-race-bias-nyome-nicholas-williams> [<https://perma.cc/ZBJ5-8TTD>].

2. The images were posted on Instagram by both Nicholas-Williams and the photographer, Alexandra Cameron. *Id.*

3. *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1158 (N.D. Cal. 2020), *aff’d*, 857 F. App’x 907 (9th Cir. 2021). Mishiyev’s suit against YouTube was dismissed in April 2020. As of this writing, he is appealing.

to the lawsuit—smaller and independent creators “are deliberately left out in the cold.”<sup>4</sup>

Suspension of accounts, and the removal of allegedly unwarranted content by platforms, are becoming routine matters.<sup>5</sup> Driven by profits, platforms seek to ensure that their digital services are aligned both with users’ expectations and with the interests of advertisers. Steps taken in response to mounting pressure to tackle disinformation during the Covid-19 health crisis<sup>6</sup> and the U.S. 2020 presidential election are very recent examples.<sup>7</sup> For instance, Facebook removed a video posted by the Trump campaign in which he claimed children are “virtually immune” to Covid-19, on the grounds that the video violated the platform’s policy banning false claims about the coronavirus.<sup>8</sup> More recently, Facebook and Twitter terminated former President Trump’s account following postings praising the Capitol riot.<sup>9</sup>

While suspension and removal decisions by platforms could trigger fundamental constitutional challenges, such as the fierce debate over the suspension of the Twitter and Facebook accounts of former U.S. President Donald Trump,<sup>10</sup> many such decisions may also raise important challenges to

4. Class Action Complaint & Demand for Jury Trial at 5, *Schneider v. YouTube, LLC*, No. 5:20-cv-4423 (N.D. Cal. July 2, 2020), ECF No. 1.

5. Guy Rosen, *Community Standards Enforcement Report, May 2020 Edition*, META (May 12, 2020), <https://about.fb.com/news/2020/05/community-standards-enforcement-report-may-2020> [<https://perma.cc/5ZFZ-RMM6>]; Andrew Hutchinson, *Facebook Publishes Latest Update on Content Removals, Fake Accounts, Government Requests and More*, SOCIALMEDIATODAY (May 12, 2020), <https://www.socialmediatoday.com/news/facebook-publishes-latest-update-on-content-removals-fake-accounts-govern/577830> [<https://perma.cc/JGD5-88XD>]; Michelle Toh, *Facebook, Google and Twitter Crack Down on Fake Coronavirus ‘Cures’ and Other Misinformation*, CNN BUS. (Feb. 3, 2020, 9:11 AM), <https://edition.cnn.com/2020/01/31/tech/facebook-twitter-google-coronavirus-misinformation/index.html> [<https://perma.cc/22FQ-UQXT>].

6. Joan Donovan, *Here’s how Social Media can Combat the Coronavirus ‘Infodemic’*, MIT TECH. REV. (Mar. 17, 2020), <https://www.technologyreview.com/2020/03/17/905279/facebook-writer-social-media-infodemic-misinformation> [<https://perma.cc/VR3Y-5KQA>]; Nikolaj Nielsen, *Tech Giants Must Stop Covid-19 ‘Infodemic’, Say Doctors*, EUOBSERVER (May 7, 2020, 12:00 PM), <https://euobserver.com/coronavirus/148281> [<https://perma.cc/5EQX-KMRW>].

7. Donie O’Sullivan, *Facebook Considers Banning Political Ads in Days Before US Election*, CNN BUS. (July 10, 2020, 9:14 PM), <https://edition.cnn.com/2020/07/10/tech/facebook-political-ads-ban/index.html> [<https://perma.cc/22LY-H3KK>]; Rebecca Bellan, *Americans Don’t Trust Tech Platforms to Prevent Misuse in the 2020 Elections*, FORBES (Feb. 26, 2020, 3:28 PM), <https://www.forbes.com/sites/rebeccabellan/2020/02/26/americans-dont-trust-tech-platforms-to-prevent-misuse-in-the-2020-elections/#67145c655d49> [<https://perma.cc/6XMT-KMWU>].

8. Todd Spangler, *Facebook, Twitter Pull Down Trump Videos Claiming Kids Are ‘Immune’ From COVID-19*, VARIETY (Aug. 5, 2020, 4:08 PM), <https://variety.com/2020/digital/news/facebook-deletes-trump-post-claiming-kids-are-immune-from-covid-19-1234726916> [<https://perma.cc/G7J4-BKTK>].

9. Jessica Guynn, *President Trump Permanently Banned from Twitter Over Risk He Could Incite Violence*, USA TODAY (Jan. 8, 2021, 6:34 PM), <https://www.usatoday.com/story/tech/2021/01/08/twitter-permanently-bans-president-trump/6603578002> [<https://perma.cc/ZL22-6EA4>].

10. Evelyn Douek, *Facebook Has Referred Trump’s Suspension to Its Oversight Board. Now What?*, LAWFARE (Jan. 21, 2021, 12:16 PM), <https://www.lawfareblog.com/facebook-has-referred-trumps-suspension-its-oversight-board-now-what> [<https://perma.cc/5F5Q-68WU>]. The termination of elected

private law.

From a private law perspective, content moderation and account suspension disputes are governed by the contract between the stakeholders in social media, which is typically composed of the Terms of Services (“ToS”) and community guidelines.<sup>11</sup> Users of digital platforms may consist of small businesses, professional or amateur creators, political activists, and individuals with vested interests in online communications and exchanges. Cutting their livelihood, removal and suspension decisions may carry irreparable harms. An indefinite suspension of a user account disconnects users and their followers, thus depriving content-creating users of financial rewards or reputational gains for which they have labored.<sup>12</sup>

Users of social media platforms are important stakeholders in the platform economy. The economic value in social media is generated through the intermingling of interdependent users. Platforms operating in multisided markets<sup>13</sup> harvest data on users and extract revenues from selling users’ profiles for targeted advertising, or other data-driven products and services.<sup>14</sup> Within this economic ecosystem, users play multiple roles. They are both consumers of services supplied by the platforms (under a vertical contract), and also providers of content, supplying added value that shapes the (horizontal) expectations of other users. Users’ content and interactions attract additional users and deepen their engagement, thus broadening the network and lengthening the time spent on social media. Consequently, while social media platforms generate their profits from advertising, it is users who provide the bricks from which the platforms build their business model. Notwithstanding the interdependency between platforms and users, platforms’ business interests may shift over time, in ways that may not align

officials raises important challenges for democracy which might not be adequately captured by contract law. Likewise, the contractual approach advanced by this Article, which focuses on the mutual contractual expectations of network members *in* social media, might not be suitable for addressing disputes between a social media platform and a web hosting services or an App Store. See *Parler LLC v. Amazon Web Servs., Inc.*, 514 F. Supp. 3d 1261, 1264–66 (W.D. Wash. 2021). These instances will therefore remain outside the scope of this paper.

11. Rory Van Loo, *Federal Rules of Platform Procedure*, 88 U. CHI. L. REV. 829, 842 (2021); Michael Karanicolas, *Too Long; Didn’t Read: Finding Meaning in Platforms’ Terms of Service Agreements*, 52 U. TOL. L. REV. 1, 15–20 (2021). ToS and community guidelines are first and foremost contracts, even if arguably they also may also resemble bills of rights. EDOARDO CELESTE, *TERMS OF SERVICE AND BILLS OF RIGHTS: NEW MECHANISMS OF CONSTITUTIONALISATION IN THE SOCIAL MEDIA ENVIRONMENT?* 2 (2018), <http://doras.dcu.ie/24696/1/E.%20Celeste%20-%20Terms%20of%20Service%20and%20Bills%20of%20Rights%20-%20AM.pdf> [<https://perma.cc/3YJB-T79>].

12. See *infra* Section II.B.

13. DAVID S. EVANS & RICHARD SCHMALENSEE, *MATCHMAKERS: THE NEW ECONOMICS OF MULTISIDED PLATFORMS* 98 (2016).

14. Jean-Charles Rochet & Jean Tirole, *Two-Sided Markets: A Progress Report*, 37 RAND J. ECON. 645, 650 (2006); Max Freedman, *How Businesses Are Collecting Data (And What They’re Doing With It)*, BUS. NEWS DAILY (Dec. 3, 2021), <https://www.businessnewsdaily.com/10625-businesses-collecting-data.html> [<https://perma.cc/8S6T-8UKZ>].

with users' interests and expectations, nor with the common goals agreed upon in the contract. This is especially the case when a handful of social media platforms dominate the online conversation, undermining the mitigating power of competitive pressures.<sup>15</sup>

In removing content or suspending accounts, platforms exercise discretionary powers conferred under boilerplate contracts. Is there any limit to platforms' discretionary power to terminate accounts or remove content?

Currently, under U.S. law, users cannot do much—legally—to protect their rights and interests on social media. From the perspective of constitutional law, platforms are treated as private actors, and so requiring them to reinstate users' content is viewed as a violation of their own First Amendment rights.<sup>16</sup> From a private law perspective, tort and contract law fail to address platforms' powers in removing content. Platforms currently enjoy broad immunity against liability for harms caused by the content they host under the Communications Decency Act ("CDA").<sup>17</sup> They are also subject to limited liability under the system of notice and takedown established by the Digital Millennium Copyright Act ("DMCA").<sup>18</sup> Consequently, so long as they presumably moderate in "good faith," and comply with takedown procedures in the case of alleged copyright infringements, they are immune from liability.<sup>19</sup>

Similarly, users' claims that content removal or account suspensions are a breach of contract were often dismissed.<sup>20</sup> As we further demonstrate below, courts conceive the contractual relationship between users and platforms as dyadic, namely involving two contracting parties.<sup>21</sup> To determine whether one party had breached its contractual obligations, courts focus on the ToS, which most often grant platforms with unlimited discretion to remove users' content or terminate their accounts. Thus, contract law is successfully invoked by platforms as a shield against lawsuits brought by users for any harm they might suffer as a result of content moderation. However, ToS alone do not reflect

---

15. See generally Elettra Bietti, *Consent as a Free Pass: Platform Power and the Limits of the Informational Turn*, 40 PACE L. REV. 310 (2019) (analyzing how notice and consent aspects of media platform's ToS provide inadequate protection to the average user).

16. See Daphne Keller, HOOVER INST., WHO DO YOU SUE? STATE AND PLATFORM HYBRID POWER OVER ONLINE SPEECH 2 (2019), [https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech\\_o.pdf](https://www.hoover.org/sites/default/files/research/docs/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech_o.pdf) [<https://perma.cc/B7SD-W6C2>].

17. Communications Decency Act of 1996 § 230, 47 U.S.C. § 230 (2018).

18. Digital Millennium Copyright Act of 1998 § 512, 17 U.S.C. § 101 (2018).

19. 47 U.S.C. § 230(c)(2)(A) (2018).

20. See, e.g., Order Issuing Alternative Writ of Mandate at 198, *Twitter, Inc., v. Superior Court ex rel. Taylor* (Cal. Ct. App. Aug. 17, 2018) (No. A154973); *Lewis v. YouTube, LLC.*, 197 Cal. Rptr. 3d 219, 220 (Ct. App. 2015); *Young v. Facebook, Inc.*, 790 F. Supp. 2d 1110, 1118–19 (N.D. Cal. 2011); *Hall v. Earthlink Network, Inc.*, No. 98 Civ. 5489(RO), 2003 WL 22990064, at \*3 (S.D.N.Y. Dec. 19, 2003).

21. See *infra* Section IV.A.

the real contractual expectations of the various stakeholders involved in social media. These mutual expectations define additional rights and obligations, beyond those affording platforms with unlimited removal power under the ToS. Interpreting the contractual relationship between platforms and users as establishing bilateral/vertical obligations only hence undermines the true intention of the contracting parties and overlooks the plethora of commitments and obligations created by such contracts to multiple stakeholders.

This Article proposes to remedy this doctrinal blind spot by introducing a contractual network approach for analyzing stakeholders' relationships on social media. This approach is based on a growing body of literature which focuses on interrelated contractual obligations among independent agents who share a common goal.<sup>22</sup> A contractual network neither reflects a formal corporate structure nor merely comprises a series of independent transactions. Rather, it is a complex system of interrelated contracts, enabling coordination without vertical integration.<sup>23</sup>

We argue that user-platform contracts in social media are contractual networks. This is based on the coordination and interdependence among different communities of users of social media, the collaborative nature of social media interactions, and the shared goals.<sup>24</sup> Indeed, social media platforms can be described as a hub-and-spoke networks, where the platform—the hub—provides the technological, business, and legal infrastructure for exchange among users, while users—the spokes—offer complementary content and services. The resulting web of interdependent economic actors constitutes a contractual network.

By framing the contractual relationship between platforms and their users as a contractual network, this Article sets the ground for users to legally contest content removal decisions that fail to meet their legitimate expectations. Within this framework, courts are called to go beyond bilateral contracts and look at the complexity characterizing the relationship between users and platforms. Precisely, courts should consider whether exercising the power to remove or suspend accounts meets the contractual expectations of the networks' members as communities and advances the particular network's common goals.

Applying a contractual network approach to social media contracts could help courts to overcome the existing blind spot in contract law, which allows

---

22. See generally Alan Schwartz & Robert E. Scott, *Third-Party Beneficiaries and Contractual Networks*, 7 J. LEGAL ANALYSIS 325 (2015) (arguing that courts third-party claims against contract members are valid if third parties incur substantial-enough reliance losses and courts can put a price on these losses); GUNTHER TEUBNER, *NETWORKS AS CONNECTED CONTRACTS* (Gunther Teubner & Michelle Everson eds., Michelle Everson trans., 2011) (explaining the concept of "connected contracts" and exploring whether a contractual network, as a whole, can be held liable for damages to third parties).

23. See *infra* note 230 and accompanying text.

24. See *infra* notes 278–82 and accompanying text.

platforms to shelter behind contractual formalities. This could have three interrelated benefits. First, it would enable users to restrain platforms' discretion and safeguard their private interests. Second, it could enhance the common goals and principles of the network. Social media contracts define the values and principles of their communities, which could also carry important implications to human rights.<sup>25</sup> Third, this approach to contract interpretation may facilitate a bottom-up check on content moderation via private ordering, thus increasing platforms' accountability. Specifically, if users could effectively raise contractual claims against platforms and hold them accountable for capricious, biased, or unfair removal decisions, they could pressure platforms to align content moderation policies with the shared interests of the community of users.<sup>26</sup>

At the same time, by underscoring the true interrelated nature of multiple user-platform agreements, the contractual network approach to platform contracts blurs the distinction between public and private. It offers an intermediate position between public and private law where the different stakeholders in social media are bound by contractual principles and values established by shared community guidelines and ToS. Contract law could thus offer a decentralized and diversified check over platforms' content moderation practices and further incentivize platforms to comply with the public interest.

The Article proceeds as follows. Part II describes existing content moderation practices, and the implications of platforms' decisions to remove content or suspend users' accounts. It further describes the different pressures which shape the speech norms applied by platforms, and maps top-down and bottom-up legal strategies for enhancing platforms' accountability when performing content moderation. Part III analyzes the legal barriers facing users who seek to claim damages based on removal decisions by platforms. It explains how the statutory immunities of tort law and the liberal principles of contract law leave users with practically no relief. We further demonstrate how the courts' narrow interpretation of platforms' contracts overlooks the horizontal relationships between users of social media, and, therefore, fails to provide adequate relief to users. In Part IV, we outline contractual network theory and its legal implications. Part V applies the contractual network approach to social media and analyzes the ramifications for users' rights with respect to content moderation. The implications and limitations of our proposal are outlined in the conclusion.

---

25. See JAMILA VENTURINI, LUIZA LOUZADA, MARILIA MACIEL, NICOLO ZINGALES, KONSTANTINOS STYLIANOU & LUCA BELLI, *TERMS OF SERVICE AND HUMAN RIGHTS: AN ANALYSIS OF ONLINE PLATFORM CONTRACTS* 22–26 (Flávio Jardim & Cibeli Hirsch trans., 2d ed. 2016).

26. See *infra* notes 118–23 and accompanying text.



## II. UNFOLDING THE PRIVATE DIMENSIONS OF CONTENT MODERATION

Social media platforms facilitate a global exchange of content generated by users, at a gigantic scale. Unlike the mass media of the 20th century, which was based on content that was centrally produced, professionally edited, and distributed by the publishing and entertainment industries, social media relies on content generated by users of all sorts: amateurs and professionals, individuals and corporations, non-profit organizations, and even governments. However, the stream of content and the contractual terms of its online exchange are heavily governed by the platforms, which apply a wide range of content moderation practices.<sup>27</sup>

The literature concerning content moderation by platforms has largely focused on its implications for public values, such as free speech, equality, trust, and accountability.<sup>28</sup>

However, content moderation also affects the private interests of platform users. In this part of the paper, we unfold these interests to set the groundwork for users' legal claims against platforms.

### A. CONTENT MODERATION BY SOCIAL MEDIA PLATFORMS

Content moderation refers to practices aimed at classifying content posted by users and determining the conditions under which it may be published online. These practices include "the screening, evaluation, categorization, approval or removal/hiding of online content according to relevant communications and publishing policies . . . to support and enforce positive communications behaviour online, and to minimize aggression and

27. Kate Klonick, *The New Governors: The People, Rules, and Process Governing Online Speech*, 131 HARV. L. REV. 1598, 1630–47 (2018).

28. See generally Giovanni De Gregorio, *Democratising Online Content Moderation: A Constitutional Framework*, 36 COMPUT. L. SEC. REV. 1 (2020), <https://www.sciencedirect.com/science/article/pii/S0267364919303851> [<https://perma.cc/W55T-DTQ9>] (arguing that traditional Free Speech principles are inadequate to protect users against the lack of transparency and accountability associated with social media platform content moderation); Hannah Bloch-Wehba, *Global Platform Governance: Private Power in the Shadow of the State*, 72 SMU L. REV. 27 (2019) (arguing that because social media platforms engage in global monitoring, they essentially act as rulemaking and adjudicatory bodies that need to be transparent with their users); Natali Helberger, Jo Pierson & Thomas Poell, *Governing Online Platforms: From Contested to Cooperative Responsibility*, 34 INFO. SOC'Y 1 (2018) (arguing that public platforms, public institutions, and platform users should all work together to shape the public values embodied by these institutions); Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353 (2018) (describing Facebook's content moderation practices as loosely related to First Amendment rights, rushed, ad hoc, and incoherent); UNITED NATIONS INTERNET GOVERNANCE F., PLATFORM REGULATIONS: HOW PLATFORMS ARE REGULATED AND HOW THEY REGULATE US (Luca Belli & Nicolo Zingales eds., 2017), [https://juliareda.eu/wp-content/uploads/2019/09/Reda2017\\_Platform-regulations-how-platforms-are-regulated-and-how-they-regulate-us3.pdf](https://juliareda.eu/wp-content/uploads/2019/09/Reda2017_Platform-regulations-how-platforms-are-regulated-and-how-they-regulate-us3.pdf) [<https://perma.cc/ZSD3-GMPM>] (suggesting that the ability of social media platform's the moderate media limits the effectiveness of users' fundamental rights directly and indirectly).

anti-social behaviour.”<sup>29</sup> As observed by Grimmelmann, content moderation practices are “the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse.”<sup>30</sup> User-generated content is the main draw for other users on social media platforms. It stimulates engagement and exchange, and thus further catalyzes traffic on the platform. At the same time, however, users’ content may also create a burden for platforms when it involves illegal or otherwise harmful speech.<sup>31</sup> Content moderation policies implemented by platforms are hence at the core of how social media operates.<sup>32</sup>

Content moderation may take different shapes and forms. Some content moderation practices seek to optimize the matching of content with those users most likely to view and potentially respond to it—for example, the process by which Facebook organizes users’ news feeds, or YouTube’s recommendation system.<sup>33</sup> Other content moderation practices are intended to ensure that content complies with appropriate norms, either internal (i.e., community guidelines) or external (i.e., regulatory restraints), leading to filtering, blocking, downgrading, or removal of online content.<sup>34</sup> Platforms initially deployed human moderators to identify unwarranted content, but with the amount of content growing exponentially, platforms were forced to supplement, and even replace, human review with automated systems.<sup>35</sup> All major platforms today deploy automated measures to filter unwarranted content before it is published, to block access to such content, or to remove it from the platform.<sup>36</sup>

29. Terry Flew, Fiona Martin & Nicolas Suzor, *Internet Regulation as Media Policy: Rethinking the Question of Digital Communication Platform Governance*, 10 J. DIGIT. MEDIA & POL’Y 33, 40 (2019).

30. James Grimmelmann, *The Virtues of Moderation*, 17 YALE J.L. & TECH. 42, 47 (2015) (emphasis omitted).

31. Maayan Perel, *Enjoining Non-Liable Platforms*, 34 HARV. J.L. & TECH. 1, 20 (2020).

32. TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA 6 (2018).

33. See, e.g., Karen Hao, *YouTube is Experimenting with Ways to Make its Algorithm Even More Addictive*, MIT TECH. REV. (Sept. 27, 2019), <https://www.technologyreview.com/2019/09/27/132829/youtube-algorithm-gets-more-addictive> [<https://perma.cc/NN6M-5ZTJ>].

34. Niva Elkin-Koren & Maayan Perel, *Separation of Functions for AI: Restraining Speech Regulation by Online Platforms*, 24 LEWIS & CLARK L. REV. 857, 877–81 (2020).

