

# Book Review:

## *Custodians of the Internet*

Gillespie, Tarleton  
New Haven, CT: Yale University Press, 2018.

**Sarah Whitcomb Laiola<sup>1</sup>**  
COASTAL CAROLINA UNIVERSITY  
slozierl [AT] coastal.edu



A tweet promoting ten different plants as forms of “natural birth control” went relatively viral on June 11, 2019. As it circulated, doctors, public health officials, journalists, botanists, and even natural/holistic medicine practitioners attacked the original sharer, on the grounds that, first, nothing on the list was an effective form of birth control, and, second, some of the plants could be harmful, even fatal, to women. These attacks were often accompanied by an injunction to remove the tweet, many of which tagged Twitter in the responses, a rhetorical move that calls on the platform to moderate its users’ content and stop the spread of dangerous misinformation while implying that it is Twitter’s responsibility as a social media platform to take this corrective action. By the following day, the tweet had become “unavailable.” Some questions, however, remain: Where did the tweet go? Did the original user remove it? Did Twitter? Has it actually been removed? Or has it simply been filtered out of my feed?

This story of social media platforms and the control they can or should exert on the content that circulates through them is one that concerns most social media users, academics, critics, developers, and lawmakers today. The inability to provide a simple answer to the role platforms should play in moderating their users' content is at the core of contemporary public debates about everything from the spread of misinformation, conspiracy theories, and "fake news," to the rise of extremist politics and radicalization, to concerns about the erosion of free speech, the free press, and other ideological cornerstones of Western democracy. It is also the problem space that Tarleton Gillespie's *Custodians of the Internet* occupies.

If the book does not explicitly argue that platforms *should* moderate their content, it is because Gillespie opens with the position—encapsulated by his first chapter's title—that "All Platforms [already] Moderate."<sup>2</sup> Expanding from the apparent simplicity of this statement, Gillespie argues that "Moderation is not an ancillary aspect of what platforms do. It is essential, constitutional, definitional. Not only can platforms not survive without moderation, they are not platforms without it."<sup>3</sup> This statement is the core thesis that recurs and develops, with slight variations, throughout the book. By the end, Gillespie asserts that "moderation is the essence of platforms, it is the commodity they offer,"<sup>4</sup> a position that contradicts how we tend to conceptualize the relationship between platforms and content. Usually, we understand the platform as a neutral intermediary through which the content travels and that plays no inherent or predetermined role in controlling, policing, affecting, or otherwise moderating the content. Far from suggesting that this cultural misconception is the fault of uninformed or uncritical users, Gillespie's thesis is accompanied by the complicating reminder that moderation

must be largely disavowed, hidden, in part to maintain the illusion of an open platform and in part to avoid legal and cultural responsibility. Platforms face what may be an irreconcilable contradiction: they are represented as mere conduits and they are premised on making choices for what users see and say.<sup>5</sup>

In the pages that unfold, Gillespie examines and provokes this potential "irreconcilable contradiction."

The second chapter, "The Myth of the Neutral Platform," provides the cultural and legal origin story for the platform's position as a "mere conduit" and explores the limitations that telecommunications and media law face in governing platforms, as they no longer operate within established legal and cultural categories. Beginning at the early, pre-platform days of the web, Gillespie takes readers through the advent of Section 230 of US telecommunications law—"Safe Harbor"—which categorizes platforms as media conduits or "intermediaries" that connect us to information rather than produce that information, thereby giving them "the right, but not the responsibility" to moderate their users' content.<sup>6</sup> The problem, Gillespie reveals, is that platforms are no longer just intermediaries, but instead operate as a hybrid that is both conduit and producer:

the moment that a platform begins to select some content over others, based not on a judgment of relevance to a search query but in the spirit of enhancing the value of the experience and keeping users on the site, it has become a hybrid.<sup>7</sup>

In its clear articulation of both legislative and computational complexities, this chapter stands out as one of the most effective for understanding how and why US legislative bodies continually fail to hold entities like Google, Facebook, and Twitter accountable for the problematic information that they host and circulate.

In chapters 3, 4, 5, and 7, Gillespie examines various ways that platforms do attempt to moderate the content that circulates on their sites. Chapter 3 delves into the world of community guidelines, documents that, Gillespie argues, “make clear the central contradiction of moderation that platform creators must attempt to reconcile, but never quite can.”<sup>8</sup> This contradiction is, of course, that social media platforms are meant to “embody the freedom of the web”—a position that is incompatible with moderation—but they must also offer something “better than the chaos of the open web”—a position that requires moderation.<sup>9</sup> In this chapter, he examines the ways community guidelines rhetorically establish an atmosphere of the platform, the content most prevalent across community guidelines, and why these guidelines matter, despite, as Gillespie reminds us throughout the chapter, existing primarily as rhetorical gestures rather than as guidelines for arbitration.

