

St. John's University School of Law

## St. John's Law Scholarship Repository

---

Faculty Publications

---

2020

### The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression

Kate Klonick

*St. John's University School of Law*

Follow this and additional works at: [https://scholarship.law.stjohns.edu/faculty\\_publications](https://scholarship.law.stjohns.edu/faculty_publications)



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

---

This Feature is brought to you for free and open access by St. John's Law Scholarship Repository. It has been accepted for inclusion in Faculty Publications by an authorized administrator of St. John's Law Scholarship Repository. For more information, please contact [selbyc@stjohns.edu](mailto:selbyc@stjohns.edu).

KATE KLONICK

## The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression

**ABSTRACT.** For a decade and a half, Facebook has dominated the landscape of digital social networks, becoming one of the most powerful arbiters of online speech. Twenty-four hours a day, seven days a week, over two billion users leverage the platform to post, share, discuss, react to, and access content from all over the globe. Through a system of semipublic rules called “Community Standards,” Facebook has created a body of “laws” and a system of governance that dictate what users may say on the platform. In recent years, as this intricately built system to dispatch the company’s immense private power over the public right of speech has become more visible, Facebook has experienced intense pressure to become more accountable, transparent, and democratic, not only in how it creates its fundamental policies for speech but also in how it enforces them.

In November 2018, after years of entreaty from the press, advocacy groups, and users, CEO and founder Mark Zuckerberg announced that Facebook would construct an independent oversight body to be researched, created, and launched within the year. The express purpose of this body was to serve as an appellate review system for user content and to make content-moderation policy recommendations to Facebook. This Feature empirically documents the creation of this institution, now called the Facebook Oversight Board. The Board is a historic endeavor both in scope and scale.

The Feature traces the events and influences that led to Facebook’s decision to create the Oversight Board. It details the year-long process of creating the Board, relying on hundreds of hours of interviews and embedded research with the Governance Team charged with researching, planning, and building this new institution.

The creation of the Oversight Board and its aims are a novel articulation of internet governance. This Feature illuminates the future implications of the new institution for global freedom of expression. Using the lens of adjudication, it analyzes what the Board is, what the Board means to users, and what the Board means for industry and governments. Ultimately, the Feature concludes that the Facebook Oversight Board has great potential to set new precedent for user participation in private platforms’ governance and a user right to procedure in content moderation.



**AUTHOR.** Assistant Professor of Law, St. John's University School of Law; Affiliate Fellow, Yale Law School Information Society Project. I am grateful to Jack Balkin, John Barrett, Hannah Bloch-Wehba, Molly Brady, Julie Cohen, Catherine Duryea, Noah Feldman, Carrie Goldberg, James Grimmelman, Andrew Gruen, Nikolas Guggenberger, Scott Hershovitz, Gus Hurwitz, Daphne Keller, Jeff Kosseff, Anita Krishnakumar, Mike Masnick, Andrew McLaughlin, Peggy McGuinness, Mike Perino, Nate Persily, Robert Post, Nicholson Price, Martin Redish, Katie Reisner, Anna Roberts, Sarah Roberts, Jennifer Rothman, John Samples, Scott Shapiro, Jeremy Sheff, Alex Stamos, Will Thomas, Jillian York, and Jonathan Zittrain for helpful guidance, feedback, and conversations on this topic and early drafts. Exceptional gratitude to evelyn douek, who served as my external critic and colleague on this project and whose own writing and insights were incredibly influential, and to my family Jon Shea, Evelyn Frazee, and Tom Klonick for their essential love and boundless support. This project benefited from feedback through presentation at Data & Society, University of Haifa, Hans-Bredow-Institut, London School of Economics, New York University, Notre Dame Law School, Cornell Tech, and Georgetown University Law Center. This project would not have been possible without, and was funded entirely by, individual research grants from the Knight Foundation, Charles Koch Foundation, and MacArthur Foundation, and it owes a great debt to Brent Bomkamp, Zabee Fenelon, Leah Ferentino, Danika Johnson, Jisu Kim, John Mixon, and Patricia Vargas for their hard work as research assistants. Finally, a huge thank you to the staff of this Volume of the *Yale Law Journal* for their hard work in editing an ever-changing draft amidst a period of great instability and uncertainty. I have never accepted any payment, funding, or support from Facebook and conducted this research free from any nondisclosure agreement.



## FEATURE CONTENTS

INTRODUCTION	2422
I. BEFORE THE BOARD: FACEBOOK'S HISTORY WITH CONTENT MODERATION	2427
A. How Facebook Moderates Online Speech	2428
B. Who Controlled the Policies Behind Moderating User Content at Facebook	2435
C. Outside Influence on Facebook	2439
II. INSTITUTION BUILDING: CREATING THE OVERSIGHT BOARD	2448
A. Phase I: Global Consultation and Recruiting	2451
B. Phase II: Structure and Design	2457
1. Board Composition, Selection, and Removal of Members	2458
2. Board Authority and Power	2462
3. Appellate Procedure	2464
4. Independence Through Structure and a Binding Charter	2465
C. Phase III: Implementation	2466
1. The Trust	2467
2. The Bylaws	2469
a. User Appeal	2470
b. Facebook Case Submissions	2473
c. Transparency and Amendment	2473
III. A NEW PRIVATE-PUBLIC PARTNERSHIP TO GOVERN ONLINE SPEECH	2474
A. What Is the Board? Adjudication and Due Process	2476
1. Fundamental Rights	2477
2. Transparency	2479
3. Independence	2481
a. Jurisdictional Independence	2481
b. Intellectual Independence	2484
c. Financial Independence	2486



B. What Does the Board Mean for Global Users?	2487
1. The Pessimists	2488
2. The Realists	2490
3. The Optimists	2491
C. Impact on Industry, Government, and Global Speech	2492
1. Industry	2493
2. Government	2494
3. Global Free Expression	2496
<b>CONCLUSION</b>	<b>2499</b>

## INTRODUCTION

On May 23, 2019, a video of Nancy Pelosi, Speaker of the United States House of Representatives, suddenly began circulating online. Supposedly footage from a speech given the night before at a Center for American Progress event, the video seemed to show Pelosi “garbled and warped,” and as the clip quickly spread, commenters described her as “drunk” and a “babbling mess.”<sup>1</sup> Former New York City mayor and Republican operative Rudy Giuliani tweeted a link to the video stating, “What is wrong with Nancy Pelosi? Her speech pattern is bizarre”<sup>2</sup> as the clip went viral, particularly through conservative websites.<sup>3</sup> Though concern was momentarily raised about a “deep fake,” within a few hours it was established that in fact the video had simply been slowed down to seventy-five percent speed, causing a deepening of Pelosi’s voice, seemingly exaggerated pauses, and slurring,<sup>4</sup> all resulting in Pelosi appearing mentally unwell or intoxicated. The quick spread of this inauthentic depiction of Pelosi resulted in calls on social-media platforms to remove the video. YouTube was the first platform to comply, stating that “the Pelosi videos violated company policies and have been removed.”<sup>5</sup> But the video had not gone viral on YouTube in nearly the same way as it had on Facebook, and Facebook refused to remove it. In an interview with CNN’s Anderson Cooper, Monika Bickert—Facebook’s Vice President of Global Policy Management—explained that “if there were misinformation that was, let’s say, tied to an ongoing riot or the threat of some physical violence somewhere in the world, we would work with safety organizations on the

---

1. Drew Harwell, *Faked Pelosi Videos, Slowed to Make Her Appear Drunk, Spread Across Social Media*, WASH. POST (May 24, 2019), <https://www.washingtonpost.com/technology/2019/05/23/faked-pelosi-videos-slowed-make-her-appear-drunk-spread-across-social-media> [https://perma.cc/3S2H-M5QN].

2. *Id.* The tweet has since been deleted. *Id.*

3. On the Facebook page “Politics WatchDog,” the video “had been viewed more than 2 million times by [the evening of May 23], been shared more than 45,000 times, and garnered 23,000 comments.” *Id.*

4. *Id.* The video was not a “deep fake” under the traditional definition. “Deep fakes” use machine learning and newly accessible CGI technology to create videos from whole cloth of people saying and doing things they never did. See Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1758 (2019).

5. Harwell, *supra* note 1.

ground to confirm falsity and the link to violence and then we actually would remove that misinformation.”<sup>6</sup>

Bickert’s description of the considerations of authenticity, safety, and importance to political discourse in Facebook’s decision not to remove the video relates to the so-called Values of the site, which the platform balances when determining whether to remove content posted by users.<sup>7</sup> But the Values are just the principles supporting an extensive system to review all user-generated content, not only viral videos of political figures. Specifically, Facebook’s Values – “[v]oice,” offset by authenticity, safety, privacy, and dignity – inform a much more detailed and specific set of rules to regulate content called “Community Standards.”<sup>8</sup> The implementation of these Standards has resulted in the build-out of a massive system of governance for the screening, reporting, reviewing, and removal of content that violates the rules of the site – a process known as content moderation.<sup>9</sup> As a private platform, Facebook can and does maintain exclusive control over the rules and enforcement used to moderate speech on the site. Legal protections for online platforms, as well as American free-speech norms, influenced the early development of Facebook’s content policies.<sup>10</sup>

For many Americans, Facebook’s response to the Pelosi video was perhaps their first moment of understanding the power of private platforms to review speech, though in the last five years these issues have received greater and greater attention. As Facebook has become a ubiquitous and necessary presence in public discourse, public awareness of its censorial control has steadily increased, as have demands for greater accountability and transparency to users about how such

- 
6. *Cooper Grills Facebook VP for Keeping Pelosi Video Up*, CNN (May 25, 2019), <https://edition.cnn.com/videos/tech/2019/05/25/facebook-monika-bickert-pelosi-video-cooper-intv-sot-ac360-vpx.cnn> [<https://perma.cc/DG3Z-Y2J4>].
  7. Monika Bickert, *Updating the Values that Inform Our Community Standards*, FACEBOOK NEWSROOM (Sept. 12, 2019), <https://newsroom.fb.com/news/2019/09/updating-the-values-that-inform-our-community-standards> [<https://perma.cc/2UGB-VBKT>].
  8. *Id.*
  9. See Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1601 n.11 (2018) (“I use the terms ‘moderate,’ ‘curate,’ and sometimes ‘regulate’ to describe the behavior of these private platforms in both keeping up and taking down user-generated content. I use these terms rather than using the term ‘censor,’ which evokes the ideas of only removal of material and various practices of culturally expressive discipline or control.”).
  10. Perhaps the most significant of these legal protections is Section 230 of the Communications Decency Act, 47 U.S.C. § 230 (2018), and its subsequently broad interpretation by the courts. See *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997) (noting that one of the purposes of intermediary immunity in Section 230 was to protect the free speech of platform users).

decisions are made. Part I of this Feature describes the history and trajectory of Facebook's content moderation, discussing the primitive years of content moderation on the platform, its evolution to a more systematic set of internal rules, and the eventual development of those rules into its outward-facing Community Standards.

The integration of the details of those once-secret internal content-moderation rules into the external Community Standards was an essential, if long overdue, moment for Facebook following years of calls from academics, civil society, media, and users demanding more transparency from the platform. But while these new public Community Standards introduced some level of transparency into the black box of Facebook's speech governance system, they also underscored Facebook's overall lack of user participation in creating, updating, and enforcing the rules dictating their speech online. As advocates aligned anew on content-moderation principles,<sup>11</sup> experts published books documenting the system and pushing for reform,<sup>12</sup> and the press reflected users' outcry,<sup>13</sup> the topic

- 
11. See, e.g., David Kaye (Special Rapporteur), *Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶¶ 3, 45, 70, U.N. Doc. A/HRC/38/35 (Apr. 6, 2018) (calling on companies to align their speech codes with standards embodied in international human-rights law, particularly the International Covenant on Civil and Political Rights (ICCPR)); THE SANTA CLARA PRINCIPLES ON TRANSPARENCY AND ACCOUNTABILITY IN CONTENT MODERATION, <https://santaclaraprinciples.org> [<https://perma.cc/7L88-76NL>] [hereinafter SANTA CLARA PRINCIPLES].
  12. The last few years have seen a number of scholars publish foundational books on the intersection of private platform governance, online speech, and democracy. In order of publication date: REBECCA MACKINNON, *CONSENT OF THE NETWORKED* (2012); DANIELLE KEATS CITRON, *HATE CRIMES IN CYBERSPACE* (2014); ZEYNEP TUFEKCI, *TWITTER AND TEAR GAS: THE POWER AND FRAGILITY OF NETWORKED PROTEST* (2017); TARLETON GILLESPIE, *CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA* (2018); SIVA VAIDHYANATHAN, *ANTISOCIAL MEDIA: HOW FACEBOOK DISCONNECTS US AND UNDERMINES DEMOCRACY* (2018); JULIE E. COHEN, *BETWEEN TRUTH AND POWER* (2019); MIKE GODWIN, *THE SPLINTERS OF OUR DISCONTENT* (2019); DAVID KAYE, *SPEECH POLICE: THE GLOBAL STRUGGLE TO GOVERN THE INTERNET* (2019); JEFF KOSSEFF, *THE TWENTY-SIX WORDS THAT CREATED THE INTERNET* (2019); SARAH T. ROBERTS, *BEHIND THE SCREEN: CONTENT MODERATION IN THE SHADOWS OF SOCIAL MEDIA* (2019); NICOLAS P. SUZOR, *LAWLESS: THE SECRET RULES THAT GOVERN OUR DIGITAL LIVES* (2019); SHOSHANA ZUBOFF, *THE AGE OF SURVEILLANCE CAPITALISM* (2019).
  13. See, e.g., Julia Angwin et al., *Breaking the Black Box: What Facebook Knows About You*, PROPUBLICA: MACHINE BIAS (Sept. 28, 2016), <https://www.propublica.org/article/breaking-the-black-box-what-facebook-knows-about-you> [<https://perma.cc/VJQ4-26VC>]; Julia Angwin & Hannes Grassegger, *Facebook's Secret Censorship Rules Protect White Men from Hate Speech but Not Black Children*, PROPUBLICA: MACHINE BIAS (June 28, 2017), <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms> [<https://perma.cc/M4LT-8YWZ>]; Catherine Buni & Soraya Chemaly, *The Secret Rules of the Internet: The Murky History of Moderation, and How It's Shaping the Future of Free Speech*, VERGE (Apr. 13, 2016),



of accountability in online speech rose into the public consciousness. In April 2018, Facebook began to indicate openly that it was taking the new public reaction seriously.<sup>14</sup> CEO and founder Mark Zuckerberg stated in an interview that one could “imagine some sort of structure, almost like a Supreme Court, that is made up of independent folks who don’t work for Facebook, who ultimately make the final judgment call on what should be acceptable speech in a community that reflects the social norms and values of people all around the world.”<sup>15</sup>

This brings us to the primary focus of this Feature. Following his spring 2018 statement about a “Supreme Court”-like structure, and in response to longstanding and increasingly vocal criticism demanding user accountability, Zuckerberg announced in November 2018 that Facebook would create an “Independent Governance and Oversight” committee by the close of 2019 to advise on content policy and listen to user appeals on content decisions.<sup>16</sup> Part II documents the creation of this novel institution, relying on exclusive embedded access to the internal Facebook team charged with creating the Oversight Board, interviews with the Board’s architects, and primary-source documents.<sup>17</sup> It seeks to

---

<https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech> [<https://perma.cc/A256-296B>]; Adrian Chen, *The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed*, WIRED (Oct. 23, 2014, 6:30 AM), <https://www.wired.com/2014/10/content-moderation> [<https://perma.cc/UL37-NBX4>]; Nick Hopkins, *Revealed: Facebook’s Internal Rulebook on Sex, Terrorism and Violence*, GUARDIAN (May 21, 2017), <https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence> [<https://perma.cc/SG28-QH7D>]; Jeffrey Rosen, *Google’s Gatekeepers*, N.Y. TIMES MAG. (Nov. 28, 2008), <https://www.nytimes.com/2008/11/30/magazine/30google-t.html> [<https://perma.cc/C3F3-ACM2>].

14. Monika Bickert, *Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process*, FACEBOOK NEWSROOM (Apr. 24, 2018), <https://newsroom.fb.com/news/2018/04/comprehensive-community-standards> [<https://perma.cc/CTQ4-FU28>].
15. Ezra Klein, *Mark Zuckerberg on Facebook’s Hardest Year, and What Comes Next*, VOX (Apr. 2, 2018, 6:00 AM EST), <https://www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge> [<https://perma.cc/4GAX-LTZU>].
16. Mark Zuckerberg, *A Blueprint for Content Governance and Enforcement*, FACEBOOK (Nov. 15, 2018), <https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634> [<https://perma.cc/XP5S-2NU6>].
17. My conditions for this access were that all conversations were free from any nondisclosure agreement and that I must be allowed to tape all interviews. I received no funding from Facebook and was funded entirely by individual research grants from the Knight Foundation, Charles Koch Institute, and MacArthur Foundation. This level of access is somewhat unprecedented and understandably raises questions as to Facebook’s motivations for allowing an individual such a level of insight into their process. When I asked this question to Facebook directly, I was told that I was the only academic or reporter who asked to embed with the Governance Team and thus the only one granted access. While this may in part be true, it seems undeniable other motives likely exist. To this I can only speculate, but I believe the motivations were centered around concerns regarding the Board’s legitimacy. To that end, my

lend transparency to Facebook's internal deliberations in deciding how to design and implement the Board.

Underscoring this descriptive analysis, however, is a core question: what motivates and incentivizes Facebook to create such an oversight body? The potential answers to this question reflect a mix of idealism, skepticism, and cynicism. From the most dewy-eyed perspective, Facebook's decision to create the Board is a noble moment of goodwill. Sceptically seen, the creation of the Board is a giant public-relations move: an effort to give the appearance of caring about users and their input on Facebook's speech policies in order to bolster trust in the company. More cynically, once the Board is up and running, it might also be a convenient scapegoat for controversial content-moderation decisions—like the Pelosi video decision—that Facebook currently is forced to absorb alone. The availability of an independent board that can take on review of such decisions provides an easy escape from scrutiny: “Don't blame Facebook for this decision—it was the Board's!” Beyond these considerations, it is also imperative to remember the current regulatory pressures bearing down on Facebook and other major platforms. The European Union has led the charge in pushing back on private platforms through privacy measures like the General Data Protection Regulation and litigation through the European Court of Justice.<sup>18</sup> Even the United States, hamstrung in directly regulating content and privacy by First Amendment concerns, has turned to discussions of antitrust regulation to break up the “monopolies” of big tech as well as an empowered Federal Communications Commission (FCC) to levy fines for user-privacy violations.

All of these factors might boil down to a central issue: economics. Fines from Europe and the FCC and threats of splitting Facebook's products risk diminishing Facebook's bottom line. Though the Board does not directly resolve these concerns, its creation might show a good-faith effort by Facebook not to hoard its power and instead involve its users in critical decision-making—which might prove useful in the looming legal battles. The Board also speaks more broadly to

---

reporting, research, and publication of this Feature could be leveraged to gain elite buy-in on the project and to demonstrate the platform's commitment to transparency and accountability on a project that was centrally *about* transparency and accountability. To the best of my ability, I have tried to remain aware and vigilant not only of Facebook's motivations for allowing me to observe this process but also of not hinging my research or writing on this access.

18. See, e.g., Jennifer Daskal & Kate Klonick, Opinion, *When a Politician Is Called a 'Lousy Traitor,' Should Facebook Censor It?*, N.Y. TIMES (June 27, 2019), <https://www.nytimes.com/2019/06/27/opinion/facebook-censorship-speech-law.html> [<https://perma.cc/C7BJ-4P77>]; Daphne Keller, *Dolphins in the Net: Internet Content Filters and the Advocate General's Glawisch-nig-Piesczek v. Facebook Ireland Opinion*, STAN. CTR. FOR INTERNET & SOC'Y (Sept. 24, 2019), <https://cyberlaw.stanford.edu/files/Dolphins-in-the-Net-AG-Analysis.pdf> [<https://perma.cc/Z7UG-AMTT>]; Adam Satariano, *G.D.P.R., A New Privacy Law, Makes Europe World's Leading Tech Watchdog*, N.Y. TIMES (May 24, 2018), <https://www.nytimes.com/2018/05/24/technology/europe-gdpr-privacy.html> [<https://perma.cc/FDP6-U4NV>].

the core of Facebook's economic model: its ability to draw and keep users on its site in order to view advertising. Users' decisions to use Facebook depend on their trust of the platform. Facebook's creation of the Oversight Board is an investment in building user trust, which is a long-term strategy for continued economic growth.

Facebook thus has myriad incentives to make the Board a meaningful adjudicatory body – or, more cynically, to make the Board *appear* to be a meaningful adjudicatory body. To do this, the Board must involve users in content moderation and have jurisdictional, intellectual, and financial independence from Facebook. Whether the Board achieves this through its founding documents is the subject of the final Part of this Feature. That Part explores four remaining questions: what the Board is vis-à-vis existing legal models; the power of the procedural right it grants to users and the structures that make it independent from Facebook; what the Board might mean for users; and finally, what the Board might mean for industry and government.

Ultimately, however, the success or failure of the Board does not change its impact on history. The mere experiment, and the lessons taken from it, will have an indelible influence on the future of global online speech and user participation in transnational internet governance.

## **I. BEFORE THE BOARD: FACEBOOK'S HISTORY WITH CONTENT MODERATION**

It is impossible to talk about the Oversight Board without first discussing the history of content moderation at Facebook. Content moderation is the industry term for a platform's review of user-generated content posted on its site and the corresponding decision to keep it up or take it down.<sup>19</sup> Content moderation – though for years hidden from public view within private platforms and not always conceived in those terms – has existed as long as the internet and presented a seemingly impossible problem: how to impose a global set of rules on speech without a global set of norms on speech and how to do that at scale.<sup>20</sup> Understanding the history of industry struggle with content moderation is essential to understanding the Oversight Board, which – while perhaps the most elaborate mechanism designed to respond to the seemingly intractable quandary of governing online speech on a global scale – is only the latest in a long series of attempts.

---

19. See SUZOR, *supra* note 12, at 15-17.

20. Not all systems of content moderation are created equal. Facebook is unique in its sophisticated system of governance for removing and moderating user speech. See ROBERTS, *supra* note 12, at 4-6.

This Part unfolds the history of content moderation at Facebook and external calls for transparency and accountability. It then discusses the various industry-wide attempts to create mechanisms for user accountability in online speech platforms, concluding with the series of events culminating in Facebook's decision to create the Oversight Board.

*A. How Facebook Moderates Online Speech*

Before diving into the private governance system that Facebook created to moderate users' speech, it is necessary to take a moment to understand the laws that underlie that development.<sup>21</sup> Though Facebook is now a global platform with over two billion users, at its inception in 2004 it was a website of static pages open only to students with harvard.edu email addresses.<sup>22</sup> Facebook, like other speech platforms, was and is a private company that may police its users' speech without constraint by the First Amendment.<sup>23</sup> But early in the history of the internet, it was unclear if platforms' policing of speech on their sites would also make them civilly liable as publishers.<sup>24</sup> This was resolved when Section 230 of the Communications Decency Act, passed in 1996, immunized platforms

- 
21. Facebook's content-moderation governance system is one of a multitude of transnational governance systems that have "multiplied" in the last fifty years but that use the same governance terminology to describe various governance forms. See COHEN, *supra* note 12, at 202-37. Cohen's theoretical analysis of these various governance literatures is excellent, and Part III explores it in more detail as it relates to Facebook and the Oversight Board.
  22. DAVID KIRKPATRICK, THE FACEBOOK EFFECT: THE INSIDE STORY OF THE COMPANY THAT IS CONNECTING THE WORLD 82-83 (2010); *Company Info*, FACEBOOK NEWSROOM, <https://newsroom.fb.com/company-info> [<https://perma.cc/WRP3-PYWA>].
  23. See Klonick, *supra* note 9, at 1601-02. The question whether the First Amendment restricts how the government uses social media presents a different and important new doctrinal challenge that others have expertly analyzed and that is currently the subject of active litigation. See *Knight First Amendment Inst. at Columbia Univ. v. Trump*, 928 F.3d 226 (2d Cir. 2019) (affirming the district court's holding that President Trump's blocking users from his Twitter account violated the First Amendment); Lyrissa Lidsky, *Public Forum 2.0*, 91 B.U. L. REV. 1975 (2011) (addressing government actors' use of social media and the impact on the public-forum doctrine); Helen Norton & Danielle Keats Citron, *Government Speech 2.0*, 87 DENVER U. L. REV. 899 (2010) (discussing the ways in which the government-speech doctrine might be altered in light of many governments' increased dependence on social-media platforms); cf. Eric Goldman, *Why Section 230 Is Better Than the First Amendment*, 95 NOTRE DAME L. REV. REFLECTION 33, 33 (2019) (arguing that Section 230 "provides defendants with more substantive and procedural benefits than the First Amendment").
  24. See KOSSEFF, *supra* note 12, at 45.

from civil liability for speech posted by users.<sup>25</sup> This allowed platforms to moderate content on their sites and paved the way for Facebook to develop its own set of semipublic rules called “Community Standards” alongside a robust system for updating and enforcing them.<sup>26</sup>

With over 2.5 billion active users per month,<sup>27</sup> Facebook cannot and does not proactively police all violations of its rules.<sup>28</sup> When a user uploads a piece of content like a photo or prerecorded video, “ex ante” automated detection screens it.<sup>29</sup> Though there have recently been steps forward in using artificial intelligence to screen for things like extremism and hate speech,<sup>30</sup> the vast majority of this system works by matching the uploaded content against a database of already-

---

25. 47 U.S.C. § 230 (2018). Despite the broad shield for platforms, Section 230 expressly states that no internet entity has immunity from federal criminal law, intellectual property law, or communications privacy law. *Id.* § 230(e).

26. See KOSSEFF, *supra* note 12, at 244-45; Goldman, *supra* note 23.

27. Press Release, Facebook, Facebook Reports Fourth Quarter and Full Year 2019 Results (Jan. 29, 2020), <https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Fourth-Quarter-and-Full-Year-2019-Results/default.aspx> [https://perma.cc/R3T2-2G3P]. An often overlooked distinction is the difference between individual pieces of content that a user uploads or posts—status updates, photos, videos—and what Facebook dubs “complex objects,” such as pages or groups. For the purposes of this Feature, and because it is the only thing under consideration in the initial scope of the Board, the term “content” will refer only to the former. For greater detail, see Section II.B *infra*.