35. See Maayan Perel & Niva Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, 19 STAN. TECH. L. REV. 473, 478–79 (2016).

36. See, e.g., Robert Gorwa, Reuben Binns & Christian Katzenbach, *Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance*, BIG DATA & SOC’Y, Jan.–June 2020, at 3, <https://journals.sagepub.com/doi/pdf/10.1177/2053951719897945> [<https://perma.cc/K3XU-AN8C>]; Katie Canales, *Mark Zuckerberg Said Content Moderation Requires ‘Nuances’ that Consider the Intent Behind a Post, but also Highlighted Facebook’s Reliance on AI to do That Job*, INSIDER (Mar. 25, 2021, 4:59 PM), <https://www.businessinsider.com/zuckerberg-nuances-content-moderation-ai-misinformation-hearing-2021-3> [<https://perma.cc/32U3-NELH>]; Sara Harrison, *Twitter and Instagram Unveil New Ways to Combat Hate—Again*, WIRED (July 11, 2019, 7:00 AM), <https://www.wired.com/story/twitter-instagram-unveil-new-ways-combat-hate-again> [<https://perma.cc/4JB2-J69N>].

Platforms may also suspend or terminate the accounts of users who they believe violate the platform's content guidelines.<sup>37</sup> Finally, content moderation may take the form of monetary sanctions.<sup>38</sup> Platforms such as YouTube incentivize the sharing of content on the platform by offering some users a share in the advertising revenues associated with their content.<sup>39</sup> Demonetizing content, namely, excluding particular content from the revenue-sharing arrangement, keeps the content available online, but cuts off ad revenues that would otherwise flow to the user.<sup>40</sup>

Taken together, these content moderation practices affect online discourse by enabling or silencing some speech or speakers, or by restricting the spread of certain expressions. These practices shape the public sphere by influencing what expressions members of the public can encounter and debate, thus carrying crucial implications for democratic deliberation and free speech.<sup>41</sup> At the same time, however, these practices are also situated at the core of the platform economy, with implications for the commercial interests of many different stakeholders, as discussed below.

#### B. PRIVATE IMPLICATIONS OF CONTENT MODERATION

Social media platforms play a dual role. They offer users a business infrastructure for networking, organizing, creating, marketing, and selling products and services. At the same time, platforms also constitute a public sphere and a social forum, where users can express their opinions, shape their identity, build their social relationships, and organize for collective action. In this capacity, content moderation policies may carry potential consequences not only for users of the platforms, but also for the offline communities to which they belong.

The removal of content or account suspensions may carry harmful consequences for amateur or professional creators of content who seek to disseminate their content for financial and reputational gains, as well as for potential consumers and audience. Indeed, users who run business activities on social media can suffer high losses from removal of their content or suspension of their accounts. A classic example is the so-called influencers,<sup>42</sup>

---

37. Eric Goldman, *Online User Account Termination and 47 U.S.C. §230(c)(2)*, 2 U.C. IRVINE L. REV. 659, 670–72 (2012).

38. See Henning Grosse Ruse-Khan, *Automated Copyright Enforcement Online: From Blocking to Monetization of User-Generated Content* 8 (Program on Info. Just. & Intell. Prop., Working Paper No. 51, 2020), <https://digitalcommons.wcl.american.edu/cgi/viewcontent.cgi?article=1053&context=research> [<https://perma.cc/HW2M-EMU3>] (arguing that on YouTube demonetizing user content has prevailed over other methods of content moderation, such as blocking content).

39. *E.g.*, through YouTube's Partner Program ("YPP").

40. Robyn Caplan & Tarleton Gillespie, *Tiered Governance and Demonetization: The Shifting Terms of Labor and Compensation in the Platform Economy*, SOC. MEDIA + SOC'Y, Apr.–June 2020, at 1–3.

41. See generally De Gregorio, *supra* note 28 (discussing the impact of content moderation of online content and democratic expression).

42. See generally THE REGULATION OF SOCIAL MEDIA INFLUENCERS (Catalina Goanta & Sofia

or cultural entrepreneurs as termed by Cunningham and Craig, who join social media as amateurs and harness network affordances to build sustainable communities of followers.<sup>43</sup> In some cases, influencers and other content creators may rely on social media monetization as their sole source of income, meaning that removal or suspension decisions by platforms could inflict significant financial harm. The case of Erik Mishiyev (“DJ Short-E”) demonstrates this point. Mishiyev began running channels on YouTube hosting music and celebrity interviews in 2007.<sup>44</sup> In the following years, he amassed over 110 million views and over 250,000 subscribers.<sup>45</sup> Between 2012 and 2018, Mishiyev earned over \$300,000 through YouTube’s Partner Program, which allows users to earn income from their online content by matching ads to their sites’ target viewers.<sup>46</sup> This arrangement was Mishiyev’s primary source of income.<sup>47</sup> Mishiyev won a dispute with YouTube over copyright claims in 2016, but then found that YouTube was allegedly failing to notify his subscribers of new content, causing his revenue to drop.<sup>48</sup> In 2019, YouTube terminated Mishiyev’s account and removed all his videos on the basis of new copyright infringement claims.<sup>49</sup> Mishiyev disputed the copyright strikes, claiming that YouTube terminated his accounts in retaliation after he informed the platform that he would be filing a lawsuit for loss of earnings following YouTube’s alleged failure to promote his earlier content.<sup>50</sup>

The harm to users whose content is removed by social media platforms is not limited to the direct financial effects of losing advertising revenues. For creators, the suspension of accounts may also entail long-term reputational damage, impairing the ability to reach their audience and sustain the community of followers which the creator has built on the platform.<sup>51</sup> For instance, in 2006 a user named Jan Lewis created the YouTube channel “bulbheadmyass,” where she posted videos of her band, the Remington Riders.<sup>52</sup> In five years, she garnered about 500,000 views and numerous positive comments, all of which were deleted after YouTube suspended her

---

Ranchordás eds., 2020) (discussing social influencers and their impact on social media as well as how to regulate them).

43. STUART CUNNINGHAM & DAVID CRAIG, *SOCIAL MEDIA ENTERTAINMENT: THE NEW INTERSECTION OF HOLLYWOOD AND SILICON VALLEY* 1, 12 (2019).

44. *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1156 (N.D. Cal. 2020), *aff’d*, 857 F. App’x 907 (9th Cir. 2021).

45. *Id.*

46. *Id.*

47. *Id.*

48. *Id.* at 1157.

49. *Id.* at 1156.

50. Amended Complaint & Demand for Jury Trial ¶¶ 31–34, *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154 (N.D. Cal. 2020) (No. 19-cv-05422).

51. *Lewis v. YouTube, LLC.*, 197 Cal. Rptr. 3d 219, 222 (Ct. App. 2015).

52. *Id.* at 221.

account and deleted her channel in 2012.<sup>53</sup> The videos were uploaded without any expectations of commercial gain. As described by the court in the proceedings following her lawsuit: “she did not sell the videos or audio versions of the music; and the Remington Riders did not perform in public. Her sole reward” for “hundreds of hours and thousands of dollars” spent to produce her videos “was ‘the acclaim that she received from the YouTube community and the opportunity to make new friends.’”<sup>54</sup> Ultimately YouTube reinstated Lewis’s account, but without restoring the videos, comments, view counts, and followers that had accumulated on the old channel—a decision which Lewis argued led to reputational loss and diminished popularity.<sup>55</sup> Lewis’s case was based on the argument that users’ investment in building communities of followers over time has important economic and social value for all such users, while also benefiting the platforms by enabling them to generate ad revenue.<sup>56</sup> Nonetheless, a California appeals court found for YouTube.<sup>57</sup>

Another type of private harm generated by content removal and private ‘censorship’ relates to the silencing of certain opinions. Consider, for instance, the removal of content posted by PragerU, a non-profit organization headed by Dennis Prager. This organization is known for sharing conservative ideas in five-minute videos with titles like “Are 1 in 5 Women Raped at College?”<sup>58</sup> and “Why Isn’t Communism as Hated as Nazism?”<sup>59</sup> Many PragerU videos, some of them viewed two million times, have been restricted or deleted by YouTube.<sup>60</sup> Claims of silencing were also raised against YouTube by LGBTQ creators, arguing that YouTube has effectively silenced their content by demonetizing it, since it has classified this content as sexual, and therefore unsuited for advertising by major brands.<sup>61</sup>

53. *Id.*

54. *Id.*

55. *Id.* at 222.

56. *Id.*

57. *Id.* at 220.

58. *Are 1 in 5 Women Raped at College?*, PRAGERU (Apr. 11, 2016), <https://www.prageru.com/video/are-1-in-5-women-raped-at-college> [<https://perma.cc/XTK4-4AND>]; PragerU, *Are 1 in 5 Women Raped at College?*, YOUTUBE (Apr. 11, 2016), <https://www.youtube.com/watch?v=KomzQL5oI-w> [<https://perma.cc/LGD4-93U6>].

59. *Why Isn’t Communism as Hated as Nazism?*, PRAGERU (May 1, 2017), <https://www.prageru.com/video/why-isnt-communism-as-hated-as-nazism> [<https://perma.cc/V4ZA-ECSP>]; PragerU, *Why Isn’t Communism as Hated as Nazism?*, YOUTUBE (May 1, 2017), <https://www.youtube.com/watch?v=nUGkKKAogDs> [<https://perma.cc/8XRE-Z8C5>].

60. Sara Harrison, *No One’s Happy with YouTube’s Content Moderation Policies*, WIRED: BUS. (Aug. 28, 2019, 7:00 AM), <https://www.wired.com/story/no-ones-happy-youtubes-content-moderation> [<https://perma.cc/2A2L-3DT8>].

61. Sal Bardo, *YouTube Continues to Restrict LGBTQ Content*, HUFFPOST: QUEER VOICES (Jan. 17, 2018), [https://www.huffingtonpost.com/entry/youtube-continues-to-restrictlgbtq-content\\_us\\_5a5e6628e4b03ed177016e90](https://www.huffingtonpost.com/entry/youtube-continues-to-restrictlgbtq-content_us_5a5e6628e4b03ed177016e90) [<http://perma.cc/U3BA-E33N>].

C. THE FORCES THAT SHAPE CONTENT MODERATION

In performing content moderation, social media platforms apply a complex set of norms through several instruments of governance, most typically contracts, guidelines, and algorithmic design.<sup>62</sup> As recently demonstrated by Caplan and Gillespie, YouTubers, creators and activists are all subject to the complexity of YouTube's tiered governance strategy, whereby rules regarding the monetization of content set incentives and shape the availability of content.<sup>63</sup>

Content moderation policies are typically defined in contracts drafted by the platforms.<sup>64</sup> ToS often incorporate by reference the content guidelines, which specify in more detail the platforms' norms regarding permissible use of content on the platform (e.g., Facebook's Community Standards or YouTube's Community Guidelines).<sup>65</sup> Users who accept the platform's ToS enter an agreement, where they are required to adhere to these norms when using the platform to share content. These speech norms are driven by consumption, commercial interests, social norms, liability rules, and regulatory duties, where each set of norms may interact with the others. Content moderation norms are shaped in response to two types of pressures: bottom-up market and top-down regulatory pressures, as explained below.

1. Bottom-Up Pressures

Market pressures are often derived from the platforms' core business model—a multi-sided model where the platforms' profit is generated by selling user attention and user data to advertisers.<sup>66</sup> This business model is based on generating data on users and extracting revenues from selling users' profiles and information for targeted advertising, or for producing other data-driven products and services.<sup>67</sup> Multi-sided platforms, such as Facebook or YouTube, offer free services to users while charging a price above cost to advertisers.<sup>68</sup> Indeed, the unique possibilities for profiling users according to their preferences and habits make social media platforms highly attractive for advertising and marketing by different businesses and organizations. As aptly observed by Roberts, users' content can be considered “the currency by which users are engaged as consumers and producers on social media sites.”<sup>69</sup> Thus,

---

62. See *infra* notes 203–07 and accompanying text.

63. Caplan & Gillespie, *supra* note 40, at 2.

64. See Karanicolas, *supra* note 11, at 15–20; Van Loo, *supra* note 11, at 830–32.

65. See, e.g., *Facebook Community Standards*, FACEBOOK, <https://www.facebook.com/communitystandards> [<https://perma.cc/B7QZ-Z6YY>]; *Community Guidelines*, YOUTUBE, <https://www.youtube.com/howyoutubeworks/policies/community-guidelines> [<https://perma.cc/XQV5-WC6T>].

66. EVANS & SCHMALENSEE, *supra* note 13, at 252.

67. Freedman, *supra* note 14.

68. Rochet & Tirole, *supra* note 14, at 645–46.

69. Sarah T. Roberts, *Digital Detritus: 'Error' and the Logic of Opacity in Social Media Content Moderation*, FIRST MONDAY (Mar. 2018), <https://firstmonday.org/ojs/index.php/fm/article/view/8283/6649> [<https://perma.cc/Q72B-3VQF>].

platforms deploy content moderation to maintain a wide community of users, and to encourage user engagement (in terms of both the number of users and time spent on the platform) with content posted by other users, thereby increasing advertising revenues.<sup>70</sup>

In multi-sided markets, where users' attention is sold to advertisers, the user's attention is often considered a product.<sup>71</sup> Yet on social media platforms, users themselves are also generators of content, thereby adding a core value to the platforms' business operation. Large-scale availability of user content on the platform attracts more users and keeps existing users on the platform for longer periods of time, generating more revenues for the platform.

This model, however, engenders a fundamental tension: content made available on the platform may attract some users, thus increasing online traffic, but the same content may alienate other users and decrease their engagement online. Content moderation norms are hence driven by consumption, with the aim of maximizing content that will "keep the userbase actively engaged as content producers,"<sup>72</sup> while not antagonizing a sufficiently large number of other users to reduce engagement. Towards this end, platforms may seek to ensure a safe environment for users, free of any expressions or opinions deemed to be harmful or offensive, where users can feel comfortable sharing and consuming content.<sup>73</sup>

Presumably, in acting to address conflicting preferences and interests of users regarding content, platforms aim to optimize the overall shared preferences of their respective user base, thus reflecting the collective preference of the platform's community. However, the success of this strategy depends on whether users have a voice, and can effectively express it in their contractual relationship with the platform.

The market pressures that help shape content moderation norms applied by platforms are not only internal, but also external. Here, the aim is to protect the platforms' corporate identity by demonstrating their commitment to values which are widely endorsed. In 2018, for example, Facebook agreed to an independent audit into its policies and practices at the behest of civil rights leaders and some members of Congress.<sup>74</sup> In 2020, while that audit was still running, many platforms—including Facebook—hastened to prove their social commitment to combating disinformation in response to the "Stop

---

70. Mathew Ingram, *How Google and Facebook Have Taken Over the Digital Ad Industry*, FORTUNE (Jan. 4, 2017, 10:30 AM), <https://fortune.com/2017/01/04/google-facebook-ad-industry> [<https://perma.cc/2CY3-2HgU>].

71. TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR HEADS 92 (2016).

72. Roberts, *supra* note 69.

73. GILLESPIE, *supra* note 32, at 16.

74. LAURA W. MURPHY & MEGAN CACACE FACEBOOK'S CIVIL RIGHTS AUDIT – FINAL REPORT 3 (2020) [hereinafter FACEBOOK CIVIL RIGHTS AUDIT], <https://about.fb.com/wp-content/uploads/2020/07/Civil-Rights-Audit-Final-Report.pdf> [<https://perma.cc/929F-5BC7>].

Hate for Profit” campaign launched by a coalition of advocacy groups.<sup>75</sup>

Public outrage following a few widely reported incidents where harmful content was disseminated online also put mounting pressure on platforms to act against allegedly harmful content shared through their services. A notable example was Facebook’s failure to remove live footage of a terrorist massacre in New Zealand, which went viral; the ensuing outcry led Facebook to tighten its live-stream policies and introduce a “one-strike” rule that temporarily restricts access for users found to be in violation.<sup>76</sup>

Besides, pressure has also arisen around privacy concerns, as in the case of the Cambridge Analytica fiasco, which became the subject of hearings in the U.S. Congress.<sup>77</sup> That episode led Facebook to introduce “a three-step plan to prevent platform abuse,” namely: (1) auditing applications showing “suspicious activity and banning any developer that has misused personally identifiable information;” (2) “restrict[ing] developer’s [sic] access to data;” and (3) “add[ing] a new tool at the top of user’s [sic] news feeds that enables users to see which applications they have used[,] and [to] revoke a third-party application’s data access.”<sup>78</sup>

Another type of external market pressure that may push platforms to engage in content moderation is organized pressure by advertisers, who constitute the primary source of revenue for platforms such as Facebook and YouTube. YouTube’s “Adpocalypse” in 2017, for instance, was a backlash from major advertisers following a story by an investigative journalist which revealed that advertisements by major brands were posted on YouTube alongside terrorist propaganda and hate speech.<sup>79</sup> YouTube responded by updating its filters and introducing a new policy on content, allowing advertisers to delete certain categories of content from their inventory, which has resulted in the demonetization of such content.<sup>80</sup> However, this new policy has not only affected bad actors, but has also led to the removal of some content with a

---

75. Afdhel Aziz, *Facebook Ad Boycott Campaign ‘Stop Hate For Profit’ Gathers Momentum and Scale: Inside the Movement for Change*, FORBES (June 24, 2020, 10:35 AM), <https://www.forbes.com/sites/afdhelaziz/2020/06/24/facebook-ad-boycott-campaign-stop-hate-for-profit-gathers-momentum-and-scale-inside-the-movement-for-change/?sh=26556do16687> [https://perma.cc/R4JS-55DM].

76. Hamza Shaban, *Facebook to Reexamine How Livestream Videos are Flagged after Christchurch Shooting*, WASH. POST (Mar. 21, 2019), <https://www.washingtonpost.com/technology/2019/03/21/facebook-reexamine-how-recently-live-videos-are-flagged-after-christchurch-shooting> [https://perma.cc/6GCP-SQNT]; Amy Gunia, *Facebook Tightens Live-Stream Rules in Response to the Christchurch Massacre*, TIME (May 15, 2019, 4:19 AM), <https://time.com/5589478/facebook-livestream-rules-new-zealand-christchurch-attack> [https://perma.cc/BV57-SA67].

77. See Bietti, *supra* note 15, at 335.

78. *In re Facebook - Cambridge Analytica*, EPIC.ORG (2021), <https://epic.org/privacy/facebook/cambridge-analytica> [https://perma.cc/T84X-C56N].

79. Rachel Dunphy, *Can YouTube Survive the Adpocalypse?*, N.Y. MAG. (Dec. 28, 2017), <http://nymag.com/intelligencer/2017/12/can-youtube-survive-the-adpocalypse.html> [https://perma.cc/Q2RE-H9DD].

80. Caplan & Gillespie, *supra* note 40, at 9.



progressive agenda, which was intended to denounce violence.<sup>81</sup>

The “Stop Hate for Profit” campaign mentioned earlier was launched by a coalition of advocacy groups, but it led quickly to a boycott by advertisers over Facebook’s handling of hate speech and disinformation during the Covid-19 crisis<sup>82</sup> and prior presidential elections in the United States. Despite much publicity, however, this campaign had little impact on Facebook’s content moderation policies.<sup>83</sup> Facebook has made some public statements stressing its commitment to address hate speech and “fake news,” but in practice Facebook chose to put its focus on the audit of its civil rights policies and practices mentioned above.<sup>84</sup>

These recent examples demonstrate how platforms’ commercial interests may be influenced by social norms. In particular, pressure from advertisers may directly affect the revenues of platforms, and therefore could force platforms to modify their content moderation policies to align with the interests and preferences of a wider public (the brands’ consumers). Yet, due to the dominance of platforms like Facebook and Instagram (which is owned by Facebook) over the online conversation, the power of advertisers is in fact limited. During the “Stop Hate for Profit” boycott, many small advertisers (which contribute the majority of Facebook’s total ad revenue) were unable to suspend their advertising on Facebook and Instagram, and many of the large brands that explicitly participated in the boycott announced they would resume spending in August.<sup>85</sup> Moreover, to the extent that such pressures are successful, they give disproportionate influence to a few powerful brands in shaping platforms’ content moderation policies. Furthermore, the value choices and trade-offs implicit in such modifications of policies are opaque and not subject to public deliberation.

## 2. Top-Down Pressures

Nonetheless, market pressures are not alone in shaping content moderation

81. Lucas Shaw, *YouTube Advertising Crackdown Puts Some Creators Out of Work*, BLOOMBERG (Dec. 8, 2017, 12:08 PM), <https://www.bloomberg.com/news/articles/2017-12-08/youtube-advertising-crackdown-puts-some-creators-out-of-work> [<https://perma.cc/HH7N-XEBg>].

82. Nancy Scola, *Inside the Ad Boycott That Has Facebook on the Defensive*, POLITICO (July 3, 2020, 3:15 PM), <https://www.politico.com/news/magazine/2020/07/03/activists-advertising-boycott-facebook-348528> [<https://perma.cc/ZX3K-7RSR>].

83. Joan E Greve & Martin Pengelly, *Twitter Limits Donald Trump Jr’s Account for Posting Covid-19 Misinformation*, GUARDIAN (July 28, 2020, 4:23 PM), <https://www.theguardian.com/us-news/2020/jul/28/donald-trump-jr-twitter-restricted-hydroxychloroquine> [<https://perma.cc/gWMR-DV2L>].

