ESSAYS

FREE SPEECH IS A TRIANGLE

*Jack M. Balkin**

The vision of free expression that characterized much of the twentieth century is inadequate to protect free expression today.

The twentieth century featured a dyadic or dualist model of speech regulation with two basic kinds of players: territorial governments on the one hand, and speakers on the other. The twenty-first-century model is pluralist, with multiple players. It is easiest to think of it as a triangle. On one corner are nation-states and the European Union. On the second corner are privately owned internet-infrastructure companies, including social media companies, search engines, broadband providers, and electronic payment systems. On the third corner are many different kinds of speakers, legacy media, civil-society organizations, hackers, and trolls.

The practical ability to speak in the digital world emerges from the struggle for power between these various forces, with "old-school," "new-school," and private regulation directed at speakers, and both nation-states and civil-society organizations pressuring infrastructure owners to regulate speech.

This configuration creates three problems. First, nation-states try to pressure digital companies through new-school speech regulation, creating problems of collateral censorship and digital prior restraint. Second, social media companies create complex systems of private governance and private bureaucracy that govern end users arbitrarily and without due process and transparency. Third, end users are vulnerable to digital surveillance and manipulation.

This Essay describes how nation-states should and should not regulate the digital infrastructure consistent with the values of freedom of speech and press. Different models of regulation are appropriate for different parts of the digital infrastructure: Basic internet services should be open to all, while social media companies should be treated as information fiduciaries toward their end users. Governments can implement all of these reforms—properly designed—consistent with constitutional guarantees of free speech and free press.

---

* Knight Professor of Constitutional Law and the First Amendment, Yale Law School. Many thanks to Jameel Jaffer, Daphne Keller, Maggie McKinley, David Pozen, and Tim Wu for their comments on a previous draft.

INTRODUCTION

Free speech is a triangle. The conception of free expression—and of the dangers to free expression—that characterized much of the nineteenth and twentieth centuries concerned whether nation-states and their political subdivisions would censor or regulate the speech of people living within their borders. That picture still describes many important free speech problems, yet it is increasingly outmoded and inadequate to protect free expression today. In the early twenty-first century, freedom of speech increasingly depends on a third group of players: a privately owned infrastructure of digital communication composed of firms that support and govern the digital public sphere that people use to communicate.

Consider a few recent speech controversies. The first is the European Union's "right to be forgotten." It requires search engine companies (essentially Google) to eliminate certain newspaper articles from their search results.[1] A second is the recently passed German law known

1. See Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, 2014 E.C.R. 317, ¶ 94; Robert C. Post, Data Privacy and Dignitary Privacy: Google Spain,

as NetzDG.[2] It requires social media companies to take down many different kinds of speech, including hate speech, within twenty-four hours of a complaint.[3] A third is the concern about fake news propagating through social media sites.[4] A fourth is the decision by various internet companies—following the Charlottesville march in August 2017—to block, censor, or otherwise refuse to do business with various neo-Nazi and hate sites.[5] Each of these controversies concerns the new structure of speech regulation in the digital age.

The twentieth century featured a *dualist* or *dyadic* system of speech regulation.[6] In the dualist model, there are essentially two players: the nation-state on the one hand and the speaker on the other. Nation-states regulated many different kinds of speakers and mass media of all kinds, including publishing houses, movie houses, newspapers, radio stations, and television stations.

---

the Right to Be Forgotten, and the Construction of the Public Sphere, 67 Duke L.J. 981, 986 (2018).

2. See Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken [Netzwerkdurchsetzungsgesetz—NetzDG] [Network Enforcement Act], Sept. 1, 2017, Bundesgesetzblatt, Teil I [BGBl I] at 3352 (Ger.), https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG_engl.pdf?__blob=publicationFile&v=2 [https://perma.cc/W2B8-JWHT].