Chapter 4 turns to the problem of scale, noting that the problem of content moderation, as such, is not new, but that the scale at which platforms operate today, the sheer amount of content that they support, is. In this chapter, Gillespie examines three imperfect solutions that different platforms have taken to address content moderation at scale: editorial review, a solution favored by Apple’s App Store, where the platform reviews content before releasing it; Community Flagging, a solution favored by YouTube and Twitter, where the platform enables its users to police and report one another; and Automatic Detection, a solution used by Google and Facebook that employs machine learning to detect problematic content. In chapters 5 and 7, Gillespie examines two additional options that platforms have taken for handling content moderation at scale from chapter 4: the option of employing humans to moderate and make determinations about the platform’s content and the option of filtering for specific audiences rather than removing the content altogether, respectively.

In chapter 6, the book takes a turn from broad, general discussions of history, policy, and practice to focus on a case study of Facebook’s moderation of breastfeeding photos. In this chapter, Gillespie takes care in his writing to represent the experiences of women who have been involved with Facebook’s moderation of breastfeeding photos, and although the ethnographic history he sketches is thorough, readers interested in the topic of this case study may find his forced neutrality frustrating, even if it is appropriate for the book’s overall purpose. Counteracting this potential frustration is the chapter’s new and valuable approach to engaging the effects of content moderation on the humans who use social media platforms. Often public debate about content moderation focuses on the effect that moderation (or not) has on users, but rarely does it focus on the effects this practice has on the users whose content has been flagged or removed. Gillespie’s focus on the experiences of new moms whose content is stuck in the purgatory of moderation is therefore both a fresh and important take on the ways content moderation affects human users.

After seven chapters examining the complex problems of platforms and moderation, it begins to seem as if the problem of platforms and effective content moderation is a problem that is impossible to solve. From this vantage point, the book’s eighth and final chapter, “What

Platforms Are, and What They Should Be,” appears as a breath of fresh air, as Gillespie offers suggestions for improvement to platform developers. These suggestions appear as a series of imperatives to “design for deliberate and actionable transparency,”<sup>10</sup> “distribute the agency of moderation, not just the work,”<sup>11</sup> “protect users as they move across platforms,”<sup>12</sup> “reject the economics of popularity,”<sup>13</sup> and “put real diversity behind the platform.”<sup>14</sup> Beyond these design imperatives, what ultimately makes this chapter the most powerful of the book is Gillespie’s turn away from platform design to a reiteration of his call for “a thorough, public discussion about the social responsibility of platforms”<sup>15</sup> that includes

a very different understanding of the role of “custodian”—not where platforms quietly clean up our mess, but where they take up guardianship of the unresolvable tensions of public discourse, hand back with care the agency for addressing these tensions to users, and responsibly support the process with the necessary tools, data, and insights.<sup>16</sup>

Alongside platforms’ growth into more responsible custodians, Gillespie complicates our own positionality, suggesting that “perhaps we must begin to be . . . the custodians of the custodians.”<sup>17</sup>

Overall, Gillespie’s study of platforms and the role they play in content moderation are an important contribution to (new) media studies and platform studies, sitting comfortably and offering additional critical insights alongside contemporaries like Safiya Noble’s *Algorithms of Oppression* (New York University Press, 2018), Siva Vaidyanathan’s *Anti-Social Network* (Oxford University Press, 2018), and Sarah T. Roberts’ forthcoming *Behind the Screen* (Yale University Press, 2019), which addresses the many costs of the human labor that support commercial content moderation, a topic Gillespie touches on briefly in chapter 5. Beyond its importance for academic study, as we continue to reel from the 2016 presidential election, facing the rise of fascism and other extremist politics while scrambling to reinstate some semblance of (Western) democracy and public discourse, the book cannot be more timely for a popular audience, and Gillespie’s highly readable prose makes *Custodians of the Internet* a critical must-read for anyone trying to operate in and get a handle on our increasingly social media-saturated world.

---

<sup>1</sup> Sarah Whitcomb Laiola is an assistant professor of Digital Culture and Design at Coastal Carolina University. She received her PhD in English from the University of California, Riverside, and specializes in new media poetics, contemporary digital cultures, and critical race and gender studies. Her recent peer-reviewed publications appear in *Criticism* (60.2), *American Quarterly* (70.3), and *Television and New Media* (18.5).

<sup>2</sup> Tarleton Gillespie, *Custodians of the Internet* (New Haven, CT: Yale University Press, 2018), 1.

<sup>3</sup> *Ibid.*, 21.

<sup>4</sup> *Ibid.*, 207.

<sup>5</sup> *Ibid.*, 21.

<sup>6</sup> *Ibid.*, 44.

<sup>7</sup> *Ibid.*, 43.

<sup>8</sup> Ibid., 47.

<sup>9</sup> Ibid., 47.

<sup>10</sup> Ibid., 198.

<sup>11</sup> Ibid., 199.

<sup>12</sup> Ibid., 200.

<sup>13</sup> Ibid., 201.

<sup>14</sup> Ibid., 201.

<sup>15</sup> Ibid., 206.

<sup>16</sup> Ibid., 211.

<sup>17</sup> Ibid., 212.