28. See Klonick, *supra* note 9, at 1635 (“Content moderation happens at many levels. It can happen before content is actually published on the site, as with ex ante moderation, or after content is published, as with ex post moderation. These methods can be either *reactive*, in which moderators passively assess content and update software only after others bring the content to their attention, or *proactive*, in which teams of moderators actively seek out published content for removal. Additionally, these processes can be *automatically* made by software or *manually* made by humans.”).

29. *Id.* at 1636-37; see also Sarah T. Roberts, *Abuse*, in UNCERTAIN ARCHIVES (Nanna Bonde Thylstrup et al. eds., forthcoming 2021) (manuscript at 7-8), <https://escholarship.org/content/qtzmp4715x/qtzmp4715x.pdf?t=pufbnf> [https://perma.cc/2VDV-RPXX] (explaining how Facebook uses PhotoDNA to screen images posted to their platform).

30. See, e.g., *Facebook: Transparency and Use of Consumer Data: Hearing Before the H. Comm. on Energy & Commerce*, 115th Cong. 84 (Apr. 11, 2018) (testimony of Mark Zuckerberg, CEO, Facebook) (“And, in addition to that, we have a number of AI tools that we are developing, like the ones that I had mentioned, that can proactively go flag the content.”); *Community Standards Enforcement Report*, FACEBOOK (Nov. 2019), <https://transparency.facebook.com/community-standards-enforcement> [https://perma.cc/6KQ8-4R2Z] (explaining process and removal statistics for different types of content, including hate speech, child nudity, and sexual exploitation).

known illegal or impermissible content using what is known as “hash technology.”<sup>31</sup> Other forms of ex ante moderation employ users’ locations to moderate the content they see. Known as “geo-blocking,” these types of moderation employ users’ internet protocol (IP) addresses, which can be roughly linked to geographical boundaries and thus allow platforms to distinguish content that might be legal to post in one country but illegal to post or view in another.<sup>32</sup>

These automatic takedown mechanisms work hand in glove with manual identification of content from the site.<sup>33</sup> For example, in reaction to public outcry around terrorist and extremist content hosted on the site, Facebook created a proactive<sup>34</sup> team of moderators to look for and delete such content, pages, and groups.<sup>35</sup> Once identified by Facebook’s moderators, the content can be hashed

- 
31. This methodology was pioneered by technology called PhotoDNA, invented by Professor Hany Farid and Microsoft to address the problems of child pornography. Tracy Ith, *Microsoft’s PhotoDNA: Protecting Children and Businesses in the Cloud*, MICROSOFT NEWS (July 15, 2015), <https://news.microsoft.com/features/microsofts-photodna-protecting-children-and-businesses-in-the-cloud> [<https://perma.cc/H7F7-KSB7>]; see also Kate Klonick, *Inside the Team at Facebook that Dealt with the Christchurch Shooting*, NEW YORKER (Apr. 25, 2019), <https://www.newyorker.com/news/news-desk/inside-the-team-at-facebook-that-dealt-with-the-christchurch-shooting> [<https://perma.cc/UW3Z-5S7Z>] (“To remove videos or photos, platforms use ‘hash’ technology, which was originally developed to combat the spread of child pornography online. Hashing works like fingerprinting for online content: whenever authorities discover, say, a video depicting sex with a minor, they take a unique set of pixels from it and use that to create a numerical identification tag, or hash. The hash is then placed in a database, and, when a user uploads a new video, a matching system automatically (and almost instantly) screens it against the database and blocks it if it’s a match. Besides child pornography, hash technology is also used to prevent the unauthorized use of copyrighted material, and over the last two and a half years it has been increasingly used to respond to the viral spread of extremist content, such as ISIS-recruitment videos or white-nationalist propaganda, though advocates concerned with the threat of censorship complain that tech companies have been opaque about how posts get added to the database.”).
  32. Geo-blocking typically is done following a request from a government notifying a platform that a certain type of posted content violates its local laws. For a thorough and excellent explanation of geo-blocking, see JACK GOLDSMITH & TIM WU, *WHO CONTROLS THE INTERNET? ILLUSIONS OF A BORDERLESS WORLD* 1-10 (2006).
  33. See Hannah Bloch-Wehba, *Automation in Moderation*, 52 CORNELL INT’L L.J. (forthcoming 2020), <https://ssrn.com/abstract=3521619> [<https://perma.cc/66NP-WE6Y>]. For a high-level discussion of the problems evinced by this dynamic, see Tim Wu, *Will Artificial Intelligence Eat the Common Law? The Rise of Hybrid Social-Ordering Systems*, 119 COLUM. L. REV. 2001 (2020).
  34. By *proactive*, I mean proactive from the perspective of the platform through its “proactive” action to use its own employees to look for and remove content violating the Community Standards.
  35. Natalie Andrews & Deepa Seetharaman, *Facebook Steps up Efforts Against Terrorism*, WALL ST. J. (Feb. 11, 2016, 7:39 PM ET), <http://www.wsj.com/articles/facebook-steps-up-efforts-against-terrorism-1455237595> [<https://perma.cc/JVU3-6XRN>]. Though not the focus of this

and compared against future uploaded content, not just on Facebook but also on other sites using the same technology.<sup>36</sup> Besides using Facebook employees, this cooperation between human moderation and automatic ex ante moderation can also involve Facebook users. For example, in identifying and controlling spam, Facebook can often algorithmically identify new spam postings based on the behavior of the poster, the proliferation on the site, and the targeting of the content. But Facebook also informs its algorithms and automatic takedowns with information gathered from users reactively<sup>37</sup> and manually<sup>38</sup> reporting spam. Thus, the manual and automatic systems work together to iteratively develop a database of content to be automatically removed from the site.<sup>39</sup>

Human decision-making is an essential part of speech moderation, especially as Facebook increasingly uses automated methods to both initially identify and adjudicate content.<sup>40</sup> Historically, Facebook's proactive moderation was largely

---

Feature, platform censorship of speech at the behest or encouragement of governments raises questions of collateral censorship and state-action doctrine. See Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035 (2017).

36. In 2016, Facebook, Microsoft, Twitter, and YouTube announced they would work together to create a shared industry database of online terrorist content. *Partnering to Help Curb Spread of Online Terrorist Content*, FACEBOOK (Dec. 5, 2016), <https://about.fb.com/news/2016/12/partnering-to-help-curb-spread-of-online-terrorist-content> [https://perma.cc/8DFG-U2VB]. This database later became Global Internet Forum to Counter Terrorism. *Global Internet Forum to Counter Terrorism: Evolving an Institution*, GLOBAL INTERNET F. TO COUNTER TERRORISM, <https://gifct.org/about> [https://perma.cc/895Q-Q88S].
37. By *reactively*, I mean reactive from the perspective of the platform through its “reactive” reliance on “ex post flagging” by platform users in order to allow “review by human content moderators against internal guidelines.” Klonick, *supra* note 9, at 1638; see Kate Crawford & Tarleton Gillespie, *What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint*, 18 NEW MEDIA & SOC'Y 410, 411 (2016).
38. See, e.g., James Parsons, *Facebook's War Continues Against Fake Profiles and Bots*, HUFFPOST (Dec. 6, 2017), [https://www.huffpost.com/entry/facebooks-war-continues-against-fake-profiles-and-bots\\_b\\_6914282](https://www.huffpost.com/entry/facebooks-war-continues-against-fake-profiles-and-bots_b_6914282) [https://perma.cc/3X6E-7AJ7].
39. See Bloch-Wehba, *supra* note 33; Robert Gorwa et al., *Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance*, 7 BIG DATA & SOC'Y 1 (2020).
40. As discussed *infra* Section III.C.3, the platform's ability to automatically censor has long-discussed risks. Because it happens prepublication, ex ante content moderation is the type of prior restraint with which scholars like Jack Balkin are concerned. See Jack M. Balkin, *Free Speech Is a Triangle*, 118 COLUM. L. REV. 2011 (2018) (describing how prior-restraint regimes censor without judicial analysis on whether or not speech is protected); see also Bloch-Wehba, *supra* note 33; Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 76 GEO. WASH. L. REV. 986, 1003-05 (2008) (describing the lack of incentives in the Digital Millennium Copyright Act's notice-and-takedown provisions for platforms to in-

confined to certain kinds of extreme content and limited by the nascent state of video- and photo-recognition technology.<sup>41</sup> The platform predominantly depended on users to flag violating speech.<sup>42</sup> This is quickly changing.<sup>43</sup> In the third quarter of 2019, nearly 100% of spam and fake accounts, 98.4% of adult nudity and sexual violations, and 98.6% of graphic and violent content on Facebook were found and flagged by the site before users notified the platform.<sup>44</sup> In other categories of banned material, proactive moderation is less effective. Facebook found and removed hate speech before users reported it only 80.2% of the time and only 16.1% of the time in cases of bullying and harassment.<sup>45</sup>

Whether already identified by Facebook or not, Facebook users flag millions of pieces of content worldwide every week.<sup>46</sup> Users have many reasons for flagging content, and much of what they flag does not violate the Community Standards. Rather, a vast majority of flagged content reflects personal opinion, conflict between groups or users, or even abuse of the flagging system to harass other users.<sup>47</sup> Facebook's reporting process attempts not only to guide users toward categorizing which aspect of the Community Standards is being violated

---

investigate before removal and the resulting risk of “suppress[ing] critical speech as well as copyright infringement”); Annemarie Bridy & Daphne Keller, U.S. Copyright Office Section 512 Study: Comments in Response to Notice of Inquiry (Mar. 30, 2016), <https://ssrn.com/abstract=2757197> [<https://perma.cc/XC6F-DRW8>].

41. See James Vincent, *AI Won't Relieve the Misery of Facebook's Human Moderators*, VERGE (Feb. 27, 2019), <https://www.theverge.com/2019/2/27/18242724/facebook-moderation-ai-artificial-intelligence-platforms> [<https://perma.cc/S7GU-H9QZ>].
42. Flagging is the tool provided by Facebook and other platforms to allow users to report potentially offensive content. The use of flagging by social media is both “practical” in that it allows platforms to enlist users in policing content that the platforms cannot proactively moderate on their own and “legitimizing” in that it gives users a voice—or at the very least the appearance of a voice—in questioning the content presented to them on the platform. Crawford & Gillespie, *supra* note 37, at 411-12.
43. See SUZOR, *supra* note 12, at 15-17.
44. *Community Standards Enforcement Report*, *supra* note 30 (looking at the categories “Spam,” “Fake Accounts,” “Adult Nudity and Sexual Activity,” and “Violent and Graphic Content” for July to September 2019, under the subcategory “Of the violating content we actioned, how much did we find before users reported it?”).
45. *Id.* (looking at the categories “Hate Speech” and “Bullying and Harassment” for July to September 2019, during which 2,200 pieces of hate-speech content that were removed were restored without an appeal and 169,700 were restored after user appeal).
46. *Understanding the Community Standards Enforcement Guide*, FACEBOOK (Nov. 13, 2019), <https://transparency.facebook.com/community-standards-enforcement/guide#section5> [<https://perma.cc/2GQS-25WL>].
47. Radiolab: *The Trust Engineers*, WNYC STUDIOS (Feb. 9, 2015), <https://www.wnycstudios.org/podcasts/radiolab/articles/trust-engineers> [<https://perma.cc/52VZ-23UF>].



to expedite review but also to urge users toward self-resolution of disputes that likely fall outside the Standards' purview.<sup>48</sup>

Once flagged by a user, the individual piece of content is stripped of user-identifying elements and placed into an internal tool that “queues” flagged content to be reviewed by a human content moderator.<sup>49</sup> These moderators are hired directly as Facebook employees or indirectly as contractors.<sup>50</sup> Working in front of computers in call centers all over the world, moderators are trained to determine if content violates the Community Standards.<sup>51</sup> If a moderator determines that reported content is in violation, it is removed from the site; all content not found in violation remains published on the platform.<sup>52</sup> In contrast to notice-and-takedown regimes, in the interim between flagging and the moderator's decision, the flagged piece of content stays up. If the content is not removed, the user to whom it belongs will never know it was even flagged. If it is removed, however, users will receive a notification that content they posted violated the Community Standards of Facebook and was removed.

Errors – both in allowing speech that should be removed to remain up, and removing speech that should be allowed – occur. Some of these errors – for instance, the removal of the Terror of War photo<sup>53</sup> – were high-profile. Initially, users could not appeal a decision on a piece of content.<sup>54</sup> Facebook's appeals systems only existed for suspended accounts and removed pages.<sup>55</sup> Those who had

- 
48. See Alexei Oreskovic, *Facebook Reporting Guide Shows How Site Is Policed (Infographic)*, HUFFPOST (Aug. 20, 2012), [https://www.huffpost.com/entry/facebook-reporting-guide\\_n\\_1610917](https://www.huffpost.com/entry/facebook-reporting-guide_n_1610917) [<https://perma.cc/4LHU-BLGD>].
49. See generally ROBERTS, *supra* note 12; SUZOR, *supra* note 12, at 15-17.
50. See Adrian Chen, *Inside Facebook's Outsourced Anti-Porn and Gore Brigade, Where "Camel Toes" Are More Offensive Than "Crushed Heads,"* GAWKER (Feb. 16, 2012, 3:45 PM), <http://gawker.com/5885714/inside-facebooks-outsourced-anti-porn-and-gore-brigade-where-camel-toes-are-more-offensive-than-crushed-heads> [<https://perma.cc/HU7H-972C>].
51. See generally ROBERTS, *supra* note 12; SUZOR, *supra* note 12, at 15-17.
52. While it might remain on the platform, content might still be censored in other ways. Down-ranking or interstitial screens are two ways that unfavorable content can remain uncensored but nonetheless “erased.” This practice is colloquially referred to as shadow banning. evelyn douek & Kate Klonick, *Facebook Releases an Update on Its Oversight Board: Many Questions, Few Answers*, LAWFARE (June 27, 2019, 3:41 PM), <https://www.lawfareblog.com/facebook-releases-update-its-oversight-board-many-questions-few-answers> [<https://perma.cc/R2JJ-VGBK>].
53. See *infra* Section I.B.
54. Marjorie Heins, *The Brave New World of Social Media Censorship*, 127 HARV. L. REV. F. 325, 326 (2014) (describing Facebook's internal appeals process as “mysterious at best” and their internal policies as secret).
55. See Klonick, *supra* note 9, at 1648.

their content mistakenly removed had little recourse unless they knew someone at Facebook or Instagram, were able to leverage a connected civil-society organization to their cause, or could rally media attention. In this way, the democratizing forces of internet speech in many ways simply recapitulated real-world power. This changed in 2018 when Facebook began offering expanded appeals for some types of removed content.<sup>56</sup>

Since instituting expanded appeals, Facebook has also started issuing quarterly transparency reports that provide insight into how content is being removed, appealed, and restored. In the second and third quarters of 2019,<sup>57</sup> for example, 4.28 billion pieces of content were removed from the site.<sup>58</sup> Of these, users appealed 40.9 million takedowns – less than 1% of the total removals<sup>59</sup> – and Facebook restored 10.1 million pieces of content, or approximately 24.7% of removals.<sup>60</sup>

This is the process of content moderation on Facebook: the *how*. While still a mystery to many users, this knowledge is now publicly available and relatively well known. But for over a decade, it was largely an unquantified phenomenon, and purposefully so.<sup>61</sup> Besides the public-facing Community Standards, which

---

56. Ian Wren, *Facebook Updates Community Standards, Expands Appeals Process*, NAT'L PUB. RADIO (Apr. 24, 2018, 5:01 AM), <https://www.npr.org/2018/04/24/605107093/facebook-updates-community-standards-expands-appeals-process> [<https://perma.cc/PFD8-PLL9>].

57. *Community Standards Enforcement Report*, *supra* note 30. Q2 is measured from April through June, Q3 from July to September. *Id.*

58. *Id.* This number considers the categories outlined in the *Community Standards Enforcement Report*, including Adult Nudity and Sexual Activity, Bullying and Harassment, Child Nudity and Sexual Exploitation of Children, Hate Speech, Regulated Goods: Drugs, Regulated Goods: Firearms, Spam, Terrorist Propaganda, Violent and Graphic Content, and Suicide and Self Injury. Notably, it excludes “Fake Accounts.”

59. *Id.* Calculated based on the total amount of appealed content (except fake accounts) for Q3 and Q2, divided by the total number of content removals (except fake accounts) for Q3 and Q2.

60. *Id.* Calculated based on the total amount of appealed content restored (except fake accounts) for Q3 and Q2, divided by the total number of content removals (except fake accounts) for Q3 and Q2. These numbers may risk obscuring the varied treatment across types of content, as different categories of speech have different rates of restoration following appeal. The most dramatic statistics are those demonstrating how often content is actually restored after appeal. Across categories, only three types of speech are restored to Facebook after user appeal over 1% of the time: adult nudity and sexual activity (3.05%), bullying and harassment (2.65%), and hate speech (2.52%). *Id.* Calculated based on the total number per category of appealed content restoration for Q2 and Q3, divided by the total amount of all removed content for Q2 and Q3. The other categories of speech have significantly different rates: 0.15% of child nudity and sexual exploitation, 0.19% of spam, 0.51% of regulated goods: firearms, 0.46% of terrorist propaganda, and 0.27% of suicide and self-injury are restored after appeal.

61. SUZOR, *supra* note 12, at 16.

were broad and vague, the specific rules and mechanisms by which Facebook removed or kept content on the site were deliberately opaque and carefully guarded. As will be discussed in the following two Sections, it took years of public agitation to generate inquiry not only into the *how* of content moderation, but also into *who* was creating the content-moderation rules and policies. Mechanics aside, the realization that Facebook—for the most part alone, and for the most part through a few people headquartered in Silicon Valley—was the sole creator and decider of the rules governing online speech struck many as deeply problematic. Small groups of civil-society members, academics, and members of the press had long been aware of how content moderation functioned. But as Facebook became an increasingly vital part of public discourse, demands for transparency and accountability on these issues became inescapable. The next two Sections discuss first the internal players at Facebook involved in creating the content rules and then the outside influencers who came to play a vital role.

### *B. Who Controlled the Policies Behind Moderating User Content at Facebook*

The individual moderators who review the flagged content described above are just the final stage of enforcement. Much like a policewoman who arrives to investigate a complaint and must assess whether illegal activity has occurred, content moderators are presented with reported user content and tasked with determining if it violates the site's terms.<sup>62</sup> Those terms, however, are set by individuals much higher up in the company.

From 2004 to early 2010, Facebook not only lacked a robust team for removing problematic content but had no real content-moderation policy to speak of.<sup>63</sup> Rather, platform-wide moderation was conducted by a team of a few dozen people, who were themselves guided by a one-page document of things to delete—like “Hitler and naked people”—and a general platform ethos of “if it makes you feel bad in your gut, then go ahead and take it down.”<sup>64</sup> It was not until late 2008, when Dave Willner joined the Trust and Safety team, that the external content-

---

62. See generally KAYE, *supra* note 12.

63. As late as November 2009, Facebook had no public policy or “Community Standards” at all. See Klonick, *supra* note 9, at 1620; Telephone Interview with Dave Willner, Former Head of Content Policy, Facebook, and Charlotte Willner, Former Safety Manager, User Operations, Facebook (Mar. 23, 2016) [hereinafter Telephone Interview with Dave Willner and Charlotte Willner]. All interviews are on file with the author.

64. Telephone Interview with Dave Willner and Charlotte Willner, *supra* note 63.

moderation policies and the underlying rules for use by content moderators began to take shape.<sup>65</sup> Willner, supervised by Jud Hoffman (who led the Trust and Safety team), eventually wrote a public-facing policy for the platform—the “Community Standards”—and an intricate 15,000-word Wiki-styled document—the “Abuse Standards”—that was internal to the site and spelled out the enforcement of the Community Standards.<sup>66</sup>

The era of Facebook content moderation under Willner and Hoffman coincided with the platform’s transition from a monolithic culture and region with similar norms to a global platform—a transition with which Facebook struggled. Until 2008, the “if it makes you feel bad, take it down” standard could be easily understood and enforced by Facebook’s small community of content moderators and was largely consistent with the values of the relatively small and homogenous “community” of then-existing Facebook users.<sup>67</sup> But as the site’s users diversified and became more global, and as the increased scale of user content demanded a more global and diverse army of content moderators, standards were no longer a feasible approach. Instead, Willner and Hoffman attempted to create granular, precise “rules” based on things “you could see in the content” rather than nonobservable values, feelings, or subjective reactions.<sup>68</sup>

Despite careful construction for “objective” enforcement, the Community Standards were by no means neutral, reflective of the norms of global society, or even reflective of the norms of Facebook users.<sup>69</sup> Rather, the early rules reflected the norms of the drafters: Americans “trained and acculturated in American free speech norms and First Amendment law” like Hoffman and Willner.<sup>70</sup> These cultural conflicts were an ever-present reality in enforcing the Community Standards worldwide.<sup>71</sup> In trying to resolve clashes between the Western values that informed Facebook’s rules and the values of local communities in which Facebook operated, Hoffman tried to turn to Facebook’s mission to “[g]ive people

---

65. See Telephone Interview with Jud Hoffman, Former Glob. Policy Manager, Facebook (Jan. 22, 2016); Telephone Interview with Dave Willner and Charlotte Willner, *supra* note 63.

66. See Klonick, *supra* note 9, at 1631-34.

67. Despite Facebook’s long-standing global presence, Willner describes users during his time at the platform as still relatively homogenous—“mostly American college students.” This rapidly changed as mobile technology improved and international access developed. *Id.* at 1633 (quoting Dave Willner and Charlotte Willner).

68. *Id.*

69. See *infra* notes 317-320 and accompanying text.

70. See Klonick, *supra* note 9, at 1621.

71. *Id.*

the power to build community and bring the world closer together”<sup>72</sup> and avoid adopting “wholesale . . . a kind of U.S. jurisprudence free expression approach.”<sup>73</sup> But even the company’s fundamental mission was not culturally neutral – nor could it have been. As Willner described, “The idea that the world should be more open and connected is not something that, for example, North Korea agrees with.”<sup>74</sup>

Despite this predisposition towards freedom of expression, Facebook’s Community Standards – which banned pornography,<sup>75</sup> hate speech,<sup>76</sup> graphic violence,<sup>77</sup> and torture of animals<sup>78</sup> – were significantly less permissive than First Amendment doctrine. Indeed, as public pressure increased over the last ten years and as European nations, like Germany, forced Facebook to comply with their national speech laws,<sup>79</sup> Facebook’s Community Standards became more restrictive and more similar to European standards. Largely absent from these Euro-American Facebook rules, however, is evidence of input from the Global South, a diverse set of nations and peoples that now comprises the majority of Facebook users.<sup>80</sup>

---

72. *About*, FACEBOOK, <https://www.facebook.com/pg/facebook/about> [<https://perma.cc/3ZV5-MECX>].

73. Klonick, *supra* note 9, at 1621 (quoting Jud Hoffman).

74. *Id.* at 1621-22 (quoting Dave Willner and Charlotte Willner). Willner evokes North Korea’s isolationist stance here as a means of illustrating how even the most basic of premises, “making the world open and connected,” does not have universal global uptake as a value or norm.

75. *Cf.* *Miller v. California*, 413 U.S. 15, 21, 24 (1973) (changing the previous definition of obscenity from that which is “utterly without redeeming social value,” to that which lacks “serious literary, artistic, political, or scientific value,” and thus creating a higher standard for finding material obscene and unprotected by the First Amendment).

76. *Cf.* *Snyder v. Phelps*, 562 U.S. 443, 445, 458 (2011) (finding that the First Amendment protects “outrageous,” “offensive,” and “targeted” speech when conducted in a public place on a matter of public concern).

77. *Cf.* *Brown v. Entm’t Merchs. Ass’n*, 564 U.S. 786, 805 (2011) (holding a law against violent video games invalid under the First Amendment as the games did not constitute obscenity under the *Miller* test).

78. *Cf.* *United States v. Stevens*, 559 U.S. 460, 468-72 (2010) (holding that depictions of animal cruelty are not categorically unprotected by the First Amendment).

79. *See* *Netzwerkdurchsetzungsgesetz [NetzDG]* [Network Enforcement Act], Sept. 1, 2017, BGBl I at 3352 (Ger.). Commonly known as “NetzDG,” this German law requires internet platforms to remove “manifestly illegal” speech within twenty-four hours of flagging, where “manifestly illegal” is defined by German law, or face fines of up to five million euros.

80. *See* J. Clement, *Leading Countries Based on Facebook Audience Size as of April 2020*, STATISTA (Apr. 24, 2020), <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users> [<https://perma.cc/S377-C6KY>].

In April 2018, after over a decade of hiding the internal rules behind the public-facing Community Standards, Facebook released one entirely public version.<sup>81</sup> The rules' public release was significant because of what this moment of transparency seemed to promise. During the fourteen years that Facebook's content moderation operated in secret, changes—good or bad, controversial or not—occurred without public knowledge or reaction. The move to transparency operated as a tacit acknowledgment that Facebook users had a right to know the rules governing them and voice their reactions. But Facebook's reveal was not only insufficient to stem outrage; it in fact fueled further critique. The public's discovery that a small cadre of people headquartered in Silicon Valley were the sole creators and deciders on rules governing this vital global platform for online speech<sup>82</sup> and that, although rules existed, their operation lacked core ideas of procedure and process added fuel to long-standing comparisons between Facebook and a feudal state, kingdom, or dictatorship.<sup>83</sup>

The rise of public awareness and demands for legality around platform content-moderation governance empowered civil society, academics, and journalists who had long been working on these issues. The more the platform emerged as a vital element of public discourse,<sup>84</sup> the louder and more frequent the demands

---

81. Bickert, *supra* note 14.

82. Jeff Rosen, *The Delete Squad*, NEW REPUBLIC (Apr. 29, 2013), <https://newrepublic.com/article/113045/free-speech-internet-silicon-valley-making-rules> [<https://perma.cc/JB95-5LTV>].