3. See infra notes 90–92 and accompanying text.

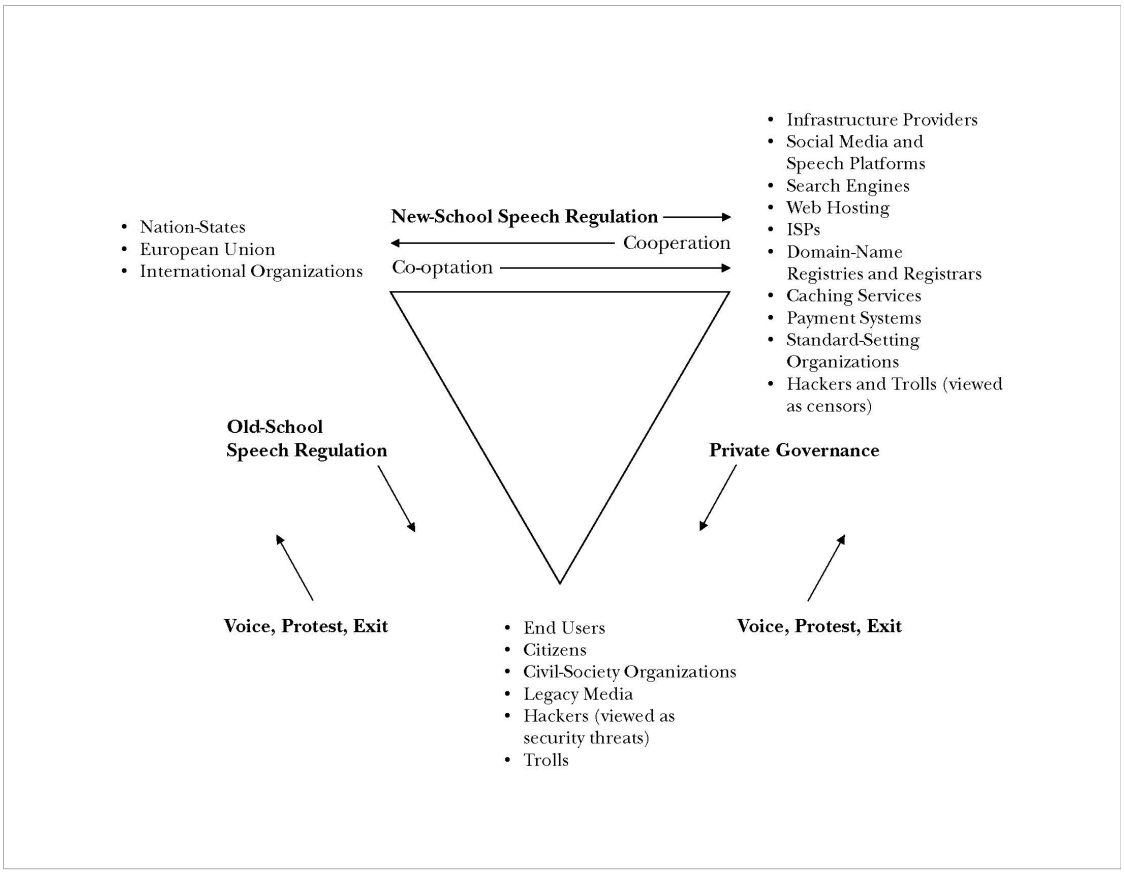
4. See, e.g., Info. Soc'y Project & The Floyd Abrams Inst. for Freedom of Expression, Fighting Fake News: Workshop Report 3 (2017), http://law.yale.edu/system/files/area/center/isp/documents/fighting_fake_news_-_workshop_report.pdf (on file with the *Columbia Law Review*); Mark Verstraete, Derek E. Bambauer & Jane R. Bambauer, Identifying and Countering Fake News 4 (Univ. of Ariz. Legal Studies, Discussion Paper No. 17-15, 2017), https://ssrn.com/abstract=3007971 (on file with the *Columbia Law Review*).

5. See, e.g., Elizabeth Flock, Spotify Has Removed White Power Music from Its Platform. But It's Still Available on Dozens of Other Sites, PBS Newshour (Aug. 18, 2017), https://www.pbs.org/newshour/art/spotify-removed-white-power-music-platform-still-available-dozens-sites [https://perma.cc/B6P3-98V3] ("In the wake of the white nationalist rally and ensuing violence in Charlottesville last weekend, Spotify announced it would remove music that promotes white nationalism from its libraries . . . ."); Kerry Flynn, After Charlottesville, Tech Companies Are Forced to Take Action Against Hate Speech, Mashable (Aug. 16, 2017), http://mashable.com/2017/08/16/after-charlottesville-tech-companies-action-nazis/#kxrJzxU9pOqP [https://perma.cc/9RJ9-5SUV] ("Facebook, Google, Spotify, Uber, Squarespace, and a variety of other tech companies are taking action to curb the use of their platforms and services by far-right organizations.").

6. Jack M. Balkin, Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation, 51 U.C. Davis L. Rev. 1149, 1187 (2018) [hereinafter Balkin, Algorithmic Society] ("The twentieth century model is a *dyadic* model: the state is on one side, speakers and publishers are on the other."); see also Jack M. Balkin, Old-School/New-School Speech Regulation, 127 Harv. L. Rev. 2296, 2298 (2014) [hereinafter Balkin, Old-School/New-School] ("Traditional or 'old-school' techniques of speech regulation have generally employed criminal penalties, civil damages, and injunctions to regulate individual speakers and publishers.").

The twenty-first-century model is *pluralist*, with many different players.[7] For ease of exposition, we might consider it as a triangle.

FIGURE 1: THE PLURALIST MODEL OF SPEECH REGULATION[8]



On one corner of the triangle are nation-states, states, municipalities, and supranational organizations like the European Union.

On the second corner of the triangle are internet-infrastructure companies. These include social media companies, search engines, internet service providers (ISPs), web-hosting services, Domain Name System (DNS) registrars and registries, cyber-defense and caching services (such as Cloudflare and Akamai), and payment systems (such as PayPal, Mastercard, and Visa). Each of these elements of the internet infrastructure is important, if not crucial, to people's practical ability to

7. See Balkin, Algorithmic Society, supra note 6, at 1189–91.

8. Id. at 1189 diagram 1.

speak. In most countries, this internet infrastructure, or important parts of it, are privately owned.[9]

On the third corner of the triangle, at the very bottom, we have speakers and legacy media, including mass-media organizations, protesters, civil-society organizations, hackers, and trolls. Although both states and infrastructure owners regulate their speech, they are sometimes able to influence states and infrastructure owners through social activism and protest.[10]

Nation-states regulate speakers and legacy mass media through *old-school speech regulation*. Nation-states regulate and attempt to co-opt and coerce internet infrastructure through *new-school speech regulation*. Finally, the internet infrastructure regulates private speakers and legacy media through techniques of *private governance*.

This is the new structure of speech regulation in the early twenty-first century, and debates about the rights of online free expression must grapple with that structure. To understand how this new system works, we must understand the distinction between old- and new-school speech regulation, explained in Part I, and the emerging system of private governance, discussed in Part II. Parts III and IV offer proposals for protecting freedom of speech in the changed environment. A brief conclusion follows.