84. See generally FACEBOOK CIVIL RIGHTS AUDIT, *supra* note 74 (investigating Facebook’s policies and practices to improve the way Facebook impacts civil rights).

85. Paurav Shukla, *Big Advertisers Are Boycotting Facebook But It’s Not Enough To #StopHateforProfit – Here’s Why*, CONVERSATION (July 30, 2020, 5:24 AM), <https://theconversation.com/big-advertisers-are-boycotting-facebook-but-its-not-enough-to-stop-hate-for-profit-heres-why-143505> [<https://perma.cc/3Q8B-NGLAJ>].

norms. Content moderation also responds to top-down regulatory pressures, reflecting risk management considerations. Essentially, to minimize their potential liability for carrying illegal content posted by users (e.g., child pornography, materials inciting violence or terrorist acts, or infringements of intellectual property), platforms may opt to remove illegal content to avoid legal exposure.<sup>86</sup> Yet as the line between legal and illegal content is not always clear, platforms may choose to remove, block access to, or filter out borderline or questionable content in order to lower their risk. The upshot is that platforms may end up removing or blocking legitimate content, leading to what has been defined as collateral censorship.<sup>87</sup>

This trend has been strengthened in recent years due to increasing pressure on the platforms to remove illegal content or otherwise to degrade content which is viewed by some stakeholders as illegitimate.<sup>88</sup> The gatekeeping nature of platforms, with their capacity to disable, remove, or block controversial content, makes them ideal partners for law enforcement.<sup>89</sup>

Applying a top-down approach to increase accountability for content moderation comes in different forms. In recent years, there have been different proposals for regulating social media across the Atlantic.<sup>90</sup> Many countries have considered, or even introduced, legal interventions that would strictly regulate content moderation by social media platforms, requiring platforms to undertake active steps to address disinformation, and to remove unlawful speech or hate speech.<sup>91</sup> Even in the United States, there have been several proposals to amend Section 230 of the CDA, which provides platforms with legislative protection from removing offensive content.<sup>92</sup> Such proposals

---

86. Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2314 (2014). The Digital Millennium Copyright Act (DMCA), for instance, encourages platforms to act expeditiously to takedown content claimed to be infringing by offering them a safe harbor which protects them from liability for acts of copyright infringements committed by their users. See *infra* Section III.B.

87. Felix T. Wu, *Collateral Censorship and the Limits of Intermediary Immunity*, 87 NOTRE DAME L. REV. 293, 295–96 (2011); J.M. Balkin, *Free Speech and Hostile Environments*, 99 COLUM. L. REV. 2295, 2296 (1999).

88. Niva Elkin-Koren, Yifat Nahmias & Maayan Perel, *Is It Time to Abolish Safe Harbor? When Rhetoric Clouds Policy Goals*, 31 STAN. L. & POL'Y REV. 1, 13 (2020).

89. *Infra* note 99 and accompanying text.

90. See, e.g., *Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, at 2, COM (2020) 825 final (the recent proposal of the Union to adopt the Digital Services Act).

91. ADRIAN SHAHBAZ, FREEDOM HOUSE, FREEDOM ON THE NET 2018: THE RISE OF DIGITAL AUTHORITARIANISM 13 (2018), [https://freedomhouse.org/sites/default/files/FOTN\\_2018\\_Final.pdf](https://freedomhouse.org/sites/default/files/FOTN_2018_Final.pdf) [<https://perma.cc/B7R7-86Q6>]; KAREN KORNBLUH, ELLEN P. GOODMAN & ELI WEINER, SAFEGUARDING DIGITAL DEMOCRACY: DIGITAL INNOVATION AND DEMOCRACY INITIATIVE ROADMAP 30 (2020), [https://www.gmfus.org/sites/default/files/Safeguarding%20Democracy%20oagaint%20Disinformation\\_v7.pdf](https://www.gmfus.org/sites/default/files/Safeguarding%20Democracy%20oagaint%20Disinformation_v7.pdf) [<https://perma.cc/Q4T6-B4AY>].

92. See Stop the Censorship Act, H.R. 4027, 116th Cong. (2019); Biased Algorithm Deterrence Act of 2019, H.R. 492, 116th Cong. (2019); Stop the Censorship Act of 2020, H.R. 7808, 116th Cong. (2020); Protect Speech Act, H.R. 8517, 116th Cong. (2020); Limiting Section

have been received with deep suspicion by U.S. policymakers, who are concerned about granting the government discretion over content or otherwise authorizing the government to limit the expression facilitated by platforms.<sup>93</sup> Nonetheless, at the state level, other legislation have tried to reduce platform discretion. Florida, for instance, introduced a bill in response to the discretionary deplatforming of Donald Trump.<sup>94</sup> The constitutionality of this bill was challenged in court.<sup>95</sup> This aversion of any governmental intervention in speech reflects a fundamental tenet of U.S. constitutional law, which heavily relies on the public/private divide. The Constitution restrains the use of governmental power, not that of private actors.<sup>96</sup> In fact, freedom of expression as a constitutional right ensures that any governmental attempt to regulate speech would be subject to strict scrutiny under the First Amendment.<sup>97</sup>

There are several reasons to be cautious about top-down regulatory efforts to restrain platforms' power to moderate content. One reason to limit regulatory intervention in content moderation is a deep distrust towards the government, and the concern that it could exercise powers against dissidents and shut down opponents' speech.<sup>98</sup> Another reason for caution is the concern that digital platforms, which rank among the largest companies in the world, could capture the governmental process and tilt regulatory policies to serve their interests. A related concern is that governments may have a vested interest in harnessing the capabilities of platforms to govern online speech, so as to achieve particular goals while circumventing constitutional barriers.<sup>99</sup> Governments that seek to block content under the radar of the judiciary are increasingly issuing removal requests based on platforms'

230 Immunity to Good Samaritans Act, H.R. 8596, 116th Cong. (2020); Protecting Americans from Dangerous Algorithms Act, H.R. 8636, 116th Cong. (2020); Break Up Big Tech Act of 2020, H.R. 8922, 116th Cong. (2020); Curbing Abuse and Saving Expression in Technology Act, H.R. 285, 117th Cong. (2021); Platform Accountability and Consumer Transparency Act, S. 4066, 116th Cong. (2020); 47 U.S.C. § 230 (2018).

93. KORNBLUH, GOODMAN & WEINER, *supra* note 91, at 9.

94. S.B. 7072, 2021 Leg., Reg. Sess. (Fla. 2021).

95. See generally *Netchoice, LLC v. Moody*, No. 4:21CV220, 2021 WL 2690876 (N.D. Fla. June 30, 2021) (challenging the constitutionality of Florida's Senate Bill 7072).

96. U.S. CONST. amend. I ("Congress shall make no law . . . abridging the freedom of speech . . .") (emphasis added).

97. See, e.g., *Sable Commc'ns of Cal., Inc. v. FCC*, 492 U.S. 115, 126 (1989).

98. See Nicolas Suzor, *Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms*, SOC. MEDIA + SOC'Y, July–Sept. 2018, at 4 (arguing that the opposition to interference from state actors reflects a concern with "demands from various governments to collect and disclose information on the activities of individuals, to remove or block access to prohibited information, and to engineer networks and technologies in ways that facilitate surveillance and law enforcement").

99. See Niva Elkin-Koren & Eldar Haber, *Governance by Proxy: Cyber Challenges to Civil Liberties*, 82 BROOK. L. REV. 105, 107 (2016); Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035, 1050–51 (2018); Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA. J.L. & TECH. 1, 48–54 (2003).

community guidelines, which often define objectionable speech more broadly than the illegal speech defined by law.<sup>100</sup>

A recent (May 2020) executive order issued by former President Trump regarding content moderation by digital platforms highlights some of the risks involved in governmental intervention in the right to free speech and, more generally, the constitutional freedoms of the private sector.<sup>101</sup> The order recognized the dominant position of online platforms in controlling online speech and driving public opinion,<sup>102</sup> asserting that platforms “wield immense, if not unprecedented, power to shape the interpretation of public events; to censor, delete, or disappear information; and to control what people see or do not see.”<sup>103</sup> To address these issues, the order requires executive departments and agencies to implement several policies that would advance a narrow interpretation of the immunity accorded to online platforms under Section 230 of the Communications Decency Act,<sup>104</sup> and to explore additional policies that would financially sanction digital platforms which fail to comply with the law.<sup>105</sup>

Courts have already underlined the limits of this executive order.<sup>106</sup> Likewise, executive orders banning the Chinese-owned platforms TikTok and WeChat have been challenged before courts.<sup>107</sup> These executive orders have been revoked by President Biden.<sup>108</sup> Still, they show the constitutional tension

---

100. For instance, the Facebook NetzDG Transparency Report, filed under the German Network Enforcement Act (NetzDG), demonstrates that during the first six months of the law, Facebook removed 1,704 items of content based on 886 NetzDG legal notices, while removing millions of items during the same period based on its Community Guidelines reporting system. FACEBOOK, NETZDG TRANSPARENCY REPORT 2–3 (2018), [https://about.fb.com/wp-content/uploads/2018/07/facebook\\_netzdg\\_july\\_2018\\_english-1.pdf](https://about.fb.com/wp-content/uploads/2018/07/facebook_netzdg_july_2018_english-1.pdf) [<https://perma.cc/HCN6-8gWL>].

101. Exec. Order No. 13,925, 85 Fed. Reg. 34,079 (May 28, 2020) [hereinafter Executive Order on Preventing Online Censorship]. See Evelyn Douek, *Trump Is a Problem That Twitter Cannot Fix*, ATLANTIC (May 27, 2020), <https://www.theatlantic.com/ideas/archive/2020/05/twitter-cant-change-who-the-president-is/612133> [<https://perma.cc/54TM-XZDg>].

102. “Online platforms” are defined by Section 7 as “any website or application that allows users to create and share content or engage in social networking, or any general search engine.” Executive Order on Preventing Online Censorship, *supra* note 101.

103. *Id.* § 1.

104. *Id.* § 2.

105. *Id.* §§ 4–6. The EO encourages the DOJ and FTC to take action against platforms on the basis of false marketing statements (Section 4), instructs the AG to draft federal legislation to advance the EO (Section 6), and encourages state AGs to investigate how state laws can be used against platforms, and to develop model state legislation (Section 5). *Id.*

106. *Rock the Vote v. Trump*, No. 20-cv-06021, 2020 WL 6342927, at \*7 (N.D. Cal. Oct. 29, 2020); *Gomez v. Zuckenburg*, No. 5:20-CV-633, 2020 WL 7684956, at \*2 (N.D.N.Y. July 23, 2020).

107. *Marland v. Trump*, 498 F. Supp. 3d 624, 628–29 (E.D. Pa. 2020); *TikTok Inc. v. Trump*, 490 F. Supp. 3d 73, 79 (D.D.C. 2020); *U.S. WeChat Users All. v. Trump*, 488 F. Supp. 3d 912, 916 (N.D. Cal. 2020).

108. David Shepardson, *Biden Revokes Trump Order that Sought to Limit Social Media Firms’ Protections*, REUTERS (May 17, 2021, 4:02 AM), <https://www.reuters.com/technology/biden-revokes-trump-order-that-sought-limit-social-media-firms-protections-2021-05-15> [<https://perma.cc/SYM9-X6CM>].

coming from top-down regulation of online platforms.

These pressures, however, are not always enough to make platforms more responsible in how they moderate users' content. Indeed, it is clear that some pressures on platforms—e.g., liability concerns arising from top-down regulation versus free-speech and civil rights concerns—may at times contradict each other.

Next, we map the legal interventions employed to hold platforms accountable for content moderation and analyze their limits.

#### D. A BOTTOM-UP CHECK ON CONTENT MODERATION BY PLATFORMS

Platforms possess largely unrestrained discretionary power in content moderation, which may carry serious implications for individual creators, speakers, subscribers, and the public at large. As governors of other people's speech, platforms are arguably expected to advance public welfare.<sup>109</sup> Yet, as commercial actors, which derive revenues from facilitating users' speech, their commercial interests often conflict with their governing roles.<sup>110</sup> Without accountability, the removal of expressions from the public sphere may silence some speakers (e.g., social activists or political opponents), and may deprive the public of access to legitimate speech.<sup>111</sup> Therefore, platforms need to be made more accountable to the people who are affected by their content moderation policies.

The ability of online platforms to define the standards of free speech on a global scale is indeed a critical issue with great public implications. However, this does not mean that executive power is a suitable instrument to address the quasi-public powers vested in the hands of profit-driven platforms. Former President Trump's executive orders illustrate two main reasons why governmental regulation of speech could produce negative effects for the public sphere.

First, top-down regulation could increase the liability risk faced by digital platforms in exercising their moderative discretion.<sup>112</sup> If platforms are held liable for failure to appropriately curb questionable online content, they are likely to take more aggressive steps to remove content. As the risk of liability

---

109. See Perel & Elkin-Koren, *supra* note 35, at 531 (discussing “algorithmic governance and how it intersects with conventional proxies of accountability”).

110. Rory Van Loo, *Rise of the Digital Regulator*, 66 DUKE L.J. 1267, 1328 (2017) (“[Digital intermediaries] may inefficiently exploit consumers and constrain choice.”); Klönick, *supra* note 27, at 1668.

111. See Sarah Myers West, *Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms*, 20(11) NEW MEDIA & SOC'Y 4366, 4380 (2018) (“[C]ontent moderation systems remove content at massive levels of scale, but do not do much to educate users about where they went wrong.”).

112. Consider, for instance, the potential ramifications of the movement to revoke the immunity accorded to digital platforms under Section 230 Communication Decency Act, which is becoming a matter of bipartisan consensus in the United States Without such a legal shield, platforms are likely to be reluctant to leave any controversial content, are likely to side with corporate and governmental complaints about users' content based on the increased risk of litigation, and overall are likely to become more aggressive about removing content.

for content rises, platforms might become increasingly reluctant to enable user-generated content at all, and may even become more actively involved in producing content themselves.<sup>113</sup>

Second, fearing governmental sanctions, platforms may opt to comply with the government's stance on divisive issues. Thus, rather than encouraging competition among platforms, and facilitating pluralist approaches to content moderation, such top-down measures may undermine diversity in the platform economy.

Indeed, Trump's Executive Order on Preventing Online Censorship reflects a constitutional paradox.<sup>114</sup> It aims to protect the First Amendment and democratic values, striving to "foster and protect diverse viewpoints in today's digital communications environment where all Americans can and should have a voice . . . and encourage standards and tools to protect and preserve the integrity and openness of American discourse and freedom of expression."<sup>115</sup> Yet such blunt interference in free speech on behalf of the executive branch, using legislative measures,<sup>116</sup> runs contrary to the very purpose of the First Amendment—namely, to shield free expression from governmental intervention.<sup>117</sup>

Alternatively, another way to hold platforms more accountable for content moderation practices is through self-governance and private ordering, i.e., a bottom-up check.<sup>118</sup> While top-down regulation often takes a one-size-fits-all approach, applying a generally applicable standard uniformly, private ordering leaves room for more diversity and exploration. Enabling pluralism is the underlying principle of the liberal view of free expression, and the constitutional shield from governmental intervention is designed to ensure sufficient space for such private exploration. Platforms, as private actors that facilitate private forums, are exempted from undertaking any public duties, and therefore are not subject to constitutional claims by users regarding removal of their content.<sup>119</sup> The liberal view of the public/private divide thus keeps the state away from content moderation by private actors, leaving it to private ordering to tackle the accountability of social media platforms.<sup>120</sup> Yet the current

---

113. Elkin-Koren, Nahmias & Perel, *supra* note 88, at 6 n.19.

114. Executive Order on Preventing Online Censorship, *supra* note 101.

115. *Id.* § 1.

116. The executive order has been widely criticized for appropriating legislative competence over platforms' immunity. See, e.g., Tim Wu, *Trump's Response to Twitter is Unconstitutional Harassment*, N.Y. TIMES (June 2, 2020), <https://www.nytimes.com/2020/06/02/opinion/trump-twitter-executive-order.html> [<https://perma.cc/XC6Z-XDF4>].

117. See *Manhattan Cmty. Access Corp. v. Halleck*, 139 S. Ct. 1921, 1926 (2019) (holding that the Free Speech Clause of the First Amendment of the U.S. Constitution prohibits only governmental, not private, abridgment of speech).

118. See Suzor, *supra* note 98, at 4; CUNNINGHAM & CRAIG, *supra* note 43, at 75.

119. See *infra* notes 130–31 and accompanying text.

120. See generally, e.g., Balkin, *supra* note 86 (discussing the need for better cooperation between public and private actors in regulating online speech); Neil Weinstock Netanel, *New*

legal situation does not sufficiently mitigate the power exercised by platforms in their quasi-public roles as moderators of online discourse.<sup>121</sup>

Platforms would have to become more accountable for removing content or blocking accounts if the varied stakeholders who are the immediate subjects of moderation—namely the users of social media platforms—encourage it (through what we named “private ordering”). Indeed, the growing concern over the dominant power acquired by a handful of large commercial players in the platform economy, and the (justifiable) desire to restrain the excessive power wielded by platforms, should not obscure the significant role of users as stakeholders who could shape the governance of discourse on social media. Disillusioned by the collapse of earlier hopes that the collaborative and participatory nature of online discourse would enhance political freedom,<sup>122</sup> many critics have emphasized the purely commercial nature of online discourse, arguing that social media platforms have now turned into standard corporate players within the media industry.<sup>123</sup> Yet what is often missing from such descriptions is a better understanding of the complexity of stakeholders acting in this space, and the voice of users, speakers, activists, creators, co-creators, and collaborators who act and connect on social media platforms.

Cunningham and Craig have offered a more comprehensive view of the social media ecosystem, which they refer to as “[S]ocial [M]edia [E]ntertainment” (“SME”).<sup>124</sup> They explain how creators, or “social entrepreneurs,” extract different types of rewards from platforms, and are able to harness multiple platforms to build communities of followers which they convert into commercial value.<sup>125</sup> More recently, they have demonstrated how these stakeholders also attempt to shape the ecosystem in which they operate.<sup>126</sup> Indeed, they argue that “while there may be a greater tendency toward oligopoly in platform

*Media in Old Bottles? Barron’s Contextual First Amendment and Copyright in the Digital Age*, 76 GEO. WASH. L. REV. 952 (2008) (discussing the importance of communication in online speech regulation); Gregory P. Magarian, *The First Amendment, the Public-Private Distinction, and Nongovernmental Suppression of Wartime Political Debate*, 73 GEO. WASH. L. REV. 101, 104 (2004) (illustrating how “the public-private distinction in constitutional law” prevents the state from addressing censorship by private actors).

121. See Giovanni De Gregorio, *From Constitutional Freedoms to the Power of the Platforms: Protecting Fundamental Rights Online in the Algorithmic Society*, 11 EUR. J. LEGAL STUD. 65, 85–89 (2019).

122. MANUEL CASTELLS, *COMMUNICATION POWER* 87–88 (2009); CLAY SHIRKY, *HERE COMES EVERYBODY: THE POWER OF ORGANIZING WITHOUT ORGANIZATIONS* 171 (2008); YOCHAI BENKLER, *THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM* 176–77 (2006).

123. JOSÉ VAN DIJCK, *THE CULTURE OF CONNECTIVITY: A CRITICAL HISTORY OF SOCIAL MEDIA* 24–25 (2013).

124. CUNNINGHAM & CRAIG, *supra* note 43, at 4–6.

125. *Id.* at 11–13, 16.

126. Stuart Cunningham & David Craig, *Creator Governance in Social Media Entertainment*, SOC. MEDIA + SOC’Y, Oct.–Dec. 2019, at 1, 2, <https://journals.sagepub.com/doi/pdf/10.1177/2056305119883428> [<https://perma.cc/HVN8-gJ8C>].

capitalism, there is also an expansive opportunity for peer-to-peer, horizontal and potentially also democratic voices and self-expression.”<sup>127</sup> Far from content production by traditional media, which is shaped by corporate business strategy, the content generated by users of social media facilitate more diverse outlets which are subject to multiple tiers of governance.<sup>128</sup>

Users of social media platforms are important stakeholders in the platform economy with vested interests and expectations. They are critical pieces of platforms’ business model, providing services, content and data which are feeding platforms’ profits. Yet, currently the law overlooks their vested interests. Enabling users to effectively express their interests in a legally binding way may bring platforms’ actions into alignment with the community goals agreed upon by users.

Specifically, if users were able to hold platforms accountable under the formal and informal norms underlying online exchange, and could claim damages against platforms when such exchanges fall short of their legitimate expectations, platforms would be obliged to account for users’ interests when applying content moderation policies. This would help to align content moderation policies with the common interests of the platforms’ users, rather than simply reflecting the narrow commercial interests of platforms.

Within this framework, we believe that contract law can play an important role in supplementing top-down regulation with inclusiveness, diversity, and breadth. However, so far, users remain unsuccessful in claiming that a platform’s content moderation practices violate contracts. That is due to a misinterpretation of the nature of social media agreements, as we discuss next in Part III.