83. See, e.g., Doug Cantor, *Mark Zuckerberg May Be the CEO of Facebook, but He's the King of Twitter*, INC. (May 9, 2014), <https://www.inc.com/doug-cantor/mark-zuckerberg-may-be-the-ceo-of-facebook-but-hes-the-king-of-twitter.html> [<https://perma.cc/4W8G-CUBX>]; Laurence Dodds, *Can Mark Zuckerberg's 'Supreme Court' End Facebook's Era of Absolute Monarchy?*, TELEGRAPH (June 24, 2019, 8:12 AM), <https://www.telegraph.co.uk/technology/2019/02/18/mark-zuckerbergs-supreme-court-would-end-facebooks-era-absolute> [<https://perma.cc/AUA6-9Y3X>]; Henry Farrell et al., *Mark Zuckerberg Runs a Nation-State, and He's the King*, VOX (Apr. 10, 2018, 7:44 AM EDT), <https://www.vox.com/the-big-idea/2018/4/9/17214752/zuckerberg-facebook-power-regulation-data-privacy-control-political-theory-data-breach-king> [<https://perma.cc/LTB5-A74E>]. But see Anupam Chander, *Facebookistan*, 90 N.C. L. REV. 1807 (2012).

84. Many have argued that internet companies and platforms are more than “vital”; they are inescapable. See K. Sabeel Rahman, *The New Utilities: Private Power, Social Infrastructure, and the Revival of the Public Utility Concept*, 39 CARDOZO L. REV. 1621 (2018); TIM WU, *THE ATTENTION MERCHANTS* (2016); Lina Khan, Note, *Amazon's Antitrust Paradox*, 126 YALE L.J. 710 (2017). The concept of information fiduciaries links the broader conversation around platform dominance over commerce to platform dominance over speech. Compare Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183, 1185-87 (2016), and Jack M. Balkin & Jonathan Zittrain, *A Grand Bargain to Make Tech Companies Trustworthy*, ATLANTIC (Oct. 3, 2016), <https://www.theatlantic.com/technology/archive/2016/10/information-fiduciary/502346> [<https://perma.cc/MRA9-L77H>], with Lina M. Khan &

from “outside” groups for transparency and accountability became.<sup>85</sup> The next Section turns to how outside actors reacted to and influenced Facebook’s governance regime.

### C. Outside Influence on Facebook

Although details about Facebook’s content-moderation rules and how they were enforced may have been hidden from those outside the platform for years, that does not mean they were not free from external influence. To the contrary, many of the policy and enforcement decisions made by architects inside Facebook came about in reaction to external pressures from civil society, governments, the media, or users.<sup>86</sup> Of these, negative media coverage had arguably the most powerful impact.<sup>87</sup>

Perhaps the internally most significant moment of journalistic impact was the scandal accompanying Facebook’s removal of a photograph called “The Terror of War.”<sup>88</sup> In September 2016, a prominent Norwegian writer, Tom Egeland, posted a graphic but award-winning historical picture to Facebook.<sup>89</sup> Taken by photographer Nick Ut in the midst of the Vietnam War, the black-and-white photo depicts a naked nine-year-old girl running and screaming down a dirt

---

David E. Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497 (2019). Additionally, in an ongoing project, the Electronic Frontier Foundation has worked to document and present evidence of the negative psychological impact that leaving—either by choice or by banning—certain social media platforms can have on users. See *Submit Report*, ONLINE-CENSORSHIP.ORG, <https://onlinecensorship.org/submit-report> [<https://perma.cc/25NK-LGA2>] (offering a platform for users to report erroneous or unjust account deactivations). These studies support the theory Lawrence Lessig describes in *Code: Version 2.0*, in which he proffers that leaving an internet platform is more difficult and costly than expected. See LAWRENCE LESSIG, *CODE: VERSION 2.0*, at 288–90 (2006); cf. David G. Post, *Anarchy, State, and the Internet: An Essay on Law-Making in Cyberspace*, 1995 J. ONLINE L., art. 3, ¶ 42 (imagining the internet as a free market complete with user exit).

85. See, e.g., Helen Nissenbaum, *Accountability in a Computerized Society*, 2 SCI. & ENG. ETHICS 25 (1996); Nicolas P. Suzor et al., *What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Content Moderation*, 13 INT’L J. COMM. 1526 (2019).
86. See Klonick, *supra* note 9, at 1648–57.
87. See *id.* at 1652–55.
88. See Telephone Interview with Peter Stern, Head of Prod. Policy Stakeholder Engagement, Facebook (Mar. 7, 2018).
89. See Kjetil Malkenes Hovland & Deepa Seetharaman, *Facebook Backs Down on Censoring ‘Napalm Girl’ Photo*, WALL ST. J. (Sept. 9, 2016, 3:07 PM ET), <http://www.wsj.com/articles/norway-accuses-facebook-of-censorship-over-deleted-photo-of-napalm-girl-1473428032> [<https://perma.cc/CA5S-FWMW>].

street in Trang Bang, Vietnam, following a napalm attack on the city.<sup>90</sup> Though often colloquially referred to as “Napalm Girl,” the title of the photograph is “The Terror of War,” an era-defining piece of photojournalism for its depiction of the horror and violence of the Vietnam War.<sup>91</sup> Despite its political and cultural significance, the photo was removed for violating Facebook’s Community Standards.<sup>92</sup> In addition to removing the picture and accompanying post by Egeland, Facebook suspended Egeland’s account. He immediately posted separately complaining of censorship.<sup>93</sup> Egeland’s profile as a well-known author caused the removal of the historic photo to receive its own news coverage, epitomized by the Norwegian newspaper *Aftenposten*’s publication of a “letter” to Mark Zuckerberg on its front page calling for Facebook to take a stand against censorship.<sup>94</sup> The controversy was significant and, for perhaps the first time, crystallized in the public eye the platform’s broad power to censor not just things like hate speech or harassment but content of great cultural and historical import. Mere hours after the controversy ignited, Facebook’s Chief Operating Officer Sheryl Sandberg issued a statement saying that the platform “[doesn’t] always get it right” and promising to take a close look at the policies that had resulted in the takedown of the photo.<sup>95</sup>

Facebook acknowledges that the public reaction to the “Terror of War” takedown set off an internal reckoning with its content-moderation policies.<sup>96</sup> “After the ‘Terror of War’ controversy, we realized that we had to create new rules for imagery that we’d normally want to disallow, but for context reasons that

---

90. Nick Ut, *The Terror of War* (photograph), <http://100photos.time.com/photos/nick-ut-terror-war> [<https://perma.cc/YJ59-9ZTU>].

91. Mark Scott & Mike Isaac, *Facebook Restores Iconic Vietnam War Photo It Censored for Nudity*, N.Y. TIMES (Sept. 9, 2016), <https://www.nytimes.com/2016/09/10/technology/facebook-vietnam-war-photo-nudity.html> [<https://perma.cc/JY27-LWZD>].

92. The photo was likely removed because of involuntary nudity of a minor, not because it was child pornography. See Hovland & Seetharaman, *supra* note 89; Scott & Isaac, *supra* note 91.

93. Hovland & Seetharaman, *supra* note 89.

94. *Id.*

95. See Claire Zillman, *Sheryl Sandberg Apologizes for Facebook’s ‘Napalm Girl’ Incident*, TIME (Sept. 13, 2016), <http://time.com/4489370/sheryl-sandberg-napalm-girl-apology> [<https://perma.cc/9KVJ-AN5Q>]. Facebook also issued a press release highlighting its commitment to “allowing more items that people find newsworthy, significant, or important to the public interest—even if they might otherwise violate our standards.” Joel Kaplan & Justin Osofsky, *Input from Community and Partners on Our Community Standards*, FACEBOOK NEWSROOM (Oct. 21, 2016), <https://newsroom.fb.com/news/2016/10/input-from-community-and-partners-on-our-community-standards> [<https://perma.cc/JCY5-2VJY>].

96. See Kate Klonick, *Facebook v. Sullivan*, KNIGHT FIRST AMEND. INST. (Oct. 1, 2018), <https://knightcolumbia.org/content/facebook-v-sullivan> [<https://perma.cc/Z8KL-XQ45>].



policy doesn't work," explained Peter Stern, who is now head of Product Policy Stakeholder Engagement at Facebook.<sup>97</sup> But the "Terror of War" moment also illuminated something else that had perhaps been lost in the institutional transitions within the content policy team: Facebook could not simply solve the problems presented by content moderation through vigorous content *removal*; rather, a more nuanced approach was required in order to satisfy users and escape media controversy. Leaders at the company began to fully understand that content-moderation policy and enforcement were not just a "customer service" or "trust and safety issue" – as a communication platform, content moderation was *the product* of Facebook.<sup>98</sup>

In addition to individual high-profile moments of censorship or failure to remove unsafe content, a number of more general controversies generated negative press for the company and eroded user trust. The first grew out of Facebook's perceived harms to democracy and elections following the 2016 U.S. presidential election and "Brexit," the British referendum to leave the European Union.<sup>99</sup> Allegations that foreign interference through fake news and disinform-

---

97. Telephone Interview with Peter Stern, *supra* note 88.

98. See GILLESPIE, *supra* note 12, at 13, 47 ("And moderation is, in many ways, the commodity that platforms offer."). There are many competing formulations of what the product of Facebook is. Many argue that Facebook's product is advertising and users pay to use Facebook through the collection of personal data. See, e.g., ZUBOFF, *supra* note 12 (arguing that platforms' capitalist purpose is to surveil users in order to collect and commoditize their data for the purpose of selling advertising). But in a microeconomics consumer-theory framework, a product is a good or service to meet the desire or need of a customer. In this sense, curation of user-generated content is the service and product Facebook is supplying to its users, who do not pay for it in money, but instead in data. Thus, data is not the product but in fact the currency of the system of governance around content moderation. Because content curation is the product, it can also be a source of market differentiation between platforms, except if those platforms become so different in kind and dominant in their unique cultures as to preclude exit. See ALBERT O. HIRSCHMAN, EXIT, VOICE, AND LOYALTY: RESPONSES TO DECLINE IN FIRMS, ORGANIZATIONS, AND STATES 1-29 (1970).

99. See Philip Bump, *All the Ways Trump's Campaign Was Aided by Facebook, Ranked by Importance*, WASH. POST (Mar. 22, 2018, 2:19 PM EST), <https://www.washingtonpost.com/news/politics/wp/2018/03/22/all-the-ways-trumps-campaign-was-aided-by-facebook-ranked-by-importance> [<https://perma.cc/AKH8-QK3U>]; Jonathan Freedland, *Why Is the BBC Downplaying the Facebook Brexit Scandal?*, GUARDIAN (July 11, 2018, 7:50 AM EDT), <https://www.theguardian.com/commentisfree/2018/jul/11/bbc-downplaying-facebook-brexit-scandal-bias> [<https://perma.cc/5576-MF6Z>]; Alexis C. Madrigal, *What Facebook Did to American Democracy*, ATLANTIC (Oct. 12, 2017), <https://www.theatlantic.com/technology/archive/2017/10/what-facebook-did/542502/> [<https://perma.cc/3K92-AKRF>]; Peter Overby, *Facebook Acknowledges Russian Ads in 2016 Election. Will Investigations Follow?*, NAT'L PUB. RADIO (Sept. 8, 2017, 12:21 PM ET), <https://www.npr.org/2017/09/08/549284183/facebook-acknowledges-russian-ads-in-2016-election-will-investigations-follow> [<https://>

mation campaigns affected the outcome of those democratic processes dominated the news cycle and public opinion for months.<sup>100</sup> The second major moment came with the “leak” of user data through the third-party consulting firm Cambridge Analytica.<sup>101</sup> In a press release on March 16, 2018, Facebook acknowledged that the company, run from the United Kingdom, had improperly gained access and used eighty-seven million Facebook users’ data, with potential effects on the American election.<sup>102</sup> Third and most recently, Facebook was accused of augmenting civil unrest and perhaps even causing the death of thousands of Rohingya Muslims, a persecuted ethnic group in Myanmar, when majority groups used Facebook’s “Free Basics” program and messaging applications (including WhatsApp) to mobilize ethnic cleansing mobs.<sup>103</sup>

The scale of these issues and resultant public outcry triggered new pressure from Western democratic governments. From 2017 to 2019, the U.S. Senate held

---

perma.cc/6N9F-R8LW]; Olivia Solon, *Facebook’s Failure: Did Fake News and Polarized Politics Get Trump Elected?*, GUARDIAN (Nov. 10, 2016, 5:59 PM EST), <https://www.theguardian.com/technology/2016/nov/10/facebook-fake-news-election-conspiracy-theories> [https://perma.cc/HG74-LRVQ]; Carole Cadwalladr, *Facebook’s Role in Brexit—And the Threat to Democracy*, TED Talk (Apr. 2019) (transcript available at [https://www.ted.com/talks/carole\\_cadwalladr\\_facebook\\_s\\_role\\_in\\_brexit\\_and\\_the\\_threat\\_to\\_democracy/transcript](https://www.ted.com/talks/carole_cadwalladr_facebook_s_role_in_brexit_and_the_threat_to_democracy/transcript) [https://perma.cc/49XS-V224]). For an excellent discussion of the potential for “digital gerrymandering,” see Jonathan Zittrain, *Engineering an Election*, 127 HARV. L. REV. F. 335 (2014).

100. See *supra* note 99 and accompanying text.

101. See Carole Cadwalladr & Emma Graham-Harrison, *Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach*, GUARDIAN (Mar. 17, 2018, 6:03 PM EDT), <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> [https://perma.cc/FDB6-G6GP]; Nicholas Confessore, *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*, N.Y. TIMES (Apr. 4, 2018), <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html> [https://perma.cc/3BUF-K6NH]; Issie Lapowsky, *How Cambridge Analytica Sparked the Great Privacy Awakening*, WIRED (Mar. 17, 2019, 7:00 AM), <https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening> [https://perma.cc/Z249-Q5Y5].

102. Press Release, Paul Grewal, VP & Deputy Gen. Counsel, Facebook, *Suspending Cambridge Analytica and SCL Group from Facebook* (Mar. 16, 2018), <https://newsroom.fb.com/news/2018/03/suspending-cambridge-analytica> [https://perma.cc/NJ6K-MSEY].

103. For a contextualized and excellent summary of Facebook’s implication in the conflict in Myanmar, see evelyn douek, *Facebook’s Role in the Genocide in Myanmar: New Reporting Complicates the Narrative*, LAWFARE (Oct. 22, 2018, 9:01 AM), <https://www.lawfareblog.com/facebook-roles-genocide-myanmar-new-reporting-complicates-narrative> [https://perma.cc/JY49-NVKC]; and evelyn douek, *Why Were Members of Congress Asking Mark Zuckerberg About Myanmar? A Primer*, LAWFARE (Apr. 26, 2018, 7:00 AM), <https://www.lawfareblog.com/why-were-members-congress-asking-mark-zuckerberg-about-myanmar-primer> [https://perma.cc/UHS5-9LTU].

five hearings on the use of technology and its impact on speech, elections, terrorism, and privacy.<sup>104</sup> These included Mark Zuckerberg's testimony before the Senate following the Cambridge Analytica scandal,<sup>105</sup> as well as COO Sheryl Sandberg's testimony following the 2016 U.S. presidential election regarding foreign election interference on social media.<sup>106</sup> During the same period, the House of Representatives held four hearings, centrally focused on platform filtering practices<sup>107</sup> and the exploration of bias against American conservatives

- 
104. *Foreign Influence Operations' Use of Social Media Platforms (Company Witnesses): Open Hearing Before the S. Select Comm. on Intelligence*, 115th Cong. (2018); *Foreign Influence Operations' Use of Social Media Platforms (Third Party Expert Witnesses): Open Hearing Before the S. Select Comm. on Intelligence*, 115th Cong. (2018); *Facebook, Social Media Privacy, and the Use and Abuse of Data: Hearing Before the S. Comm. on Commerce, Sci., & Transp. and the S. Comm. on the Judiciary*, 115th Cong. (2018); *Terrorism and Social Media: #isbigtechdoingenough?: Hearing Before the S. Comm. on Commerce, Sci., & Transp.*, 115th Cong. (2018); *Policy Response to the Russian Interference in the 2016 U.S. Election: Open Hearing Before the S. Select Comm. on Intelligence*, 115th Cong. (2017).
105. *Facebook: Transparency and Use of Consumer Data: Hearing Before the H. Comm. on Energy & Commerce*, *supra* note 30. For media coverage, see Camila Domonoske, *Mark Zuckerberg Tells Senate: Election Security Is an 'Arms Race'*, NAT'L PUB. RADIO (Apr. 10, 2018, 2:30 PM ET), <https://www.npr.org/sections/thetwo-way/2018/04/10/599808766/i-m-responsible-for-what-happens-at-facebook-mark-zuckerberg-will-tell-senate> [<https://perma.cc/PR6Z-SWZV>]; and Emily Stewart, *Lawmakers Seem Confused About What Facebook Does – And How to Fix It*, VOX (Apr. 10, 2018, 7:50 PM EDT), <https://www.vox.com/policy-and-politics/2018/4/10/17222062/mark-zuckerberg-testimony-graham-facebook-regulations> [<https://perma.cc/6L8A-PJ6S>].
106. *Foreign Influence Operations' Use of Social Media Platforms (Company Witnesses)*, *supra* note 104. For media coverage, see Farhad Manjoo, *What Jack Dorsey and Sheryl Sandberg Taught Congress and Vice Versa*, N.Y. TIMES (Sept. 6, 2018), <https://www.nytimes.com/2018/09/06/technology/jack-dorsey-sheryl-sandberg-congress-hearings.html> [<https://perma.cc/53Y9-PJH8>]; and Ian Sherr, *Facebook, Twitter Face Senate Scrutiny Over Data, Foreign Election Interference*, CNET (Sept. 5, 2018, 9:43 AM PDT), <https://www.cnet.com/news/watch-live-facebook-twitter-testify-at-senate-panel-wednesday-google-a-no-show> [<https://perma.cc/U46F-N7VD>].
107. *Facebook, Google and Twitter: Examining the Content Filtering Practices of Social Media Giants: Hearing Before the H. Comm. on the Judiciary*, 115th Cong. (2018); *Filtering Practices of Social Media Platforms: Hearing Before the H. Comm. on the Judiciary*, 115th Cong. (2018). For media coverage, see Tony Romm, *Republicans Accused Facebook, Google and Twitter of Bias. Democrats Called the Meeting 'Dumb.'*, WASH. POST (July 17, 2018, 2:03 PM), <https://www.washingtonpost.com/technology/2018/07/17/republicans-accused-facebook-google-twitter-bias-democrats-called-hearing-dumb> [<https://perma.cc/B3N7-8RCN>]; and David Shepardson, *Social Media Companies Defend Filtering Practices Before Congress*, REUTERS (July 17, 2018, 1:15 PM), <https://www.reuters.com/article/us-usa-congress-socialmedia/social-media-companies-defend-filtering-practices-before-congress-idUSKBN1K729F> [<https://perma.cc/Q2SG-5W7E>]. See also *Algorithms: How Company's Decisions About Data and Content Impact Customers: Joint Hearing Before the Subcomm. on Comm'ns & Tech. and the Subcomm. on Dig. Commerce & Consumer Prot. of the H. Comm. on Energy & Commerce*, 115th Cong. (2017).

through content moderation.<sup>108</sup> The United Kingdom also held four lengthy inquiries into the influence of technology.<sup>109</sup>

Intergovernmental organizations applied a similar level of scrutiny. Six weeks after his congressional testimony, Mark Zuckerberg testified before the

---

For media coverage, see Roger Yu, *GOP Wrangles Tech Companies, Net Neutrality into Algorithm Scrutiny*, BLOOMBERG (Nov. 29, 2019, 10:58 AM), <https://news.bloomberglaw.com/tech-and-telecom-law/gop-wrangles-tech-companies-net-neutrality-into-algorithm-scrutiny> [<https://perma.cc/UW7W-TCMN>].

108. *Twitter: Transparency and Accountability: Hearing Before the H. Comm. on Energy & Commerce*, 115th Cong. (2018). For media coverage, see Barbra Ortutay, *Twitter CEO Jack Dorsey Keeps His Cool Before Congress*, ASSOCIATED PRESS (Sept. 5, 2018), <https://apnews.com/653ae098f935478d8a3d3edoob7543f9> [<https://perma.cc/9PME-Z452>]; and Sara Salinas, *Jack Dorsey to Congress: 'Twitter Does Not Use Political Ideology to Make Any Decisions'*, CNBC (Sept. 4, 2018, 1:49 PM), <https://www.cnbc.com/2018/09/04/jack-dorsey-to-congress-full-written-testimony-on-political-bias.html> [<https://perma.cc/4DGG-M3U6>]. See also *Facebook: Transparency and the Use of Consumer Data: Hearing Before H. Comm. on Energy & Commerce*, 115th Cong. (2018). For media coverage, see Julia Carrie Wong, *Mark Zuckerberg Faces Tough Questions in Two-Day Congressional Testimony—As It Happened*, GUARDIAN (Apr. 11, 2018, 3:37 PM), <https://www.theguardian.com/technology/live/2018/apr/11/mark-zuckerberg-testimony-live-updates-house-congress-cambridge-analytica> [<https://perma.cc/A3TF-6A9W>]; and Tony Romm, *Facebook's Zuckerberg Just Survived 10 Hours of Questioning by Congress*, WASH. POST (Apr. 11, 2018, 6:30 PM), <https://www.washingtonpost.com/news/the-switch/wp/2018/04/11/zuckerberg-facebook-hearing-congress-house-testimony> [<https://perma.cc/6M83-UBDC>].
109. DIGITAL, CULTURE, MEDIA AND SPORT COMMITTEE, DISINFORMATION AND 'FAKE NEWS': INTERIM REPORT, 2017–19, HC 363 (UK); SCIENCE AND TECHNOLOGY COMMITTEE, DIGITAL GOVERNMENT, 2017–19, HC 1455 (UK); DIGITAL, CULTURE, MEDIA, AND SPORT COMMITTEE, IMMERSIVE AND ADDICTIVE TECHNOLOGIES, 2017–19, HC 1846 (UK); EDUCATION COMMITTEE, FOURTH INDUSTRIAL REVOLUTION, 2017–19, HC 21 (UK). For media coverage, see Mark D'Arcy, *What's Hot on Committee Corridor?*, BBC (Oct. 10, 2018), <https://www.bbc.com/news/uk-politics-parliaments-45809876> [<https://perma.cc/ZQQ7-3JYR>]; Derek du Preez, *Digital Government Inquiry Concludes 'Momentum Has Been Lost and GDS Lost Its Way'*, DIGINOMICA (July 9, 2019), <https://diginomica.com/digital-government-inquiry-concludes-momentum-has-been-lost-and-gds-lost-its-way> [<https://perma.cc/BA9Y-CLFQ>]; *Parliamentary Committee Says Gov't Digital Approach Has Lost Impetus*, GOV'T COMPUTING (July 10, 2019), <https://www.governmentcomputing.com/central-government/news/science-technology-committee-digital-inquiry> [<https://perma.cc/QM8A-QSDA>]; and James Vincent, *The UK Invited a Robot to 'Give Evidence' in Parliament for Attention*, VERGE (Oct. 12, 2018, 11:45 AM EDT), <https://www.theverge.com/2018/10/12/17967752/uk-parliament-pepper-robot-invited-evidence-select-committee> [<https://perma.cc/Q8A7-RZRR>].