I. OLD-SCHOOL AND NEW-SCHOOL SPEECH REGULATION

In traditional or old-school speech regulation, nation-states use threats of fines, penalties, imprisonment, or other forms of punishment or retribution to regulate or control the speech of individuals, associations, and media companies.[11] As noted above, this conception is dyadic. In this traditional conception, freedom of speech and press simply means being free of old-school speech regulation.

Old-school speech regulation still exists around the world. But digital free speech has created new problems for which old-school methods are inadequate. The early twenty-first century has developed new methods for controlling digital speech. This is the "new school" of speech regulation, and because of its ascension, freedom of speech today requires far more than freedom from old-school speech regulation.

Whereas old-school regulation is directed at speakers, new-school speech regulation is directed at the internet infrastructure.[12] Nation-states (or supranational entities like the European Union) attempt to regulate, threaten, coerce, or co-opt elements of the internet infrastructure in order to get the infrastructure to surveil, police, and control

9. See id. at 1188.

10. See id. at 1188–90.

11. Id. at 1174; Balkin, Old-School/New-School, supra note 6, at 2298.

12. Balkin, Old-School/New-School, supra note 6, at 2298.

speakers.[13] In essence, nation-states attempt to get the privately owned infrastructure to do their work for them.

Consider the free speech controversies mentioned in the Introduction.[14] Germany's NetzDG is aimed at search engines and social media companies to limit forbidden speech.[15] The European Union's "right to be forgotten" is directed (in part) at search engines in order to make it hard for people to discover embarrassing stories in newspapers. Calls for government regulation to prevent fake news demand that social media companies—and other parts of the internet infrastructure—take steps to limit the publication and distribution of false stories among end users. Following the Charlottesville protests, neo-Nazi sites were hampered or blocked not by states and municipal governments but by private-infrastructure owners.

Although nation-states continue to regulate speech directly through old-school methods, they increasingly depend on new-school speech regulation—attempting to coerce or co-opt private owners of digital infrastructure to regulate the speech of private actors. For this reason, new-school speech regulation affects the practical ability to speak every bit as much as old-school speech regulation.

A. *Collateral Censorship and Digital Prior Restraint*

New-school speech regulation poses two central problems for freedom of speech. First, it usually involves some form of *collateral censorship*. Second, it raises many of the same problems as *prior restraint*, except that the restraint is performed by private bureaucrats and algorithms in the service of the state.

1. *Collateral Censorship.* — Collateral censorship occurs when the state targets entity A to control the speech of another entity, B.[16] The state tells

---

13. See Balkin, Algorithmic Society, supra note 6, at 1179–82; Balkin, Old-School/New-School, supra note 6, at 2324–29.

14. See supra notes 1–5 and accompanying text.

15. See Germany Starts Enforcing Hate Speech Law, BBC (Jan. 1, 2018), https://www.bbc.com/news/technology-42510868 [https://perma.cc/2UAA-BHR7]; infra notes 90–91.

16. J.M. Balkin, Free Speech and Hostile Environments, 99 Colum. L. Rev. 2295, 2298 (1999); Balkin, Old-School/New-School, supra note 6, at 2309; see also Christina Mulligan, Technological Intermediaries and Freedom of the Press, 66 SMU L. Rev. 157, 165–66 (2013) (arguing that collateral censorship threatens freedom of the press); Felix T. Wu, Collateral Censorship and the Limits of Intermediary Immunity, 87 Notre Dame L. Rev. 293, 299–304 (2011) (arguing that intermediary immunity should be tailored to the problem of collateral censorship).

Professor Michael Meyerson coined the term "collateral censorship." See Michael I. Meyerson, Authors, Editors, and Uncommon Carriers: Identifying the "Speaker" Within the New Media, 71 Notre Dame L. Rev. 79, 118 (1995) (defining collateral censorship as "the silencing by a private party of the communication of others"); see also Seth F. Kreimer, Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link, 155 U. Pa. L. Rev. 11, 16 (2006) (coining the terms "proxy censorship" and "censorship by proxy").

A: Locate and block or censor B, or else we will punish or fine you. In effect, collateral censorship attempts to harness a private organization to regulate speech on the state's behalf.

Collateral censorship does not raise special problems for freedom of expression when A and B are part of the same entity or firm that produces the expression (for example, when B is A's employee), or when A has a traditional editorial or publishing relationship to B. Defamation law holds newspapers liable for what their reporters write and their advertisers advertise, and it holds book publishers liable for their authors' defamation.[17] When A is B's book publisher, when B works for A, or when B advertises in A's newspaper, the law assumes that A has a vested interest in defending and protecting the speech produced by B that A edits and publishes.

We cannot make the same assumption, however, when A is part of the internet infrastructure and B is one of the countless number of people who use A's services to communicate with others. Then A and B are not in the same relationship as the newspaper and its reporters, the publishing house and its authors, or the magazine and its advertisers.

In these cases, A's incentives are somewhat different. Told by the state that it must censor or block speakers like B, A will err on the side of caution.[18] It will tend to overblock or overfilter content, discarding the wheat with the chaff. In addition, A will be more likely to take down speech that anyone objects to or that it fears someone might object to. Because there are so many speakers (and because A wants to make the vast majority of its end users feel comfortable), denying access to a very small number of speakers will not damage A's business model, whereas repeated imposition of government liability for the speech of total strangers might seriously hinder its ability to do business.

2. *Digital Prior Restraint.* — Imposing liability on infrastructure providers unless they surveil and block speech, or remove speech that

---

17. See Restatement (Second) Of Torts § 578 (Am. Law Inst. 1977) ("Except as to those who only deliver or transmit defamation published by a third person, one who repeats or otherwise republishes defamatory matter is subject to liability as if he had originally published it.").