### III. THE LEGAL BARRIERS TO CONTESTING CONTENT REMOVALS BY PLATFORMS

Can users claim damages caused by unjustified content removal or account termination decisions? Can they require platforms to reinstate their content or reactivate their accounts? In this section, we demonstrate that vertical claims based on allegations that a platform’s content removal actions violated one’s rights—whether based on constitutional grounds, tort law, or contract law—are bound to fail.<sup>129</sup> As we show below, the legal framework that

---

127. *Id.* at 3.

128. Niva Elkin-Koren, *Governing Access to User-Generated Content: The Changing Nature of Private Ordering in Digital Networks*, in GOVERNANCE, REGULATIONS AND POWERS ON THE INTERNET 318, 332 (Eric Brousseau, Meryem Marzouki & Cécile Méadel eds., 2012).

129. See *Belknap v. Alphabet, Inc.*, 504 F. Supp. 3d 1156, 1159–61 (D. Or. 2020); *Doe v. Google LLC*, No. 20-cv-07502-BLF, 2020 WL 6460548, at \*3–5 (N.D. Cal. Nov. 3, 2020); *Jones v. Twitter, Inc.*, No. RDB-20-1963, 2020 WL 6263412, at \*3–5 (D. Md. Oct. 23, 2020); *Zimmerman v. Facebook, Inc.*, No. 19-cv-04591-VC, 2020 WL 5877863, at \*1–2 (N.D. Cal. Oct. 2, 2020); *Brikman v. Twitter, Inc.*, No. 19-cv-5143, 2020 WL 5594637, at \*2–5 (E.D.N.Y. Sept. 17, 2020); *Enhanced Athlete Inc. v. Google LLC*, 479 F. Supp. 3d 824, 834 (N.D. Cal. 2020).



applies to the vertical relationship between platforms and users, and its current implementation by the courts, provides insufficient remedies for users seeking to challenge the varied implications of content removals.<sup>130</sup> We then argue that this problem could be corrected if courts adopt a broader view and consider the different horizontal relationships underlying the operation of social media platforms, as suggested by contractual network theory.

#### A. CONSTITUTIONAL LAW BARRIERS

Illegitimate removals of content from social media could impede users' right to express their opinions freely, and deprive the public at large from accessing diverse information online. Indeed, any decision to remove content or terminate an account may result in the silencing of speakers—something that inevitably affects people's constitutional rights and shapes the online discourse. Accordingly, it seems that public law, which aims to safeguard fundamental rights and democratic principles, including free speech and the free flow of information, should offer an appropriate legal framework for addressing such content removal conflicts.

Unsurprisingly, some users rely on their constitutional rights, primarily freedom of expression, when seeking remedies against the misuse of power in content removal or account termination cases.<sup>131</sup> However, constitutional law hands a clear and unrestricted role, as well as a legal shield, to social media platforms in deciding the conditions for content removal, leaving users with no remedies under constitutional law. Specifically, the free speech clause of the First Amendment constitutes the first barrier to users seeking to contest allegedly unjustifiable removals.<sup>132</sup>

In numerous cases, courts have considered platforms as private companies that function as private forums, hence denying any attempt to hold them to judicial scrutiny under the First Amendment. For instance, in *Johnson v. Twitter*, the court stressed that Twitter is a private sector entity whose services are dependent on agreements with users.<sup>133</sup> After Twitter blocked the account of Charles C. Johnson because he allegedly asked for donations to help “take

130. See discussion *supra* Part II.

131. In the field of constitutional law, see, e.g., *Langdon v. Google, Inc.*, 474 F. Supp. 2d 622, 631–32 (D. Del. 2007); *Estavillo v. Sony Comput. Ent. Am.*, No. C-09-03007 RMW, 2009 WL 3072887, at \*1 (N.D. Cal. Sept. 22, 2009); *Green v. AOL*, 318 F.3d 465, 472 (3d Cir. 2003); *Cyber Promotions, Inc. v. AOL*, 948 F. Supp. 436, 438 (E.D. Pa. 1996). In the field of private law, see, e.g., *Order Issuing Alternative Writ of Mandate at 194, Twitter, Inc., v. Superior Ct. ex rel. Taylor*, No. CGC-18-564460 (Cal. Ct. App. Aug. 17, 2018); *Hall v. Earthlink Network, Inc.*, No. 98 Civ. 5489(RO), 2003 WL 22990064, at \*2–3 (S.D.N.Y. Dec. 19, 2003); *Young v. Facebook, Inc.*, 790 F. Supp. 2d 1110, 1118–19 (N.D. Cal. 2011).

132. U.S. CONST. amend. I (“Congress shall make no law . . . abridging the freedom of speech . . .”).

133. Tentative Rulings for Department 503 at 30, *Johnson v. Twitter, Inc.*, No. 18CECG00078 (Cal. Sup. Ct. June 6, 2018).

out” civil rights activist DeRay Mckesson,<sup>134</sup> Johnson tried to contest his account termination in court.<sup>135</sup> He argued that social media such as Twitter “are the modern version of the old public square” and that “parties should be able to freely express their views, without social media companies monopolizing what types of speech may be expressed on their platforms.”<sup>136</sup> The court, however, disagreed, explaining that Twitter has the “First Amendment right to exercise independent editorial control over the content on its platform,”<sup>137</sup> and termination of the account “is an editorial decision regarding how to present content.”<sup>138</sup>

Likewise, in the Prager University (“PragerU”) case,<sup>139</sup> YouTube tagged several dozen of PragerU’s videos as appropriate for the platform’s “restricted mode,” which meant they would not be available to any users accessing YouTube under this setting (including libraries, schools, and businesses).<sup>140</sup> “YouTube [further] ‘demonetized’ some of PragerU’s videos,” so that the plaintiff could not collect advertising revenues.<sup>141</sup> PragerU challenged the subjective and discretionary tagging of its videos by YouTube, claiming that this violated the non-profit’s constitutional rights.<sup>142</sup> The Court of Appeals of the Ninth Circuit disagreed, explaining that the First Amendment framework does not apply to YouTube, and therefore the removal of content according to YouTube’s internal policies is not subject to the constitutional safeguards of free speech.<sup>143</sup> The court held that “[d]espite YouTube’s ubiquity and its role as a public-facing platform, it remains a private forum, not a public forum subject to judicial scrutiny under the First Amendment.”<sup>144</sup> Even if content removals such as those experienced by PragerU harm users, the court explained, they have agreed to be bound by YouTube’s ToS.<sup>145</sup> Therefore, they cannot claim a violation of their constitutional right to free speech.<sup>146</sup>

Even the “state action doctrine,” which theoretically leaves room for subjecting private actors who perform public functions to constitutional scrutiny, has been applied in such a way as to reinforce the constitutional shield protecting social media platforms. Also known as the horizontal effect

---

134. *Id.* at 31.

135. *Id.* at 27.

136. *Id.* at 29.

137. *Id.* at 30.

138. *Id.*

139. *Prager Univ. v. Google LLC*, 951 F.3d 991, 991 (9th Cir. 2020).

140. *Id.* at 996.

141. *Id.*

142. *Id.*

143. *Id.* at 999.

144. *Id.* at 995.

145. *Id.* at 999.

146. *Id.*

of fundamental rights,<sup>147</sup> the state action doctrine holds that the Constitution generally applies only to governmental conduct, and does not prohibit the deprivation of constitutional rights by private actors.<sup>148</sup> Lower courts have stressed in different decisions that private actors, including social media platforms, do not qualify “as state actors subject to First Amendment scrutiny merely because they hold out and operate their private property as a forum for expression of diverse points of view.”<sup>149</sup> The fact that social media platforms allow the use of their network by the public is not enough to subject them to First Amendment safeguards. While private actors might be held liable for violating the constitutional right to free speech under the state action doctrine if they act on behalf of the government or perform a function that is normally implemented by the government,<sup>150</sup> the operation of a public forum for speech is not a traditional, exclusive public function, and therefore it is not bound by governmental constraints on speech.<sup>151</sup>

For instance, in the case of *Tulsi Now v. Google*, the plaintiff claimed that the temporary suspension of a verified political advertising account for several hours shortly after a Democratic primary debate was a violation of the plaintiff’s First Amendment rights.<sup>152</sup> Specifically, it was argued that Google had become a state actor because it provided advertising services surrounding the 2020 presidential election, and “that, by regulating political advertising on its own platform, Google exercised the traditional government function of regulating elections.”<sup>153</sup> The court, however, rejected this argument, noting that the plaintiff failed to establish “how Google’s regulation of its own platform is in any way equivalent to a government[] regulation of an election,”

---

147. See Stephen Gardbaum, *The “Horizontal Effect” of Constitutional Rights*, 102 MICH. L. REV. 387, 388 (2003) (“These alternatives refer to whether constitutional rights regulate only the conduct of governmental actors in their dealings with private individuals (vertical) or also relations between private individuals (horizontal).”); Mark Tushnet, *The Issue of State Action/Horizontal Effect in Comparative Constitutional Law*, 1 INT’L J. CONST. L. 79, 92 (2003) (“[I]f horizontality is understood as a response to the threat to liberty posed by concentrated private power, the solution is to require that all private actors conform to the norms applicable to governmental actors.”).

148. See, e.g., Erwin Chemerinsky, *Rethinking State Action*, 80 NW. U. L. REV. 503, 507 (1985); Lillian BeVier & John Harrison, *The State Action Principle and Its Critics*, 96 VA. L. REV. 1767, 1769 (2010). The Civil Rights Cases are usually credited with being the origin of the state action requirement. See *Civil Rights Cases*, 109 U.S. 3, 11 (1883).

149. *Prager Univ. v. Google, Inc.*, No. 17-CV-06064, 2018 WL 1471939, at \*8 (N.D. Cal. Mar. 26, 2018); *Cyber Promotions, Inc. v. AOL*, 948 F. Supp. 436, 456 (E.D. Pa. 1996); *Estavillo v. Sony Comput. Ent. Am.*, No. C-09-03007, 2009 WL 3072887, at \*1-2 (N.D. Cal. Sept. 22, 2009).

150. *Prager Univ.*, 951 F.3d at 997.

151. See generally *Manhattan Cmty. Access Corp. v. Halleck*, 139 S. Ct. 1921 (2019) (holding that the discretion to limit free speech of a non-profit corporation designated by New York City to run a public access television network does not violate First Amendment rights because the TV station is not exercising a power traditionally and exclusively performed by the government).

152. *Tulsi Now, Inc. v. Google, LLC*, No. 2:19-cv-06444, 2020 WL 4353686, at \*1 (C.D. Cal. Mar. 3, 2020).

153. *Id.*

and that “[t]o the extent Google ‘regulates’ anything, it regulates its own private speech and platform.”<sup>154</sup> Similarly, in the case of *Prager University*, the court stressed that “private property does not ‘lose its private character merely because the public is generally invited to use it for designated purposes’ [and] YouTube may be a paradigmatic public square on the Internet, but it is ‘not transformed’ into a state actor solely by ‘provid[ing] a forum for speech.’”<sup>155</sup>

Accordingly, it appears that the vertical approach to constitutional rights leaves social media platforms free to moderate content without infringing users’ constitutional rights. The result is that no matter what harm users face when their content is removed, they will fail to claim a violation of their constitutional right to free speech vis-à-vis social media.

### B. TORT IMMUNITY AND EXEMPTIONS FROM LIABILITY

It is not only constitutional law that forecloses vertical allegations by users against platforms that remove their content or terminate their accounts. In fact, when users seek to hold platforms liable for the harm they incur as a result of removal decisions, their claims are most often barred by statutory immunities concerning tort liability. The CDA and DMCA are the two pillars of the U.S. civil law shield, which exempt social media platforms from tort liability when they make editorial decisions on third-party content. Despite some differences between the two legal instruments,<sup>156</sup> both statutes consider social media platforms as extraneous to the unlawful content posted by users. From users’ perspective, these provisions put a barrier in the way of legal remedies against harms incurred when social media platforms remove their content.

Congress passed Section 230 of the CDA at the end of the last century primarily to ensure the development of the digital environment,<sup>157</sup> making it one of “the most important protections of free expression in the United States in the digital age.”<sup>158</sup> The legal (and political) choice was to introduce a system based on an exemption from liability for computer services which merely host third-party content. Specifically, Section 230 protects platforms from any liability for harm caused by content posted by their users, including

154. *Id.* at \*2.

155. *Prager Univ.*, 951 F.3d at 997 (alteration in original) (citations omitted) (quoting *Lloyd Corp. v. Tanner*, 407 U.S. 551, 569 (1972); *Manhattan Cmty. Access Corp.*, 139 S. Ct. at 1930).

156. Salil K. Mehra & Marketa Trimble, *Secondary Liability, ISP Immunity, and Incumbent Entrenchment*, 62 AM.J. COMPAR. L. 685, 686–88 (2014).

157. Before the adoption of the CDA, some cases had already shown that online intermediaries would not have been able to develop new digital services without legal protection from liability due to the vast range of claims concerning their liability for editing third-party content. *See, e.g.*, *Cubby, Inc. v. CompuServe Inc.*, 776 F. Supp. 135, 139–42 (S.D.N.Y. 1991); *Stratton Oakmont, Inc. v. Prodigy Servs. Co.*, 1995 WL 323710, at \*2, \*5 (N.Y. Sup. Ct. May 24, 1995) (*superseded by statute as recognized in Shiamili v. Real Est. Grp. of N.Y., Inc.*, 952 N.E.2d 1011 (N.Y. 2011)).

158. Balkin, *supra* note 86, at 2313.

harm caused by their actions in monitoring or moderating their services.<sup>159</sup> Hence, to the extent that users' claims are based on social media actions as publishers of content, courts would likely dismiss them.

The case of Meghan Murphy is a paradigmatic example.<sup>160</sup> Murphy, a "feminist writer and journalist," posted on Twitter "a series of [t]weets regarding a person named Hailey Heartless, a self-identified transsexual whose legal name is Lisa Kreut, that referred to that person as a 'white man.'"<sup>161</sup> After asking Murphy to delete some of this content, Twitter suspended Murphy's account, claiming she violated its hateful conduct policy, which banned "misgendering" transgender individuals (i.e., referring to someone using pronouns or other terms that do not reflect the gender with which they identify).<sup>162</sup> Murphy acknowledged that the hateful conduct policy had been amended to prohibit "targeting individuals with repeated slurs, tropes or other content that intends to dehumanize, degrade or reinforce negative or harmful stereotypes about a protected category. This includes targeted misgendering or deadnaming of transgender individuals."<sup>163</sup> Murphy contended, however, "that Twitter failed to provide adequate notice to her or other users of that [amendment to its policy], and improperly applied it retroactively to her."<sup>164</sup> In her complaint, Murphy raised a contractual cause of action, claiming that Twitter breached its own User Agreement "by failing to provide Murphy with 30 days advance notice of the changes to its Hateful Conduct Policy, by retroactively applying the amended policy to Murphy, and by permanently suspending her account although she did not violate the Terms of Service, Rules or policies."<sup>165</sup> The court disagreed. It held that the complaint was barred by Section 230 of the CDA, and more precisely, that Twitter was acting legitimately in its capacity as a publisher when it suspended Murphy's account.<sup>166</sup> Specifically, "there [was] no dispute that Twitter is a 'provider . . . of an interactive computer service,'"<sup>167</sup> and "that

159. 47 U.S.C. § 230(c)(1) (2018) ("No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."); *see also* Klayman v. Zuckerberg, 910 F. Supp. 2d 314, 317–18 (D.D.C. 2012).

160. *See generally* Order Denying Special Motion to Strike the Complaint Under California Code of Civil Procedure Section 425.16 and Sustaining Demurrer to Complaint Without Leave to Amend, *Murphy v. Twitter, Inc.*, No. CGC-19-573712 (Cal. Super. Ct. June 12, 2019), <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2968&context=historical> [<https://perma.cc/28BY-FHXH>] (discussing a motion by Twitter to strike Plaintiff's complaint that among other things, alleges Twitter violated their User Agreement).

161. *Id.* at 2.

162. *Id.*

163. *Id.* at 3 (quoting *Hateful Conduct Policy*, TWITTER (Oct. 2018), <https://web.archive.org/web/20181012054519/https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> [<https://perma.cc/YQ6E-HB7P>]).

164. *Id.*

165. *Id.* at 5.

166. *Id.* at 11–16.

167. *Id.* at 11 (quoting 47 U.S.C. § 230(c)(1) (2018)).

Murphy's Tweets [were] 'information provided by another information content provider.'<sup>168</sup> The dispute between the parties centered on whether Murphy sought to impose liability on Twitter in its capacity as a publisher—a capacity under which Section 230 specifically precludes courts from entertaining claims of liability.<sup>169</sup> The court answered positively, finding that the actions of Twitter, namely suspending or banning users' accounts and enforcing policies governing the permissible scope of content in those accounts—were all actions within the traditional scope of a publisher's role.<sup>170</sup> According to the court, the fact that Murphy pleaded a contractual cause of action had no bearing on the matter, because "what matters is whether the cause of action inherently requires the court to treat the defendant as the 'publisher or speaker' of content provided by another."<sup>171</sup>

But there are exceptions. Courts have consistently held that for the purpose of classifying platforms as publishers, there is no difference between actively deciding to post content and to remove it.<sup>172</sup> This rationale does not apply, however, when platforms are acting not in their capacity as publishers of someone else's content, but as speakers.<sup>173</sup> For instance, in *Fair Housing Council of San Fernando Valley v. Roommates.com*,<sup>174</sup> the court found that Roommates.com was the creator of the content because it required subscribers to create profiles and answer personal questions about themselves and their preferences.<sup>175</sup> Since Roommates.com became "much more than a passive transmitter of information provided by others[,]” the court found it liable for violating the federal Fair Housing Act and California housing-discrimination laws.<sup>176</sup> Likewise, the court rejected Facebook's Section 230 defense in *Fraleay v. Facebook*,<sup>177</sup> where users sued Facebook for using their profile pictures in ads, claiming a right-of-publicity violation. The court ruled

---

168. *Id.* (citing 47 U.S.C. § 230(c)(1), (f)(3)).

169. *Id.* at 11.

170. *Id.* at 11–16.

171. *Id.* at 15 (quoting *Cross v. Facebook, Inc.*, 222 Cal. Rptr. 3d 250, 264 (Ct. App. 2017)).

172. See *Mezey v. Twitter, Inc.*, No. 1:18-cv-21069, 2018 WL 5306769, at \*1–2 (S.D. Fla. July 19, 2018); *Cohen v. Facebook, Inc.*, 252 F. Supp. 3d 140, 155–58 (E.D.N.Y. 2017); *Cross*, 222 Cal. Rptr. 3d at 264; *Fields v. Twitter, Inc.*, 217 F. Supp. 3d 1116, 1120–24 (N.D. Cal. 2016); *Riggs v. MySpace, Inc.*, 444 F. App'x 986, 987 (9th Cir. 2011); *Barrett v. Rosenthal*, 146 P.3d 510, 518–20 (Cal. 2006) (holding that suspensions or deletions of users' accounts are considered publishing activity under Section 230).

173. In *Demetriades*, the court held that Section 230 did not bar the plaintiff from holding Yelp liable for making public statements about Yelp (praising the quality of its reviews or claiming that it runs an effective "filter" for unreliable reviews which results in "the most trusted reviews"). *Demetriades v. Yelp, Inc.*, 175 Cal. Rptr. 3d 131, 143–44 (Ct. App. 2014).

174. *Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1165 (9th Cir. 2008) (en banc) ("[T]he party responsible for putting information online may be subject to liability, even if the information originated with a user.").

175. *Id.* at 1161.

176. *Id.* at 1166.

177. *Fraleay v. Facebook, Inc.*, 830 F. Supp. 2d 785, 801–03 (N.D. Cal. 2011).

in this case that the platform had not been accused “of publishing tortious content, but . . . of creating and developing commercial content that violates [the plaintiffs’] statutory right of publicity.”<sup>178</sup> Similarly, Snapchat was held liable for applying a graphic filter over a user’s photo, creating something new.<sup>179</sup> Since this was Snapchat’s private conduct, the Court held that CDA immunity did not apply.<sup>180</sup>

The CDA further provides immunity for good faith actions aimed at restricting access to (or availability of) content considered “obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.”<sup>181</sup> This provision allows social media platforms to remove content discretionarily without taking into account the impact of such removals on the public interest. The only limit governing such removals is the requirement of maintaining “good faith.”<sup>182</sup>

Of course, this limit is effectively meaningless as good faith is measured subjectively, and there is no impartial or objective way to ascertain the real reasons or logic behind a specific act of content removal.<sup>183</sup>

The CDA is not the only statutory provision that bars users from using the law to hold platforms liable for their harms. The DMCA’s safe harbor system fills in the gap concerning copyrighted content, which is excluded from the scope of the CDA.<sup>184</sup> This legislation reflected a compromise between platforms and holders of copyrights, who hoped to harness the platforms’ technological capabilities to benefit from their exclusive rights online.<sup>185</sup> On the one hand, the challenges raised by online piracy and the consequent economic losses experienced by copyright owners led them to demand a shift in the burden of monitoring the use of their creations. On the other hand, online intermediaries, wishing to maintain their passive role in relation to content creation, wanted to minimize barriers to the free flow of information online, which is the pillar of their business model.<sup>186</sup>

178. *Id.* at 801.