European Parliament – the body’s first of a three-part hearing series on the Cambridge Analytica scandal.<sup>110</sup> The European Parliament,<sup>111</sup> the European Parliament’s center-left political group,<sup>112</sup> and the European Commission<sup>113</sup> also paid attention to the dissemination of fake news on social media. Meanwhile, the UN has investigated the interface between social media and human rights,<sup>114</sup> counter-terrorism,<sup>115</sup> and speech,<sup>116</sup> with a special focus on Facebook as a platform for hate speech directed at Myanmar’s Rohingya Muslim minority.<sup>117</sup>

- 
110. Comm. on Civil Liberties, Justice & Home Affairs, *Hearing on the Facebook/Cambridge Analytica Case Part I*, EUR. PARLIAMENT, <https://www.europarl.europa.eu/committees/en/events-hearings.html?id=20180529CHE04141> [<https://perma.cc/4Y5T-B7K4>]; Comm. on Civil Liberties, Justice & Home Affairs, *Hearing on the Facebook/Cambridge Analytica Case Part II*, EUR. PARLIAMENT, <https://www.europarl.europa.eu/committees/en/events-hearings.html?id=20180613CHE04342> [<https://perma.cc/3GXF-VGK9>]; Comm. on Civil Liberties, Justice & Home Affairs, *Hearing on the Facebook/Cambridge Analytica Case Part III*, EUR. PARLIAMENT, <https://www.europarl.europa.eu/committees/en/events-hearings.html?id=20180626CHE04501> [<https://perma.cc/3C34-KFLN>].
111. Comm. on Culture & Educ., *Scientific and Academic Culture to Counter Radicalism and Fake News*, EUR. PARLIAMENT, <https://www.europarl.europa.eu/committees/en/events-hearings.html?id=20171127CHE03001> [<https://perma.cc/63HS-K8WZ>].
112. *S&D Joint Extremism/Digital Europe Working Group Conference - Fake News: Political and Legal Challenges*, SOCIALISTS & DEMOCRATS GROUP (Sept. 6, 2017), <https://www.socialistsanddemocrats.eu/events/sd-joint-extremismdigital-europe-working-group-conference-fake-news-political-and-legal> [<https://perma.cc/H27W-A75C>].
113. *Countering Online Disinformation - Towards a More Transparent, Credible and Diverse Digital Media Ecosystem*, EUR. COMMISSION (Jan. 29, 2019), <https://ec.europa.eu/digital-single-market/en/news/countering-online-disinformation-towards-more-transparent-credible-and-diverse-digital-media> [<https://perma.cc/TPF5-LPWH>].
114. Human Rights Council, Rep. of the Special Rapporteur on the Rights to Freedom of Peaceful Assembly and of Association 2, U.N. Doc. A/HRC/41/41 (May 17, 2019), <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G19/141/02/PDF/G1914102.pdf?OpenElement> [<https://perma.cc/8NXC-EDVF>] (“Facebook, Twitter and YouTube have become the gatekeepers to people’s ability to enjoy the rights of peaceful assembly and of association . . .”); Human Rights Council, Rep. of the Independent International Fact-Finding Mission on Myanmar 14, U.N. Doc. A/HRC/39/64 (Sept. 12, 2018), [https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A\\_HRC\\_39\\_64.pdf](https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf) [<https://perma.cc/B7B8-38NX>] [hereinafter *Myanmar Report*].
115. Counter-Terrorism Comm., *Open Meeting of the Counter-Terrorism Committee on #CTNarratives*, U.N. SECURITY COUNCIL (May 11, 2018), <https://www.un.org/sc/ctc/news/event/open-meeting-counter-terrorism-committee-ctnarratives> [<https://perma.cc/J2QX-7Q6S>].
116. Kaye, *supra* note 11.
117. *Myanmar Report*, *supra* note 114, at 14 (“Facebook has been a useful instrument for those seeking to spread hate, in a context where, for most users, Facebook is the Internet.”). For media reports, see Alastair Jamieson, *U.N. Says Facebook ‘Slow’ to Respond to Myanmar ‘Genocide’ Against Rohingya*, NBC NEWS (Aug. 27, 2018, 10:36 AM), <https://www.nbcnews.com/news/world/u-n-says-facebook-slow-respond-myanmar-genocide-against-rohingya>

Though the volume of government attention to the public outcry over Facebook's various scandals evinces notions of democratic accountability and process, their regulatory power to change these platforms is limited. The transnational operation of technology platforms has moved such entities beyond traditional notions of regulatory accountability—particularly where speech and content moderation are concerned. As Jack Balkin has described, public rights like free expression are now no longer “dyadic”—the classic struggle between government censorship and citizens—but “pluralist” in that private platforms now allow users to circumvent government intervention, and platforms can largely continue to operate successfully despite any one government's regulation.<sup>118</sup> In addition, the ability to regulate platforms has been hampered in the United States by First Amendment limitations on government restrictions of speech. This is notably not true of the European Union, which has recently passed comprehensive privacy regulation and whose courts have issued sweeping rulings to rein in the power of private platforms.<sup>119</sup>

Before this heightened government interest in platform regulation, civil-society groups were actively involved in discussions to construct better platform governance through policy advocacy, litigation, or independent research.<sup>120</sup> One of the earliest influencers in this area was the Anti-Defamation League (ADL), which established itself as an early advocate for thoughtful content policy that balanced hate speech against freedom of expression.<sup>121</sup> In comparison, the Electronic Frontier Foundation (EFF) has historically taken a more watchdog-like role than a collaborative one with platforms.<sup>122</sup> Between 2014 and 2019, a num-

---

-n904056 [<https://perma.cc/S772-ACRK>]; Casey Newton, *One Easy Thing Facebook Should Do in Myanmar*, VERGE (Nov. 10, 2018, 6:00 AM), <https://www.theverge.com/2018/11/10/18080962/facebook-myanmar-report-bsr-united-nations-hate-speech> [<https://perma.cc/6C7M-UHTC>].

118. Balkin, *supra* note 40, at 2013-14.

119. See generally NetzDG, *supra* note 79; Council Regulation 2016/679, 2016 O.J. (L 119) 1, arts. 22, 13, 14, 15 (EU).

120. See Klönick, *supra* note 9, at 1655-57.

121. *Id.* at 1655.

122. For example, in 2012 EFF launched [onlinecensorship.org](http://onlinecensorship.org), a public portal for users to report platform censorship. *Id.* at 1656-57; see also *What We Do*, [ONLINECENSORSHIP.ORG](http://ONLINECENSORSHIP.ORG), <https://onlinecensorship.org/about/what-we-do> [<https://perma.cc/8AJK-AV4R>]. As described in its mission statement, “Onlinecensorship.org seeks to encourage companies to operate with greater transparency and accountability toward their users as they make decisions that regulate speech”—principles that align with a pro-freedom of expression sentiment. See Klönick, *supra* note 9, at 1656-57; *What We Do*, *supra* note 122.

ber of other civil-society groups have started initiatives to address platform regulation of online speech.<sup>123</sup> This influence has now coalesced into Facebook’s “Trusted Partner” program which puts nongovernmental organizations into dialogue with the platform around important issues of speech.<sup>124</sup>

The sum of this outside pressure – from media, government, and civil society – eventually led to change at Facebook. In addition to more transparency in the rules, Facebook started to dedicate more staff to their enforcement. In the wake of Facebook’s much criticized failure to quickly remove a murderer’s video in April 2017,<sup>125</sup> Zuckerberg announced the company would hire 3,000 additional content moderators,<sup>126</sup> a number that later grew to 7,500.<sup>127</sup> The fevered media and government response to Facebook’s harm to elections led to the creation of Social Science One, a research organization dedicated to looking at the effects of social media on democracy.<sup>128</sup> Finally, the tragedy in Myanmar resulted directly in code changes to the product to prevent future use of the Messenger services to create and encourage violent mob behavior, as well as the hiring of

---

123. The Center for Democracy and Technology (CDT), for example, was a group that splintered off from EFF to focus on policy efforts in Washington, D.C. Other more general tech and speech advocacy groups like the Electronic Privacy Information Center, WITNESS, New America’s Open Technology Institute, AccessNow, Committee to Protect Journalists, and the ACLU all began making online speech and platform governance part of their work either through the creation of new roles or new advocacy directives. In contrast to policy advocacy, a number of these groups also worked for platform policy change through litigation against platforms and governments. *Reno v. ACLU*, for example, formed the foundational support for Section 230 of the CDA allowing – for better or worse – the proliferation of internet platform self-regulation. 521 U.S. 844 (1997). Today, new organizations like the Knight First Amendment Institute leverage the broadening role of platforms in speech to take aim at government.

124. *Trusted Partner Form*, FACEBOOK, [https://www.facebook.com/help/contact/1791786581083703?helpref=faq\\_content](https://www.facebook.com/help/contact/1791786581083703?helpref=faq_content) [https://perma.cc/GY47-EV4Q].

125. Klonick, *supra* note 31; Emily Dreyfuss, *Facebook Streams a Murder, and Must Now Face Itself*, WIRED (Apr. 16, 2017, 9:26 PM), <https://www.wired.com/2017/04/facebook-live-murder-steve-stephens> [https://perma.cc/U4KW-Z6J6].

126. Alex Heath, *Facebook Will Hire 3,000 More Moderators to Keep Deaths and Crimes from Being Streamed*, BUS. INSIDER (May 3, 2017, 10:35 AM), <http://www.businessinsider.com/facebook-to-hire-3000-moderators-to-keep-suicides-from-being-streamed-2017-5> [https://perma.cc/M3CZ-TXBH].

127. Alexis C. Madrigal, *Inside Facebook’s Fast-Growing Content-Moderation Effort*, ATLANTIC (Feb. 7, 2018), <https://www.theatlantic.com/technology/archive/2018/02/what-facebook-told-insiders-about-how-it-moderates-posts/552632/> [https://perma.cc/D69M-Q7XJ].

128. *Our Facebook Partnership*, SOC. SCI. ONE, <https://socialscience.one/our-facebook-partnership> [https://perma.cc/7XYW-W87F].

over 100 Burmese-fluent content moderators to recognize flagged hate speech.<sup>129</sup>

As will be discussed more extensively in Part III, Facebook's vulnerability to collective public action from media, government, and civil society is an important statement on the platform's accountability.<sup>130</sup> But it is a participatory and market-based model of accountability, not one based on guarantees of procedural rights or democratic votes. So far, users' power to influence Facebook comes from haphazard methods of collective complaint and expectations for response. Because disappointment of these expectations hurts Facebook's bottom line and long-term viability, Facebook faces a formidable challenge: it needs to preserve private power over its policy and product, while ensuring long-term engagement by meeting user expectations of accountability. The following Part details Facebook's efforts to meet this challenge through a new public-private governance regime: the Oversight Board.

## II. INSTITUTION BUILDING: CREATING THE OVERSIGHT BOARD

In the wake of revelations about Facebook's role in the 2016 U.S. presidential election, Mark Zuckerberg hosted a number of dinners for different groups of academics and advocates at his home to discuss what Facebook could do better.<sup>131</sup> These salons seem to have been the beginning of acculturating the CEO to the broader conversation around private governance in content moderation and privacy.

The Santa Clara Principles on Transparency and Accountability in Content Moderation were born in this new moment of leverage. Drafted by a diverse group of content-moderation academics — Irina Raicu, Sarah T. Roberts, Nic Suzor, and Sarah Myers West — and joined by EFF, ACLU Foundation of Northern California, CDT, and New America's Open Technology Institute, the Principles

---

129. Sarah Su, *Update on Myanmar*, FACEBOOK NEWSROOM (Aug. 15, 2018), <https://newsroom.fb.com/news/2018/08/update-on-myanmar> [<https://perma.cc/S46U-RBGW>].

130. See Jürgen Habermas, *Political Communication in Media Society: Does Democracy Still Enjoy an Epistemic Dimension? The Impact of Normative Theory on Empirical Research*, 16 COMM. THEORY 411, 419 (2006); Robert Post, *Participatory Democracy as a Theory of Free Speech: A Reply*, 97 VA. L. REV. 617, 624 (2011) ("Public opinion could not create democratic legitimacy if it were merely the voice of the loudest or the most violent . . . Public opinion can therefore serve the cause of democratic legitimacy only if it is at least partially formed in compliance with the civility rules that constitute reason and debate.").

131. Interview with Sarah T. Roberts (May 30, 2019) (describing attending one of several such dinners at the home of Mark Zuckerberg in 2017).



declared three “initial steps” the companies must take in order to provide “meaningful due process” on platforms.<sup>132</sup> The first steps addressed issues of transparency relating to content-moderation statistics and notice to users whose content is removed, but the core of the initiative lay in the final principle. The Third Santa Clara Principle urged online speech platforms to create “meaningful opportunity for timely appeal” for their users and in the long term consider “independent external review processes.”<sup>133</sup>

The focus on the idea of an independent external appeals process or an independent multistakeholder body in order to create accountability for online speech platforms was important but not new.<sup>134</sup> In January 2018, Noah Feldman, a constitutional-law professor, wrote a brief memo spelling out the details for a “Facebook Supreme Court,” the virtues such a tribunal would have to embody, and how it might solve both Facebook’s and Facebook’s users’ problems.<sup>135</sup> In getting buy-in for his idea, Feldman had something many others did not: a close friendship with Facebook COO Sheryl Sandberg. Feldman passed the memo to Sandberg, who passed it to Zuckerberg.<sup>136</sup>

As it turns out, as both the Santa Clara Principles and Feldman were lobbying for the idea of an external independent appellate body, the concept was also in the works at the company.<sup>137</sup> In April 2018, these ideas began to surface in public talking points from Facebook’s leadership. Monika Bickert, Facebook’s head of Global Policy Management, stated that the platform was “going to build out the ability for people to appeal our decisions” because “[w]e believe giving

---

132. See SANTA CLARA PRINCIPLES, *supra* note 11.

133. *Id.*

134. MacKinnon was one of the first to document and identify the potential of stakeholder influence in private internet governance. See MACKINNON, *supra* note 12, at 243-48; see also Klonick, *supra* note 9, at 1665 (“[T]he lack of an appeals system for individual users and the open acknowledgment of different treatment and rule sets for powerful users over others reveal that a fair opportunity to participate is not currently a prioritized part of platform moderation systems.”).

135. Memorandum from Noah Feldman on a Supreme Court for Facebook to Facebook (Jan. 30, 2018) (on file with author); Skype Interview with Noah Feldman, Professor, Harvard Law School (May 30, 2019).

136. See Mark Sullivan, *Exclusive: The Harvard Professor Behind Facebook’s Oversight Board Defends Its Role*, FAST COMPANY (July 8, 2019), <https://www.fastcompany.com/90373102/exclusive-the-harvard-professor-behind-facebooks-oversight-board-defends-its-role> [https://perma.cc/STJ9-WMFH].

137. Interview with Brent Harris, Head of Strategic Initiatives (Governance Team), Facebook, in Menlo Park, Cal. (June 6, 2019) (describing how building the Governance Team started in January 2018).

people a voice in the process is another essential component of building a fair system.”<sup>138</sup> That month in an interview with *Vox*’s Ezra Klein, Zuckerberg stated:

You can imagine some sort of structure, almost like a Supreme Court, that is made up of independent folks who don’t work for Facebook, who ultimately make the final judgment call on what should be acceptable speech in a community that reflects the social norms and values of people all around the world.<sup>139</sup>

Seven months later, Zuckerberg made the announcement official. In “A Blueprint for Content Governance and Enforcement,” the CEO acknowledged that he “increasingly [has] come to believe that Facebook should not make so many important decisions about free expression and safety on [its] own.”<sup>140</sup> Though he did not call it a Supreme Court, Zuckerberg described the creation of an independent oversight committee that would help govern content.<sup>141</sup> In creating such a committee, Zuckerberg promised to listen to outside advice on a host of difficult questions:

This is an incredibly important undertaking—and we’re still in the early stages of defining how this will work in practice. Starting today, we’re beginning a consultation period to address the hardest questions, such as: how are members of the body selected? How do we ensure their independence from Facebook, but also their commitment to the principles they must uphold? How do people petition this body? How does the body pick which cases to hear from potentially millions of requests?<sup>142</sup>

The statement closed with an aggressive goal: that such an independent body would be researched and created by the end of 2019.<sup>143</sup>

The thorny questions Zuckerberg posed in his statement were just a small sample of the many difficult problems to be addressed in creating the Oversight Board. Over the course of the next sixteen months, a small but steadily growing team of people inside Facebook worked on finding answers to them. This Part

---

138. Bickert, *supra* note 14.

139. Klein, *supra* note 15.

140. Zuckerberg, *supra* note 16.

141. See evelyn douek, *Facebook’s New ‘Supreme Court’ Could Revolutionize Online Speech*, LAWFARE (Nov. 19, 2018, 3:09 PM), <https://www.lawfareblog.com/facebooks-new-supreme-court-could-revolutionize-online-speech> [<https://perma.cc/F3XY-PGQ8>].

142. Zuckerberg, *supra* note 16.

143. *Id.*

describes the institution-building process of creating the Oversight Board. Section II.A discusses the Board’s conceptual development: the global consultation process, its mechanics, and the resultant report on its findings published by Facebook in June 2019. Given the input from this consultancy, Section II.B describes how the team then made decisions and tradeoffs in designing the Board and writing its founding documents, culminating in the publication of the Board’s Charter in the fall of 2019. Finally, Section II.C discusses the implementation of the Trust and Limited Liability Corporation (LLC) to insulate the Board and the publication of the Bylaws, which detail the procedural mechanisms of the Board. This descriptive account lays the groundwork for Part III’s normative assessment of the Board’s impact on users, industry, governments, and global free expression.

*A. Phase I: Global Consultation and Recruiting*

In January 2019, Facebook’s Vice President of Global Affairs and Communications, Nick Clegg, followed up on Zuckerberg’s November 2018 announcement with specifics about how the Board would operate, what its potential scope might be, and a draft of its Charter.<sup>144</sup> Clegg’s post also detailed the timeline and mechanisms by which Facebook would run the first phase of its process of creating the Board. This “Global Consultation” process would reach out to stakeholders and users worldwide to survey what people thought the Oversight Board should look like, accomplish, and address. It would take place over the first two quarters of 2019 and include a series of workshops hosted by Facebook to “convene experts and organizations who work on a range of issues such as free expression, technology and democracy, procedural fairness and human rights.”<sup>145</sup> These workshops were to be held in Singapore, Delhi, Nairobi, Berlin, New York, and Mexico City, but in an effort to not “hand-pick[.]” a small group of global experts, Clegg detailed that additional outreach efforts would take place in other ways, such as by small group consultation, individual meetings, and a public portal for submitting proposals about the Board.<sup>146</sup>

Reaction to Zuckerberg’s and Clegg’s announcements was mixed. Some worried that an Oversight Board would not only create more of the same types of

---

144. Nick Clegg, *Charting a Course for an Oversight Board for Content Decisions*, FACEBOOK NEWSROOM (Jan. 28, 2019), <https://newsroom.fb.com/news/2019/01/oversight-board> [<https://perma.cc/Q998-VPUA>].

145. *Id.*

146. *See id.*; see also Email from Zoe Darmé, Member, Governance Team, Facebook, to author (Sept. 14, 2019, 16:59 EST) (on file with author) (listing the exact dates of workshops and other workshop locations).

ensorship of user speech—particularly “borderline” content—but also provide a show of self-governance that would block governments from regulating Facebook.<sup>147</sup> Others were less concerned with borderline speech decisions and instead saw the Oversight Board as a public-relations gambit that might never become a reality or, if it did, would be more soundbite than substance.<sup>148</sup> Cynically, there was also the realization that, should such an independent body come to fruition, the Board would reflect not the noble aim of creating accountability and responsible governance that so many had called for, but Facebook’s best interests.<sup>149</sup> Specifically, some worried that having an independent body set some of Facebook’s most contentious and high-impact policies might end up serving as a scapegoat, allowing the platform to divert responsibility in the face of negative public reactions.<sup>150</sup>

Before discussing the external consultation phase, it is useful to have an idea of what was happening internally at Facebook during those first six months. Before Zuckerberg publicly announced the creation of the Board in November 2018, Facebook had already begun to hire a team to make the idea a reality.<sup>151</sup> This team came to be called the Governance and Strategic Initiatives Team, or Governance Team.<sup>152</sup> Its director was a young consultant, Brent Harris, who was hired from outside Facebook in January 2018.<sup>153</sup> The Governance team led by

---

147. See, e.g., Dipayan Ghosh, *Facebook’s Oversight Board Is Not Enough*, HARV. BUS. REV. (Oct. 16, 2019), <https://hbr.org/2019/10/facebooks-oversight-board-is-not-enough> [<https://perma.cc/LYJ4-TP8C>]; *Color of Change: Facebook Retaliated Against Protests by Pushing Anti-Semitic, Anti-Black Narratives*, DEMOCRACYNOW (Nov. 16, 2018), [https://www.democracynow.org/2018/11/16/color\\_of\\_change\\_facebook\\_retaliated\\_against](https://www.democracynow.org/2018/11/16/color_of_change_facebook_retaliated_against) [<https://perma.cc/DVK4-ECJS>] (transcript of interview of Siva Vaidhyanathan about the Oversight Board).

148. See, e.g., Conor Friedersdorf, *The Speech That Facebook Plans to Punish*, ATLANTIC (Dec. 11, 2018), <https://www.theatlantic.com/ideas/archive/2018/12/facebook-punish-censorship/577654> [<https://perma.cc/VWW8-TCBS>]; Zeynep Tufekci, *Why Zuckerberg’s 14-Year Apology Tour Hasn’t Fixed Facebook*, WIRED (Apr. 6, 2018, 3:32 PM), <https://www.wired.com/story/why-zuckerberg-15-year-apology-tour-hasnt-fixed-facebook> [<https://perma.cc/9ZVB-527Y>].

149. See, e.g., Josh Constine, *Toothless: Facebook Proposes a Weak Oversight Board*, TECHCRUNCH (Jan. 28, 2020, 4:33 PM EST), <https://techcrunch.com/2020/01/28/under-consideration> [<https://perma.cc/523P-L3UR>].

150. *Id.*

151. Email from Carolyn Glanville, Policy Communications (Governance Board), Facebook, to author (Sept. 19, 2019, 9:04 PM EST) (on file with author) (listing the team members’ start dates and their previous roles); see also Interview with Brent Harris, *supra* note 137 (describing his own start on the team in January 2018).

152. Email from Carolyn Glanville, *supra* note 151.

153. *Id.*

Harris has around ten direct employees as well as dozens of “cross-functional” employees, who work on the Governance Team and in other areas of Facebook.<sup>154</sup> Harris and the Governance Team report to Clegg, who in turn reports directly to Zuckerberg.<sup>155</sup> At Facebook, a company that thickly layers its management, having only one reporting level between the Governance Team and Zuckerberg implied a major commitment to the project.

Within the Governance Team, work on building the Board was initially split into four “tracks” – or “subject focuses.”<sup>156</sup> Track 1 focused on the selection of the Board members. While information about what users wanted an Oversight Board to look like was gathered from the consultancy process, those involved with Track 1 prepped the back end of the selection process. Working with outside consultants, they developed a list of qualities desired in Board members, a vetting procedure for candidates, and a process by which users could “nominate” members.<sup>157</sup> Track 2’s role was to decide on the powers, scope, and structure of the Board, which would be laid out in the founding documents.<sup>158</sup> One of the most difficult aspects was how to build a sufficiently independent adjudicatory body to provide oversight for Facebook, if Facebook was the creator of that Board. This problem arose in even basic questions of procedure, such as how the Board should review content decision cases. As one Team member described:

Let’s start with . . . how Facebook will present those cases to the Board, so what kind of information is in the cases, how much information we give[] them, what is their basis of the review of the decision, are we asking them to do a de novo review or is it in deference to Facebook’s earlier decision, . . . how are the panels selected that finally hear the cases, are we going for regional diversity, are we going for more of a global panel . . . [t]hen the deliberation process, what are the timelines—how

---

154. See Email from Carolyn Glanville, Policy Communications (Governance Board), Facebook, to author (May 25, 2020, 9:09 PM EST) (on file with author).

155. See Kurt Wagner, *Facebook Found Its New Public Face: Former U.K. Deputy Prime Minister Nick Clegg*, VOX (Oct. 19, 2018, 11:32 AM EDT), <https://www.vox.com/2018/10/19/17999968/facebook-nick-clegg-hire-deputy-prime-minister-elliott-schrage> [https://perma.cc/2LDB-76FH]; Email from Carolyn Glanville, *supra* note 154.

156. Interview with Kristen Murdock, Abigail Bridgman, McKenzie Thomas, and Emily Terwelp, Members of Governance Team, Facebook, in Menlo Park, Cal. (June 3, 2019) (describing the different tracks of “members, decisions, support, and go to market”).

157. Interview with Fariba Yassaee and McKenzie Thomas, Members of Governance Team, Facebook in Menlo Park, Cal. (June 5, 2019) (describing the purpose of Track 1).

158. Interview with Kristen Murdock et al., *supra* note 156 (describing the purpose of Track 2).

long does the case move from the time it is sourced, selected, to hearing and deliberation and how does that work.<sup>159</sup>

Track 3 dealt with the administrative side of the Board, setting up the secretariat to staff the Board, creating a framework to research and give local context on an appeal to the Board through outside experts, designing the “tooling that the Board is going to need, functionally, to be able to sift through the volume of cases,” which included how the Board would be able to interact with Facebook while retaining independence.<sup>160</sup> Perhaps most saliently, the major task of Track 3 was to tackle the issue of governance:

[H]ow will the Board govern itself? What is the Board set up functionally to do as [it relates to] the Charter? How does the Board govern itself from a governance perspective in its Bylaws? And . . . how does that all work together to create the institution that is the Board and its staff?<sup>161</sup>

The fourth and final Track was the “go to market” team, tasked with implementing everything decided in Tracks 1, 2, and 3.<sup>162</sup> This team focused on the launch of the Oversight Board, from creating a curriculum for training its members to having a flagship event that could both orient and welcome the Board.<sup>163</sup>

While these Tracks worked internally at Facebook, a separate outreach team prepared to launch the global consulting process and interface between Facebook and the public.<sup>164</sup> This meant designing and planning the six global workshops. Crucially for user participation, it also involved inviting participants and collecting feedback. User participation was not limited to these workshops, however. The outreach team also met one-on-one with experts who could not attend the workshops, created one-day workshops for small groups of stakeholders, and constructed an online portal to solicit feedback from all global users.<sup>165</sup>

---

159. *Id.* (quoting Abigail Bridgman).

160. Interview with Heather Moore and McKenzie Thomas, Members of Governance Team, Facebook in Menlo Park, Cal. (June 5, 2019) (describing the purpose of Track 3).

161. *Id.* (quoting Heather Moore).

162. Interview with Kristen Murdock et al., *supra* note 156.

163. Interview with Zoe Darmé, McKenzie Thomas, and Fariba Yassaee, Members of Governance Team, Facebook in Menlo Park, Cal. (June 4, 2019) (discussing syncing outreach and member recruiting with Track 4); Interview with Fariba Yassaee, Kendall, Greg, and Eric, Members of Governance Team, Facebook in Menlo Park, Cal. (June 4, 2019) (discussing organization and planning of member training).

164. Email from Carolyn Glanville, *supra* note 151.

165. Interview with Zoe Darmé et al., *supra* note 163; Interview with Zoe Darmé and McKenzie Thomas, Members, Governance Team, Facebook, in Berlin, Ger. (June 24, 2019).

The first Global Workshop took place in Singapore on February 20 and 21, 2019.<sup>166</sup> The decision to start in the Global South was a purposeful one, said one member of the team, since “the rules for Facebook had always come from an American and Western tradition, which was part of the problem. We wanted to do our best to anchor ourselves to the concerns of people that were outside that culture.”<sup>167</sup> But at the first workshop, both sides had no idea what to plan for or expect. Facebook had invited around forty people and had not organized a presentation or structure for the discussion. “I figured they would just tell us what they thought of the Board,” said one organizer, “but it turns out they showed up expecting to be told what the Board was, or might be, and they had no idea how to give feedback without that.”<sup>168</sup> The result was a retooling of the format for the following day’s meetings that included an introduction for workshop participants about the Board’s purpose and a set of group exercises.