In fact, the landmark case of *New York Times Co. v. Sullivan* involved a kind of collateral censorship. Alabama sought to hold the New York Times liable for a political advertisement, "Heed Their Rising Voices," which complained about police misconduct against civil rights demonstrators. N.Y. Times Co. v. Sullivan, 376 U.S. 254, 256 (1964). Although the Supreme Court created a constitutional privilege to protect the *New York Times*, it did not question the traditional rule of publisher liability; rather, it assumed without discussion that newspapers would exercise their traditional editorial functions with respect to advertisements published within their pages. See id. at 286–88; see also Cantrell v. Forest City Publ'g Co., 419 U.S. 245, 253–54 (1974) (approving a jury charge that permitted the imposition of vicarious liability upon a publisher for the knowing falsehoods written by its staff writer).

18. Balkin, Old-School/New-School, supra note 6, at 2309; see also Kreimer, supra note 16, at 28–29; Wu, supra note 16, at 300–01.

others complain about, has many features of a prior restraint, although technically it is not identical to a classic prior restraint.[19]

Administrative prior restraints deny people the right to speak without a full judicial determination of whether their speech is protected or unprotected and without the procedural protections of the Bill of Rights.[20] In addition, administrative prior restraints place the burden of inertia on the speaker and the benefit of inertia on the government.[21] People have to get someone else's permission before they can speak (or speak again, if the order comes down after they have begun to publish copies). Administrative review acts as a bottleneck to free speech; nothing will happen until the bureaucrat gets around to deciding, and the decision, when it occurs, may happen in secret with no transparency or due process.[22]

Many of these problems also occur when internet-infrastructure companies block, filter, or take down content. If end users are blocked, or their speech is taken down, they do not get to speak until somebody in the infrastructure company decides that they have permission. This blocking or removal occurs without any judicial determination of whether their speech is protected or unprotected, without any Bill of Rights protections, without any due process rights to a hearing before the action is taken, or indeed, without any obligation to consider and resolve end-user objections promptly.[23] Rather, some company functionary or bureaucrat—or algorithm—decides whether and when they get to speak.[24]

---

19. Balkin, Algorithmic Society, supra note 6, at 1177–79; Balkin, Old-School/New-School, supra note 6, at 2299, 2309–10, 2318–20.

Mention prior restraints and most lawyers will think of judicial injunctions like the injunction that the Nixon Administration sought against the publication of the Pentagon Papers in New York Times Co. v. United States, 403 U.S. 713 (1971) (per curiam). However, the concept of prior restraints is much older; it originally concerned prior restraints by executive authorities against those who owned and operated printing presses. See Fredrick Seaton Siebert, Freedom of the Press in England 1476–1776: The Rise and Decline of Government Control 21–30 (1952) (discussing the history of administrative prior restraint); Philip Hamburger, The Development of the Law of Seditious Libel and the Control of the Press, 37 Stan. L. Rev. 661, 673 (1985) (explaining how licensing systems allowed the Crown to control the use of printing presses). Like the internet in our own day, the printing press was a powerful technology of mass distribution and therefore feared by the state, which sought to control its dangers.

20. Balkin, Old-School/New-School, supra note 6, at 2316–17; see also Thomas I. Emerson, The Doctrine of Prior Restraint, 20 Law & Contemp. Probs. 648, 657–58 (1955).

21. Balkin, Old-School/New-School, supra note 6, at 2316–17; see also Emerson, supra note 20, at 657.

22. Balkin, Old-School/New-School, supra note 6, at 2316–17; see also Emerson, supra note 20, at 657–58.

23. Balkin, Algorithmic Society, supra note 6, at 1196–98; see also Balkin, Old-School/New-School, supra note 6, at 2318–19 (explaining how governments and cooperating private companies filter and block content without affording speakers due process).

24. See James Grimmelmann, The Virtues of Moderation, 17 Yale J.L. & Tech. 42, 63–65 (2015) (describing cost and efficiency advantages of moderation by computer code); Kate Klonick, The New Governors: The People, Rules, and Processes Governing Online

In this way, our twenty-first-century digital world has recreated the prior restraints of the sixteenth and seventeenth centuries, offering a twenty-first-century version of administrative prior restraint. There are two important differences, of course. First, although some content is blocked at the outset, other content is removed after appearing for a brief period of time.[25] Second, the restraint is not at the hands of government bureaucrats, but at the hands of privately owned companies who act to avoid threats of liability by nation-states.[26]

B. *Public–Private Cooperation and Co-optation*

This leads to the next key feature of new-school speech regulation: public–private cooperation and co-optation. Governments attempt to coax, cajole, or coerce private-infrastructure owners to do their bidding and to help them surveil and regulate speech.[27] Public–private cooperation—or co-optation—is a natural consequence of new-school speech regulation.

First, the technical capacities of infrastructure owners for identifying and removing content far outstrip those of most countries; hence it is easier to get private companies to perform these tasks for the government.[28]

Second, new-school speech regulation often depends on data surveillance—or else is in aid of data surveillance—because many methods of speech regulation require some ability to know what end users are doing.[29] Owners of private infrastructure are essential to effective data collection and surveillance; indeed, the very same infrastructure that makes broad participation in free expression possible is also the infrastructure that facilitates widespread digital surveillance.[30]

Third, complementary incentives drive nation-states to develop new-school speech regulation and private-infrastructure owners to cooperate

---

Speech, 131 Harv. L. Rev. 1598, 1635–48 (2018) [hereinafter Klonick, New Governors] (describing bureaucracies at Facebook, YouTube, and Twitter); Katrin Bennhold, Germany Acts to Tame Facebook, Learning from Its Own History of Hate, N.Y. Times (May 19, 2018), https://www.nytimes.com/2018/05/19/technology/facebook-deletion-center-germany.html (on file with the *Columbia Law Review*) (describing Facebook's bureaucratic operations in Germany, organized to enforce Germany's NetzDG law). These bureaucrats apply rules and filters to regulate content, either ex ante (preventing publication of content uploaded to the site) or ex post (taking down content that has already been published).