179. *Maynard v. Snapchat, Inc.*, 816 S.E.2d 77, 79–81 (Ga. Ct. App. 2018); *see also* *Maynard v. Snapchat, Inc.*, 851 S.E.2d 128, 130 (Ga. Ct. App. 2020) (“Snapchat created and distributed a feature within its application, known as the Speed Filter, that allows Snapchat users to record their speed and overlay that speed onto a Snapchat photo or video.”).

180. *Maynard*, 816 S.E.2d at 136.

181. 47 U.S.C. § 230(c)(2)(A) (2018).

182. *Id.* (“No [platform] . . . shall be held liable on account of — (A) any action voluntarily taken in good faith to restrict access to or availability of material that the [platform] . . . considers . . . [in any way] objectionable, whether or not such material is constitutionally protected.”).

183. Perel & Elkin-Koren, *supra* note 35, at 501–02.

184. Communications Decency Act of 1996, Pub. L. No. 104-104, § 230(d)(2), 110 Stat. 56, 139 (1996) (codified as amended at 47 U.S.C. § 230(e)(2)).

185. For a comprehensive description and analysis of the passage of the DMCA, see JESSICA LITMAN, *DIGITAL COPYRIGHT* 35–69 (2006).

186. Niva Elkin-Koren, *After Twenty Years: Revisiting Copyright Liability of Online Intermediaries*, in *THE EVOLUTION AND EQUILIBRIUM OF COPYRIGHT IN THE DIGITAL AGE* 29, 35–38 (Susy Frankel & Daniel Gervais eds., 2014).

In contrast to the CDA, the DMCA does not provide an absolute exemption from copyright liability, but it shields service providers that host, store, route, or transmit user-generated content from liability as long as they meet certain conditions. Under the DMCA, service providers hosting “[i]nformation [r]esiding on [s]ystems or [n]etworks at [the] [d]irection of [u]sers” must meet three conditions in order to avoid liability for infringing content.<sup>187</sup> First, they must not have either (a) actual knowledge that material on the system or network, or an activity using the material, is infringing,<sup>188</sup> or (b) “aware[ness] of facts or circumstances from which infringing activity is apparent.”<sup>189</sup> Second, they must not be in receipt of any “financial benefit directly attributable to the infringing activity, in a case [where] the service provider has the right and ability to control such activity.”<sup>190</sup>

Finally, “upon notification of claimed infringement . . . , [the service provider must] respond[] expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity.”<sup>191</sup> Only if all three of those conditions are met is the service provider not held liable for copyright infringements.

While the CDA does not provide any remedies for users whose content is removed, the DMCA establishes certain safeguards against arbitrary or unfair content removal. Most notably, the filing of a notice of copyright infringement must be founded on “a good-faith belief that [the targeted] use . . . [was] not authorized by the copyright owner, its agent, or the law.”<sup>192</sup> Any party who files a takedown notice without such good-faith belief might be liable for damages.<sup>193</sup> The scope of this provision was addressed in *Lenz v. Universal Music Corp.*,<sup>194</sup> which held that right holders must consider whether potentially infringing material is permitted under fair use provisions before issuing a takedown notice.<sup>195</sup> Hence, users can seek damages where they suspect that the DMCA notice-and-takedown regime was misused by right holders in an attempt to remove content that was authorized by law.<sup>196</sup>

Furthermore, under the DMCA, if they satisfy two general requirements relating to standard technical measures and removal of repeat infringers, social media platforms are not liable for removal in good faith if they comply

---

187. Digital Millennium Copyright Act, Pub. L. No. 105-304, § 512(c), 112 Stat. 2860, 2879–80 (1998) (codified as amended at 17 U.S.C. § 512(c)).

188. *Id.* § 512(c)(1)(A)(i).

189. *Id.* § 512(c)(1)(A)(ii).

190. *Id.* § 512(c)(1)(B).

191. *Id.* § 512(c)(3).

192. *Id.* § 512(c)(3)(A)(v).

193. *Id.* § 512(f).

194. *Lenz v. Universal Music Corp.*, 815 F.3d 1145, 1153 (9th Cir. 2016), *cert. denied*, 137 S. Ct. 416 (2016).

195. *Id.* at 1157.

196. Niva Elkin-Koren, *Fair Use by Design*, 64 UCLA L. REV. 1082, 1091–92 (2017).



with additional cumulative conditions. First, service providers must promptly “take[] reasonable steps to . . . notify the subscriber that [they have] removed or disabled access to the material.”<sup>197</sup> Second, “upon receipt of a counter notification . . . , [service providers must—again] promptly[—]provide[] the person who [submitted] the notification . . . with a copy of the counter notification, and inform[] that person that it will replace the removed material or cease disabling access to it in 10 business days.”<sup>198</sup> Third, service providers must then indeed replace (or cease disabling access to) the removed material within 10 to 14 business days “following receipt of the counter notice, unless” the person who submitted the original notification has filed a copyright infringement claim relating to the content “on the service provider’s system or network.”<sup>199</sup>

Failure to comply with these procedural safeguards could trigger legal liability on the part of social media platforms and allow users to contest content removals.<sup>200</sup> Nevertheless, although scholars have underlined the potential for abuse of the DMCA removal system,<sup>201</sup> in practice, users hardly ever rely on these procedural safeguards. Moreover, as we have argued before, the turn to algorithmic copyright enforcement, which includes the use of automated mechanisms both for notification and takedown purposes, further reduces the effectiveness of the DMCA to guard against erroneous removals.<sup>202</sup> Indeed, today massive volumes of material are automatically detected by algorithms and removed from public circulation unless explicitly authorized by the right holder, making it nearly impossible to review the legitimacy of removal decisions on a case-by-case basis.

### C. TERMS OF SERVICE AND BOILERPLATE CONTRACTS

Besides the external statutory immunities that bar users’ attempts to seek remedies for harms suffered due to the removal of their content or the termination of their accounts, users’ claims are also blocked by the internal contractual setting of platforms’ policies. Indeed, platforms seek to govern

197. Digital Millennium Copyright Act § 512 (g)(2)(A).

198. *Id.* § 512 (g)(2)(B).

199. *Id.* § 512 (g)(2)(C).

200. Though of course these safeguards do not address platforms’ powers to control access to content through the organization of content, recommendation algorithms, and blocking users’ accounts.

201. See, e.g., John Tehranian, *The New ©ensorship*, 101 IOWA L. REV. 245, 282–83 (2015); Daniel Seng, *The State of the Discordant Union: An Empirical Analysis of DMCA Takedown Notices*, 18 VA. J.L. & TECH. 369, 427 (2014); Wendy Seltzer, *Free Speech Unmoored in Copyright’s Safe Harbor: Chilling Effects of the DMCA on the First Amendment*, 24 HARV. J.L. & TECH. 171, 178–81 (2010); Jeffrey Cobia, Note, *The Digital Millennium Copyright Act Takedown Notice Procedure: Misuses, Abuses, and Shortcomings of the Process*, 10 MINN. J.L. SCI. & TECH. 387, 395 (2009); Jennifer M. Urban & Laura Quilter, *Efficient Process or “Chilling Effects”? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA COMPUT. & HIGH TECH. L.J. 621, 687–88 (2006).

202. Elkin-Koren, *supra* note 196, at 1093; Perel & Elkin-Koren, *supra* note 35, at 502.

their relationship with users by ToS which further incorporate additional legal documents.<sup>203</sup> Typically, additional policies would include rules of conduct or, documents resembling a bill of rights, such as Facebook’s community standards,<sup>204</sup> or YouTube Community Guidelines,<sup>205</sup> and a Data Policy or Privacy Policy.<sup>206</sup>

The ToS incorporate these documents by reference, stating that a violation of such terms and policies also constitutes a violation of the ToS.<sup>207</sup>

Courts have broadly addressed cases where such contracts have been used as legal barriers to consumer complaints.<sup>208</sup> That is, when drafting their ToS, platforms often shield themselves against different forms of responsibility through broad disclaimers of liability as well as specific provisions. In the context of content removal, these might involve contractual provisions that grant platforms explicit authority to remove content for any reason, at their own discretion.<sup>209</sup>

The aforementioned case of Jan Lewis is demonstrative.<sup>210</sup> In 2012, YouTube deleted Lewis’s channels without sending her any notice or explanation, claiming that she had violated the terms of use that prohibited users from “collect[ing] or harvest[ing] any personally identifiable information, including account names, from the Service . . . [and] us[ing] the communication systems provided by the Service (e.g., comments, email) for any commercial solicitation purposes.”<sup>211</sup> As noted earlier, YouTube reinstated Lewis’s account, but without restoring the videos and follower-generated information (comments, view counts) that had accumulated on the old channel.<sup>212</sup> Lewis filed a complaint

203. See for instance, *Terms and Policies*, FACEBOOK, <https://www.facebook.com/policies> [<https://perma.cc/Q2JS-44LV>].

204. *Facebook Community Standards*, *supra* note 65.

205. *Community Guidelines*, *supra* note 65.

206. *Data Policy*, FACEBOOK (Jan. 11, 2021), <https://www.facebook.com/policy.php> [<https://perma.cc/B2K2-PR2K>]; *Privacy Policy*, GOOGLE (July 1, 2021), <https://policies.google.com/privacy?hl=en-US> [<https://perma.cc/W2RJ-YXSJ>].

207. See *Terms of Service*, FACEBOOK (Oct. 22, 2020), at 3.2, <https://www.facebook.com/legal/terms> [<https://perma.cc/39XL-5MAC>] (allowing removal of content in violation of the Community Standards); see *id.* at 4.2 (allowing suspension or termination of accounts if Facebook determines that a user “clearly, seriously or repeatedly breached” the Community Standards).

208. See, e.g., *AT&T Mobility, LLC v. Concepcion*, 563 U.S. 333, 336–38 (2011); *Fiser v. Dell Comput. Corp.*, 188 P.3d 1215, 1222 (N.M. 2008); *Bouley v. Quizno’s Master LLC (In re Bouley)*, 503 B.R. 524, 528–29 (Bankr. D.N.H. 2013); *Kirton v. Fields*, 997 So. 2d 349, 351 (Fla. 2008); *City of Santa Barbara v. Superior Ct.*, 161 P.3d 1095, 1097–98 (Cal. 2007); *Plattner v. Edge Sols. Inc.*, No. 03-CV-2646, 2004 WL 1575557, at \*1 (N.D. Ill. Apr. 1, 2004); *Swain v. Auto Servs., Inc.*, 128 S.W.3d 103, 108 (Mo. Ct. App. 2003); *Wilder v. Absorption Corp.*, 107 S.W.3d 181, 185 (Ky. 2003); *Hagedorn v. Veritas Software Corp.*, 250 F. Supp. 2d 857, 862 (S.D. Ohio 2002).

209. See, e.g., *Twitter Terms of Service*, TWITTER (Aug. 19, 2021), <https://twitter.com/en/tos> [<https://perma.cc/G8RL-H8QR>] (“We reserve the right to remove Content that violates the User Agreement, including for example, copyright or trademark violations or other intellectual property misappropriation, impersonation, unlawful conduct, or harassment.”).

210. *Lewis v. YouTube, LLC*, 197 Cal. Rptr. 3d 219, 221–22 (Ct. App. 2015).

211. *Id.* at 222.

212. *Id.* at 226.

for breach of contract and damages, seeking a court order restoring her deleted channel to its condition prior to its removal.<sup>213</sup> The court, however, refused to grant the order, explaining that “there is no provision in the Terms of Service that requires YouTube to maintain particular content on the Service or at a particular location on the Service.”<sup>214</sup> The court continued, “[t]here is also no provision in the Terms of Service pursuant to which YouTube is obligated to display view counts or comments associated with videos. There is nothing in the Terms of Service even suggesting that YouTube is a storage site for users’ content.”<sup>215</sup>

The court also rejected Lewis’s claim for damages, on the grounds that these were prohibited under the limitation of liability provision in YouTube’s ToS.<sup>216</sup>

Addressing this limitation of liability clause, the court stated that such clauses were generally “appropriate when one party is offering a service for free to the public.”<sup>217</sup> Lewis argued that this clause should not apply to her case because “she has not alleged that there were any errors or omissions in any content, but rather a deletion of her content without prior notice.”<sup>218</sup> The court did not agree. Pointing at the language of the ToS, which broadly defines “content” to include “the text, software, scripts, graphics, photos, sounds, music, videos, audiovisual combinations, interactive features and other materials you may view on, access through, or contribute to the Service,”<sup>219</sup> the court concluded that “the limitation of liability clause encompassed Lewis’s claim that YouTube wrongfully failed to include her videos, the number of views of these videos, and the comments on the videos by other YouTube visitors on its Web site.”<sup>220</sup> In other words, the court held that the deletion of content falls under the rubric of “omissions,” and therefore, that YouTube’s limitation of liability does in fact apply to Lewis’s claim.

The case of Mishiyev is another example.<sup>221</sup> Recall that YouTube claimed to have terminated Mishiyev’s account following repeated copyright violations.<sup>222</sup> In his complaint, Mishiyev contended that, in fact, YouTube had terminated his account out of retaliation for his threatened lawsuit after YouTube failed

---

213. *Id.* at 225.

214. *Id.*

215. *Id.* at 225–26.

216. *Id.* at 225.

217. *Id.* at 224 (citing *Markborough Cal., Inc. v. Superior Ct.*, 277 Cal. Rptr. 919, 925 (Ct. App. 1991)).

218. *Id.*

219. *Id.* at 225.

220. *Id.*

221. *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1156 (N.D. Cal. 2020), *aff’d*, 857 F. App’x 907 (9th Cir. 2021).

222. *Id.*

to properly distribute his new videos to his subscribers.<sup>223</sup> Mishiyev alleged six claims, including breach of contract, on the grounds “that YouTube breached the parties’ agreement by declining to restore plaintiff’s videos after YouTube removed them based on allegations of copyright infringement.”<sup>224</sup> The court found that “YouTube’s ‘Terms of Service’ agreement governed the terminated relationship. The agreement vested YouTube with significant control over the operation of its service, including the ability to remove uploaded content.”<sup>225</sup> Specifically, the court pointed at YouTube’s power to discretionally remove content that infringed upon another’s intellectual property rights, as expressed in its ToS.<sup>226</sup> Even concerning YouTube’s handling of the counter-notices filed by users seeking to rebut such removals, YouTube retains full discretion to decide whether to send a copy of the counter-notice to the complainant, and whether to replace (or restore access to) the removed content if no infringement suit is filed by the complainant within 14 days.<sup>227</sup> Based on these provisions, the court concluded that YouTube was authorized to decline to restore Mishiyev’s videos after it removed them based on allegations of copyright infringement.<sup>228</sup> As to the counter-notices filed by Mishiyev, the court emphasized that “YouTube did not agree to act as a neutral processor of notices and counter-notices,” but rather “retained control to evaluate counter-notices and infringement on its own.”<sup>229</sup>

The broad discretion exercised by platforms under their ToS, against which users largely stand powerless, are a consequence of the platforms’ reliance on so-called “standard contracts,” “boilerplate contracts,” or “adhesion contracts,” all characterized by the use of standard language not based on negotiations between the contractual parties.<sup>230</sup> In contrast to traditional contracts, which are founded on the mutual consent of the parties to agreed terms, contracts of adhesion are based on a different logic meant to facilitate the closure of vast numbers of agreements every day.<sup>231</sup> Specifically, standards decided by one entity become the rule for many under a take-it-or-leave-it

---

223. *Id.*

224. *Id.* at 1159.

225. *Id.* at 1156.

226. *Id.*

227. *Id.* at 1157.

228. *Id.* at 1161.

229. *Id.* at 1159.

230. *See generally* Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 HARV. L. REV. 1173 (1983) (discussing the use of contracts of adhesion in business practices); Friedrich Kessler, *Contracts of Adhesion—Some Thoughts About Freedom of Contract*, 43 COLUM. L. REV. 629, 632 (1943). The expression “contract of adhesion” was already used at the beginning of the nineteenth century. *See* Edwin W. Patterson, *The Delivery of a Life-Insurance Policy*, 33 HARV. L. REV. 198, 222 (1919).

231. *See generally* W. David Slawson, *Standard Form Contracts and Democratic Control of Lawmaking Power*, 84 HARV. L. REV. 529 (1971) (describing contracts of adhesion as being coercive).

formula. In capitalist societies, consumers occasionally find themselves needing to enter into contracts to access certain products and services. Often, they agree to waive certain rights and freedoms for the sake of entering into these agreements.<sup>232</sup> Similarly, ToS set the rules of the relationship in the digital environment, from the e-commerce marketplace and social media platforms all the way down to navigation through a news website.<sup>233</sup> The dominant role of online platforms, and the critical services they provide in the modern economy, leave almost no room for users to object to these ToS.<sup>234</sup>

Notwithstanding the imbalanced negotiation powers of users and platforms, courts have not found obstacles to recognizing the enforceability of boilerplate contracts so long as their terms are clearly presented to users.<sup>235</sup> Indeed, even the “unconscionability doctrine,” which has been conceived as a way to mitigate “unconscientious bargains” resulting in an imbalance of contractual power,<sup>236</sup> mostly fails to assure fairness in contracts of adhesion. This is mainly because there is no consensus view concerning what can contractually “shock the conscience.”<sup>237</sup>

For instance, in *Song Fi v. Google*, YouTube was not considered liable for breach of contract deriving from its removal of a video published by the plaintiff for violating its terms of use.<sup>238</sup> The plaintiff argued that YouTube’s ToS was an unconscionable contract from a procedural and substantive

232. See generally Brittany Scott, Note, *Waiving Goodbye to First Amendment Protections: First Amendment Waiver by Contract*, 46 HASTINGS CONST. L.Q. 451 (2019) (arguing that rights to free speech can be waived); Alan E. Garfield, *Promises of Silence: Contract Law and Freedom of Speech*, 83 CORNELL L. REV. 261, 264–66 (1998) (providing examples of contracts waiving one’s right to free speech). Note that the Supreme Court has not banned waivers of constitutional rights including First Amendment speech rights, including not through contractual arrangements. See *Curtis Publ’g Co. v. Butts*, 388 U.S. 130, 135 (1967); *Johnson v. Zerbst*, 304 U.S. 458, 464–65 (1938) (explaining one’s ability to waive their Constitutional rights under the Sixth Amendment); *Cohen v. Cowles Media Co.*, 501 U.S. 663, 671–72 (1991); *Snepp v. United States*, 444 U.S. 507, 515–16 (1980).

233. See generally NANCY S. KIM, WRAP CONTRACTS: FOUNDATIONS AND RAMIFICATIONS (2013) (explaining “wrap contracts” and how they affect the parties involved).

234. See generally Yannis Bakos, Florencia Marotta-Wurgler & David R. Trossen, *Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts*, 43 J. LEGAL STUD. 1 (2014) (discussing a study they performed that yielded results showing that most people who purchase software do not read, or read only a small portion of, the license agreement).

235. See *Hancock v. AT&T Co.*, 701 F.3d 1248, 1256 (10th Cir. 2012); *Specht v. Netscape Commc’ns Corp.*, 306 F.3d 17, 32–35 (2d Cir. 2002); *Serrano v. Cablevision Sys. Corp.*, 863 F. Supp. 2d 157, 164 (E.D.N.Y. 2012).

236. *Chesterfield v. Janssen*, 28 Eng. Rep. 82, 100 (Eng. 1750); see also *Hume v. United States*, 132 U.S. 406, 411 (1889) (holding that unconscionable contracts, defined as contracts that no man would enter into, are void).

237. *Bragg v. Linden Rsch., Inc.*, 487 F. Supp. 2d 593, 605 (E.D. Pa. 2007); *Comb v. PayPal, Inc.*, 218 F. Supp. 2d 1165, 1172 (N.D. Cal. 2002); see M. P. Ellinghaus, *In Defense of Unconscionability*, 78 YALE L.J. 757, 757–58 (1969). But see Robert L. Oakley, *Fairness in Electronic Contracting: Minimum Standards for Non-Negotiated Contracts*, 42 HOUS. L. REV. 1041, 1061–65 (2005).

238. *Song fi, Inc. v. Google Inc.*, 72 F. Supp. 3d 53, 56 (D.D.C. 2014).

standpoint.<sup>239</sup> More precisely, the plaintiffs—mainly Song Fi, Inc., a music production company, and Rasta Rock Opera, a music group—argued that their small size and YouTube’s market power meant they had no alternative but to accept YouTube’s terms and conditions if they wished to publish their video online.<sup>240</sup> The court rejected these arguments on the grounds that YouTube was not the only platform available to share videos, as users could also publish content on independent websites.<sup>241</sup> In addition, the court held that the advantage offered by YouTube in making its hosting service free meant that the plaintiffs could not claim YouTube’s terms of service—even its absolute discretion in removing content—were unconscionable.<sup>242</sup> Similarly, in *Feldman v. Google*,<sup>243</sup> the liability limitations, including warranty disclaimers, in Google’s AdWords agreement were considered enforceable.<sup>244</sup>

To conclude, users whose content is removed by social media platforms have very little recourse in efforts to contest the removal. This is due either to external legal barriers (e.g., constitutional bars, or safe harbor provisions under the CDA and DMCA), or to contractual bars (e.g., limitation of liability clauses, broad removal discretion in boilerplate ToS). While overcoming the constitutional shield, or bypassing Section 230, require controversial regulatory interventions, this Article turns to suggest to overcome the barriers of contract law by way of conceiving the contractual relationships underlying social media platforms differently.