The format developed in Singapore became the standard programming for all global workshops going forward.<sup>169</sup> At each event, forty to fifty invited stakeholders, experts, and journalists convened over the course of two days. Participants were assigned seats around a nine-person table. At each seat was an iPad that was pre-loaded with case-study documents and the Facebook Community Standards. After a morning introduction to content moderation and the Oversight Board, the remainder of the two days was spent addressing two hypothetical cases that the Board might see. In these simulations, each workshop table was a notional Oversight Board “panel.” Participants were asked to read the “case for review,” and then, using the Community Standards as a guide, discuss with their group whether the content should stay up or come down. Tables voted individually and then a full-group poll was taken, followed by discussion. As the workshops moved around the globe – Delhi later in February, Nairobi in March, Mexico City and New York City in May, and Berlin in June – the outreach team for Facebook gathered careful notes from the conversations, polls, and group breakout sessions.<sup>170</sup>

In addition to these larger forums, the outreach team also organized a series of smaller roundtables to reach communities and groups of experts that could not attend the Global Workshops. The first of these occurred before the announcement of the Board and was held at Facebook. It included legal scholars,

---

166. Email from Zoe Darmé, *supra* note 146.

167. Interview with Zoe Darmé and McKenzie Thomas, *supra* note 165.

168. *Id.* (quoting McKenzie Thomas).

169. *Id.*

170. Email from Zoe Darmé, *supra* note 146.

human-rights experts, and technical specialists at Facebook headquarters in October 2018.<sup>171</sup> Similar one-day stakeholder events were held all over the world, ranging from Taipei to Washington, D.C. to Istanbul, and even one town hall in Norman, Oklahoma.<sup>172</sup> In total, 650 people from eighty-eight countries attended the workshops, roundtables, and town halls.<sup>173</sup> Moreover, 1,206 people participated in the online questionnaire about the Board that was open to the public for responses between April 1 and May 27, 2019.<sup>174</sup>

User participation in determining what the Board should be was not the only goal of the Global Consultancy process. The events also served as potential sources of recruitment of members for the Board.<sup>175</sup> A Facebook employee, often from the Governance Team, oversaw the simulations at each table. As participants deliberated, Facebook employees looked for various qualities and attributes that early research in member selection had deemed relevant: was a participant open to changing their mind through conversation;<sup>176</sup> did a participant listen to others; did a participant express good insight into a particular issue;

- 
171. Telephone Interview with Robert Post, Professor, Yale Law School (Feb. 28, 2020) (discussing his attendance at the meeting and describing participants and nature of discussion).
172. Interview with Zoe Darmé, Peter Stern, and McKenzie Thomas, Facebook, in Menlo Park, Cal. (June 7, 2019); Brent Harris, *Global Feedback and Input on the Facebook Oversight Board for Content Decisions*, FACEBOOK NEWSROOM (June 27, 2019), <https://newsroom.fb.com/news/2019/06/global-feedback-on-oversight-board> [<https://perma.cc/946H-ZM6A>]; see also Zoe Mentel Darmé et al., *Global Feedback & Input on the Facebook Oversight Board for Content Decisions*, FACEBOOK 41 (June 27, 2019), <https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf> [<https://perma.cc/JC7U-GCD8>] [hereinafter Facebook Consultancy Report]; *Global Feedback & Input on the Facebook Oversight Board for Content Decisions: Appendix*, FACEBOOK, <https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-appendix.pdf> [<https://perma.cc/3UN7-S8ZD>] [hereinafter Facebook Consultancy Report Appendices].
173. Facebook Consultancy Report, *supra* note 172, at 12. There was also one-on-one consultation with more than 250 people about the Oversight Board, but little has been made available about how those people were sourced, paid, or how their feedback influenced the creation of the Board.
174. The survey was available in thirteen languages and had quantitative and qualitative elements. The first part was a series of multiple-choice questions about the Charter, while the second part consisted of open essay-style questions about who should serve on the Board, how the Board should hear cases and deliberate, and what governance structure would result in the most independent outcomes. See *id.* at 15.
175. Interview with Zoe Darmé et al., *supra* note 163.
176. This particular metric was also quantitatively measured. Before deliberating with their table, participants were asked to use the application on their iPad to vote on whether they would keep up or remove content. After a few hours of deliberation, they were asked to vote again. Email from Zoe Darmé, Member, Governance Team, Facebook (June 18, 2019, 11:15 PM EST) (on file with author).



how did a participant deal with opinions hostile to their own?<sup>177</sup> While nominations of people to serve on the Board came through a variety of means, including an online public portal, the workshops provided critical feedback as the Governance Team moved into the membership selection phase. By the conclusion of the process, the Governance Team ran outreach on and sought feedback from well over two thousand individuals around the world, making it perhaps one of the largest global recruitment and stakeholder participation processes in history.

The results of the Global Consultancy were published in a forty-four-page executive report with a 180-page appendix on June 27, 2019.<sup>178</sup> As one might expect from such an extensive global survey, there was no clear consensus on what users wanted from an Oversight Board, but a few trends emerged.<sup>179</sup> For instance, the majority of those who gave feedback agreed that a set of substantive values should underpin the Charter and purpose of the Board, though there was significant disagreement about what those values should be.<sup>180</sup> Respondents also strongly favored the Board having the ability to influence policy, with ninety-five percent of respondents favoring the idea that the Board could make policy recommendations to Facebook.<sup>181</sup> Finally, while many felt that diversity should be a major priority for the Board and the people working for it, there was little unanimity about how to implement that idea on a global scale.<sup>182</sup>

These results, while scattered, provided a number of signals about what values and principles to aim for as the Governance Team turned to the next phase: drafting the Charter of the Oversight Board.

### *B. Phase II: Structure and Design*

In early January 2019, a draft Charter was announced, listing some of the big questions that would have to be addressed by the final document.<sup>183</sup> The Oversight Board Charter was imagined as, and ultimately became, a constitution-like

---

177. Interview with Brent Harris, *supra* note 137; Interview with Fariba Yassae et al., *supra* note 163.

178. See Facebook Consultancy Report, *supra* note 172.

179. douek & Klonick, *supra* note 52.

180. *Id.*

181. *Id.*

182. *Id.*

183. *Draft Charter: An Oversight Board for Content Decisions*, FACEBOOK, <https://about.fb.com/wp-content/uploads/2019/01/draft-charter-oversight-board-for-content-decisions-2.pdf> [https://perma.cc/KL7H-2RUD].

document that laid out the structural relationship between Facebook, the Oversight Board, and the Trust that would sit between them. On September 17, 2019, the Governance Team released the final Charter.<sup>184</sup> The nine-page foundational document was split into seven articles: Members, Authority to Review, Procedures for Review, Implementation, Governance, Amendments and Bylaws, and Compliance with Law.<sup>185</sup> This Section examines those provisions while adding context from interviews and observations at Facebook while the decisions were being made. It discusses four broad areas that involve the interplay of all provisions: Board Composition, Selection, and Removal of Members; Board Authority and Power; Appellate Procedure; and Structural Independence.

### 1. Board Composition, Selection, and Removal of Members

From the beginning of the Oversight Board project, once certain structural decisions were made, the final core question would be who Board members were and how they were selected.<sup>186</sup> This included substantive questions about what qualifications members should possess and whom they should represent, as well as procedural problems about how members would be chosen and removed.<sup>187</sup>

Feedback from the consulting period established that people overwhelmingly favored a diverse Board that represented “as many segments of society as possible” and incorporated a “multidisciplinary, multi-stakeholder approach.”<sup>188</sup> The discussion of what diversity meant was lengthy and included discussions about gender, race, sexual orientation, geographic location, education, age, disability, wealth, and political views.<sup>189</sup> Additionally, much time was spent on members’ necessary experience and qualifications. Many felt that Board members should be drawn from elite sectors of the global population, while others

---

184. Brent Harris, *Establishing Structure and Governance for an Independent Oversight Board*, FACEBOOK NEWSROOM (Sept. 17, 2019), <https://about.fb.com/news/2019/09/oversight-board-structure> [https://perma.cc/BY6N-JAB8].

185. *Oversight Board Charter*, FACEBOOK, [https://about.fb.com/wp-content/uploads/2019/09/oversight\\_board\\_charter.pdf](https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf) [https://perma.cc/H53N-PHP4].

186. Interview with Brent Harris, *supra* note 137.

187. Interview with Fariba Yassaee, Member, Governance Team, Facebook, in Menlo Park, Cal. (Aug. 8, 2019); Interview with Heather Moore, Brent Harris, and Louis Chang, Members of Governance Team, Facebook, in Menlo Park, Cal. (Sept. 11, 2019); Observation at Charter Workshop for Outside Stakeholders, Facebook, in Menlo Park, Cal. (Aug. 8, 2019) (Chatham House Rule discussion around risks of various Board selection mechanisms).

188. Facebook Consultancy Report, *supra* note 172, at 20.

189. *Id.* at 19–20.

“stress[ed] that the Board ‘doesn’t need people with public profiles.’”<sup>190</sup> Feedback on subject-matter expertise of Board members generally fell into two camps: those who felt Board members should be versed in a specific subject matter like human rights, freedom of expression, law, or technology, and those who favored requiring more general decision-making expertise.<sup>191</sup>

The final decision in the Charter of how to balance professional qualifications and diversity pointed toward an adjudicatory rather than a representative model:

[M]embers must possess and exhibit a broad range of *knowledge*, competencies, diversity, and *expertise* . . . [and] must have demonstrated *experience* at deliberating thoughtfully and as an open-minded contributor on a team; be *skilled at making and explaining decisions* . . . ; and have *familiarity* with matters relating to digital content and governance, including free expression, civic discourse, safety, privacy and technology.<sup>192</sup>

The Charter also instructs the Board to split into subsections, termed panels, when reviewing cases.<sup>193</sup> Conversations about the panels have operated on the assumption that each will be comprised of five Board members, but the Charter does not specify a number.<sup>194</sup> The Charter does specify, however, that each panel will contain “at least one member from the region” where the case arose.<sup>195</sup>

This decision to mandate a “local representative” was heavily debated. In early June 2019, a cross-function meeting of over twenty people discussed the feasibility of including a mandatory local member on each panel.<sup>196</sup> “What do we mean by local? One person from the [United States] can’t represent all of the [United States]. So what do we mean? Does it have to be someone simply versed in the cultural issues at stake?” asked one participant.<sup>197</sup> Others worried about

190. *Id.* at 19.

191. Somewhat predictably, in subject-matter suggestions, respondents tended to strongly favor the importance of people in their own area of expertise. For example, human-rights scholars insisted that the Board be comprised primarily by those who were experts in human rights; lawyers felt the board should be made up of lawyers or judges; and engineers and technologists deemed deep technical understanding necessary. *See id.*

192. *Oversight Board Charter*, *supra* note 185, art. 1, § 2 (emphasis added).

193. *Id.*

194. *Id.*

195. *Id.*

196. Observation of Cross Function Meeting with Members of Governance Team, Facebook, in Menlo Park, Cal. (June 5, 2019).

197. *Id.* (quoting Heather Moore).

the impracticality of looping in a regional representative given the speed at which decisions would have to be made. Finally, others emphasized that such concerns were no different for geography than for any other cultural factors.<sup>198</sup> By the end of that June meeting, the idea of a local panel member seemed to be put aside, but in early August it reemerged as a central element of the Charter.<sup>199</sup> In follow-up questions, a Governance Team member explained that the idea returned because it “was something we heard over and over again in talking to stakeholders. It was really important to people.”<sup>200</sup> The drafters imagined a “scale up” solution to the problem of finding a regional panel member—defining the “region” broadly—rather than a “scale down” solution that would attempt to find the perfect local representative for each case.<sup>201</sup>

The Charter also outlines the selection process for members. Initial selection presented a chicken-and-egg problem. Though the Board could eventually self-select, the question of how to select initial Board members was difficult. If Facebook selected the initial Board, it undermined its goal of having the Board be fully independent. A selection committee to choose the initial Board members was also no solution, merely perpetuating the same problem of Facebook “selecting the selectors.”<sup>202</sup> Direct voting was not only practically infeasible across languages and cultures, but also unlikely to lead to a productive outcome.<sup>203</sup> Ultimately, the Governance Team decided that Facebook’s initial selection of members was inevitable but endeavored to reduce its impact by selecting an initial cadre, who would then select additional members. In an August workshop about the Charter, a few experts expressed deep skepticism about this approach, argu-

---

198. *Id.*

199. *Compare id.* (describing consensus vote “leaning more towards not regional” representation on panels), *with* Oversight Board Charter, *supra* note 185, art. 3, § 2.

200. Observation at Charter Workshop for Outside Stakeholders, *supra* note 187 (quoting Zoe Darmé).

201. *Id.*; *see also* Oversight Board Bylaws, FACEBOOK (Jan. 2020), [https://about.fb.com/wp-content/uploads/2020/01/Bylaws\\_v6.pdf](https://about.fb.com/wp-content/uploads/2020/01/Bylaws_v6.pdf) [<https://perma.cc/JK4W-YGMK>].

202. Interview with Zoe Darmé et al., *supra* note 163 (debating various methods of Board membership selection).

203. *See, e.g.,* Katie Rogers, *Boaty McBoatface: What You Get When You Let the Internet Decide*, N.Y. TIMES (Mar. 21, 2016), <https://www.nytimes.com/2016/03/22/world/europe/boaty-mcboatface-what-you-get-when-you-let-the-internet-decide.html> [<https://perma.cc/YW52-AHZ5>].

ing that without more direction the Board would become “chaos” and that research in this area showed that people are drawn to selecting “people most like them,” thus threatening the Board’s diversity.<sup>204</sup>

Despite these concerns, the Charter announced that the initial Board would be comprised of at least eleven members, with the potential to grow to “likely” forty,<sup>205</sup> and that Facebook would select the first Board members (“Co-Chairs”), who would then work with Facebook to select the remaining members.<sup>206</sup> Internally, the Governance Team referred to this as Phase 1, a period in which Facebook had the most direct involvement in helping to vet and decide on Board members.<sup>207</sup> Phase 2 of member selection would begin after the official launch of the Board and the selection of this initial group. Originally, it was imagined that this process would be largely independent from Facebook, with the Board Hiring Committee selecting “candidates to serve as board members based on a review of the candidates’ qualifications,” and with recommendations for Board members coming from Facebook and the public.<sup>208</sup> Difficulty in selecting Board members and delays have changed this dynamic.<sup>209</sup> Phase 2 selection of Board members actively involves Facebook and the partners hired to assist with vetting, even though this is not mentioned in the Charter.<sup>210</sup> Presumably, after the Board reaches forty members, or Phase 3, Facebook’s role is planned to return to the minimal one set out in the Charter.<sup>211</sup>

One of the final questions regarding Board composition was how a Board member’s term would end. The Charter specifies that members will serve for three-year terms, for a maximum of three terms.<sup>212</sup> Outside this limitation, the Charter states only that “trustees may remove a member before the expiration of

---

204. Observation at Charter Workshop for Outside Stakeholders, *supra* note 187; see generally LAUREN A. RIVERA, PEDIGREE: HOW ELITE STUDENTS GET ELITE JOBS (2015) (describing how elite institutions seek out people from similarly privileged backgrounds).

205. *Oversight Board Charter*, *supra* note 185, art. 1, § 1.

206. *Id.* art. 1, § 8. Trustee approval is necessary to officially appoint the members. *Id.*

207. Blue Jeans Interview with Fariba Yassae, Member, Governance Team, Facebook (Mar. 6, 2020).

208. *Oversight Board Charter*, *supra* note 185, art. 1, § 8.

209. Originally, the Board was slated to be announced one year from Zuckerberg’s announcement of the project in December 2019, but this date was pushed back, first to January 2020, then to March, then to May of that year. At this writing, the Board is expected to publicly announce members in mid-May 2020. Implications of this are discussed in more detail *infra* Part III.

210. Blue Jeans Interview with Fariba Yassae, *supra* note 207.

211. *Id.*

212. *Oversight Board Charter*, *supra* note 185, art. 1, § 3.

their term for violations of the code of conduct, but they may not remove a member due to content decisions they have made.”<sup>213</sup> The Charter does not specify a process for removal.<sup>214</sup> Generally, a detailed procedure on removal of government members is a vital part of ensuring accountability in governance.<sup>215</sup> Though many in the Governance Team argued for a more detailed removal process in the Charter, the ultimate decision was to place this process in the Bylaws.

## 2. Board Authority and Power

From the outset, the creation of an external independent Oversight Board raised questions of authority and power. Which content would fall under the Board’s purview and what power it might wield over Facebook were ever present in the minds of the Governance Team and also of outside observers.<sup>216</sup> When it was released, the Charter was intentionally vague on many of these issues. Internally, the Governance Team and those in higher positions were actively debating exactly what should fall under the Board’s jurisdiction, how cases would rise to their attention, what principles would guide decision-making, and what Facebook’s responsibilities would be following a Board decision. The Charter’s final language left many of these questions open, to be resolved by the Bylaws.<sup>217</sup>

The Charter set out a broad potential subject-matter jurisdiction for the Board, mentioning little about *what* content would be eligible or ineligible for review and primarily focusing on *who* could submit cases.<sup>218</sup> Excluding content

---

213. *Id.* art. 1, § 8.

214. More details are made clear in the *Oversight Board Bylaws*, *supra* note 201.

215. For an excellent analysis of the power of removal clauses in federal systems, see Saikrishna Prakash, *Removal and Tenure in Office*, 92 VA. L. REV. 1779 (2006).

216. This is particularly true of evelyn douek, who has analyzed the issue of the Board’s power closely. For an excellent analysis, see evelyn douek, *Facebook’s “Oversight Board:” Move Fast with Stable Infrastructure and Humility*, 21 N.C. J.L. & TECH. 1 (2019); evelyn douek, *How Much Power Did Facebook Give Its Oversight Board?*, LAWFARE (Sept. 25, 2019, 8:47 AM), <https://www.lawfareblog.com/how-much-power-did-facebook-give-its-oversight-board> [<https://perma.cc/6ARB-QR99>].

217. As will be discussed *infra*, the Bylaws sometimes limit the language of the Charter. Compare *Oversight Board Charter*, *supra* note 185, art. 2, § 1, with *Oversight Board Bylaws*, *supra* note 201, art. 2, § 1.2 (limiting and perhaps attempting to remove certain types of content from the discretion of the Board). In conversations leading up to the final version of the Charter, many Governance Team members struggled with balancing where to place specifics on scope and authority, sometimes deciding after long discussions to resolve such issues in the Bylaws. Interview with Heather Moore et al., *supra* note 187.

218. *Oversight Board Charter*, *supra* note 185, art. 2, § 1.

that was removed in compliance with local laws<sup>219</sup> and following an exhaustion of appeals through Facebook, “a request for review can be submitted to the board by either the original poster of the content or a person who previously submitted the content to Facebook for review.”<sup>220</sup> This statement is significant, as it implies that the Board has the authority to review not only content that is *removed* (“original poster of the content”) but content that is *kept up* (“person who previously submitted” content for review). As will be discussed later, this promise of reviewing removal and nonremoval of content is one that Facebook circumscribed in later documents, but this commitment in the Charter should be a meaningful point of leverage as the Board seeks to expand its powers in the future.

It is not only users who can seek redress with the Board. The Charter also sets forth a mechanism by which Facebook itself can ask the Board to review certain cases, “including additional questions related to the treatment of content beyond whether the content should be allowed or removed completely.”<sup>221</sup> The Board will have the ability to refuse Facebook’s requests for review, just as it has the “discretion to choose which requests it will review and decide upon.”<sup>222</sup> But the Charter also imagines a specialized process in which Facebook can get an “automatic and expedited review” in the case of “exceptional circumstances, including when content could result in urgent real world consequences.”<sup>223</sup>

Once the Board selects a case for review, the question becomes which documents and principles will guide the decision. The Charter determines that Board members should “review content decisions and determine whether they were consistent with Facebook’s content policies and values.” Facebook’s Values, updated in August 2019, balance “voice” with four secondary concerns: safety, privacy, authenticity, and dignity.<sup>224</sup> Content policies refer mainly to Facebook’s Community Standards. The Charter also specifies that any prior decisions by the Board on a piece of content will have “highly persuasive” precedential value.<sup>225</sup> Substantial debate took place within the Governance Team around how binding the Board’s decisions should be on the Board itself. Many were concerned that creating a binding precedential structure would limit the Board, while others

---

219. *Id.* art. 7.

220. *Id.* art. 2, §1.

221. *Id.*

222. *Id.*

223. *Id.* art. 3, § 7.2.

224. See Bickert, *supra* note 7.

225. *Oversight Board Charter*, *supra* note 185, art. 2, § 2.

worried that not establishing a strong preference for precedential decisions could damage consistency and fairness.<sup>226</sup>

A final question related to the authority of the Board concerns Facebook's obligations in reacting to Board decisions. As to a decision on any single piece of content, the Charter states that the Board's decision is "binding" and "Facebook will implement it promptly, unless implementation of a resolution could violate the law."<sup>227</sup> This obligation is strictly limited to the precise piece of content decided upon and the decision is *not* binding on any identical or similar content. Additionally, the Charter indicates that policy recommendations by the Board are "advisory" on Facebook, and that Facebook must "transparently communicat[e] about actions taken as a result."<sup>228</sup> This latter obligation to respond to the Board's policy recommendations creates a weak-form review over Facebook that could potentially become powerful by creating a public system of accountability.<sup>229</sup>

### 3. *Appellate Procedure*

The section of the Charter that highlights the procedural process for Board review sheds little light on what it will ultimately look like from the user perspective, something better accomplished by the Bylaws that were published a few months later. Rather than lay out any specific procedures, the Charter instead introduces broad guarantees. These include various commitments not only to transparency<sup>230</sup> but also to privacy and confidentiality.<sup>231</sup> The Charter, furthermore, gives direction on the number of members needed to reach a decision (a majority, though consensus is encouraged) and what a panel should include in its written decision.<sup>232</sup> It also introduces the administrative operation of the Board, which will encompass full-time staff to "support" the Board, including

---

226. Observation at Charter Workshop for Outside Stakeholders, *supra* note 187.

227. *Oversight Board Charter*, *supra* note 185, art. 4.

228. *Id.*

229. See douek, *LAWFARE*, *supra* note 216. See generally MARK TUSHNET, *WEAK COURTS, STRONG RIGHTS* (2008); Rosalind Dixon, *Weak-Form Judicial Review and American Exceptionalism* (Univ. of Chi. Law Sch., Pub. Law & Legal Theory, Working Paper No. 348, 2011), <https://ssrn.com/abstract=1833743> [<https://perma.cc/SSE3-KA3V>].

230. *Oversight Board Charter*, *supra* note 185, art. 3, § 6 (committing to publicly communicate Board decisions on a website and database).

231. *Id.* art. 3, §§ 2-3 (describing panels as anonymous and discussing development of case files in keeping with privacy and legal restrictions).

232. *Id.* art. 3, § 4.



“reviewing case submissions and coordinating outside research and statements for selected cases.”<sup>233</sup>

#### 4. *Independence Through Structure and a Binding Charter*

One of the most frequent recommendations from the Global Consultancy period was to ensure the Board’s independence from Facebook. The extent to which this recommendation was implemented will be evaluated in Part III. This Section aims only to describe the commitments in the Charter related to that goal. These commitments include the discretion given to the Board in case and member selection, the governance relationship between the Oversight Board, the Trust, and Facebook, and the structure around Charter amendments.

The ability of the Board to set its own docket is an essential part of its intellectual independence as an oversight body. Just as it would be problematic for, say, the U.S. Department of Justice to determine the Supreme Court’s case selection, Facebook’s ability to direct which cases the Board hears would create an obvious conflict, given that Facebook is the defendant in all actions before the Board.<sup>234</sup> To avoid such conflict, the Charter states that the Board will select its cases from those appealed by users.<sup>235</sup> However, the Charter also envisions “Special Procedures” in “exceptional circumstances” wherein the Board would “automatic[ally]” review a case at Facebook’s request.<sup>236</sup>

A second fundamental part of the Board’s intellectual independence is its ability to select its membership. The Charter states that after the announcement of initial members, the remaining members will be selected internally by the Board.<sup>237</sup> But as previously discussed,<sup>238</sup> the selection process of the initial cohort of members has increasingly been handled by Facebook. It is unclear if this precedent of Facebook’s initial involvement will forever taint the process and put in place long-term mechanisms that compromise members’ ability to fairly adjudicate.

---

<sup>233</sup>. *Id.* art. 3, § 1.

<sup>234</sup>. Facebook could shield itself from scrutiny through control of the Board’s docket. *See, e.g.*, Stephen B. Burbank, *The Architecture of Judicial Independence*, 72 S. CAL. L. REV. 315, 318–26 (1999); Amanda L. Tyler, *Setting the Supreme Court’s Agenda: Is There a Place for Certification?*, 78 GEO. WASH. L. REV. 1310, 1310–12 (2010).

<sup>235</sup>. *Oversight Board Charter*, *supra* note 185, art. 2, § 1.

<sup>236</sup>. *Id.* art. 3, § 7.

<sup>237</sup>. *Id.* art. 1, § 4; Blue Jeans Interview with Fariba Yassaee, *supra* note 207.

<sup>238</sup>. *See supra* Section II.B.1.

Regarding the governance relationship between the Oversight Board, Trust, and Facebook, the Charter states that the “board, the trust and Facebook will work together to fulfill the charter and the board’s purpose.”<sup>239</sup> The Board is to “review content and issue reasoned, public decisions within the bounds of this charter . . . [and] provide advisory opinions on Facebook’s content policies.”<sup>240</sup> The Trust is to fund the Board’s budget and appoint and remove members.<sup>241</sup> And Facebook is to “commit to the board’s independent oversight on content decisions and the implementation of those decisions,” funding the Trust, and appointing trustees.<sup>242</sup> Moreover, it is to “contract for services” with the Board.<sup>243</sup> The Charter thus assigns Facebook a broad but vague role.

Finally, the Charter promotes independence by giving the Board a role in the process of amending the Charter. Amendment, according to the Charter, requires the approval of a majority of the individual trustees, a majority of the Board, and the agreement of Facebook.<sup>244</sup> The decision to include Facebook in the amendment process is problematic. On the one hand, it addresses practical concerns: the Board and the Trust might otherwise decide to amend aspects of the Charter in ways that are practically or legally infeasible for Facebook. On the other hand, the necessity to procure Facebook’s agreement could prevent any meaningful amendment from taking place. Moreover, Facebook’s amendment power is magnified by the fact that two influential documents lie outside the Board’s control: the Community Standards and the substantive Values that underlie them. The Charter binds the Board to interpret these documents, and it is unclear what would happen if Facebook were to unilaterally alter them. Presumably, the Board would be beholden to Facebook’s changing guidelines.