As a result of public pressure and media coverage, Facebook recently released some of its guidelines for content moderation. Julia Carrie Wong & Olivia Solon, Facebook Releases Content Moderation Guidelines—Rules Long Kept Secret, Guardian (Apr. 24, 2018), https://www.theguardian.com/technology/2018/apr/24/facebook-releases-content-moderation-guidelines-secret-rules [https://perma.cc/F3A3-LYFX].

25. See Klonick, New Governors, supra note 24, at 1635.

26. Balkin, Algorithmic Society, supra note 6, at 1176.

27. See id. at 1179–80; Balkin, Old-School/New-School, supra note 6, at 2324 ("Public/private cooperation and co-optation are hallmarks of new-school speech regulation.").

28. Balkin, Algorithmic Society, supra note 6, at 1175.

29. Balkin, Old-School/New-School, supra note 6, at 2304–05.

30. Id. at 2297.

with this regulation, whether grudgingly or willingly.[31] It is usually easier for nation-states to regulate the infrastructure operator or owner than to locate and regulate individual speakers: There may be too many speakers, they may be anonymous or not even people, they may be difficult to find, or they may be located outside of the nation-state's jurisdiction.[32] Conversely, infrastructure providers are usually easier to locate, and most have good reasons to be receptive to state pressure.[33] They want to make money, and they want to expand their markets to reach customers within the nation-state's jurisdiction.[34] Even if infrastructure companies strongly believe in civil liberties and would rather not abridge the speech of their customers and end users, they may nevertheless conclude that cooperating with nation-states better furthers their profit-making goals.[35]

Fourth, market incentives and repeated public–private interactions have also driven the development of private governance and new-school speech regulation. Infrastructure owners' technical capacities for surveillance and control continue to grow over time, not only because of market competition and demands from business partners but also as a result of continual political pressure from nation-states and the European Union.[36] The more powerful infrastructure operators become, and the greater their capacity for governance of large populations of end users, the more valuable targets they become for new-school speech regulation.

The result is a burgeoning dialectic of governing power and public–private cooperation. Private-infrastructure companies develop ever greater governing capacities.[37] Nation-states attempt to co-opt these capacities through coercion or threats of regulation. This, in turn, causes increased development of governing, surveilling, and regulatory capacities. And this, in turn, makes private-infrastructure owners even more tempting targets for government pressure—because private companies can no longer pretend that they cannot actually do what governments want them to do.[38]

This dialectic encourages new-school speech regulation, making it ever more important to nation-states as a method of surveilling, regulating, and controlling forbidden speech and conduct on the internet. This dialectic was not so obvious in the early days of the internet, before the rise of social media companies, when surveillance and filtering techniques were far more primitive. But as technology companies grew,

---

31. Balkin, Algorithmic Society, supra note 6, at 1180–81.

32. Balkin, Old-School/New-School, supra note 6, at 2338.

33. Id. at 2305.

34. Balkin, Algorithmic Society, supra note 6, at 1179–80, 1182.

35. See Balkin, Old-School/New-School, supra note 6, at 2329 (describing pressure placed on private enterprises to stop doing business with WikiLeaks).

36. See Balkin, Algorithmic Society, supra note 6, at 1180–81.

37. Investments in capacity will depend on a company's place in the digital infrastructure. Search-engine companies like Google and social media companies like Facebook may invest far more in surveillance and control technologies than DNS registrars. Id. at 1182.

38. Id. at 1180–82.

expanded internationally, and became ever more technically proficient, nation-states began to demand more and more from them.[39]

II. PRIVATE GOVERNANCE AND PRIVATE BUREAUCRACY

A. *Private Governance*

Technology companies' ever-expanding capacities for private surveillance and control lead naturally to viewing them as a new form of private governance. By this I mean that we should think of private-infrastructure owners—and especially social media companies—as governing online speakers, communities, and populations, rather than thinking of them as merely facilitating or hindering digital communication.[40] Instead of viewing digital-infrastructure companies as mere conduits or platforms, we should recognize them as the governors of social spaces.

Professor Kate Klonick has developed this idea in her study of the emergence of internal bureaucracies in social media companies such as Facebook, Twitter, and YouTube.[41] She explains how the concept of community governance, and the creation of large global bureaucracies, emerged almost by accident as social media companies sought to enforce their terms-of-service agreements and had to respond to pressure from various nation-states to control or curb speech that these countries regarded as illegal or undesirable.[42] Faced with an unruly and unpredictable collection of all types of people from around the world (not to mention agents of various nation-states), these companies learned that they had to govern—that is, promulgate and enforce the values and norms that their communities stood for.[43] They did so through a combination of contract (that is, terms of service or end-user license agreements) and code.[44] Over time, social media companies, which originally thought of themselves only as technology companies, accepted their role as community governors and developed elaborate bureaucracies, which are effectively governance structures.[45]