#### D. THE BLIND SPOT OF CONTRACT LAW

Contract law may offer a complementary framework for negotiating content moderation norms and nudging platforms to attend to the interests of wider circles of stakeholders. As we show below, the contractual claims of users against platforms have been vigorously denied due to the narrow perspective of courts, which overlooks the full complexity of the contractual connections underlying social media. Courts are currently blind to the horizontal relationships of collaboration driven by a common contractual goal between users of social media platforms, and this inhibits their ability to recognize the various interests of the stakeholders involved.

The current approach taken by courts to analyze legal disputes between platforms and users assumes a vertical relationship. Figure 1 illustrates this

---

239. *Id.* at 61–64.

240. *Id.*

241. *Id.* at 62–63.

242. *Id.* at 64.

243. *See generally* *Feldman v. Google, Inc.*, 513 F. Supp. 2d 229 (E.D. Pa. 2007) (discussing the enforceability of an AdWords Agreement).

244. *Id.* at 242–43. In *Grace v. eBay Inc.*, 16 Cal. Rptr. 3d 192, 195 (Ct. App. 2004), *appeal dismissed*, 101 P.3d 509 (Cal. 2004), an intermediate appellate court in California upheld the enforceability of a release in eBay’s User Agreement, holding that the contractual language precluded claims for liability.

pattern whereby courts focus on the contractual relationship between A, the principal user (e.g., a cultural entrepreneur or a political activist) whose content has been removed, such as Murphy, Lewis, or Mishiyevev, and the platform removing the content. But they fail to spot the additional dimension of the contract connecting different users (A–B).

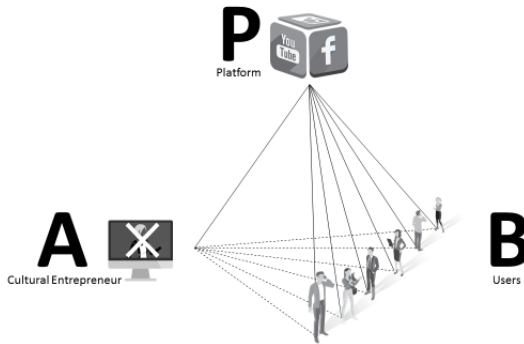


Figure 1: The Current Approach to Platform–User Disputes

#### IV. CONTRACTUAL NETWORK THEORY

Courts have so far taken a formalistic approach to platform contracts, overlooking both the economic contribution of cultural entrepreneurs, and the network relationships among stakeholders in the platform economy which are furthering their mutual goal.<sup>245</sup> Consequently, rather than reflecting the real expectations of the stakeholders involved, contract law is currently invoked by the platforms as a shield against lawsuits brought by users seeking redress against harms caused as a result of content moderation. The bilateral perspective of contract law thus fails to remedy the asymmetry of power between platforms and their users. The doctrinal interpretation of platforms' contracts should be informed by the interconnected expectations of a wider range of stakeholders to fully account for the varied implications of content removals. For contract law to provide successfully a supplementary instrument to reduce private harms generated by arbitrary content removals, the following Parts propose to apply the contractual network theory to social media.

Accordingly, in the discussion below, we first present the contractual form of contractual networks, which essentially considers multiple horizontal complexities, going beyond binary contract classifications. Then, to set the groundwork for conceiving social media platforms as contractual networks, we turn to discuss the legal implications of incorporating a contractual network perspective in the legal analysis of contracts.

245. See *infra* notes 310–26 and accompanying text.

## A. BEYOND BINARY CONTRACT CLASSIFICATION

Under current legal doctrine, platforms have little incentive to mitigate the harms generated by content moderation, thus failing to consider the true nature of user-platform transactions and consequently failing to facilitate the performance of efficient networks.<sup>246</sup> This calls for rethinking the legal analysis of the contracts between platforms and users. The blind spot of contract law conceals the true nature of the cultural entrepreneurship phenomenon, which relies on a complex set of coordinated economic activities by various stakeholders. We therefore propose to fix this blind spot by applying an analytical framework based on contractual network theory to analyze the relationships between the different stakeholders in social media.

The contractual network is situated between bilateral contracts and organizational governance.<sup>247</sup> Coordinated economic activities can be generally classified into two types of categories: market exchange through independent transactions (governed by contract law) and management by firms (governed by corporate law). Contracts facilitate market transactions by enabling individuals to freely undertake mutual obligations in a legally binding manner, conferring remedies upon the breaching party. As Ronald Coase showed in his theory of the firm, when the transaction costs of coordinating economic activity through market exchanges are high, the parties will opt to integrate that activity in firms, enabling coordination by command and control.<sup>248</sup>

However, this binary classification between market exchange, on the one hand, and corporate management of economic activity, on the other, fails to fully capture the variety and complexity of human interactions and coordination. While bilateral contracts may neatly describe the exchange of products and services, they do not adequately apply to collaborative activity for a common purpose.

In recent years, a growing body of literature has identified patterns of behavior that are not neatly captured by standard contract law. Gunther Teubner introduced the legal concept of contractual networks to describe a novel market phenomenon—a business network which reflects neither a formal corporate structure nor a series of independent transactions.<sup>249</sup> A

---

246. See Schwartz & Scott, *supra* note 22, at 326 (arguing that the goal of courts in this context should be “to facilitate the founding and performance of efficient networks”).

247. According to Collins, the network features include multi-party informal arrangements between separate legal entities, involving intensive co-operation between interdependent parties, where the success of each independent actor in the long term “depends on and will be maximised by the success of the production operation as a whole.” Hugh Collins, *Introduction to Networks as Connected Contracts*, in NETWORKS AS CONNECTED CONTRACTS 1, 11 (Gunther Teubner & Michelle Everson eds., Michelle Everson trans., 2011).

248. R. H. COASE, THE FIRM, THE MARKET, AND THE LAW 44 (1988) (“[A] firm will tend to expand until the costs of organizing an extra transaction within the firm become equal to the costs of carrying out the same transaction by means of an exchange on the open market.”) (quoting R.H. Coase, *The Nature of the Firm*, 4 *ECONOMICA* 386, 395 (1937)).

249. TEUBNER, *supra* note 22, at 129–30.



contractual network, he argued, consists of a pattern of interrelated contracts among independent agents, which enables coordination without vertical integration into a single firm.<sup>250</sup>

The core features of the contractual network are cooperation or coordination (typically in pursuit of a shared goal), and interdependency.<sup>251</sup> As such, whenever contracts combine exchange with ongoing informal coordination, we move outside contractual bilateralism and enter a framework of increasing complexity. Collaboration and interdependency entail complexity because contractual networks need to rely on some form of governance to organize the functioning of the network. Such governance is based on norms complementing the contractual relationships shared by participants in the network.<sup>252</sup> Note that coordinating a network also involves administrative costs. Contractual networks thus occur when the benefits of coordination, which would not be possible outside the network, outweigh such costs.<sup>253</sup>

Contractual networks are often interdependent (contractual) efforts of collaboration in the long term, driven by a shared objective.<sup>254</sup> They are not structured for small activities or short-term projects. Indeed, they are usually characterized by stability and a long-term relationship. This requires the exercise of monitoring and enforcement activities as well as the definition of a general framework defining the rights and obligations of parties within the

250. Collins, *supra* note 247, at 1. The term ‘contractual network’ as defined by Teubner refers to economic co-operation between businesses, specifically “multilateral long-term business relationships constructed, at least in part, through contracts, though without an overarching formal business association that binds the parties together.” *Id.* at 13. This definition is, indeed, limited to business relationships, yet as applied in this Article it may also extend to economic and even social activities by individuals and small businesses. As revealed during the advertising boycott of Facebook, small businesses are responsible for a big chunk of Facebook’s revenues, and they are also more dependent on the social media platform. Brian Fung, *The Hard Truth About the Facebook Ad Boycott: Nothing Matters but Zuckerberg*, CNN BUS. (June 26, 2020), <https://edition.cnn.com/2020/06/26/tech/facebook-boycott/index.html> [<https://perma.cc/525S-SMFM>].

251. Teubner looks at networks as an instance of “contractual business cooperation.” TEUBNER, *supra* note 22, at 172. It is the common goal which connects all the contracts in the network while maintaining the autonomy of each party. Indeed, even if it is possible to define a common goal pursued by the contractual network, a network contract does not lead to a new organizational structure. Nor does the common goal become the exclusive goal of the networks’ members. Indeed, while a network exists for the purpose of mitigating complexity and achieving a shared purpose in the long run, each member maintains its interests. This is why contractual networks are characterized by mutual obligations among the members driven by the principle of loyalty, precluding an outcome where the personal interests of members would lead to the destruction of the network’s goal.

252. Collins, *supra* note 247, at 10–11.

253. Besides, the costs of the network would be compensated by the falling costs of organization and monitoring in the long term.

254. See generally Stefan Grundmann, Fabrizio Cafaggi & Giuseppe Vettori, *The Contractual Basis of Long-Term Organization – The Overall Architecture*, in THE ORGANIZATIONAL CONTRACT: FROM EXCHANGE TO LONG-TERM NETWORK COOPERATION IN EUROPEAN CONTRACT LAW 3 (Stefan Grundmann, Fabrizio Cafaggi & Giuseppe Vettori eds., 2013) (introducing organizational contracts in European law).

network.<sup>255</sup> In a more analytical way, Cafaggi has clarified that a network is characterized by: “(1) a strong collective interest to pursue (2) a common objective, and (3) a high level of interdependence among the contracts and the activities performed through contracts.”<sup>256</sup> These characteristics make the network a new form of relationship defined by rules governing and connecting a set of bilateral contracts. According to Amstutz, “conflicts in a contractual network come under the contract the rules of which, in the specific case, ensure the functionality of the network as such.”<sup>257</sup> Hence, in a network contract, it is necessary to pay more attention to the constellation (i.e., the network) rather than the single star (i.e., the bilateral contract).<sup>258</sup>

Contractual networks are therefore hybrid forms of organizations consisting of multilateral contracts between markets and hierarchies, exchange and organization, and involving both contract and corporate law.<sup>259</sup> According to Cafaggi, “[n]etworks differ from market contract[s] because the participants are not impersonal agents, but well[-]identified players chosen on the basis of resource complementarities.”<sup>260</sup> Additionally, “[t]hey permit resource bundling that markets are unable to achieve. They differ from hierarchies because enterprises are autonomous and legally independent even if they may be economically dependent.”<sup>261</sup> In other words, networks can be conceptualized as a hybrid form characterized by the sum of relational contracts based on the collaboration of members (e.g., association) but still formally organized by a series of bilateral contracts.<sup>262</sup>

“Contractual network” is a fuzzy term.<sup>263</sup> Networks entail a combination of economic, social, and legal dynamics characterized by complexity. The network demands diverse methods for organizing the social relationships between the market and different component hierarchies, yet the concept of

255. See Robert C. Feenstra, *Integration of Trade and Disintegration of Production in the Global Economy*, 12 J. ECON. PERSPS. 31, 47 (1998).

256. Fabrizio Cafaggi, *Contractual Networks and Contract Theory: A Research Agenda for European Contract Law*, in CONTRACTUAL NETWORKS, INTER-FIRM COOPERATION AND ECONOMIC GROWTH 66, 74 (Fabrizio Cafaggi ed., 2011).

257. Marc Amstutz, *The Constitution of Contractual Networks*, in NETWORKS: LEGAL ISSUES OF MULTILATERAL CO-OPERATION 309, 342 (Marc Amstutz & Gunther Teubner eds., 2009) (emphasis omitted).

258. See generally *id.* (arguing that the rise of ‘networks’ in corporations poses risks towards networks and third parties and warrants re-examining existing contract and corporation laws to attempt to minimize such risks).

259. Fabrizio Cafaggi, *Contractual Networks and The Small Business Act: Towards European Principles?*, 4 EUR. REV. CONT. L. 493, 495 (2008).

260. Cafaggi, *supra* note 259, at 493.

261. *Id.* at 496.

262. See Gunther Teubner, *In the Blind Spot: The Hybridization of Contract*, 8 THEORETICAL INQUIRIES L. 51, 53-55 (2007).

263. A network is not simply a framework contract. As observed by Schwartz and Scott, network members frequently define a framework agreement to allow other parties to join rather than signing bilateral contracts. See Schwartz & Scott, *supra* note 22, at 325-26.

a *network* does not entail hard boundaries.<sup>264</sup>

Sometimes it refers to hybrid contracts, such as franchises, which constitute networks of independent organizations characterized by collaboration and driven by a common goal.<sup>265</sup> In other cases, it may reflect a stage in the evolution of organizational design, as explained by Cafaggi: “Often enterprises start with a contractual network that is perceived as a lighter form of commitment, but which subsequently evolves into an organizational network.”<sup>266</sup> In such cases, formal bilateral contracts which initially met the needs of the parties may fail to reflect the contractual relationships which characterize an evolving network. Yet as long as the parties do not opt to create an independent legal entity to manage their legal responsibilities, such coordination does not constitute a new legal person (a firm), but continues legally to be based on a collection of bilateral contracts.<sup>267</sup>

The study of contractual networks is of great importance to legal analysis. In particular, it may enable courts to overcome a blind spot created by the rigidity of contract law, and align the legal analysis with the economic reality.<sup>268</sup> While contract law focuses on a bilateral understanding of the contractual relationship, the coordinated activity of a network often functions outside a formal contract. Contractual network theory can thus help courts analyze the informal interactions which shape the expectations of the parties involved.

Amstutz observed that:

[a]nyone tackling contractual networks will quickly find that this phenomenon cannot be grasped using traditional doctrine. The deeper reason for this lies in an emergent phenomenon: contractual networks allow new orders of expectations to arise from bilateral

264. See generally Claude Ménard, *The Economics of Hybrid Organizations*, 160 J. INSTITUTIONAL & THEORETICAL ECON. 345 (2004) (analyzing hybrid organizations’ attributes and examining their growing nature and role in a market economy).

265. See TEUBNER, *supra* note 22, at 14–15. Each franchisee is bound by contract to the franchisor, with no direct contractual relationship among the franchisees. Yet the franchisees are also bound by a common goal, in that any franchisee can damage the interests (e.g., the reputation) of the others. This will be discussed further below.

266. Fabrizio Cafaggi, *Introduction to CONTRACTUAL NETWORKS, INTER-FIRM COOPERATION AND ECONOMIC GROWTH* 1, 1 (Fabrizio Cafaggi ed., 2011).

267. Collins, *supra* note 247, at 10.

268. Likewise, complexity theory, as applied in legal analysis seeks to identify common patterns in complex systems, which are often linked to interactions between the particular components that comprise those systems. See MELANIE MITCHELL, *COMPLEXITY: A GUIDED TOUR* 12–13 (2009). Identifying such patterns could inform legal analysis by refining the underlying assumptions made by law and bringing them closer to reality. As explained by Smith, “[i]f the set of interactions of relevance to the law has some structure (it is organized complexity), this can be reflected in the law in order to optimize information costs.” Henry E. Smith, *Complexity and the Cathedral: Making Law and Economics More Calabresian*, 48 EUR. J.L. & ECON. 43, 51 (2019) (applying complexity systems theory to challenge the Calabresi and Melamed (C&M) framework to legal entitlements and liability rules).

contracts, linking several, sometimes many, actors who selectively interact with each other (as, for instance, in franchising or in inter-bank payment systems). The traditional law of contract is blind to these new expectations.<sup>269</sup>

Networks reflect informal coordination between several sets of participating parties, which are therefore often left under the radar of contract law.<sup>270</sup> Indeed, the rigidity of contract law based on autonomy and the privity of the contract could clash with the open and dynamic framework of network relationships. According to Mitchell, the familiar dynamic of the “contract-down” approach, which would leave contract law wanted in many areas, should be eliminated in favor of a “networks-up” approach in contract law.<sup>271</sup> Alternatively, rather than putting attention toward the contractual constitution of some varieties of networks, we could put our attention more broadly toward contextual factors that prompt the creation of contractual networks, as opposed to other organizational forms.<sup>272</sup>

#### B. THE LEGAL IMPLICATIONS OF CONTRACTUAL NETWORK THEORY

Given the unique attributes of contractual networks which are reflected in the collaborative efforts of the network members to pursue a common goal, what could be the legal implications of incorporating a contractual network perspective in a legal analysis of contracts?

The literature on contractual networks proposes several ways by which a network perspective could inform legal analysis. Here we focus on two sets of legal implications which emerge from a contractual network analysis.<sup>273</sup> The first relates to new legal obligations arising from the common goal of the network. The common goal shared by parties to the network is the beacon of a relationship based on trust and loyalty. For instance, Collins argues that once the interconnection between the bilateral contracts is established, it may create a new mutual duty of loyalty among the network participants to the

---

269. Amstutz, *supra* note 257, at 309.

270. The network does not need contract law per se. But some rules would lead to defining formal rules through contracts, creating a sort of hybrid contracting between formalities and informalities. This flexibility leads networks to constitute forms of private ordering, benefiting from the lower costs of collaboration rather than relying on bilateral contracts. Networks are thus often the most efficient private regulators. John Gava & Janey Greene, *Do We Need a Hybrid Law of Contract? Why Hugh Collins is Wrong and Why it Matters*, 63 CAMBRIDGE L.J. 605, 626 (2004).

271. Catherine Mitchell, *Network Commercial Relationships: What Role for Contract Law?*, in CONTRACT AND REGULATION: A HANDBOOK ON NEW METHODS OF LAW MAKING IN PRIVATE LAW 198, 232 (Roger Brownsword, Rob A.J. Van Gestel & Hans-W Micklitz, eds., 2017).

272. *Id.*

273. Contractual networks have both an internal and an external dimension, where the former concerns relationships between the network members, and the latter concerns relationships between network members and third parties. The implications of contractual networks on third parties are beyond the scope of this Article.

shared goal of the network as a whole.<sup>274</sup> Such a common goal may affect contractual interpretations, taking into account not only the self-interested behavior of each of the parties to bilateral contracts, but also the duty of loyalty of network participants to their mutual purpose. Consequently, if a party to a network behaves in a way which may undermine the common purpose of the network contract, this may amount to a breach of an implied duty of loyalty or good faith, even if it does not constitute a breach of any explicit provision of the relevant bilateral contracts.<sup>275</sup> Note that this principle applies regardless of the form taken by the network—i.e., whether it is a decentralized network or a hub-and-spokes model, where a central anchor (the hub) links multiple smaller components (spokes). In the latter case, recognition of a network dimension among the participating parties (spokes), even if implicit, may affect the legal analysis of formal contracts between the hub and the spokes.<sup>276</sup>

The second set of legal implications concerns the horizontal relationship between parties to a network.<sup>277</sup> Rather than limiting the legal analysis only to express terms in isolated bilateral contracts, this approach may disclose legally relevant interactions between different entities—interactions which may both shape and be shaped by the relationships directly involved in the contract. Scott and Schwartz, for instance, invoked this theory “to facilitate the resolution of disputes where the contract is unclear about the right of third parties to recover for contract breach.”<sup>278</sup> Where the contracting parties’ intent to allow third party recovery is unclear, Scott and Schwartz suggested how to identify and implement their goals in the network context.<sup>279</sup> If benefitting a third party is aligned with these goals, that is it is “ex ante profitable for the network contracting members to serve the potential beneficiary class to which the plaintiff belongs,”<sup>280</sup> then courts should allow that third party to recover against the formal contracting parties.

The network perspective thus may offer a framework for analyzing both interests and expectations that ought to be protected and implicit legal duties of network members. Such rights and duties may prove to be relevant when the vertical contractual relationship between the hub of a network and each of the spokes is insufficient to protect the interests of all parties. Consider, for instance, harm that is caused by the behavior of one party to a network, say a franchisee whose actions tarnish the reputation of the entire franchise. Now imagine that the franchisor is reluctant to bring a lawsuit against that

---

274. Collins, *supra* note 247, at 6.

275. *Id.* at 15.

276. See *supra* notes 268–72 and accompanying text.

277. The perspective of a contractual network may enrich legal analysis by exposing the interconnections between coordinated parties. Collins, *supra* note 247, at 42.

278. See Schwartz & Scott, *supra* note 22, at 329.

279. *Id.*

280. *Id.* at 359.

particular franchisee, for whatever reason (e.g., conflicting interests). A contractual network analysis may give rise to informal obligations among the franchisees arising from their shared interest in promoting the franchise. The contractual network approach may also help claimants to establish that a contract between parties to a network is worthy of legal protection, even in the absence of privity. Based on such an implicit contract, claimants might be able to raise legal claims such as tortious interference with contractual relations, which assume the existence of a valid contract that the defendant intentionally acted to disrupt.

## V. RECONFIGURING PLATFORM CONTRACTS

A contractual network perspective can inform legal analysis regarding the internal relationships within a network in various respects. It may help identify implicit obligations of network participants which arise from the common goals of the network; and it may confirm implicit legal duties between the parties to the network.<sup>281</sup> The following discussion demonstrates how perceiving social media platforms as contractual networks could shed light on the blind spot of contract law, offering courts a novel interpretive framework to recognize the communal expectations of platforms' users. The contractual network framing would further allow courts to consider the extent to which specific removals or suspension decisions meet the contractual expectations of the networks' members and advance their common goals.