### C. Phase III: Implementation

In the final stages of drafting the Charter, as details were removed and the document became increasingly high-level, members of the Governance Team started describing the role of the document as “putting a stake in the ground.”<sup>245</sup> The phase implied a metaphor that once “land” was claimed, details of use, ownership, or possession could be filled in later. Similarly, the role of the Charter was

---

239. *Oversight Board Charter*, *supra* note 185, art. 5.

240. *Id.* art. 5, § 1.

241. *Id.* art. 5, § 2.

242. *Id.*

243. *Id.* art. 5, § 1.

244. *Id.* art. 6, § 1.

245. Interview with Heather Moore et al., *supra* note 187 (quoting Heather Moore).

to stake out the core principles, with the rest to come later.<sup>246</sup> Accordingly, following the release of the Charter in mid-September 2019, the Governance Team began crafting the final founding documents and implementations that would support the Board and provide greater detail on processes and relationships between the three parties. This Section looks at the documents creating the Trust and the irrevocable grant to the Trust; the Bylaws and Code of Conduct; and the start of the administrative side of the Oversight Board.

### 1. *The Trust*

To be meaningfully independent, the Oversight Board must, at a minimum, have financial independence from Facebook. This point was frequently mentioned during the Global Consultancy period.<sup>247</sup> According to the Global Consultancy Report, “the most common suggestion in this regard was the establishment of a separate trust, endowment or foundation.”<sup>248</sup> The Governance Team adopted this suggestion and publicly announced in mid-September 2019 that it would create “an independent trust.”<sup>249</sup> Work on the Trust had begun months earlier, as members of the Team worked with outside consultants and Facebook’s legal division to develop a Trust agreement and create a beneficiary for the Trust—the Oversight Board and an LLC, both independent entities. On October 16, 2019, Team members traveled to Delaware to meet with the Corporate Trustee, Brown Brothers Harriman, to sign the final documents and make the Trust and Oversight Board LLC official.<sup>250</sup> These documents were released to the public on December 12, 2019, alongside an announcement that Facebook had made an “initial commitment of \$130 million” to support the Board’s operational costs and allow it to “operate for at least its first two full terms, approximately 6 years.”<sup>251</sup>

---

<sup>246</sup>. *Id.*

<sup>247</sup>. Facebook Consultancy Report, *supra* note 172, at 31.

<sup>248</sup>. *Id.*

<sup>249</sup>. Harris, *supra* note 184.

<sup>250</sup>. *Limited Liability Company Agreement: Oversight Board LLC*, FACEBOOK (Oct. 17, 2019), <https://about.fb.com/wp-content/uploads/2019/12/LLC-Agreement.pdf> [<https://perma.cc/FDW3-JXFM>] [hereinafter *Oversight Board LLC Agreement*]; *Oversight Board Trust*, FACEBOOK (Oct. 16, 2019), <https://about.fb.com/wp-content/uploads/2019/12/Trust-Agreement.pdf> [<https://perma.cc/P82R-QHQS>] [hereinafter *Oversight Board Trust Agreement*].

<sup>251</sup>. Brent Harris, *An Update on Building a Global Oversight Board*, FACEBOOK NEWSROOM (Dec. 12, 2019), <https://about.fb.com/news/2019/12/oversight-board-update> [<https://perma.cc/W925-J84G>].

According to the Trust Agreement, the Trust's stated purpose "is to facilitate the creation, funding, management, and oversight of a structure that will permit and protect the operation of an Oversight Board."<sup>252</sup> To fulfill this purpose, the Trust funds the Oversight Board LLC, which will "establish, administer and attend to the ongoing operation" of the Board.<sup>253</sup> In addition to the Corporate Trustees, there will be between three and eleven Individual Trustees, all appointed by Facebook.<sup>254</sup> Trustees are slated to serve five-year terms<sup>255</sup> with annual compensation of \$200,000.<sup>256</sup> Over two pages are devoted to tenure, review, removal, and resignation of Individual Trustees or Corporate Trustees, but none of these procedures empower any such actions by the Oversight Board itself.<sup>257</sup>

As a corporate legal document, the Oversight Board Trust Agreement is anything but boilerplate. It rests the core powers of Trustee appointment with Facebook, which might compromise the Trust's independence. But specific provisions do seem to empower Trustees. Section 2.2, for instance, states that a "vital role of the Trust by its Individual Trustees is to protect the independent judgment of the Board Members and their ability to fulfill their stated purpose."<sup>258</sup> It further declares that Facebook "has relinquished its authority over the Trust except with respect to key provisions . . . and under exceptional circumstances as a way to protect the Purpose and avoid frustrating the independent judgment of the Board."<sup>259</sup> The Agreement also gives the Trust the power to form companies.<sup>260</sup> This provision might allow the Trust and the Board to create new independent bodies as long as those entities serve the Board's goals. Finally, the Agreement makes an irrevocable grant to the Trust through the Trust Estate.<sup>261</sup> The Governance Team had long debated the exact structure of how to fund the Trust, weighing two clear but extremely different options.<sup>262</sup> The Trust could be

---

252. *Oversight Board Trust Agreement*, *supra* note 250, § 2.1.

253. *Id.*

254. *Id.* §§ 6.1-6.2.

255. *Id.* § 6.2.2(b).

256. *Id.* § 6.7.

257. *Id.* § 6.2.

258. *Id.* § 2.2.

259. *Id.*

260. *Id.* § 4.6.

261. *Id.* § 1.4.

262. Blue Jeans Interview with Heather Moore and Louis Chang, Members of Governance Team, Facebook (Mar. 4, 2020) (describing retrospectively the tradeoffs in how to structure the financial and corporate entities of the Board and Trust).

funded annually by Facebook, which would force Trustees to come back on a yearly basis for funding. This would allow annual assessment of the Board and its efficacy, but it might also harm the Board's financial independence. At the other extreme, the Governance Team discussed creating an investible endowment at the outset, allowing the Trust and the Board to be self-funding so they could run in perpetuity.<sup>263</sup> This second proposal would have eliminated the Trust's financial dependence on Facebook but could have created the problem of creating an indelible source of funding for potentially nothing, if the Board were not to succeed. The final decision was somewhere in the middle: the Trust would be funded with an irrevocable grant of \$130 million, enough to fund operations for two terms or approximately six years.<sup>264</sup>

Released alongside the Trust Agreement was the Oversight Board Limited Liability Agreement, which formally incorporated the Board as an LLC.<sup>265</sup> Much of the LLC Agreement rehashed the Charter's provisions about Board members' powers, but it also detailed the administrative side of the Board, creating, for instance, a Director of the Oversight Board. The Director is empowered to enter into service agreements and contracts on behalf of the Board; employ staff; provide for office, company, and Board member expenses; direct members' compensation; and dictate the provisions of members' contracts.<sup>266</sup> The provision of a Director of Administration added a new intermediary in the relationship between the Trustees and Board.

## 2. *The Bylaws*

The Bylaws, announced on January 28, 2020, were the Board's final founding document.<sup>267</sup> The Bylaws are split into four articles detailing the relationship between Facebook, the Trust, the Oversight Board, and the People who will be able to appeal through the Board.<sup>268</sup> The addition of "People" was significant. Though users were ostensibly the impetus for the Board and were the concern of most Governance Team deliberations, almost all of the founding documents

---

<sup>263</sup>. *Id.*

<sup>264</sup>. Harris, *supra* note 251.

<sup>265</sup>. *Oversight Board LLC Agreement*, *supra* note 250.

<sup>266</sup>. *Id.* art. 5.

<sup>267</sup>. Brent Harris, *Preparing the Way Forward for Facebook's Oversight Board*, FACEBOOK NEWSROOM (Jan. 28, 2020), <https://about.fb.com/news/2020/01/facebooks-oversight-board> [<https://perma.cc/R6JM-ZV5K>]. Included in the same document as the Bylaws is the Board's Code of Conduct. *Oversight Board Bylaws*, *supra* note 201, at Code of Conduct. The appointment of Thomas Hughes as the Director of the Oversight Board was also announced with the Bylaws. Harris, *supra*.

<sup>268</sup>. *Oversight Board Bylaws*, *supra* note 201.

focused on defining the powers and relationship of only the Board, the Trust, and Facebook. The Bylaws, for the first time, introduced and clearly defined the rights of users under the Board. This Section focuses on what the Bylaws mean for users seeking appeal through the Board. Moreover, it highlights the input that Facebook can seek from the Board. It concludes by discussing the levels of transparency in both procedure and reporting, and by reviewing the avenues of amending the founding documents.

*a. User Appeal*

According to the Charter, the Board can review a broad swath of content, but while the Bylaws reiterate the Charter's language,<sup>269</sup> they also create many exceptions to the Board's scope of review. At the Board's launch, only single-object<sup>270</sup> removals<sup>271</sup> of organic content<sup>272</sup> posted on Facebook and Instagram are eligible for review.<sup>273</sup> Within that, content decisions "pursuant to legal obligations," including those having to do with intellectual property, Facebook marketplace, fundraisers, Facebook dating, messages, and spam, are out of the scope.<sup>274</sup> The Bylaws do, however, envision a broadening of this scope, stating that "[i]n the future, people will have the opportunity to request the board's review of other enforcement actions,"<sup>275</sup> but these too are limited to specific areas that will become appealable, including pages, profiles, groups, events, advertising, and content that Facebook reviews and "ultimately allow[s] to remain on the platform."<sup>276</sup>

---

269. *Id.* art. 3, § 1; *Oversight Board Charter*, *supra* note 185, art. 4.

270. *Oversight Board Bylaws*, *supra* note 201, art. 3, § 1. The most fundamental type of organic user-generated content is a post of a photo, video, or status message—internally referred to as "simple objects." This type of content can be contrasted with "complex objects," such as a user profile, group, or page, which involve many layers of user interaction and differing levels of privacy. At launch, complex objects are not appealable to the Board.

271. As discussed in Section II.B, *supra*, the language of the Charter in art. 2, § 1 was read as creating rights of appeal for both users whose content was erroneously removed and users whose request of removal was denied. At launch, no "kept up" content will be appealable to the Board.

272. *Oversight Board Bylaws*, *supra* note 201, art. 3, § 1. Content comes from commercial advertising or "organically" from users. Commercial content includes ads or paid prioritization of content.

273. WhatsApp, Messenger, Instagram Direct, and Oculus are specifically excluded. *Id.* art. 2, § 1.2.

274. *Id.*

275. *Id.* art. 3, § 1.1.

276. *Id.*

The Bylaws also set out the following appeals process for content that falls within the Board’s limited subject-matter jurisdiction. First, a user has to exhaust Facebook’s internal appeals process. Once exhausted—and if the content stays removed—a user can request an appeal to the Oversight Board. Doing so will generate an “individualized identification number,”<sup>277</sup> which must be copied from Facebook and taken to the Oversight Board’s website.<sup>278</sup> Entering that number will generate a case submission to the Board. In addition to automatically transferring the file about the removed content to the Board, users will be able to submit additional information, such as an explanation of why they believe Facebook’s decision was incorrect, “why they believe the board should hear their case,” and why the decision could impact other Facebook users.<sup>279</sup> From the moment Facebook makes its final decision, the user and Board have ninety days to submit and process a review.<sup>280</sup>

The next stage of Board appeal primarily involves the administrative staff and Case Selection Committee members. Staff members will prepare case submissions, prioritizing cases according to the directions of the Case Selection Committee.<sup>281</sup> The Case Selection Committee, which is comprised of five Board members, will then review those cases.<sup>282</sup> A majority vote is necessary to move a case to panel review.<sup>283</sup> If a case is moved to panel review, both the user and Facebook are notified.<sup>284</sup> Then, a five-member panel is convened to hear the case.<sup>285</sup> Four members are randomly selected, and one will be a random Board member from the region “which the content primarily affects.”<sup>286</sup> Though all members’ names will be public, panels will remain anonymous.<sup>287</sup> In conducting its review, the panel can seek input from outside sources or a “pool of outside

---

277. *Id.* art. 3, § 1.2.1.

278. Blue Jeans Interview with Fay Johnson, Member of Governance Team, Facebook (Sept. 17, 2019) (describing the reference ID generated for users on Facebook that must then be taken to the Oversight Board site).

279. *Oversight Board Bylaws*, *supra* note 201, art. 3, § 1.2.

280. *Id.* art. 1, § 3.

281. *Id.* art. 1, § 3.1.2.

282. *Id.* art. 1, § 3.1.3.

283. *Id.*

284. *Id.* at art. 1, § 3.1.6.

285. *Id.*

286. *Id.* The Bylaws specify that these regions are the United States and Canada; Latin America and Caribbean; Europe; Sub-Saharan Africa; Middle East and North Africa; Central and South Asia; and Asia Pacific and Oceania. *Id.* art. 1, § 1.4.1.

287. *Id.* art. 1, § 3.1.3.

subject-matter experts” to be populated by the Board.<sup>288</sup> Before deliberation, case files will ultimately contain the statement of the user who submitted the case, a case history from Facebook, a policy rationale from Facebook, clarifying information from Facebook if requested by the Board, and any additional information gathered by the panel.<sup>289</sup> All stages of case-file preparation and advocacy are done in writing, and never in person.

Once the case file is complete and the panel is ready to deliberate, the panel will convene “privately.”<sup>290</sup> All panelists must attend and vote.<sup>291</sup> Though consensus is encouraged, it is not required, and all that is needed for a decision is a majority.<sup>292</sup> Once a decision is reached, it must be written up and must include a specific “determination on the content” as well as the reasoning and explanation for that decision.<sup>293</sup> Critically, it can also provide a policy-advisory statement for Facebook to consider.<sup>294</sup> Additionally, dissents and concurring opinions are allowed.<sup>295</sup> The panel then sends the written decision to the entire Oversight Board for review, where it must receive a majority vote in order to be formally adopted and published.<sup>296</sup> In this time, members outside the panel are welcome to ask questions, and if the decision does not receive majority approval, a new panel will be convened to rehear the case.<sup>297</sup> If a decision is approved, the Board administration will notify the persons involved and the decision will be published on the Board’s website.<sup>298</sup>

Once a decision has been written, approved, and made public, the shuttlecock returns to Facebook for implementation. Under the Bylaws, Facebook must implement the Board’s decision on the specific piece of content reviewed within seven days of the “release of the board’s decision.”<sup>299</sup> At its own discretion, Facebook can decide if there is “identical content with parallel context” implied by the Board’s decision that it can or should also take action on.<sup>300</sup> As specified by

---

288. *Id.* at art. 1, § 3.1.4.

289. *Id.* art. 1, § 3.1.5.

290. *Id.* art. 1, § 3.1.6.

291. *Id.*

292. *Oversight Board Charter*, *supra* note 185, art. 1, § 4.

293. *Oversight Board Bylaws*, *supra* note 201, art. 1, § 3.1.6.

294. *Id.* art. 1, § 3.1.7.

295. *Id.* art. 1, § 3.1.8.

296. *Id.*

297. *Id.*

298. *Id.* art. 1, § 3.2.

299. *Id.* art. 2, § 2.3.1.

300. *Id.* art. 2, § 2.3.2.



the Charter, the Board's policy recommendations are merely advisory.<sup>301</sup> Under the Bylaws, Facebook is required to return a "public response regarding any policy recommendations and follow-on actions" within thirty days of the Board's decision being received.<sup>302</sup>

*b. Facebook Case Submissions*

The Bylaws also spell out Facebook's separate avenue of appeal. Whereas the Charter only gestured to this separate avenue,<sup>303</sup> the Bylaws specify the ways in which Facebook can ask the Board for review and the extent to which the Board must listen.<sup>304</sup>

According to the Bylaws, Facebook can refer cases to the Board for review in three ways. The first is similar to users' right to appeal. Facebook can select a particular case or piece of content and request review. Here, the Board has the same discretion as with any user-presented case in determining whether to hear the case.<sup>305</sup> The second way leaves the Board with no discretion. In "exceptional" circumstances that might "result in urgent real-world consequences," the Board *must* hear Facebook's request automatically and in an expedited fashion, which is deemed to be no more than thirty days.<sup>306</sup> Facebook is then bound by the Board's decision.<sup>307</sup> Finally, separate from any specific case, Facebook can simply request policy guidance from the Board, which the Board has discretion to accept or reject.<sup>308</sup> If the Board does issue a policy recommendation, it is only advisory.<sup>309</sup>

*c. Transparency and Amendment*

Finally, the Bylaws also speak to transparency and amendment. These provisions are significant because transparency allows the Board to leverage the weight of public opinion against Facebook and because amendment procedures speak to who truly controls the documents that convey power to the Board.

---

301. *Oversight Board Charter*, *supra* note 185, art. 4.

302. *Oversight Board Bylaws*, *supra* note 201, art. 2, § 2.3.2.

303. *Oversight Board Charter*, *supra* note 185, art. 3, § 7.

304. *Oversight Board Bylaws*, *supra* note 201, art. 2, § 2.1.

305. *Id.* art. 2, § 2.1.1.

306. *Id.* art. 2, § 2.1.2.

307. *Id.*

308. *Id.* art. 2, § 2.1.3.

309. *Id.*

The Bylaws lay out a number of ways in which the Board, the Trust, and Facebook seek to be transparent. For instance, the Bylaws commit the Board to making all case decisions publicly available and archiving them in a database.<sup>310</sup> The Board must also publish an Annual Report with metrics on the number and type of cases reviewed, case submissions by region, and a report on the timeliness of decisions.<sup>311</sup> However, the Bylaws do not specify if the Board's transparency reports must include Facebook's specific case requests or requests for expedited review.<sup>312</sup>

When it comes to amending the Bylaws, Facebook retains significant power and the Board has little ability to control change to the rules outside its own administration. In matters regarding Oversight Board members, administration, and transparency reports, only a two-thirds majority of the Board is needed to amend the Bylaws.<sup>313</sup> For every other provision, Facebook must approve all changes.<sup>314</sup> This means, for instance, that the Bylaws' promise to expand the Board's scope of review in the future can only be fulfilled through an amendment to which Facebook agrees.<sup>315</sup>

### III. A NEW PRIVATE-PUBLIC PARTNERSHIP TO GOVERN ONLINE SPEECH

In early interviews discussing the body that would become the Oversight Board, Zuckerberg stated that it should "make the final judgment call on what should be acceptable speech in a community that *reflects the social norms and values of people all around the world*."<sup>316</sup> This vision is not just lofty; it is impossible. Norms and values develop in and come to define particular communities.<sup>317</sup> Facebook may wish to have one set of Community Standards that would satisfy all of its users. But a universal community with a universal set of norms underlying

---

310. *Id.* art. 2, § 2.3.2.

311. *Id.*

312. *Cf. id.* art. 1, § 4.1 (explicitly stating that policy requests and recommendations be public after the advice of the Oversight Board is received). Compare *id.* art. 1, § 4.1, with *id.* art. 2, § 2.

313. *Id.* art. 5.

314. *Id.*

315. *Id.* art. 3, §§ 1.1-1.2 (describing future appealable items, including pages, profiles, groups, events, advertising, and content reviewed by Facebook but "ultimately allowed to remain on the platform").

316. Klein, *supra* note 15 (emphasis added).

317. See Robert C. Post, *The Social Foundations of Privacy: Community and Self in the Common Law Tort*, 77 CALIF. L. REV. 957, 963-64 (1989).

those standards does not exist at the global level. And if such a global community of users were to come into existence, it would not be through Facebook *reflecting* “the social norms and values of people all around the world” in its rules but rather by Facebook *projecting* its own set of rules and values onto users.

Facebook’s projected rules and values are not neutral. For example, Facebook has a very vocal predisposition to and commitment to freedom of expression.<sup>318</sup> This is also reflected in its Values, which prioritize voice over concerns of safety, privacy, dignity, and authenticity.<sup>319</sup> Facebook’s removal of hate speech, harassment, and bullying, all of which would be permissive in a First Amendment framework, mark concessions to limiting unfettered expression.<sup>320</sup> Moreover, Facebook’s categorization of content cannot draw on universal categories because no such consensus exists. Hate speech to a Spaniard might be political free expression to a New Yorker. Obscenity to a Canadian might be art to a Korean. Fake news to an Australian might be satire to a Brazilian.

The rules that Facebook imposes on its users are thus frequently in conflict with the norms and expectations of its users. The Board might mitigate such conflict to some extent. Through policy recommendations, it can surface disagreements and suggest changes to the Community Standards. However, the Board’s idealized primary role is to simply apply rather than change Facebook’s Values and Community Standards.

This Part explores how the Board can best serve the interests of Facebook’s global users. Section III.A focuses on defining what exactly the Board is by comparing it to existing structures that are similar in process, kind, and purpose. Section III.B assesses the Board’s significance to global users and suggests ways in which the Board can grow and improve to best meet users’ needs. Finally, Section III.C looks at the Board’s long-term impact on industry standards, government, and global online speech. The language of the Charter was left inten-

---

318. See Tony Romm, *Zuckerberg: Standing for Voice and Free Expression*, WASH. POST (Oct. 17, 2019), <https://www.washingtonpost.com/technology/2019/10/17/zuckerberg-standing-voice-free-expression> [<https://perma.cc/25XT-DE4E>] (publishing a Facebook-provided text of a speech by Mark Zuckerberg at Georgetown University).

319. Bickert, *supra* note 7; *Community Standards: Introduction*, FACEBOOK, <https://www.facebook.com/communitystandards> [<https://perma.cc/7PR2-AYDD>] (outlining Facebook’s Values) [hereinafter *Facebook Values*].

320. Compare Jillian C. York, *A Complete Guide to All the Things Facebook Censors Most*, QUARTZ (June 29, 2016), <https://qz.com/719905/a-complete-guide-to-all-the-things-facebook-censors-hate-most> [<https://perma.cc/5RAS-GS2P>], with Daphne Leprince-Ringuet, *Facebook’s Approach to Content Moderation Slammed by EU Commissioners*, ZDNET (Feb. 17, 2020), <https://www.zdnet.com/article/facebooks-approach-to-content-moderation-slammed-by-eu-commissioners> [<https://perma.cc/439Y-VEFC>].

tionally vague in order to allow other platforms to adopt this structure and language for their own use. Whether this ushers in a new era of private-public governance and industry collaboration or a standardization that stifles diversity among platforms and threatens to censor speech remains unclear.

*A. What Is the Board? Adjudication and Due Process*

Currently, Facebook's governance of users' speech is an unchecked system. It is the kind of arbitrary power that, unrestrained, can trample individual rights.<sup>321</sup> This Section focuses on the analogy between the Board and a court to contextualize and understand the Board and to identify ways in which Facebook's governance might be effectively restrained by an external adjudicator.<sup>322</sup> Specifically, it focuses on courts bound by a constitution to protect individual rights.<sup>323</sup>

The analogy to courts is valuable, but also imperfect. Compared to U.S. federal courts, the Board's ability to select cases and its jurisdictional power of final authority are similar to that of the U.S. Supreme Court. The use of five-person panels that decide by majority is akin to a circuit court, though the Board's en banc approval of a decision is not required by a circuit court. Like a trial court, the Board—primarily through its administrative staff—also researches and builds cases around facts. And as with courts, parties—the appealing user and Facebook itself—submit written arguments. The Board's limited jurisdictional scope renders it similar to a court of special or exclusive jurisdiction. Specifically,

---

321. See J.W. GOUGH, *FUNDAMENTAL LAW IN ENGLISH CONSTITUTIONAL HISTORY* 207 (1955); JOHN LOCKE, *TWO TREATISES OF GOVERNMENT* 359, 398 (Peter Laslett ed., Cambridge Univ. Press 1988) (1690).

322. Similarly, DeNardis insists that we should ask what the “most effective” and ideal form of internet governance would be in a given context, and make that form as robust and rights enhancing as possible. LAURA DENARDIS, *THE GLOBAL WAR FOR INTERNET GOVERNANCE* 226 (2014).

323. See Rolf H. Weber & Shawn Gunnarson, *A Constitutional Solution for Internet Governance*, 14 *COLUM. SCI. & TECH. L. REV.* 1, 47 (2012). Imagining the Oversight Board as a court is to invite debate. This comparison is sometimes problematic for those who think of courts in formal terms as an arm of traditional nation-state *government* and those who see them as a process in *governance*. The Oversight Board is operating in the latter sense, adjudicating individual claims made on a private transnational platform, but it falls short of formal definitions as a court or legal system. Ultimately, it is unclear whether saying the Board is a *court* or *court-like* matters. See SCOTT J. SHAPIRO, *LEGALITY* 223-24 (2011) (describing “degrees of legality” in that “legality itself is not a binary property but also comes in degrees” through analogy to the United States Golf Association (USGA) and concluding that “[t]he best we can say about the USGA, therefore, is that it is like a legal system in some senses, but not in others, and leave it at that”).

it (combined with the Board’s limited transparency) suggests a comparison to administrative courts like the Social Security Administrative Court or the Foreign Intelligence Surveillance Court.

Compared to other adjudication models, the Board has a similar vision for global internet adjudication as ICANN’s Uniform Domain-Name Dispute-Resolution Policy, which sets forth a procedure for resolving domain-name disputes between registrars and customers.<sup>324</sup> And while borrowing rhetoric and procedure from public courts, the Board is at its core a private, independent arbitration system built by Facebook.<sup>325</sup>

Using the analogy of courts bound by a constitution to protect individual rights, this Section evaluates the substantive and procedural protections that the founding documents provide for users, and the extent to which the documents give the Board independence.

### 1. *Fundamental Rights*

The analogy to a constitution that guarantees substantive and procedural rights through review by an independent judiciary is crucial for understanding the founding documents that create the Board. The Charter invokes this analogy most explicitly; it was envisioned, drafted, and described as a constitutional document. But all of the founding documents inform the rights of users, the Board, and Facebook.

When it comes to the substantive guarantees for users, the Charter’s Introduction states that “[f]reedom of expression is a fundamental human right.”<sup>326</sup> But it does not expressly guarantee that right for users. Instead, it contemplates

---

324. See generally AnneMarie Bridy, *Notice and Takedown in the Domain Name System: ICANN’s Ambivalent Drift into Online Content Regulation*, 74 WASH. & LEE. L. REV. 1345 (2017) (discussing how ICANN enforces intellectual-property rights by adjudicating disputes over “domain names containing trademarked words and phrases”).