---

39. See id. at 1180–81.

40. See id. at 1194–97; Klonick, New Governors, supra note 24, at 1602–03.

41. See Klonick, New Governors, supra note 24.

42. See id. at 1618–30.

43. Balkin, Algorithmic Society, supra note 6, at 1195–97.

44. Id. at 1186–87.

45. Id. at 1181–82; see also Klonick, New Governors, supra note 24, at 1634–35 (describing the development of Facebook's complex "Community Standards" and the evolution of content moderation at Facebook and YouTube). The evolution of social media companies mirrors the experience of system administrators for online worlds in the early days of the internet. These system administrators were sometimes called "game gods" because they created and ran multiplayer online games. People occasionally abused these spaces by finding exploits in the games or harassing and trolling other players. Eventually, the game gods had to step in to govern the space, specifying what was or was not a permitted move in the game and sanctioning or expelling people who would not behave properly. See, e.g.,

The task of governing online spaces need not be wholly public spirited. It may be driven by market incentives or by the quest for economic and political power. Facebook has adopted community rules because of its business model, which requires that its space be safe, attractive, and absorbing for its billions of users around the world.[46] Social media companies cannot afford to scare off their customers because they need to capture end users' scarce attention to make money. The business model of social media companies requires vast numbers of individuals to repeatedly check the site, read the site, and post to the site so that the company can sell their scarce attention to advertisers.[47]

Companies like Facebook generate growth—and thus please the demands of their shareholders—in one of two ways: First, they can expand their membership to more people around the world. Second, they can gain a greater share of their end users' attention.[48] The first strategy offers limited possibilities for a company as large as Facebook; therefore, the second strategy begins to dominate. As Professor Tim Wu has pointed out, social media companies have an incentive to make their services addictive so that they can garner a larger share of their end users' attention.[49]

Before Roblox: An Online Rape When Cyberspace Was New, Village Voice (July 25, 2018), https://www.villagevoice.com/2018/07/25/before-roblox-an-online-rape-when-cyberspace-was-new/ [https://perma.cc/FT8E-KF9U] (reprinting Julian Dibbel, A Rape in Cyberspace, Village Voice (Dec. 23, 1993)).

46. See Balkin, Algorithmic Society, supra note 6, at 1181; see also Danielle Keats Citron & Helen Norton, Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age, 91 B.U. L. Rev. 1435, 1454–55 (2011) (arguing that intermediaries regulate speech as a matter of corporate responsibility and to protect profits); Klonick, New Governors, supra note 24, at 1625 ("Platforms create rules and systems to curate speech out of a sense of corporate social responsibility, but also, more importantly, because their economic viability depends on meeting users' speech and community norms.").

47. See Klonick, New Governors, supra note 24, at 1627 ("[T]he primary reason companies take down obscene and violent material is the threat that allowing such material poses to potential profits based in advertising revenue.").

48. See Peter Eavis, How You're Making Facebook a Money Machine, N.Y. Times: The Upshot (Apr. 29, 2016), https://www.nytimes.com/2016/04/30/upshot/how-youre-making-facebook-a-money-machine.html (on file with the *Columbia Law Review*) ("[T]hat constant lure [to check Facebook], a fix you can easily satisfy both on a phone and a desktop computer, explains why Facebook is pulling ahead of every other large technology company right now."); James B. Stewart, Facebook Has 50 Minutes of Your Time Each Day. It Wants More., N.Y. Times (May 5, 2016), https://www.nytimes.com/2016/05/06/business/facebook-bends-the-rules-of-audience-engagement-to-its-advantage.html (on file with the *Columbia Law Review*) ("Time is the best measure of engagement, and engagement correlates with advertising effectiveness . . . . And time enables Facebook to learn more about its users—their habits and interests—and thus better target its ads.").

49. See Tim Wu, The Attention Merchants: The Epic Scramble to Get Inside Our Heads 289–302 (2016) [hereinafter Wu, Attention Merchants] (describing how social media companies attempt to attract advertisers by cornering the market on attention and addicting customers); Tim Wu, Opinion, Subtle and Insidious, Technology Is Designed to Addict Us, Wash. Post (Mar. 2, 2017), https://www.washingtonpost.com/opinions/subtle-and-insidious-technology-is-designed-to-addict-us/2017/03/02/5b983ef4-fcee-11e6-99b4-9e613afeb09f_

Twentieth-century freedom of speech faced a problem of scarcity of access to media.[50] Twenty-first-century freedom of speech faces the problem of scarcity of attention.[51] The logic of scarcity of attention drives the business models of many twenty-first-century digital companies that attract end users by offering free (or subsidized) services in exchange for brokering end users' attention to advertisers.

The capitalist logic of digital media services requires continuous growth either through expansion of membership or through expansion of attention.[52] To seize attention, a social media platform must have both absorbing content *and* provide a community in which people feel safe; otherwise end users will not spend time on the site. Hence, the economic logic of advertiser-driven social media leads them to become governors of their spaces.

Moreover, to sell end users' attention to advertisers, it is necessary to know things about them so that advertising dollars are not wasted. The ability to serve different ads to different audiences requires knowledge about audiences, and thus the collection of ever-greater amounts of data about end users. The logic of digital capitalism, in other words, also drives companies toward surveillance as well as governance.[53]

The same logic of digital capitalism that leads to governance and surveillance of end users also leads to the creation of bureaucracies, which consist of the company's digital workers using easy-to-apply rules for deciding vast numbers of cases and controversies, while pushing a

---

story.html [https://perma.cc/78ND-N8SD] [hereinafter Wu, Subtle and Insidious] ("[F]or a product like Facebook, success and user addiction are the same thing.").