As we demonstrate below, application of the contractual network perspective to social media platforms makes room for users to contest allegedly illegitimate restrictions on the sharing of content (e.g., removals, demonetization, account terminations) based on obligations which arise from the nature of platforms as complex networks operating to pursue a common goal. Allowing users to challenge content removals by platforms could create an important bottom-up pressure on platforms to become more accountable for the way they implement their content moderation policies. It may also enable users to protect their reasonable expectations arising from the contractual network. This is specifically important in the context of discretionary content removals by social media platforms.

### A. PLATFORM CONTRACTS: BEYOND BILATERAL CONTRACTS

There are several reasons why the relationships between users and platforms cannot be neatly classified into the traditional categories of corporate governance or bilateral contracts. One set of reasons arises from the nature of social production, which enables the coordination of online activities without a corporate structure. The second set of reasons relates to the

---

<sup>281</sup> As noted above, a contractual network perspective may also offer a framework for analyzing the impact of the network on third parties. This is beyond the scope of the present Article. *See supra* Section IV.B.

multiple roles of users in social media platforms and the interdependency among users in generating value. We discuss these aspects of the user—platform relationship in more detail below.

As we demonstrated above,<sup>282</sup> social media platforms do not comply with the corporate model of the entertainment industry, which organizes the production of content through centralized planning. That is, platforms do not generally initiate, produce, or market original content conceived by corporate employees and brought to fruition by corporate managers.<sup>283</sup> Users of social media are not employees of the platform, who follow management instructions on what content to generate, when and how. They are autonomous agents who can choose whether and when to post content or actively engage with other users.<sup>284</sup> To some degree, dissemination of content posted by users on social media platforms is centrally managed by platforms, through a mixture of governance mechanisms. Platforms can shape how content and data are displayed, organized, and shared among users by designing the digital interfaces through which exchange occurs.<sup>285</sup> These affordances are obviously driven by economic interests. “For example, ‘like,’ ‘share,’ and ‘retweet’ not only provide a means for users to express themselves but also facilitate ranking, product recommendations, and data analytics.”<sup>286</sup> Platforms apply their corporate organizational structure to harness user—generated content, and to match content with its potential audience, using machine learning algorithms that collect and analyze data on users’ preferences.<sup>287</sup> Yet coordination is also facilitated in the absence of a strict institutional structure.<sup>288</sup> Communities of fans, followers, and so on are driven by the efforts of cultural entrepreneurs, and by the connectivity among users themselves. As explained by Cunningham and Craig: “The same network effects that accord platforms enormous power also enable better connected, networked possibilities for horizontal, grass roots peer-to-peer connectivity and communicative and organizational capability.”<sup>289</sup> Thus, the economic

282. See *supra* notes 123–27 and accompanying text.

283. Benkler describes the shift from industrial production of content to social production, which allows users to coordinate and collaborate outside the organizational structure of firms and the state. See BENKLER, *supra* note 122, at 3–5.

284. But see Mark Andrejevic, *Estranged Free Labor*, in DIGITAL LABOR: THE INTERNET AS PLAYGROUND AND FACTORY 149, 149–50 (Trebor Scholz ed., 2013).

285. Consider, as an example, the use of Trends. See, e.g., *Twitter Trends FAQ*, TWITTER, <https://help.twitter.com/en/using-twitter/twitter-trending-faqs> [<https://perma.cc/3FZQ-U3ZC>].

286. Jean-Christophe Plantin, Carl Lagoze, Paul N. Edwards & Christian Sandvig, *Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook*, 20 NEW MEDIA & SOC’Y 293, 297 (2018).

287. Elkin-Koren & Perel, *supra* note 34, at 876.

288. Platforms, as digital networks, may facilitate coordination and collaboration among users without resorting to any formal organizational structure. See generally SHIRKY, *supra* note 122 (arguing that social production occurs without any organizational structure due to the low communications cost of coordinating online).

289. Cunningham & Craig, *supra* note 126, at 3.

value in social media transactions is generated through a coordinated effort of interdependent users.

Could the platforms' relationship with users be adequately described as simply bilateral transactions with consumers? Arguably, platforms such as Facebook or YouTube enter into an independent market transaction with each user. Indeed, platforms supply each user with some services, such as hosting content and enabling them to communicate with other users of the network. YouTube's video-sharing service, for instance, allows users to upload or view videos for free, and enables certain users to monetize their videos through YouTube's Partner Program.<sup>290</sup>

To use these services, users must agree to the platform's terms of use, namely, a bilateral vertical agreement that permits them to make use of the services, subject to the platform's ToS.<sup>291</sup> Nevertheless, the relationship between users and social media platforms is not entirely a bilateral consumer transaction. Unlike standard market transactions where the parties exchange assets or services for an agreed-upon price, users in platforms pay through engagement with other participants on the platform. More precisely, users' engagement reflects two types of currency. The first of these is content. By generating and sharing content—tweeting a news story, sharing pictures, posting reflections on Facebook, or uploading music to YouTube—users contribute to the platform by attracting other users to join it.

The second type of currency, by which users “pay” for services received from social media, is their personal data (or we might say, a combination of their personal data and their attention, on the assumption that users pay at least some attention to the ads targeted toward them based on their data).<sup>292</sup> Platforms' advertising revenues and overall revenues from data collection depend on the three V's of data: volume, velocity, and variety.<sup>293</sup> To enhance the amount, range, and freshness of data collected on each user, platforms seek to increase the amount of time and attention users devote to the site.<sup>294</sup> For this purpose they require content uploaded by other users. In particular, a diverse range of content enables platforms to match users with content that best meets their preferences. Users' ability to interact with other users thus constitutes the critical fuel that generates more data and enhances the platforms' revenues.

It is apparent that the platform economy rests upon interdependency both between users and platforms, and among users themselves. To generate

---

290. *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1156 (N.D. Cal. 2020), *aff'd*, 857 F. App'x 907 (9th Cir. 2021).

291. *See supra* Section III.C.

292. *See supra* Section II.C.1.

293. JULES J. BERMAN, *PRINCIPLES OF BIG DATA: PREPARING, SHARING, AND ANALYZING COMPLEX INFORMATION* 242 (Andrea Dierna & Heather Scherer eds., 2013).

294. JACK M. BALKIN, *FIXING SOCIAL MEDIA'S GRAND BARGAIN* 2 (2018), [https://www.hoover.org/sites/default/files/research/docs/balkin\\_webready.pdf](https://www.hoover.org/sites/default/files/research/docs/balkin_webready.pdf) [<https://perma.cc/KL87-8QUE>].



profits from advertisers, platforms rely on users to engage with content in the digital space. At the level of users themselves, cultural entrepreneurs depend on the engagement of their peers to get the most out of the platform. For example, YouTubers who are accepted into the platform's Partner Program receive a share of YouTube's advertising revenues from ads associated with their content, and so their income depends on the number of users viewing their channels.<sup>295</sup> Activists need other users to read their political message or make a contribution to a shared social cause. Whether users choose to use social media platforms for social purposes, for business interests, or for political causes, they are strongly dependent on their peers' ability to communicate via the platform.

Accordingly, users of social media platforms play multiple roles. Besides being consumers of services supplied by the platform (under a vertical contract), users themselves are also providers of content and supply added value to platforms, in ways that affect the myriad interests of other users of the platform and shape their (horizontal) expectations. Since the value of usage to each user, and subsequently also to the platform, is generated by engagement among users, users should be conceived as "partners" with the platform in a contractual network that aims at a collaborative goal. Indeed, the YouTube Partner Program effectively acknowledges this point in its very name; and several other platforms, such as Medium, Facebook, Snapchat and Twitch, have followed that lead.<sup>296</sup> Yet, notwithstanding the "partnership" framing, courts currently define the rights and obligations between platforms and users solely according to a bilateral boilerplate contract where the platform reserves unlimited removal power.<sup>297</sup>

Bilateral contracts fail to capture the full extent of the varied mutual obligations, commitments and expectations which underline the critical collaboration between the partners operating within the network. The bilateral contracts into which users enter when accessing social media platforms are only the formal contractual appearance which obscures the underlying complexity of multilayered network relationships.

#### B. SOCIAL MEDIA CONTRACTS AS NETWORK CONTRACTS

Social media platforms facilitate layers of systematic interactions and mutual reliance between different stakeholders, yet the way they are treated in law is still based on the logic of ordinary consumer transactions. Interdependency between users, the collaborative nature of social media interactions, and the complexity of the economic reality behind social media interactions call for a different analytical framework for social media. Network

---

295. *YouTube Partner Program Overview & Eligibility*, YOUTUBE HELP (Aug. 2021), <https://support.google.com/youtube/answer/72851?hl=en> [<https://perma.cc/XZT5-FWG8>].

296. See Cunningham & Craig, *supra* note 126, at 7; Caplan & Gillespie, *supra* note 40, at 2.

297. See *supra* Section III.C.

contract theory could provide a theoretical basis for developing such a framework.<sup>298</sup>

A social media platform can be described as a hub-and-spoke network, where the platform—the hub—provides the technological, business, and legal infrastructure for exchange among users, while users—the spokes—offer complementary content and services. The resulting web of interdependent economic actors constitutes a contractual network.

A contractual network approach to social media may be used to uncover the relevance of content moderation not only as an instrument of monitoring and control, but also as a means to advance the network's common interests. Unlike other forms of cooperation which would lead to the creation of a new legal entity, one that would merge the economic interests of the parties involved and bear the risks of their actions, in the case of social media, both the platforms and their users remain autonomous agents.

Collaboration is just one of the primary characteristics of the contractual network and is a shared objective among users.<sup>299</sup> The parties on social media networks—whether cultural entrepreneurs seeking to build a community of fans, social activists who use the platform to organize grassroots protests, or ordinary users seeking to acquire information and connect with others—all share a common goal with platforms, namely to build, preserve and expand a wide community of users. In the case of social media, such a common goal could be the creation of connected communities able to share expressions and opinions on a global scale,<sup>300</sup> to promote public conversation,<sup>301</sup> or community trust.<sup>302</sup> A common goal could also be more specific, shared only

298. See generally THE ORGANIZATIONAL CONTRACT: FROM EXCHANGE TO LONG-TERM NETWORK COOPERATION IN EUROPEAN CONTRACT LAW (Stefan Grundmann, Fabrizio Cafaggi & Guiseppe Vettori eds., 2013) (discussing organizational contracts); TEUBNER, *supra* note 22 (arguing for interpretation of business networks as connected contracts); IMPLICIT DIMENSIONS OF CONTRACT: DISCRETE, RELATIONAL AND NETWORK CONTRACTS (David Campbell, Hugh Collins & John Wightman eds., 2003) (discussing implicit contracts created through transactional networks); Walter W. Powell, *Neither Market nor Hierarchy: Network Forms of Organization*, 12 RSCH. ORGANIZATIONAL BEHAV. 295 (1990) (discussing examples of network forms of organization).

299. Collins, *supra* note 247, at 11.

300. In the case of Facebook, the community guidelines states:

The goal of our Community Standards is to create a place for expression and give people a voice. The Facebook company wants people to be able to talk openly about the issues that matter to them, even if some may disagree or find them objectionable. In some cases, we allow content—which would otherwise go against our Standards—if it's newsworthy and in the public interest.

*Facebook Community Standards*, *supra* note 65.

301. Twitter underlines: "Twitter's purpose is to serve the public conversation. Violence, harassment and other similar types of behavior discourage people from expressing themselves, and ultimately diminish the value of global public conversation. Our rules are to ensure all people can participate in the public conversation freely and safely." *The Twitter Rules*, TWITTER HELP CTR. (2021), <https://help.twitter.com/en/rules-and-policies/twitter-rules> [<https://perma.cc/QQV8-QAAC>].

302. Taking as example YouTube, the Community guidelines underlines that: Our policies

by a subset of users in a Facebook group for instance.

In short, social media platforms are based on mutual, informal collaboration between stakeholders (platforms and different types of users) to provide, share, and consume content in the service of both common and independent goals. This mix of independent and common goals in the network effectively creates new contractual rights and obligations. We now show how this contractual network approach to social media platforms could help establish legal claims by users against unjustified content removals, while facilitating a bottom-up check over platforms' discretionary removal decisions.

### C. NETWORK ANALYSIS OF PLATFORM CONTRACTS

Users' expectations are badly represented under the dominant model by which platforms manage online content. As we demonstrated in Part III, platforms can remove content or terminate accounts without prior notice.<sup>303</sup> They can even amend their terms of use unilaterally without notifying their users in advance.<sup>304</sup> Despite the fact that social media platforms and users heavily rely on cooperation, platforms currently have almost no accountability for the impact of their decisions to remove content or to terminate a user's account, even though such decisions affect the mutual goal of the network's participants.<sup>305</sup>

Relationships between users and platforms tend to be governed by formal contracts, which define the scope of both parties' legal duties and rights.<sup>306</sup> These boilerplate ToS reflect critical power asymmetries and cannot be said to reflect users' preferences over content moderation norms.<sup>307</sup> As noted by Suzor, "the users who care deeply about how content is regulated are not as well-organized or influential on the policies of platforms."<sup>308</sup>

---

aim to make YouTube a safer community while still giving creators the freedom to share a broad range of experiences and perspectives. *Community Guidelines*, *supra* note 65.

303. *Lewis v. YouTube, LLC*, 197 Cal. Rptr. 3d 219, 225–26 (Ct. App. 2015); *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1159–60 (N.D. Cal. 2020), *aff'd*, 857 F. App'x 907 (9th Cir. 2021).

304. Order Denying Special Motion to Strike the Complaint Under California Code of Civil Procedure Section 425.16 and Sustaining Demurrer to Complaint Without Leave to Amend at 16, *Murphy v. Twitter, Inc.*, No. CGC-19-573712 (Cal. Super. Ct. June 12, 2019), <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2968&context=historical> [<https://perma.cc/28BY-FHXH>].

305. See *supra* Section II.B, *supra* note 37 and accompanying text.

306. See *supra* notes 203–07 and accompanying text.

307. See *supra* Section III.C.

308. Suzor, *supra* note 98, at 4.

A network approach to social media platforms can improve the legal analysis of these formal contractual obligations by uncovering the blind spot of contract law and recognizing the horizontal obligations between different users of a platform. Figure 2 illustrates how the contractual network approach enables us to evaluate the varied interactions facilitated by social media platforms. Where Figure 1 showed how courts focus on the vertical contractual obligations between A, the principal user and the platform, Figure 2 directs our attention to the implied obligations between A and other users of the platform (B). A could be a YouTuber who has endeavored to build a community of followers, she could be an influencer who is marketing accessories among her fellow Instagrammers, or an activist who has collaborated with other users to organize livestreams showing abusive treatment toward members of minority groups. Following Cunningham and Craig, we refer to all these users as cultural entrepreneurs.<sup>309</sup>

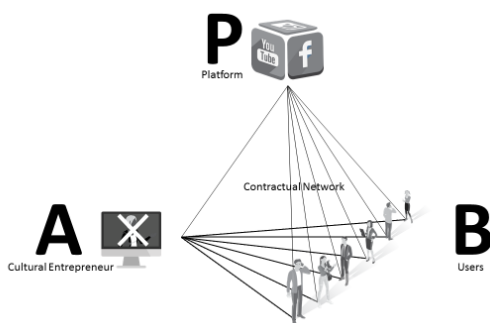


Figure 2: Social Media Platforms as Contractual Networks

Applying the contractual network approach to social media platforms could give rise to new implicit obligations on the part of both the platform and users, which derive from their mutual goal to maximize the production and sharing of content online. It could further sustain implicit horizontal legal duties between users of the platform, and it may even extend the duties of the network towards external, third parties.<sup>310</sup> Since these implicit obligations are horizontal, they could overcome the barriers discussed in Part III, which apply to users' vertical claims against platforms.

In the case of Meghan Murphy, for instance, the court approached the complainant's contractual claim as narrowly contesting Twitter's decision to remove content, and in doing so, affecting Murphy's expectations pertaining

309. CUNNINGHAM & CRAIG, *supra* note 43, at 5–6.

310. Third parties outside the network could benefit or suffer harm from network collaboration. As noted earlier, discussion of these aspects is beyond the scope of this Article.

to the management of her account by Twitter.<sup>311</sup> Obviously, such a narrow approach immediately falls under the ban of Section 230 of the CDA, which prevents vertical allegations against platforms that are not to be considered publishers of content.<sup>312</sup> Similarly, in the cases of Lewis and Mishiyev, the courts emphasized that the platforms' ToS shape their vertical obligations to their users, granting them full discretion to exercise their removal power as they see fit and exempting them from liability for any harms caused by their removal actions.<sup>313</sup>

However, courts have overlooked the horizontal dimension in online content moderation, which concerns various types of contractual relationships between users.<sup>314</sup> These horizontal relationships create expectations among users in relation to how their content will be managed which derive from platforms' role as contracting parties and go beyond their role as publishers of users' content. Importantly, these expectations, which are shaped and facilitated by platforms, are also reflected in the platforms' business model.

For instance, YouTube allows users to upload, view, and share videos for free, in exchange for a non-exclusive license to host users' videos.<sup>315</sup> When users actively engage with the YouTube community, they help to generate the traffic which allows YouTube to sell advertising to those who wish to target YouTube user groups.<sup>316</sup> The value to cultural entrepreneurs, as well as the platform itself, is based on these informal agreements among users.

Some users, such as Lewis, share their videos for the sole purpose of receiving acclaim from the YouTube community and making new friends and connections.<sup>317</sup> Other users, such as Mishiyev, also contract with the platform

311. Order Denying Special Motion to Strike the Complaint Under California Code of Civil Procedure Section 425.16 and Sustaining Demurrer to Complaint Without Leave to Amend at 5, *Murphy v. Twitter, Inc.*, No. CGC-19-573712 (Cal. Super. Ct. June 12, 2019), <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2968&context=historical> [<https://perma.cc/28BY-FHXH>]. In her first cause of action for breach of contract, she alleges that Twitter's User Agreement, which includes its Terms of Service, rules, and associated policies, constitutes a binding contract with each of its user, including Murphy, and that Twitter breached that contract by failing to provide Murphy with 30 days advance notice of the changes to its Hateful Conduct Policy, by retroactively applying the amended policy to Murphy, and by permanently suspending her account although she did not violate the Terms of Service, rules or policies. *Id.*

312. 47 U.S.C. § 230(c)(1) (2018); *Klayman v. Zuckerberg*, 910 F. Supp. 2d 314, 318 (D.D.C. 2012).

313. *Lewis v. YouTube*, 197 Cal. Rptr. 3d 219, 224–25 (Ct. App. 2015); *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1159 (N.D. Cal. 2020), *aff'd*, 857 F. App'x 907 (9th Cir. 2021).

314. See generally 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 (2021) (demonstrating that Internet services have won essentially all of the 62 lawsuits brought by terminated/removed users as of March 2021).

315. *Terms of Service*, YOUTUBE (Mar. 17, 2021), <https://www.youtube.com/static?template=terms> [<https://perma.cc/YA6X-M2J6>].

316. *Lewis*, 197 Cal. Rptr. 3d at 222.

317. *Id.* at 221.

for commercial purposes, and extract earnings from their content through YouTube's Partner Program.<sup>318</sup> Accordingly, the contractual expectations of YouTubers such as Lewis and Mishiyev—user A in Figure 2—extend beyond the vertical relationship with the platform to the horizontal connections with other users (B). The more B users click on A's videos, the greater the gains (both reputational and, potentially, financial) that ultimately accrue to A, as well as directly to YouTube. Hence, when contracting with YouTube, A also relies on "clicks" from B, which are facilitated by YouTube; and if YouTube removes content uploaded by A, it also interferes with the contractual relationship between A and B. This interference goes beyond the vertical relationship between A and the platform and should not be necessarily restricted by the legal barriers described previously.

Applying such a comprehensive network approach to YouTube's contracts may change the basic contractual analysis of the courts and lead to different outcomes. Indeed, in her complaint against YouTube for deleting her channel, Lewis claimed that she had an "inferred right" based on the ToS,<sup>319</sup> which, as stated in her complaint, "represent a binding written agreement between YouTube and its users."<sup>320</sup> According to Lewis's allegations, YouTube's ToS "incorporates an implied covenant of good faith and fair dealing such that neither party will do anything to destroy or injure the right of the other party to receive the benefits of the contract."<sup>321</sup> YouTube, according to Lewis, breached that covenant of good faith and fair dealing when it allegedly:

[D]eprived Lewis of her reasonable expectations under the Terms of Service. Those reasonable expectations included the expectation that if she faithfully complied with the Terms of Service that her channel would be maintained, that the channel would continue to include her videos and would continue to reflect the acclaim of fellow YouTube users in terms of number of views and comments.<sup>322</sup>

Put simply, in choosing to spend her time and money on creating appealing videos, Lewis relied on the implied horizontal relationship between herself, as a creator, and her followers. Under a contractual network approach, YouTube's limitation of liability<sup>323</sup> relates to the vertical relationship between YouTube and Lewis, and does not extend to horizontal contractual relationships between users. Accordingly, to the extent that Lewis met all conditions required under YouTube's ToS, Lewis's contractual claim should not be automatically barred under YouTube's ToS.

---

318. *Mishiyev*, 444 F. Supp. 3d at 1156.

319. Brief for the Appellant at 1, *Lewis*, 197 Cal. Rptr. 3d 219 (Ct. App. Oct. 14, 2014) (No. Ho41127).

320. *Id.* at 3.

321. *Lewis*, 197 Cal. Rptr. 3d at 222.