325. Attribution for this analogy to alternative dispute resolution goes to Jeff Gary. Unlike ADR, however, the Board does not foreclose litigation outside of Facebook within actual court systems in the way many ADR agreements do, and it does not address anything other than equitable remedies. See George Applebey, *What Is Alternative Dispute Resolution?*, 15 HOLDSWORTH L. REV. 20, 32 (1991). Moreover, unlike ADR’s guarantee of a procedure, the Board has discretion over whether to hear a case. See *Oversight Board Bylaws*, *supra* note 201, art. 1, § 3.1.2.

326. *Oversight Board Charter*, *supra* note 185, at Introduction.

balancing the right of expression with “authenticity, safety, privacy, and dignity.”<sup>327</sup> This formulation invokes Facebook’s Values<sup>328</sup> and also the preamble to Facebook’s Community Standards, which states: “The goal of our Community Standards has always been to create a place for expression and give people a voice. This has not and will not change.”<sup>329</sup> Though this is a strong statement of company commitment to the *principle* of free expression, it too is not a *promise of rights* to users. Nor does the preamble guarantee authenticity, safety, privacy, or dignity to users as substantive rights. Its language remains vague and noncommittal: “We want to make sure the content people are seeing is authentic”; “We are committed to making Facebook a safe place”; “We are committed to protecting personal privacy and information”; and “We believe that all people are equal in dignity and rights.”<sup>330</sup> Other founding documents, too, nowhere provide for users’ substantive rights as a constitution would.

The absence of such guarantees is important to understanding what the Board is and can be for users. First, it makes clear that the limitation on Facebook’s “arbitrary will” through the Board will not come from reviewing violations of individual users’ rights but from reviewing Facebook’s promise to uphold general principles.<sup>331</sup> Second, Facebook’s commitments are in large part still located in documents – like Facebook’s Values and Community Standards – that are amendable solely at Facebook’s discretion. Finally, though the Charter and Bylaws do not create substantive rights for users, they do articulate (albeit vaguely) Facebook’s promises to users on five critical concerns around online speech and instruct the Board to also “pay particular attention to the impact of removing content in light of human rights norms protecting free expression.”<sup>332</sup> Though far from a broad or robust adoption of human rights, this is better than nothing, especially if the Board creates a meaningful method for users to enforce such rights.

Creating a means for users to seek redress when they find that their “voice” has been silenced should be the Board’s core purpose. The Charter states that “[t]he purpose of the board is to protect free expression by making principled, independent decisions about important pieces of content and by issuing policy

---

327. *Id.*

328. *Facebook Values*, *supra* note 319.

329. *Id.*

330. *Id.*

331. It is unclear what practical difference this distinction will make for users’ rights – whether Facebook’s formulation will result in speech decisions that are “better” or “worse” than a system like that of the United States, which takes the individual-rights approach.

332. *Oversight Board Charter*, *supra* note 185, art. 2, § 2.

advisory opinions on Facebook’s content policies.”<sup>333</sup> While the scope of reviewable content is relatively narrow, especially at inception, its authority over Facebook is direct and clear. The Charter states “Facebook will commit to the board’s independent oversight on content decisions and the implementation of those decisions.”<sup>334</sup> This grant of judicial deference is crucial. But equally crucial for ensuring users’ effective redress is the Board’s ability to provide meaningful due process through transparency and independence.<sup>335</sup>

## 2. Transparency

Transparency is a prerequisite of due process because if a legal system’s rules and processes are opaque, citizens can neither avail themselves of the process, nor reliably know if they have violated a rule.<sup>336</sup> To allow for due process, citizens need (1) notice of what the rule is, (2) notice that they have allegedly violated the rule, (3) notice that a procedural system exists for review of the alleged violation, (4) notice of what that procedural system entails, and (5) notice of the ultimate decision reached. In most of these respects, Facebook and the Board provide a robust level of transparency. Since spring of 2018, Facebook has made

---

333. *Id.* at Introduction.

334. *Id.* art. 5, § 3.

335. Similarly, the U.S. Supreme Court has “stated that the core rights of due process are notice and hearing.” Martin Redish & Lawrence Marshall, *Adjudicatory Independence and the Values of Due Process*, 95 YALE L.J. 455, 475-76 (1986) (citing *Cleveland Bd. of Educ. v. Loudermill*, 470 U.S. 532, 542 (1985)). In a technological capacity, these rights have also been brilliantly contemplated early on by Danielle Keats Citron in *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1249 (2008), which argued that government administrative agencies’ implementation of technology should be subjected to due process; Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 23 (2014), which resolved that due process is necessary in providing automated credit scores; and Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 93 (2014), which addressed due process in the context of predictive privacy analytics.

336. Redish & Marshall, *supra* note 335, at 485 (quoting LON FULLER, *THE MORALITY OF LAW* 39 (1969) (“Certainly there can be no rational ground for asserting that a man can have a moral obligation to obey a legal rule that does not exist, or is kept secret from him.”)). *But see* David Pozen, *Transparency’s Ideological Drift*, 128 YALE L.J. 100, 156 (2018) (analyzing and discussing private corporations’ use of transparency to access valuable information); Lawrence Lessig, *Against Transparency*, NEW REPUBLIC (Oct. 9, 2009), <https://newrepublic.com/article/70097/against-transparency> [<https://perma.cc/QXX4-SBK9>] (discussing the drawbacks of transparency in government for individual privacy and institutional legitimacy).

its Community Standards and content policy public.<sup>337</sup> This allows users to see how the rules are enforced against them, even if there is currently no way to independently verify if the public-facing Community Standards are in fact the ones enforced on users.<sup>338</sup> Moreover, significant deliberation went into creating the Board and its rules during the Global Consultancy. Regarding the second notice requirement, Facebook provides users with both notice of rule violation and appeal. Moreover, the founding documents transparently communicate the structure and procedures of the Board and the Trust. Once the user has appealed to the Board, the Bylaws provide for notice when the appeal is accepted for review and when a decision has been reached on the case.<sup>339</sup> Perhaps most crucially, the Charter instructs panel determinations to include the final directive on the content along with “a corresponding plain language explanation of the board’s rationale.”<sup>340</sup> These decisions will be publicly posted on the Board’s website, which will operate in eighteen languages.<sup>341</sup> Finally, the Bylaws also instruct the Board to publish annual reports on the number and type of cases and the timeliness of Facebook’s implementation and response on policy advisory statements.<sup>342</sup> This amounts to significantly more transparency around procedure than Facebook had provided before.

Despite these advances, however, some areas of the Board’s process remain opaque. While the members will be public, the individual members on a given panel will serve anonymously. Furthermore, while the Case Selection committee “will document its selection criteria, as well as the volume and types of cases . . . the board has selected for review” and publish them in annual reports, the actual process will remain cloaked.<sup>343</sup>

---

337. Facebook has made a number of other significant moves toward transparency around content moderation – most significantly, the tabulation and publication of quarterly transparency reports around content removal, appeals, and restoration site-wide, and the creation of the Public Policy Forum. Bickert, *supra* note 14. While these are massive efforts, they still have limitations. The Transparency Report, for instance, does not make it particularly easy to find the actual numbers and percentages of removals, appeals, and restoration. The Public Policy Forum, while publishing reports and letting in academic observers, is still generally a closed system private to Facebook.

338. Ideally, this will be a problem that the Oversight Board is built to address.

339. *Oversight Board Bylaws*, *supra* note 201, art. 1, § 3.

340. *Oversight Board Charter*, *supra* note 185, art. 3, § 4.

341. *Oversight Board Bylaws*, *supra* note 201, art. 1, § 4.

342. *Id.*

343. *Id.* art. 1, § 2.1.3. This is not dissimilar to other similar systems of case selection. See, e.g., Sonja R. West, *The Supreme Court’s Limited Public Forum*, 73 WASH. & LEE L. REV. ONLINE 572 (2017) (describing the complicated semipublic sphere of the Supreme Court).



### 3. *Independence*

In addition to transparency, due process also requires “the participation of an independent adjudicator.”<sup>344</sup> This Section will assess the extent to which the founding documents establish the independence of the Board by using a three-part framework: (1) jurisdictional independence, (2) intellectual independence, and (3) financial independence.

#### *a. Jurisdictional Independence*

Jurisdictional independence describes the Board’s powers and obligations to others and its authority to govern itself. The Board’s authority is predicated on its creation by the Oversight Board LLC, which in turn derives authority from the Oversight Board Trust. The Trust Agreement and the Oversight Board LLC Agreement, together with the Charter and the Bylaws, define the authority, limitations, and relationships of each entity as well as basic procedures for their interaction.

Until the Board’s creation, Facebook had exclusive control over the jurisdiction of its platform—like a territorial sovereign.<sup>345</sup> But the very concept of the Board, which contemplates an independent entity with binding control within the jurisdiction of Facebook, requires Facebook to convey a portion of its authority to another entity with limited jurisdiction. Accordingly, the Trust Agreement “relinquishes [Facebook’s] authority over the Trust,” with the “goal to ensure the proper administration and structure” to allow the Board to “render its independent judgment.”<sup>346</sup> In order to ensure the establishment of the Board within this new jurisdiction, the Trust agreement tethers the purpose of the Trust to “the creation, funding, management, and oversight of a structure that will permit and protect the operation of an Oversight Board,” and also mandates the creation of the Oversight Board LLC.<sup>347</sup> Facebook’s delegation of authority to the Trust is far from absolute. Facebook “has relinquished its authority over the Trust *except with respect to key provisions stated herein and under exceptional circumstances*.”<sup>348</sup> Importantly, this clause does not specifically define *which* key provisions are limiting the Trust’s authority.

---

344. Redish & Marshall, *supra* note 335, at 475.

345. This is limited of course by its obligations under the laws of the nation-states in which it operates.

346. *Oversight Board Trust Agreement*, *supra* note 250, § 2.2.

347. *Id.* §§ 2.1, 2.3.

348. *Id.* § 2.2 (emphasis added).

In regard to the “business and affairs” of the Oversight Board LLC, the Trust has “sole and absolute discretion.”<sup>349</sup> This includes specific administrative powers, such as the ability to enter into Board member contracts and service agreements, remove and appoint members and staff, issue payment and compensation, and provide office and research expenses.<sup>350</sup> The Trust and the LLC each have administrative representatives who facilitate requests from the Board and disbursements from the Trust, respectively. For the Trust, this is the Director of the LLC Administration, who is appointed by the Corporate Manager.<sup>351</sup> For the LLC, it is the Director of the Oversight Board, who is appointed by the Individual Managers.<sup>352</sup> Critically, in regards to these administrative financial and legal powers, the Agreements have no mention of Facebook.<sup>353</sup> At least in regard to administrative matters and operation, the Board and Trust largely self-govern.

The majority of the powers granted to the Trust can only be exercised if a request is made from the Oversight Board Director, but there are some important exceptions. The most significant of these is the power of the Trustees to form “one or more companies for the purpose of effectuating the Purpose of this Trust.”<sup>354</sup> “Company” is defined broadly as “any corporate body (of whatsoever kind), partnership, limited liability company, foundation, organization.”<sup>355</sup> This is a significant power that could allow the Trust, for instance, to create a legislative body to provide more direct user representation and perhaps issue new content-moderation rules. However, exercising this power requires Facebook’s consent.<sup>356</sup> Moreover, it is unclear how this would be initiated as it requires no request by the Board or LLC for execution. It is therefore uncertain whether the Board, which is the entity most likely to want such a legislative body, even has the power to request it.

Once the Board is created, Facebook must also expressly delegate power to the Board itself, not just the Trust. The Charter and the Bylaws state that

[the] board’s resolutions of each case will be binding and Facebook will implement it promptly . . . . When a decision includes policy guidance or

---

349. *Oversight Board LLC Agreement*, *supra* note 250, § 5.1.

350. *Id.* § 5.3; see *Oversight Board Trust Agreement*, *supra* note 250, § 4.10-4.11.

351. *Oversight Board LLC Agreement*, *supra* note 250, § 5.2(b)(ii).

352. *Id.* § 5.2(b)(i).

353. This is not true for the financial management of the Trust Estate or the ability to amend founding documents, to be explained *infra*.

354. *Oversight Board Trust Agreement*, *supra* note 250, § 4.6.

355. *Id.* § 10.1.6.

356. *Id.* § 4.6.

a policy advisory opinion, Facebook will take further action by analyzing . . . [and] considering it . . . and transparently communicating about actions taken as a result [within thirty days].<sup>357</sup>

The Charter also establishes the Board's obligations to Facebook: to review content appeals brought by users, or by Facebook, and issue "reasoned, public decisions . . . [and] provide advisory opinions on Facebook's content policies."<sup>358</sup> The Board's unilateral authority to order Facebook to enforce an equitable remedy makes it seem to have authority akin to that of a constitutional court. Like the government in such a system, Facebook is always the defendant, and the court adjudicates its adherence to the value of free expression.

But just as with Facebook's creation of the Trust, the mere granting of some limited authority tells us little about how the Board can wield it. The strength of the Board's authority will be primarily dictated by the scope of material that Facebook allows the Board to review and users to appeal. Here, as we saw, the Charter and Bylaws set different expectations. Unlike the Charter's commitment to a broad scope, the Bylaws designate only individual pieces of content like posts, photos, videos, and comments that have been removed by Facebook as initially eligible for appeal.<sup>359</sup> Moreover, in order to exercise the right to appeal, the user must "have an active Facebook or Instagram account."<sup>360</sup> Other Facebook services—WhatsApp, Messenger, Instagram Direct, and Oculus—are explicitly excluded.<sup>361</sup> The Bylaws add that "[i]n the future, people will have the opportunity to request the board's review" for content reviewed by Facebook "and ultimately allowed to remain on the platform."<sup>362</sup> But it is unclear how this promise of a broader scope in the future can be enforced. It appears that such an initiative would have to be considered through amendment to the Bylaws. Here, neither the Board, nor the Trust, nor users appear to have significant recourse. Instead, the Bylaws give only Facebook the authority to amend the relevant section "after consulting with the board and the trustees."<sup>363</sup> Thus, although much content will be available for review, the Bylaws and their process of amendment

---

357. *Oversight Board Bylaws*, *supra* note 201, art. 1, § 2.3.2; *Oversight Board Charter*, *supra* note 185, art. 4.

358. *Oversight Board Charter*, *supra* note 185, art. 5, § 1.

359. *Oversight Board Bylaws*, *supra* note 201, art. 3, § 1.1.

360. *Id.* art. 2, § 1.1.2.

361. *Id.* art. 2, § 1.2.1.

362. *Id.* art. 3, § 1.1.2.

363. *Id.* art. 5, § 1.

place significant and meaningful limits on the Board's ability to self-govern and on its overall jurisdictional independence.

*b. Intellectual Independence*

A critical piece of establishing the independence of an adjudicatory body is ensuring that members are without, or are willing to set aside, personal bias on matters before them.<sup>364</sup> Intellectual independence refers to the procedural affordances made to a body to accomplish this. It is characterized by a body's freedom from external control when making decisions and its ability to avoid bias through such procedures as the selection and removal of members.

With the exception of the inaugural member-selection process, the Board's "Membership Committee" has "sole responsibility" over the recruitment, interview, and selection of new members.<sup>365</sup> The only method of outside input is a public portal through which members of the public, Facebook, and Board members can submit candidate recommendations.<sup>366</sup> As feedback from the Global Consultancy had urged,<sup>367</sup> any current or former employee of Facebook, and any person engaged in litigation with Facebook, is disqualified from candidacy to avoid conflicts of interest.<sup>368</sup> If candidates are successfully screened and interviewed by the Committee, their appointment is presented to the full Board for review, where it requires a majority vote.<sup>369</sup> The Bylaws seem to also give Trustees some oversight authority, instructing them that the appointment of new members should ensure "that the board maintains geographic balance."<sup>370</sup> Feedback from the Global Consultancy expressed strong concern about the intellectual independence of the Board due to Facebook's role in the initial selection process.<sup>371</sup> But although the selection of the initial Board included Facebook, Facebook's exclusion from the process going forward, the disqualification of current or former employees from candidacy, and the clear procedural affordances to the Board and Trust suggest that they will be able to develop and maintain intellectual independence in the long-term.

---

364. See Redish & Marshall, *supra* note 335, at 500-02.

365. *Oversight Board Bylaws*, *supra* note 201, art. 1, § 1.2.2.

366. *Id.*

367. Facebook Consultancy Report, *supra* note 172, at 20-21.

368. *Oversight Board Bylaws*, *supra* note 201, at app. A..

369. *Id.* art. § 1.2.2.

370. *Id.* art. 4, § 2.1.1.

371. Facebook Consultancy Report, *supra* note 172, at 17-19.

Ensuring the Board's intellectual independence also requires giving it significant control over the ability to remove members. Feedback from the Global Consultancy stressed the need to empower the Board in this regard and to define what would qualify as a violation of the terms of appointment.<sup>372</sup> The founding documents implement this feedback. They vest formal removal power of Board members in the Trust.<sup>373</sup> The Trust can receive removal requests from "the board, the director, or the public," a list from which Facebook is notably absent.<sup>374</sup> But it is unclear which party, upon notice of an alleged violation, is responsible for investigating the claims. On the one hand, the Bylaws give the Trust the ability to "receive, verify, and act upon requests to remove members based on violations of the code of conduct."<sup>375</sup> On the other hand, they direct that removals "require a two-thirds vote of the board . . . subject to approval of the Trustees."<sup>376</sup> Regardless, the founding documents are clear that a member may be removed before the end of a term only for violations of the Code of Conduct, and Trustees "may not remove a member due to content decisions they have made."<sup>377</sup> The Code of Conduct, published with the Bylaws, "sets forth the rules and guidelines that govern the personal and professional conduct for all board members . . . and staff of the Oversight Board."<sup>378</sup> It also recommends a method of internal reporting of suspected misconduct for Board members or staff.<sup>379</sup> Thus, the removal of members lies entirely with the Board and Trust, and not with Facebook.

Finally, the Board must be intellectually independent of Facebook not only in its control over *who* is deliberating through appointments and removals but also in its ability to decide *what* it deliberates on. Within its jurisdiction, that is, the Board must be able to review, research, and select cases for deliberation without interference from Facebook. For user appeals, a "Case Selection" committee is empowered to set and apply criteria for prioritizing cases. A majority of the

---

372. *Id.* at 21.

373. *Oversight Board Bylaws*, *supra* note 201, art. 4, § 2.1.3.

374. *Id.* art. 4, § 2.1.2.

375. *Id.*

376. *Id.* art. 1, § 1.1.2. It might be of note that this provision is found with the "Membership Committee" section of the Oversight Board section of the Bylaws.

377. *Oversight Board Charter*, *supra* note 185, art. 1, § 8.

378. *Oversight Board Bylaws*, *supra* note 201, at Code of Conduct.

379. *Id.* at Code of Conduct, art. 11 ("Members and staff must immediately report violations or suspected violations of this code of conduct to the director. . . . In the case of a violation involving the director, it should be reported to the trustees.").

Committee must agree for a case to go to panel review.<sup>380</sup> Were this the only means of the Board reviewing cases, it would be an entirely independent process. But it is not. As we saw, Facebook also has power to seek review from the Board through “Special Procedures.”<sup>381</sup> The Charter gives Facebook authority “in exceptional circumstances” to “send cases to the board for an automatic and expedited review, which the board will accept and review as quickly as possible.”<sup>382</sup> Facebook also has the ability to “request” “advisory” “policy guidance” from the Board.<sup>383</sup>

The effect of these provisions on the Board’s intellectual independence is hard to predict. Focusing on the Board’s role as an external appeals body for users, Facebook’s special powers may have no bearing on the Board’s intellectual independence. However, Facebook could co-opt the time, energy, and resources of the Board in pursuit of its own issues and at the expense of the Board’s mandate to consider users’ appeals. Whether Facebook will use its special powers in these ways remains to be seen. But on the whole, the founding documents ensure a robust level of intellectual independence.

### c. *Financial Independence*

Financial independence is another key component of an adjudicative body’s independence. The most central worry regarding the Board’s financial independence derives from the fact that Facebook provides all of the Board’s funding. If members’ compensation comes from the very entity that they are supposed to be critically reviewing, there is a strong conflict of interest. Members could easily feel obligated to provide decisions supportive of Facebook, in order to ensure their continuation on the Board. This is the central reason the Trust was created. Through the Trust Agreement, the Trust legally accepts the “Initial Trust Estate” — a gift of \$130 million — from Facebook.<sup>384</sup> This gift is irrevocable.<sup>385</sup> The

---

380. *Id.* art. 1, § 3.1.3.

381. See *Oversight Board Charter*, *supra* note 185, art. 3, § 7.

382. *Id.*

383. *Id.*

384. *Oversight Board Trust Agreement*, *supra* note 250, § 1.2; Todd Spangler, *Facebook Pledges \$130 Million to Fund Content Oversight Board as It Hits Delays*, VARIETY (Dec. 12, 2019), <https://variety.com/2019/digital/news/facebook-130-million-fund-content-oversight-board-1203434228> [<https://perma.cc/XAJ5-MWX8>].

385. *Oversight Board Trust Agreement*, *supra* note 250, §§ 1.4, 1.6. It is worth noting that, at its inception, the Trust Estate is more of a gift than an endowment. The distinction relates to who controls the management of the investment of the Estate, which at the outset is managed by

purpose of the Trust is to create, fund, manage, and oversee the “structure that will permit and protect the operation of an Oversight Board.”<sup>386</sup> This gives Trustees the power to set up the LLC and provide it with funds for Board member and staff compensation, expenses, and other financial obligations.<sup>387</sup> In addition to structurally isolating the Board from financial conflicts of interest, the Trust Agreement also establishes protections to guard against self-dealing by the Trustees.<sup>388</sup> These safeguards, together with the irrevocable grant of funding and creation of the Trust to administer those funds, give the Board meaningful financial independence.<sup>389</sup>

### B. What Does the Board Mean for Global Users?

Having discussed the nature of the Oversight Board, this Section considers how the Board, as a system of adjudication, will work for users. Predictions of the Board’s potential for users can be roughly categorized into three camps—corresponding to different groups of critics: pessimists, realists, and optimists.<sup>390</sup>

---

the Corporate Trustee. Significantly, the *Trust Agreement* describes how the Trustees gain control of the investment over time through the creation of an Investment Committee. *Id.* §§ 4.1, 8.

<sup>386.</sup> *Id.* § 2.1.

<sup>387.</sup> *Oversight Board LLC Agreement*, *supra* note 250, § 5.3.

<sup>388.</sup> *Oversight Board Trust Agreement*, *supra* note 250, §§ 3.4, 5, 6.2, 6.6-7, 6.12.

<sup>389.</sup> See Redish & Marshall, *supra* note 335, at 494-500 (describing how such insulation can be created for judges through tenure and salary protections).

<sup>390.</sup> See Klonick, *supra* note 9, at 1613-14 and accompanying footnotes. The opinions expressed in this Section come from feedback received in presenting this project at the *Fordham International Law Journal* Symposium, Fordham University School of Law (Oct. 4, 2019); Faculty Talk, Loyola Law School, Loyola Marymount University (Oct. 30, 2019); Technology Colloquium, New York University School of Law (Nov. 4, 2019); Tech & Regulations Roundtable, Scalia Law School, George Mason University (Nov. 8, 2019); First Amendment & Technology, Kline School of Law, Drexel University (Nov. 14, 2019); Program on IP & Technology Law Lecture, Notre Dame Law School (Nov. 20, 2019); Harmful Online Activity & Private Law Conference, University of Haifa (Dec. 5, 2019); Oversight Board Paper Workshop, Data & Society (Dec. 12, 2019); Oversight Board Paper Workshop, Hans-Bredow-Institut, University of Hamburg (Jan. 6, 2020); The Facebook Oversight Board, London School of Economics (Jan. 8, 2020); *Ohio State Technology Law Journal* Symposium, Ohio State University Moritz College of Law (Jan. 17, 2020); Faculty Talk, Kennedy School of Government, Harvard University (Feb. 7, 2020); Tech Law & Policy Colloquium: Julie Cohen, Georgetown University Law Center (Feb. 20, 2020); Digital Life Seminar: The Facebook Oversight Board, Cornell Tech, Cornell University (Feb. 27, 2020). Each of these meetings was held under Chatham House Rules or off the record and no direct attribution is given to any speaker or location. A

### 1. *The Pessimists*

Critics of the Board worry that it will reduce Facebook's incentives to accurately moderate content. The existence of the Board, for instance, might lead Facebook to remove more content because it can now refer frustrated users to the Board. Or it could have the opposite effect. If the Board proved to be a harsh critic, Facebook might leave up more reported content in order to remove it from the Board's jurisdiction.

From a regulatory perspective, many see Facebook's creation of the Board as a display of self-regulation in order to stave off actual government regulation. The Board might also be a purposeful distraction of public attention away from more critical technological concerns like algorithmic content management or microtargeting. Most critically, some see the Board as simply serving Facebook's business interests, describing Facebook's creation of the Board as a co-optation of principles of due process in order to shield and legitimate its capitalistic enterprises.

Perhaps the most common criticism of the Board is that it is a public-relations mechanism for Facebook to defer responsibility and does not do enough for users. This criticism goes mostly to the Board's independence. While the structure of the Board through the LLC and Trust give it robust financial independence from Facebook, its limited subject-matter jurisdiction threatens to undermine its potential impact. Additionally, while the Board's membership-selection process will eventually be entirely self-governed, it is unclear when that will occur. Both the Charter and Bylaws provide for Facebook's role in choosing the first cochair and then working with those cochair to "form the initial board," which "thereafter" will become the sole responsibility of the Membership Committee.<sup>391</sup> But they leave open when exactly the initial Board will have been created. The Charter contemplates Facebook's involvement continuing for the "remainder of the board seats,"<sup>392</sup> which is troublingly vague — especially given that the size of the Board is fluidly defined as "no less than eleven members," and when fully staffed "likely to be forty members."<sup>393</sup> Neither the Charter nor By-

---

number of critiques, hopes, and observations about the Board repeated themselves across discussions. Accordingly, the ideas expressed in this Section are an anonymized summary that reflects the discussions, questions, comments, and observations that were gathered through the outside feedback from these presentations and workshops.