50. See Red Lion Broad. Co. v. FCC, 395 U.S. 367, 390 (1969) ("Because of the scarcity of radio frequencies, the Government is permitted to put restraints on licensees in favor of others whose views should be expressed on this unique medium.").

51. See Zeynep Tufekci, Twitter and Tear Gas: The Power and Fragility of Networked Protest 271 (2017); Jack M. Balkin, Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, 79 N.Y.U. L. Rev. 1, 7 (2004) [hereinafter Balkin, Digital Speech] ("The digital revolution made a different kind of scarcity salient. It is not the scarcity of bandwidth but the scarcity of audiences, and, in particular, scarcity of audience attention."); Herbert A. Simon, Designing Organizations for an Information-Rich World, *in* Computers, Communications, and the Public Interest 37, 40 (Martin Greenberger ed., 1971) ("[A] wealth of information creates a poverty of attention.").

52. Other parts of the internet infrastructure have different business models, but all require growth over time.

53. Zeynep Tufekci, Opinion, Facebook's Surveillance Machine, N.Y. Times (Mar. 19, 2018), https://www.nytimes.com/2018/03/19/opinion/facebook-cambridge-analytica.html (on file with the *Columbia Law Review*) [hereinafter Tufekci, Facebook's Surveillance Machine] ("Facebook makes money, in other words, by profiling us and then selling our attention to advertisers, political actors and others. These are Facebook's true customers, whom it works hard to please."). These business models, and the incentives they create, are examples of what Professor Shoshana Zuboff calls "surveillance capitalism." See Shoshana Zuboff, Big Other: Surveillance Capitalism and the Prospects of an Information Civilization, 30 J. Info. Tech. 75, 75 (2015) (defining "surveillance capitalism" as a "new logic of accumulation" and a "new form of information capitalism [that] aims to predict and modify human behavior as a means to produce revenue and market control").

small number of more complicated cases up the chain of decision.[54] This follows naturally from the global nature of both the company's end users and its employees: When both content and employees come from everywhere, social media companies need simple, easily understandable, easy-to-apply rules that can be followed uniformly.[55]

Another method for lowering costs and ensuring uniformity is to substitute algorithmic for human judgment.[56] Algorithmic employees cost even less than human employees: They do not have families, they do not take coffee breaks, and they can do some—but by no means all—of the work of discovery and selection that human employees can do. Algorithms may be especially useful in ex ante blocking of content—for example, identifying child pornography or preventing the upload of content that has been digitally watermarked as copyright protected.[57]

Governance by Facebook, Twitter, and YouTube has many aspects of a nineteenth-century autocratic state, one that protects basic civil freedoms but responds to public opinion only in limited ways. The end users—akin to the citizens or subjects—are in effect unpaid laborers for the site, in the same way that anyone who uses open-source software and reports bugs is an unpaid laborer for the open-source project.[58] When end users spot bugs, make complaints, or demand new features, this helps inform the company, its bureaucrats, and its programmers how best to attract and mollify end users and keep profits flowing. Every end user is a potential reporting or surveillance device for maintaining community standards.[59] Every time end users complain about racist speech or trolling, they are in effect working for Facebook because they provide the company with information that helps it enforce its community standards.[60]

---

54. Klonick, New Governors, supra note 24, at 1638–42 (describing the structure of ex post review of content by human moderators).

55. See id. at 1632–34 (describing how social media companies moved to concrete rules that can be consistently applied because of the global diversity of their workforce); id. at 1642 (noting that "Facebook's Community Standards were applied globally, without differentiation along cultural or national boundaries").

56. Id. at 1636–37 (describing the use of algorithmic systems to protect copyright interests and to block spam); cf. Jack M. Balkin, The Path of Robotics Law, 5 Calif. L. Rev. Cir. 45, 46, 55–58 (2015), http://www.californialawreview.org/wp-content/uploads/2015/06/Balkin-Circuit.pdf [https://perma.cc/D8F6-8KSJ] (describing the "substitution effect[s]" produced by attempts to substitute robots and artificial intelligence agents for human beings.).

57. Klonick, New Governors, supra note 24, at 1636–37.

58. See 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), https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech [https://perma.cc/DQ6P-SGH6] ("[U]sers are not so much customers as uncompensated digital laborers who play dynamic and indispensable functions (despite being largely uninformed about the ways in which their labor is being used and capitalized).").

59. See id.

60. See id.

Because the governance of social media companies is generally autocratic, their governance policies are, for the most part, nontransparent and waived whenever necessary or convenient.[61] There is normally little in the way of due process for end users, much less a right to a hearing either before or immediately after sanctions are applied.[62] Companies often make special exceptions for powerful and influential actors and organizations.[63] But if the speaker is a "puny anonymit[y],"[64] it is far more likely that a social media company will sanction or ban the speaker.[65]

B. *Should Private Governance Be Private?*

Nevertheless, the best alternative to this autocracy is not the imposition of First Amendment doctrines by analogy to the public forum or the company town.[66] Of course, new-school speech regulation may violate the First Amendment—because the state has passed laws that pressure infrastructure providers to do its bidding.[67] But when we focus

---

61. See id. ("The details of moderation practices are routinely hidden from public view, siloed within companies and treated as trade secrets when it comes to users and the public.").