322. Brief for the Appellant at 6–7, *Lewis*, 197 Cal. Rptr. 3d 219 (No. Ho41127).

323. *Id.* at 28.

Similarly, in *Mishiyev*, the court addressed the complaint as an “action . . . about YouTube’s decision to terminate plaintiff’s account and disable the channels associated with it.”<sup>324</sup> However, by terminating Mishiyev’s account, YouTube also interfered with his horizontal expectations, violating the covenant of good faith on which Mishiyev relied. Indeed, Mishiyev’s participation in YouTube’s Partner Program was “his primary source of income” through which he generated “over \$300,000 between 2012 and 2018.”<sup>325</sup> YouTube’s actions with respect to Mishiyev’s account led to the loss of “new subscribers, views, future hits, performance bookings, and . . . advertising and sponsorship revenue.”<sup>326</sup> Yet in addressing Mishiyev’s third cause of action, for tortious “interference with contractual relations,” the court concluded that the plaintiff had “fail[ed] to identify any contract between himself and a third party” and that “[t]o the extent the complaint alleges that a contractual relationship arose between plaintiff and his subscribers based on the advertisement revenue the subscribers generated and YouTube shared with plaintiff, this is a legal conclusion we need not accept as true.”<sup>327</sup> As with the Lewis case, a contractual network approach would have precluded the court from automatically rejecting Mishiyev’s contractual claim.

#### D. ADDITIONAL LEGAL IMPLICATIONS

The application of the contractual network approach to social media platforms reveals the hidden contractual complexities that underlie the relationships between platforms and users. It offers an analytical framework for bypassing the rigid barriers to vertical claims by users against platforms who arbitrarily remove their content or terminate their accounts. This approach also provides courts with an interpretative angle highlighting users’ contractual expectations.

In this section we address additional legal implications arising from the contractual network approach to social media platforms. We show how the contractual network approach could open the door to implied duties and obligations that are currently concealed by the blind spot of contract law.

##### 1. Limiting the Discretionary Power of Platforms

A network perspective could give rise to claims by users that a platform’s actions, or failure to take action, compromise the shared goal of the network. This could occur in cases where the platform’s business interests clash with the common interests of the network. For example, a platform’s algorithms might encourage the sharing and visibility of objectionable content (such as

---

324. *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1156 (N.D. Cal. 2020), *aff’d*, 857 F. App’x 907 (9th Cir. 2021).

325. *Id.*

326. *Id.* at 1158.

327. *Id.* at 1158, 1161.

violent or radical content) to promote engagement by certain users, while undermining a goal shared by other stakeholders of the network.<sup>328</sup>

YouTubers (cultural entrepreneurs) could be harmed by such behavior because they rely on the safety of the YouTube brand to increase the number of their followers, and subsequently, their economic benefits.

This would hardly amount to a breach of the formal contract, since the formal language of the contract (the ToS) typically grants full discretion to the platform in making content moderation choices. A network approach, however, would call for interpreting the contract in light of the common network goal, which is the beacon of the contractual relationship in a contractual network. When interpreting the contract to determine whether the platform has committed a contractual breach, courts should thus consider whether actions by the platform (either to remove or not to remove content) undermine the common objectives of the network. Consequently, while platforms and users each preserve their independent agency, the multilateral construction of bilateral contracts may give rise to their mutual obligation to work towards promoting—and refrain from compromising—the shared goal of the network. That is, the platform's business goal of profiting from advertising revenues based on content provided by users should not take precedence over the common aim of the network—namely, to facilitate the sharing of content online.

A duty to deploy content moderation without prejudice to the interests of the network as a whole may thus arise from the network approach. Unlike good faith, such a duty does not purport to give priority to the interests of any particular user over the legitimate self-interests of the platform as reflected in the contractual language. Instead, the network approach gives priority to the implicit shared goal of the overall network.

## 2. Predictability of Content Moderation Rules

As described by numerous scholars, content moderation policies, which are unilaterally defined by platforms, are subject to frequent changes.<sup>329</sup> These changes are often inconspicuous, especially when embedded in algorithmic systems that automatically flag, classify, restrict, or remove content. Changes in what counts as legitimate speech, which are implemented either through demonetization or the blocking/removal of content, have an immediate impact on users' potential reach, with consequences for their revenues and audience. Moreover, platforms typically provide no explanation for changes in their moderation policies which lead to the restriction or removal of content, leading users to conclude that moderation decisions are arbitrary, or even capricious.

---

328. Elkin-Koren & Perel, *supra* note 34, at 879.

329. See generally Roberts, *supra* note 69 (analyzing the moderation of social media content); Caplan & Gillespie, *supra* note 40, at 7; Suzor, *supra* note 98, at 1.



The contractual network approach could establish a duty requiring platforms to avoid capricious alterations of their terms of use and provide users with proper notification in advance of any such prospective alteration. Imposing such a procedural duty on platforms derives from the mutual goal of the network. Indeed, as cultural entrepreneurs, users must be able to predict whether their content will meet with the platforms' terms of use. Otherwise, they might lose their incentive to invest resources in creating and sharing content. This will in turn impede the mutual goal of the stakeholders in the contractual network, potentially reducing not merely the prospective revenues of the platform, but also the value (whether commercial, reputational, or political) to its users.

Accordingly, in the case of *Murphy*, discussed earlier, where the plaintiff's Twitter account was terminated for violation of the platform's revised content moderation policies,<sup>330</sup> *Murphy* should have been allowed to claim that Twitter violated its implied obligation to avoid altering its policies without first notifying her. This proposed framing of *Murphy*'s allegations shifts from challenging Twitter's substantive decision to terminate an account, which is barred under Section 230 of the CDA, to contesting the procedural circumstances of such termination.

Further related to the duty to avoid capricious alteration of their terms of use, platforms should also be required to assure the clarity of their policies. Users must be able to rely on the platform's declared policies so they can endeavor to advance the shared goal of the network. Otherwise, "[t]he lack of clarity around platform policies, procedures and the values that inform them lead users to wildly different interpretations of the user experience on the same site, resulting in confusion in no small part by the platforms' own design."<sup>331</sup> Currently, platforms largely fail to describe in detail how they determine which content complies with their moderation policies. The boundaries of what is and is not legitimate content are obscured through lengthy boilerplate terms of use that leave platforms with unlimited removal discretion. This, as Roberts asserts, is mainly a strategic decision meant "to render platforms as objective in the public imagination."<sup>332</sup> Indeed, users have no way to know precisely who is in charge of moderating content on the platform, or their locations, or the constraints they face.<sup>333</sup> Users whose

---

330. Order Denying Special Motion to Strike the Complaint Under California Code of Civil Procedure Section 425.16 and Sustaining Demurrer to Complaint Without Leave to Amend at 2, *Murphy v. Twitter, Inc.*, No. CGC-19-573712 (Cal. Super. Ct. June 12, 2019), <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2968&context=historical> [<https://perma.cc/28BY-FHXX>].

331. Roberts, *supra* note 69.

332. *Id.*

333. *Id.* In recent years, following pressure from civil society advocates, the news media and the public in general, major platforms—including Facebook—have disclosed some information about their moderation practices.

content was removed from the platform are left to guess why, and this may diminish their ability to contest the removal.

If users were better informed about the platforms' content moderation policies, they could be more alert to misapplications of these policies. This would provide users with a solid basis to scrutinize whether the platforms' decisions have been made in the interest of the network as a whole. Consequently, social media platforms could become more transparent and accountable vis-à-vis the other stakeholders of the network while building their trust and strengthening their reliance, which are key factors in pursuing the common goal of the contractual network.

### 3. Due Process and the Right to an Explanation

Another important implication of applying the contractual network framework to social media contracts relates to procedural due process, and its underlying values of transparency, accuracy, participation, and fairness.<sup>334</sup> Specifically, due process implies that social media platforms should be required to notify users about removal, demonetization, or account termination decisions, explain the reasons underlying such decisions, and allow users to appeal them. Notwithstanding that social media platforms are effectively private parties, subjecting them to procedural due process should be driven by the nature of the relationships they facilitate among the partners of the network. Indeed, a different approach would derive due process obligations of social media platforms from the standards of the rule of law applicable to “digital platforms as ‘architects of public spaces.’”<sup>335</sup> Yet this Article contends that procedural due process should serve as a critical safeguard, assuring platforms apply their policies in a way that promotes the shared goal of the network, rather than their own, narrow commercial interests.<sup>336</sup>

Furthermore, providing users with notice about removals and recognizing their right to appeal will create trust among the network partners and give voice to users—an extremely significant outcome given their inferior position in the contractual relationship with social media platforms. As the case of Mishiyev discussed earlier shows, users may have good reason to be suspicious of platforms' removal incentives. Mishiyev had reasons to believe that

---

334. Martin H. Redish & Lawrence C. Marshall, *Adjudicatory Independence and the Values of Procedural Due Process*, 95 YALE L.J. 455, 478–89 (1986).

335. Suzor, *supra* note 98, at 2 (quoting Tarleton Gillespie, *Governance of and by Platforms*, in THE SAGE HANDBOOK SOCIAL MEDIA 254, 273 (Jean Burgess, Alice Marwick & Thomas Poell eds., 2017)) (“The rule of law framework provides a lens through which to evaluate the legitimacy of online governance and therefore to begin to articulate what limits societies should impose on the autonomy of platforms.”).

336. For instance, in a recent complaint filed against Adobe by Green Savannah, a reseller of genuine copies of Adobe software, the plaintiff complained that his eBay account was illegitimately terminated due to Adobe filing bogus DMCA notices with eBay in order to prevent the plaintiff from reselling its software. *See Green Savannah LLC v. Adobe Inc.*, No. 3:20-CV-05568 (N.D. Cal Aug. 11, 2020).

YouTube terminated his accounts in retaliation for his planned lawsuit, rather than copyright infringement, as YouTube asserted.<sup>337</sup> This could mean that Mishiyev could not trust YouTube to make impartial removal decisions, not tainted by YouTube's private commercial interests. YouTube reinforced Mishiyev's mistrust by ignoring many of his counter-notices, which were filed in response to the numerous copyright takedowns submitted in relation to his videos. Nevertheless, the court overlooked this issue, stressing that "[e]ven taking the retaliation allegations as true, however, the complaint fails to overcome YouTube's express right to terminate plaintiff's account for repeat copyright infringement."<sup>338</sup> The court focused on YouTube's discretionary power to remove allegedly infringing materials, stressing that "YouTube did not agree to act as a neutral processor of notices and counter-notices. YouTube retained control to evaluate counter-notices and infringement."<sup>339</sup> But, under the contractual network framework, the court could have considered YouTube's implied contractual obligations towards the network participants, which include providing users with notice and a meaningful right to appeal decisions that seem illegitimate to them. This is because facilitating procedural due process should be conceived as a necessary component in sustaining the mutual reliance of the network participants and assuring they coordinate to pursue the common goal of the network. Allowing users to appeal the platforms' decisions creates the necessary limit and check over platforms' autonomous and discretionary removal power, allowing users to express their interests and make them matter.

#### 4. Fairness and Equal Treatment

Users of social media platforms are of different types: amateur and professional, non-profit and commercial, individuals and organizations, liberals and conservatives, mainstream groups and marginalized communities. There is a growing concern that platforms do not equally apply the content moderation norms which are defined by their contracts and policies.<sup>340</sup> Users often complain that they have received differential treatment by platforms.<sup>341</sup> For instance, in their study of YouTube's monetization policy, Caplan and Gillespie described users' concerns over differential treatment by YouTube content

---

337. *Mishiyev v. Alphabet, Inc.*, 444 F. Supp. 3d 1154, 1156 (N.D. Cal. 2020), *aff'd*, 857 F. App'x 907 (9th Cir. 2021).

338. *Id.* at 1159.

339. *Id.*

340. Spandana Singh & Eliza Campbell, *Content Moderation Trends in the MENA Region: Censorship, Discrimination by Design, and Linguistic Challenges*, NEW AM. (Aug. 25, 2021), <https://www.newamerica.org/oti/blog/content-moderation-trends-in-the-mena-region-censorship-discrimination-by-design-and-linguistic-challenges> [<https://perma.cc/YD6V-ALDF>].

341. Elizabeth Dwoskin, *YouTube's Arbitrary Standards: Stars Keep Making Money Even After Breaking the Rules*, WASH. POST (Aug. 9, 2019), <https://www.washingtonpost.com/technology/2019/08/09/youtubes-arbitrary-standards-stars-keep-making-money-even-after-breaking-rules> [<https://perma.cc/4SGD-3DC9>].

moderation policies. YouTubers complained that the platform gave priority to “the interests of advertisers over the needs of creators: that established media personalities were seen as more ‘advertiser-friendly’ and were thus being treated differently by the platform . . . and that user-generated content was policed separately, more strictly, and through different mechanisms—including an over-reliance on flawed automation techniques.”<sup>342</sup>

The exercise of discretionary power under the contractual provisions which set the norms of online speech (e.g., Facebook’s Community Standards, YouTube’s Community Guidelines) have also led to differential treatment. The ToS offered by platforms indiscriminately set equal terms for users. Yet, with very little information provided by the ToS and the community standards on which content is allowed or disallowed on the platform, such platforms retain full discretion to interpret these standards arbitrarily and even inconsistently.<sup>343</sup> Thus, it has been shown that platforms often engage in selective enforcement in applying content moderation norms. For instance, users have complained of selective enforcement of copyright, which gives priority to content provided by established media partners.<sup>344</sup> A recent class action suit brought by musicians claims that YouTube only protects major record labels and studios.<sup>345</sup> The plaintiffs argue that YouTube deploys its Content ID system, which allows qualifying copyright holders to automatically manage their content and enforce their copyright, in a way that makes it available to the major music studies (from whom YouTube fears litigation) and not smaller right holders.<sup>346</sup>

The lawsuit argues that the “[d]efendants have, in effect, created a two-tiered system whereby the rights of large creators with the resources to take Defendants to court on their own are protected, while smaller and independent creators like Plaintiffs and the Class are deliberately left out in the cold.”<sup>347</sup>

Likewise, conservatives in the United States have persistently argued that platforms such as Facebook, Twitter, and YouTube are biased against conservatives, using “mass demonetization,” along with systematic filtering and removal of content, in order to stop the online spread of conservative ideology.<sup>348</sup>

---

342. Caplan & Gillespie, *supra* note 40, at 6 (citation omitted).

343. VENTURINI ET AL., *supra* note 25, at 23–24, 59.

344. Caplan & Gillespie, *supra* note 40, at 6.

345. Class Action Complaint & Demand for Jury Trial at 2, *Schneider v. YouTube, LLC*, No. 5:20-cv-4423 (N.D. Cal. July 2, 2020), ECF No. 1.

346. *Id.*

347. *Id.* at 5.

348. See for instance, Michael Nunez, *Former Facebook Workers: We Routinely Suppressed Conservative News*, GIZMODO (May 9, 2016, 9:10 AM), <https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006> [<https://perma.cc/YR9N-55CC>], cited by Senator Ted Cruz as proof of suppression of conservative news on Facebook during Mark Zuckerberg’s Senate hearings. *Transcript of Mark Zuckerberg’s Senate Hearing*, WASH. POST (Apr. 10, 2018), <https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing> [<https://perma.cc/3YA7-3F9Q>]. These allegations have been very

In the Prager University case, the plaintiffs argued that YouTube demonetized PragerU's videos and tagged its materials as unqualified for wide distribution ("restricted mode") in an effort to suppress conservative viewpoints.<sup>349</sup>

Could platforms' contracts offer a basis for challenging the unequal treatment of users? Though not a breach of any expressed terms of the contract, which give platforms full discretion over content removals, selective enforcement, biased exercise of contractual power, and unequal treatment of users could undermine the common purpose of the contractual network.

A platform's choice to discriminate against a particular user or group of users not only inflicts harm on the individual user, but may undermine the shared understanding regarding permissible speech on which all users in the network rely. Users depend on this shared understanding because it increases the predictability of norms while lowering the cost of uncertainty, attracting more users to the platform and overall producing economies of scale in advertising revenues.

Arguably, platforms should be able to create different tiers of partnerships, rewarding particular users for their contribution to increasing the platform's overall revenues. Yet, if users are required to comply with a set of norms that purport to apply generally to all users, users could reasonably expect that these norms will equally apply to all cases and users which are similarly situated. Platforms should be able to differentiate between users only by adhering to evident and visible norms which explicitly state the basis for such differentiation.

If courts were to adopt the contractual network approach and look beyond the bilateral transaction to consider the myriad of informal contractual relationships in social media, they could better protect the expectations and interests of the different stakeholders involved. Lewis, for instance, claimed that "YouTube deprived . . . her [of] reasonable expectations under the [ToS], which included the expectation that if she complied with the Terms of Service that her channel would be maintained, that the channel would continue to include her videos, and that her channel would continue to reflect the same number of views and comments by fellow YouTube users."<sup>350</sup> A network approach would enable the court to consider the consequences of unequal treatment of different users on the network's overall performance.

## 5. Damages

Another important legal consequence of applying the contractual network framework to the relationship between users and social media platforms

---

controversial, with no empirical studies to support them. See Rachel Kraus, *Once Again, There is No 'Anti-conservative' Bias on Social Media*, MASHABLE (July 28, 2020), <https://mashable.com/article/anti-conservative-bias-facebook/?europa=true> [<https://perma.cc/DR8W-DHDR>].

349. See *Prager Univ. v. Google LLC*, 951 F.3d 991, 995–96 (9th Cir. 2020).

350. *Lewis v. YouTube, LLC*, 197 Cal. Rptr. 3d 219, 222 (Ct. App. 2015).

relates to users' ability to claim damages for the harms created when platforms disrupt their horizontal contracts with their peers. Indeed, as we showed previously in Section III.C, courts ignore the contractual relationships between users, focusing merely on the vertical boilerplate ToS between the platform and each user individually.<sup>351</sup> In the case of Lewis discussed earlier, the plaintiff sought damages for "her out-of-pocket costs associated with producing the videos and the reasonable value of her time spent generating her original content and participating as a member of the YouTube community."<sup>352</sup> She claimed that even though YouTube restored her terminated account, she still suffered losses when the platform failed to restore the account's historical view counts and comments.<sup>353</sup> The court, however, held that the limitation of liability clause in YouTube's ToS encompassed this claim, and therefore refused to grant Lewis any damages.<sup>354</sup> As stressed by the court, even "[a]ssuming that YouTube has breached the Terms of Service, we conclude that Lewis cannot establish damages."<sup>355</sup> However, the damages Lewis claimed derived from her participation and coordination in the network.<sup>356</sup> As a member of the network, she relied on her ability to benefit from her horizontal relationships with other users: that is, Lewis expected that if she spent her resources to produce content and engage with others, she could extract her expected social benefit (which in this case was receiving users' acclaim and acquiring new friends). The historical view counts and comments on Lewis's account, which YouTube failed to restore, were not just facilitated by the platform, but they also and even more so reflect the outcome of Lewis's communal interactions with her peers. This contractual interdependency between the partners of the network creates mutual expectations that produce mutual obligations to adhere to them. Any violation of these obligations should justify a remedy, otherwise these obligations would be meaningless. Indeed, it is well established that there is "[n]o right without a remedy."<sup>357</sup> Hence the contractual network approach to social media has to establish, in addition to numerous layers of vertical and horizontal legal relationships, broad and comprehensive remedies that capture the full complexity of the network.

---

351. See *supra* Section III.C.

352. *Lewis*, 197 Cal. Rptr. 3d at 222–23 (quoting Brief for Appellant at 7, *Lewis*, 197 Cal. Rptr. 3d 219 (Ct. App. Oct. 14, 2014) (No. H041127)).

353. *Id.* at 221–22.

354. *Id.* at 223.

355. *Id.* at 224.

356. *Id.* at 224–25.

357. Frederick Pollock, *The Continuity of the Common Law*, 11 HARV. L. REV. 423, 424 (1898) (quoting modern maxim).

## VI. CONCLUSIONS

Content moderation by social media platforms may carry harmful implications for users, both individually and collectively. It may interfere with their business interests, damage their reputation, or diminish free speech and other fundamental rights. Content moderation may further constrain public discourse, obstruct the free flow of information, or restrict our access to information. Top-down regulation of social media, which seeks to address these implications, suffers from considerable limitations. Likewise, the narrow interpretation applied by the courts to platforms' boilerplate contracts has vitiated the role of private ordering in mitigating the platforms' discretionary power.

This Article makes a case for limiting platforms' discretionary power over content moderation by making platforms more accountable to their users. Where scarce competition means that users of these platforms have few or no alternatives if they wish to communicate with peers or followers online, effective legal standing could compensate for the lack of an exit. Accordingly, we propose contractual networks approach as a bottom-up approach to content moderation. We develop an interpretive framework under which courts could look at the contractual complexity of platforms contracts to determine the legitimacy of content removals and ensure they meet with the goals shared by different communities of users.

Holding platforms accountable for content moderation practices via private ordering has obvious advantages. It protects platforms and users against potential interferences of public actors while facilitating more diversity and exploration, enabling the emergence of different models of moderating digital content, thus supporting a more pluralist public discourse. Yet such diversity will not emerge if users are not given a voice in negotiating speech norms. Upholding the contractual expectations of users could be the first step. Supplementary steps would include the right to bring class actions and collect statutory damages for breaches of these expectations.