391. *Oversight Board Bylaws*, *supra* note 201, art. 1, § 1.2.2; *Oversight Board Charter*, *supra* note 185, art. 1, § 8.

392. *Oversight Board Charter*, *supra* note 185, art. 1, § 8.

393. *Id.* art. 2, § 1.



laws describe an event or time that will signal the end of the initial Board's creation. The Governance Team initially envisioned Facebook's role in member selection to end upon the official announcement of eleven to fifteen Board members.<sup>394</sup> But the date for that announcement has been delayed and the projected number of members has changed multiple times.<sup>395</sup> In recent conversations, Facebook officials have projected the Board's initial size to lie above fifteen and "in the twenties" and have admitted that Facebook's lack of involvement after the announcement "might not happen."<sup>396</sup> Facebook's prolonged involvement in the inaugural selection process poses significant long-term risks to the intellectual independence of the Board.<sup>397</sup> Finally, critics have also raised transparency concerns about the initial Board selection and the lack of public deliberations.<sup>398</sup>

Some of these criticisms can be addressed. For instance, the Board could remove the panels' anonymity and increase transparency and public communications by amending the Bylaws.<sup>399</sup> An amendment to the Charter or Bylaws could also change the scope of content for the Board's review. This latter amendment,

---

394. Interview with Andy Pegram, Member Governance Team, Facebook, in Menlo Park, Cal. (Oct. 21, 2019).

395. See Zuckerberg, *supra* note 16 (stating in November 2018 that an independent oversight body would be created "[i]n the next year"); Telephone Interview with Zoe Darmé, Member, Governance Team, Facebook (Apr. 3, 2020) (stating that current projected deadline for announcing Board members is early May 2020, but that the timeline changes daily); Telephone Interview with Carolyn Glanville, Commc'ns Manager, Governance Team, Facebook (Nov. 4, 2019) (providing an update that there would be no announcement of Board members in December, only an announcement of Trust, LLC documents, human-rights report, and monetary commitment, and that Board members would likely be announced in January 2020); Telephone Interview with Carolyn Glanville, Commc'ns Manager, Governance Team, Facebook (Dec. 9, 2019) (providing an update that Board members would likely not be announced until February, and instead the Bylaws and Oversight Board Administrative Director would be announced in January); Interview with Carolyn Glanville, Commc'ns Manager, Governance Team, Facebook, in Menlo Park, Cal. (Feb. 27, 2020) (stating that Board members would likely not be announced until March or April, and perhaps as late as May 2020).

396. Blue Jeans Interview with Fariba Yassae, *supra* note 207 (quoting Fariba Yassae).

397. See *Recommendations for the Facebook Content Review Board*, STAN. L. SCH. L. & POL'Y LAB 10-11 (Spring 2019), [https://www-cdn.law.stanford.edu/wp-content/uploads/2019/07/Stanford\\_Policy\\_Lab\\_Recs\\_for\\_Facebook\\_Content\\_Review\\_Board\\_FINAL.pdf](https://www-cdn.law.stanford.edu/wp-content/uploads/2019/07/Stanford_Policy_Lab_Recs_for_Facebook_Content_Review_Board_FINAL.pdf) [<https://perma.cc/N6BE-WPFQ>] [hereinafter *Stanford Law Report*].

398. *Id.* at 10, 33-38.

399. The Board can amend the Bylaws as long as they do not conflict with the Charter. *Oversight Board Bylaws*, *supra* note 201, art. 5.

however, depends on Facebook's consent.<sup>400</sup> And the same is true for other critical decisions, such as the appointment of Trustees.<sup>401</sup>

Missing from the amendment process is a critical voice: users. While the founding documents provide users with procedural rights, they do not provide them with direct accountability mechanisms. The rights of the user are only to *use* the system, not to give input on its structure, form, procedures, or underlying rules and values. Disappointment in the limits of user rights is therefore understandable. But it might be premised on an unrealistic expectation of accountability. While users might expect *democratic* accountability, a more realistic outcome is *participatory* empowerment.<sup>402</sup> The creation of the Board supports this thesis. It originated in Facebook users advocating for an independent appeals system and Facebook listening. Moreover, the process of creating the Board during six months of Global Consultancy likewise empowered users through participation.

## 2. *The Realists*

Many commenters expressed concern that whether or not the Board was a good idea, it simply would not work in practice. Their chief concern was scale.<sup>403</sup> In the second and third quarters of 2019, 30.8 million pieces of content remained down even after appeal.<sup>404</sup> This equates to approximately 170,000 pieces of content per day that would be potentially eligible for Board review.<sup>405</sup> Even if just one percent of those cases were appealed, that would still amount to around 1,700 cases a day. This is a daunting number of cases to process by a Selection Committee that forms a subset of an eleven-to-forty-person Board, even with staff support. Meaningfully processing the volume of cases submitted will be challenging—especially given the timeline of ninety days from filing to decision on appeal.<sup>406</sup>

---

400. *Id.*

401. *Oversight Board Trust Agreement*, *supra* note 250, §§ 1.5, 6.2.

402. See *Stanford Law Report*, *supra* note 397 at 11.

403. *Id.* at 19.

404. *Community Standards Enforcement Report*, *supra* note 30 (describing the calculation methodology—subtracting the total number of pieces of content appealed from the total number of pieces of content restored).

405. *Id.* (describing the calculation method, which took the number of pieces that remained down after appeal and were eligible for Board review and divided it by the number of days in half a year).

406. *Stanford Law Report*, *supra* note 397, at 19-20.

Additional practical concerns arose with regard to the system's potential manipulation, for instance by trolls coordinating to submit identical or similar requests and crashing the submission system. The Board's commitment to transparency on decisions raises complications as well. The Board will provide reasoned decisions on the cases it hears and log those decisions in a public database. It will thus operate much like a common-law court, providing people with application of rules to real-life facts. But how much transparency is required? It is tempting to suggest that the decision should contain the actual piece of content reviewed. But if the decision upholds Facebook's decision to remove the content, it seems contradictory to post a copy of that very content on the Board's website.

Such contradictions are not limited to hearing cases but also include amending the founding documents. For instance, ads are imagined as coming into the future scope of the Board. But if the only appealable decisions are on content "removed for violating the Community Standards," it is very hard to see how this would ever work in practice for a typical user, whose primary complaint would likely be that the ad itself violated Community Standards. This means that before ads can come into the scope of the Board's review it is likely that the scope will first have to be expanded to include "content reviewed . . . and ultimately allowed to remain on the platform."<sup>407</sup>

### 3. *The Optimists*

Besides raising normative and practical concerns, the Board also holds great potential for users. At the outset, it gives users more process and ability to be heard than ever before. Even with its narrow jurisdictional scope, the procedural right of review by an independent adjudicator with the authority to instruct Facebook to restore speech is significant for global freedom of expression. It creates a tool that can be leveraged by users. Groups or organizations that disagree with Facebook's Community Standards could organize to submit large numbers of appeals regarding the same kind of removed content and thus attempt to create better odds of the Board selecting its case. If the Board selects the case, the submitting group would have an opportunity to argue not only for reinstatement of the content but for why the rule is wrong. If successful, the Board might not only restore the content, but it would ideally recommend that Facebook change the policy.

By giving users a right to appeal, the Board also tackles two of the most significant inequities in the moderation of online speech: Facebook's lack of transparency around content moderation and inequities in user access to content reinstatement. This promises to hold Facebook accountable for content-

---

407. *Oversight Board Bylaws*, *supra* note 201, art. 3, § 1.1.2.

moderation inaccuracies and give users equal access to an equitable remedy. In so doing, it replaces the historically opaque private system marked by cronyism and influence. Every day, Facebook removes millions of pieces of content, and the vast majority of users have no recourse to achieve reinstatement. External message boards and third-party websites proliferate to answer users' attempts at self-help to unlock disabled accounts or reinstate removed content.<sup>408</sup> For most of Facebook's history, getting your content restored was simply a matter of knowing the right people. Users who had connections to people inside Facebook or Instagram, to civil-society members who maintained connections to the company, or to members of the media who could draw attention to the fact that they had been censored were the most successful at getting their content restored. Even the media pushback against Facebook's censorship was often not a tale of popular empowerment, but rather an example of real-world power manifesting itself online. Facebook had removed posts of the "Terror of War" photo thousands of times before that day in September 2016, when the user censored was influential and politically connected enough for the media to expend coverage. By providing every user with transparency in appeals, process, and independent review, the Board takes a small but important step in remedying this inequity.

Finally, the Board might make Facebook better at serving users. The Board's decisions and policy recommendations may help Facebook to create policies that reflect what users want to see. The Board might also improve how Facebook researches, designs, and builds its other products. The Governance Team's model of user participation in the process of creating the Board has received great interest company-wide. Beyond Facebook, too, the Board might lead to more widespread user participation in deciding how to design private systems that govern our basic human rights.

### *C. Impact on Industry, Government, and Global Speech*

The majority of this Feature has focused on the relationship between four parties: Facebook, users, the Trust, and the Board. As discussed, the creation of the Board will likely have both positive and negative effects on these entities, but it will also have a far broader impact on industry, governments, and global speech.

---

408. See, e.g., *Facebook Remove My Post Because They Look Like Spam*, MPSSOCIAL (Aug. 2017), <https://mpsocial.com/t/facebook-remove-my-post-because-they-look-like-spam/23514> [<https://perma.cc/C5RE-7B43>]; *Got Your Facebook Account Disabled?*, GETASSIST.NET (Nov. 20, 2019), <https://getassist.net/recover-disabled-locked-suspended-facebook-account> [<https://perma.cc/P2ER-KDUX>].

The Trust Agreement in particular gives significant power to Trustees to act in fulfilling their purpose of supporting the Board. It empowers them, for instance, to form companies – such as the legislative body discussed above.<sup>409</sup> Or they could create a research group of experts on content moderation or an assembly of human-rights experts to advise the Board on decisions. True, Facebook would need to agree to the creation of such bodies, but the Board could exert pressure for this to happen. Finally, as will be discussed more below, the Trust’s power to form companies could allow it to form LLCs for other private platforms that wish to use the insulation created by the Trust to set up their own oversight boards.

1. *Industry*

There are at least five possible ways in which the Board might influence how online speech platforms adjudicate content.

1. The Board has no impact at all because it fails or is never adopted outside of Facebook.
2. Platforms *replicate the process of creating* the Board or design their own processes (involving user participation) to create adjudicatory bodies.
3. Platforms contribute money to the Trust but, instead of creating their own boards, they *request to use the Facebook Oversight Board* to adjudicate appeals on their own content. In this scenario, the Board would then either have to delegate each platform’s appeal review to specific members or attempt to make all members familiar with all sets of rules and values and apply the correct set of rules to each appeal.
4. Platforms contribute money to the Trust to *set up their own Oversight Board* to adjudicate appeals on content under their own rules and values. For this to happen, a platform would make an irrevocable grant to the Trust and relinquish its authority over the Trust; the Trust would then create an LLC to manage the platform’s Oversight Board.
5. Platforms use the founding documents as models, making slight changes, and convene their own trusts and boards.

The Board’s industry impact on speech varies greatly among these potential outcomes. If the Board fails entirely or remains relevant only to Facebook, the effects will likely be limited and confined to how Facebook regulates its speech.

---

409. *Oversight Board Trust Agreement*, *supra* note 250, § 4.6.

At the other extreme, platforms may decide to replicate the participatory process used to create the Oversight Board, draft their own founding documents, and create their own trust and board. Because this latter path involves user participation, it has the highest likelihood of achieving legitimacy and accurately reflecting the interests of platform users. This method also has the highest likelihood of generating diversity among platforms by opening the possibility of new forms outside the Facebook model. But unfortunately, this is also the least likely outcome given the difficulty and expense of such a process, and given the comparative ease of simply using existing structures and documents created by Facebook's process.

Of the remaining outcomes, the third is the worst for speech. The administrative and operational complexity of having one Oversight Board applying multiple sets of platform rules and values would be overwhelming. Moreover, it vests users' power of appeal in one industry-wide board. Over time, that board's decisions might standardize rules across the industry. This would likely reduce diversification in markets for different types of speech environments.

The final two outcomes are the best for speech. They involve independent oversight boards that are specifically employed to adjudicate the appeals of their platforms. Having distinct oversight boards preserves diversity in industry speech policies and optimizes opportunities for online free expression.<sup>410</sup> Of these options, platforms using the founding documents as models is preferable to platforms simply using the Oversight Board Trust because it allows platforms more control over customizing the structures of the Trust and Board. In the long term, it could also help produce better content moderation at all levels by creating a "market of rule sets." Larger platforms that can afford to fund and create a board could make their systems available to other firms based on the scale of their content-moderation needs. In this scenario, small or nascent firms would then be able to easily and affordably choose which model of speech regulation they wanted on their platform and employ the corresponding board. The Board's founding documents were designed for this very purpose – to allow platforms to use them easily as a template for their own systems of speech governance.

## 2. Government

There are at least three possible ways in which governments might react to the Board:

---

410. As of this writing, one major global speech platform has already announced its own "advisory council." Vanessa Pappas, *Introducing the TikTok Content Advisory Council*, TIKTOK (Mar. 18, 2020), <https://newsroom.tiktok.com/en-us/introducing-the-tiktok-content-advisory-council> [<https://perma.cc/A2Z2-WXD3>].

1. Nation-states *ignore* the Oversight Board and future platform boards.
2. Nation-states *recognize* the Oversight Board or boards' decisions as precedential authority when public courts adjudicate platform speech issues.
3. Nation-states *co-opt* the Oversight Board or boards by requiring these private systems to adjudicate platform speech issues that surface in public courts.

The impact of the Oversight Board on governments is expressly limited, in that the founding documents exclude from appeal any content that is removed by Facebook in accordance with the law. Governments, however, could have a potentially strong impact on the Board. This of course would not be true if governments and courts simply ignored the Board and similar private systems. But that seems unlikely given the current political climate surrounding technology companies. One possibility, favorable to the Board, is that courts might recognize the Board's reasoning when the same or a similar content claim is raised in a legal action. Courts are unlikely to accept the Board's decision as precedential or give it comity, but they might consider it persuasive. Such recognition of the limited and private jurisdiction of the Board might be similar to courts' recognition of private alternative dispute resolution. The more problematic outcome both for global speech and the Board would be if governments were to co-opt this private system for their own ends. For example, courts might order the Board to adjudicate and enforce claims of a certain type, similar to the European Court of Justice's directive to Google to administer right-to-be-forgotten claims.<sup>411</sup> This would be especially bad for speech if a court decided to also direct the Board to administer such claims in accordance with the law in the court's jurisdiction and apply the decisions to all of Facebook. Recent cases have suggested courts' ability to use a platform to enforce their jurisdiction extraterritorially, a development that flouts national sovereignty and individual rights.<sup>412</sup>

---

411. See Case C-131/12, *Google Spain SL v. Agencia Española de Protección de Datos*, 2014 E.C.R. 317 (ruling that Google must honor a claimant's request to "delist" information from its search engines but acknowledging that the claimant's right must be balanced against concerns for access to information, and placing adjudication of that balancing test on a case-by-case basis in the hands of Google). For a rich discussion of the implications of nation-states using private companies to adjudicate and administer their rules, see Daphne Keller, *The Right Tools: Europe's Intermediary Liability Laws and the EU 2016 General Data Protection Regulation*, 33 *BERKELEY TECH. L.J.* 287 (2018).

412. See Case C-18/18, *Glawischnig-Piesczek v. Facebook Ir. Ltd.*, Judgment, 2019 E.C.R. 821; Daskal & Klonick, *supra* note 18 (discussing the written opinion of the Advocate General in

### 3. *Global Free Expression*

The impact of the Oversight Board on global free expression has been addressed and analyzed throughout this piece. What has been largely unexplored, however, is the way in which the future of online speech will affect the Board.

As this Feature goes to press, the world is experiencing a global pandemic seemingly unprecedented in human history.<sup>413</sup> Governments have ordered people to remain in their homes,<sup>414</sup> closed schools,<sup>415</sup> banned the assembly of people in large groups,<sup>416</sup> and restricted travel.<sup>417</sup> Museums, concert halls, stadiums, operas, theatres, and even parks are empty.<sup>418</sup> Courts have closed and canceled

---

Case C-18/18, *Glawischnig-Piesczek v. Facebook Ir. Ltd.*, 2019 E.C.R. 458). For excellent discussion on this topic, see Jennifer Daskal, *Borders and Bits*, 71 VAND. L. REV. 179 (2018).

413. See Ronald Bailey, *Are We Battling an Unprecedented Pandemic or Panicking at a Computer-Generated Mirage?*, REASON (Mar. 18, 2020), <https://reason.com/2020/03/18/are-we-battling-an-unprecedented-pandemic-or-panicking-at-a-computer-generated-mirage> [<https://perma.cc/R35A-DK6Q>]; Elizabeth Kolbert, *Pandemics and the Shape of Human History*, NEW YORKER (Mar. 30, 2020), <https://www.newyorker.com/magazine/2020/04/06/pandemics-and-the-shape-of-human-history> [<https://perma.cc/KB2A-DMEM>].
414. Sarah Mervosh et al., *See Which States and Cities Have Told Residents to Stay Home*, N.Y. TIMES (Apr. 7, 2020), <https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html> [<https://perma.cc/K8NH-FW4H>]; Sonia Sarkar, *Coronavirus Quarantine in India: No Tests, Stained Toilets and Broken Beds Force Some to Flee*, S. CHINA MORNING POST (Mar. 25, 2020), <https://www.scmp.com/week-asia/health-environment/article/3076896/coronavirus-quarantine-india-unhygienic-facilities> [<https://perma.cc/ZSM8-GXCB>].
415. *Map: Coronavirus and School Closures*, EDUC. WEEK (Apr. 17, 2020), <https://www.edweek.org/ew/section/multimedia/map-coronavirus-and-school-closures.html> [<https://perma.cc/5KAK-JWYZ>].
416. Katrin Bennhold & Melissa Eddy, *Germany Bans Groups of More than 2 to Stop Coronavirus as Merkel Self-Isolates*, N.Y. TIMES (Mar. 22, 2020), <https://www.nytimes.com/2020/03/22/world/europe/germany-coronavirus-budget.html> [<https://perma.cc/8Z6Z-NMQ8>].
417. Megan Specia, *What You Need to Know About Trump's European Travel Ban*, N.Y. TIMES (Mar. 12, 2020), <https://www.nytimes.com/2020/03/12/world/europe/trump-travel-ban-coronavirus.html> [<https://perma.cc/BQ8Z-2KGF>].
418. *Coronavirus-Related Closures and Canceled Events in Washington, DC*, DESTINATION DC (Apr. 17, 2020), <https://washington.org/dc-information/coronavirus-event-attraction-information> [<https://perma.cc/SH95-YMBR>]; Robin Pogrebin & Michael Cooper, *New York's Major Cultural Institutions Close in Response to Coronavirus*, N.Y. TIMES (Mar. 12, 2020), <https://www.nytimes.com/2020/03/12/arts/design/met-museum-opera-carnegie-hall-close-coronavirus.html> [<https://perma.cc/35E3-KRC3>].



civil and criminal trials and hearings.<sup>419</sup> Shops, cafes, gyms, malls, restaurants, and bars have closed doors that might never reopen.<sup>420</sup>

The realities of worldwide mass quarantine have decimated the economy and stymied government, but perhaps the most devastating and permanent effects will be social. Around the world, people have been shuttered in their homes for weeks and months, unable to gather in groups or experience social connection. Children are separated from parents, teachers are separated from students, neighbors are separated from each other. The infected, sick, and at-risk are cut off.

This real-world social isolation has exposed and deepened the essential and massive power of network communications technology. Individuals' ability to connect to each other through the internet and its platforms swiftly became a basic human right. Instead of litigating the legality of government mandates denying assembly, people sought out communities online. New Facebook Groups, Slack channels, and group texts proliferated.<sup>421</sup> Universities and grade schools turned to Zoom, WebEx, and Google Hangouts to hold regular classes.<sup>422</sup> Virtual board meetings, birthdays, dinners, and happy hours were held online. The sick and dying said goodbye through FaceTime and Skype.<sup>423</sup>

---

419. Debra Cassens Weiss, *A Slew of Federal and State Courts Suspend Trials or Close for Coronavirus Threat*, ABA J. (Mar. 18, 2020), <https://www.abajournal.com/news/article/a-slew-of-federal-and-state-courts-jump-on-the-bandwagon-suspending-trials-for-coronavirus-threat> [https://perma.cc/RXH5-3BGG].

420. Alex Gangitano, *Poll: Almost One in Four Small Businesses Are Two Months or Less Away from Closing Permanently*, USA TODAY (Apr. 3, 2020), <https://thehill.com/business-a-lobbying/business-a-lobbying/490957-poll-one-in-four-small-businesses-are-two-months-or> [https://perma.cc/7ARZ-KDRN].

421. Kara Swisher, *Opinion, How Will Tech Help in a Time of Pandemic?*, N.Y. TIMES (Feb. 27, 2020), <https://www.nytimes.com/2020/02/27/opinion/coronavirus-tech-facebook.html> [https://perma.cc/7BEX-8XV9].

422. Doug Lederman, *Will Shift to Remote Teaching Be Boon or Bane for Online Learning?*, INSIDE HIGHER ED (Mar. 18, 2020), <https://www.insidehighered.com/digital-learning/article/2020/03/18/most-teaching-going-remote-will-help-or-hurt-online-learning> [https://perma.cc/UK9R-EECY].

423. Bill Gardner, *Deathbed Goodbyes Should Be Done over Skype, Says New NHS Coronavirus Guidance*, TELEGRAPH (Mar. 17, 2020), <https://www.telegraph.co.uk/news/2020/03/17/exclusive-deathbed-goodbyes-should-done-skype-says-new-nhs-coronavirus> [https://perma.cc/8DQW-2FER]; Amir Vera, *A Woman Got to Say Goodbye to Her Mother over FaceTime Before She Died Thanks to a Nurse at This Washington Hospital*, CNN (Apr. 5, 2020), <https://www.cnn.com/2020/03/30/us/washington-nurse-coronavirus-facetime/index.html> [https://perma.cc/A8FZ-AJRG].

In the midst of the crisis, Facebook announced it would be suspending most of its human content moderation globally.<sup>424</sup> Unlike the billions of people working from home, content moderators must work on special screens and be supervised to ensure user privacy. In place of human review, Facebook announced it would do more automated content moderation.<sup>425</sup> At a normal time, this decline in meaningful human content review<sup>426</sup> on a prominent speech platform would raise questions of over- or undercensorship of speech. But in the midst of a global pandemic with a population never before so reliant on internet platforms, over- or undercensorship of speech is also a threat to public health.<sup>427</sup>

The effect of this long-term experiment without human content moderation remains to be seen. Its greatest threat is not that it gives Facebook plausible justification to remove human review but that it may groom users to accept this outcome.<sup>428</sup> If a long-term policy in the future were to deny users the right of human review, the Oversight Board would become more than just an exercise in user participation in platform governance; it would be a vital procedural element of protecting users' right to global free expression.

---

424. Alex Schultz, *Keeping Our Services Stable and Reliable During the COVID-19 Outbreak*, FACEBOOK NEWSROOM (Mar. 24, 2020), <https://about.fb.com/news/2020/03/keeping-our-apps-stable-during-covid-19> [<https://perma.cc/672W-8PLN>].

425. Kang-Xing Jin, *Keeping People Safe and Informed About the Coronavirus*, FACEBOOK NEWSROOM (Apr. 21, 2020), <https://about.fb.com/news/2020/04/coronavirus/#content-review> [<https://perma.cc/4M99-FAXQ>] (“[W]ith a reduced and remote workforce, we will now rely more on our automated systems to detect and remove violating content and disable accounts.”).

426. For an excellent discussion on the right to human decision in the midst of algorithmic adjudication, see Aziz Z. Huq, *A Right to a Human Decision*, 105 VA. L. REV. 611 (2020) (citing empirical evidence and international normative arguments for the use of human decision-makers). For discussion of human decision-making and human dignity, see generally JEREMY WALDRON, *DIGNITY, RANK, AND RIGHTS* (Meir Dan-Cohen ed., 2012), which describes dignity and its relation to the law; Martha Nussbaum, *Human Dignity and Political Entitlements*, in *HUMAN DIGNITY AND BIOETHICS: ESSAYS COMMISSIONED BY THE PRESIDENT’S COUNCIL ON BIOETHICS* (2008), which analyzes a capabilities-based theory of dignity for bioethics; and Tal Zarsky, *The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making*, 41 SCI. TECH. & HUM. VALUES 118 (2016), which analyzes fairness issues with algorithmic decision-making.

427. See, e.g., David Harsanyi, Opinion, *Facebook Fails the Coronavirus Censorship Test*, BOS. HERALD (Apr. 26, 2020), <https://www.bostonherald.com/2020/04/26/facebook-fails-the-coronavirus-censorship-test> [<https://perma.cc/S3FB-V7CA>].

428. See Bloch-Wehba, *supra* note 33.

## CONCLUSION

Global events demonstrate that freedom of speech can no longer be separated into two separate worlds – one online, the other offline. The Oversight Board marks the first platform-scaled moment of transnational internet adjudication of online speech. It signifies a step towards empowering users by involving them in private platform governance and providing them with a modicum of procedural due process.

The pace at which this area of scholarship and technology is changing is remarkable. Twenty years ago, Facebook did not yet exist; fifteen years ago, it was a fad website from a college dorm room; ten years ago, it was an emerging empire; today, it connects almost a third of the global population. Similarly, five years ago, few would have thought it possible that a private corporation would voluntarily divest itself of part of its power in order to create an independent oversight body. Like the past events that led to it, the future of the Oversight Board is impossible to predict, once humans start interpreting and understanding the documents and processes contemplated in this Feature. But even though the Board's founding documents have flaws and limitations and leave questions unanswered, they are a promising new tool for ensuring free speech around the world.