62. See id.

63. See Klonick, New Governors, supra note 24, at 1654–55 (noting that Facebook may "disproportionately favor people with power over the individual users" (footnote omitted)). For example, the President of the United States is a serial violator of the community policies of Facebook and Twitter, but neither site has yet banned him, and they appear unlikely to do so. See id. at 1655 (noting Facebook founder Mark Zuckerberg's decision to keep Trump on Facebook despite his violations of the company's hate speech policies); Doug Bolton, This Is Why Facebook Isn't Removing Donald Trump's 'Hate Speech' from the Site, Independent (Dec. 15, 2015), http://www.independent.co.uk/life-style/gadgets-and-tech/news/donaldtrump-muslim-hate-speech-facebook-a6774676.html [https://perma.cc/L5HH-DX9S] (noting special rules for Trump on Facebook); Arjun Kharpal, Why Twitter Won't Take Down Donald Trump's Tweet Which North Korea Called a 'Declaration of War,' CNBC (Sept. 26, 2017), https://www.cnbc.com/2017/09/26/donald-trump-north-korea-twitter-tweet.html [https://perma.cc/S9FT-VE9S] (noting special rules on Twitter for Trump); Deepa Seetharaman, Facebook Employees Pushed to Remove Trump's Posts as Hate Speech, Wall St. J. (Oct. 21, 2016), https://www.wsj.com/articles/facebook-employees-pushed-to-remove-trump-posts-as-hate-speech-1477075392 (on file with the *Columbia Law Review*) (noting special rules for Trump on Facebook).

64. Abrams v. United States, 250 U.S. 616, 629 (1919) (Holmes, J., dissenting) (arguing, among other things, that the Court should have reversed the Espionage Act convictions of defendants because few people would have paid attention to them).

65. See Kate Klonick, Facebook v. Sullivan 28–31 (Apr. 2018) (unpublished manuscript) (on file with the *Columbia Law Review*) (describing how Facebook has reinterpreted concepts like "public figure" and "newsworthiness" to govern its community).

66. Cf. Marsh v. Alabama, 326 U.S. 501, 505–09 (1946) (holding that a company town could not refuse access to Jehovah's Witnesses engaging in leafletting and would be treated as the effective equivalent of a government-owned public forum).

67. See, e.g., Reno v. ACLU, 521 U.S. 844, 874 (1997) (striking down a federal law that required filtering and blocking of content purportedly harmful to minors); ACLU v. Mukasey, 534 F.3d 181, 197–98 (3d Cir. 2008) (striking down the federal Child Online Protection Act, which required sites to filter, segregate, and block content); Ctr. for

on social media governance that is *not* the result of new-school speech regulation, our analysis should be different. Social media companies should recognize and protect free speech values as well as due process values in the resolution of complaints. Even so, it is generally a bad idea to hold social media spaces to the same standards as municipal governments under the First Amendment.

Imposing the same First Amendment doctrines that apply to municipalities to social media companies would quickly make these spaces far less valuable to end users, if not wholly ungovernable. First Amendment law significantly limits the ability of municipalities to regulate anonymous or pseudonymous speech in public forums;[68] yet sites may want to require real names or easily identifiable pseudonyms in order to prevent cyberbullying, harassment, and trolling. Under current First Amendment doctrine, sites might not be able to ban hate speech or other kinds of abusive and emotionally upsetting speech that make the site far less valuable for the vast majority of customers.[69] Municipalities can ban fighting words,[70] but speakers on the internet may be nowhere near the recipients of their venom so that an immediate breach of the peace is highly unlikely.[71] Although the Supreme Court has not declared the tort of intentional infliction of emotional distress unconstitutional, it has been careful to suggest that speech that causes emotional distress is protected if it discusses matters of public concern.[72]

A final problem is that, unlike municipalities, social media sites cannot levy damages or fines. They have only limited sanctions for

---

Democracy & Tech. v. Pappert, 337 F. Supp. 2d 606, 611 (E.D. Pa. 2004) (striking down a state law that required filtering of child pornography).

68. See, e.g., McIntyre v. Ohio Elections Comm'n, 514 U.S. 334, 357 (1995) (striking down a law banning distribution of anonymous campaign literature); Talley v. California, 362 U.S. 60, 64–65 (1960) (striking down a law banning handbills unless they identified the distributor's name and address).

69. See, e.g., R.A.V. v. City of St. Paul, 505 U.S. 377, 381 (1992) (striking down a city ordinance that banned fighting words directed at others on the basis of race, ethnicity, gender, or religion).

70. See Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942) (stating that fighting words constitute words that by their nature result in an immediate breach of the peace). *Chaplinsky* stated that words that "by their very utterance inflict injury" also constitute fighting words. Id. But in *R.A.V.*, the Court explained that the fighting words doctrine does not allow states to punish speech that merely causes emotional upset. *R.A.V.*, 505 U.S. at 414 (White, J., concurring in the judgment); see also United States v. Eichman, 496 U.S. 310, 318 (1990) (explaining that although "desecration of the flag is deeply offensive to many[,] . . . the same might be said . . . of virulent ethnic and religious epithets").

71. See Texas v. Johnson, 491 U.S. 397, 409 (1989) (holding that the government may not classify provocative ideas as fighting words without "careful consideration of the actual circumstances surrounding such expression," including "whether the expression 'is directed to inciting or producing imminent lawless action and is likely to incite or produce such action'" (quoting Brandenburg v. Ohio, 395 U.S. 444, 447 (1969))).

72. See Snyder v. Phelps, 562 U.S. 443, 458 (2011) ("Such speech cannot be restricted simply because it is upsetting or arouses contempt.").

misbehavior: denying access to the site, either temporarily or permanently, and removing some or all of an offender's previous content.[73] Denying access to a public forum and removing all of an end user's content as punishment for previous conduct may create problems under prior restraint doctrine. Take the example of defamation: Under current First Amendment law, the government—and hence a social media site treated as an arm of the state—might not be able to deny access to an end user and remove some or all of the end user's previous posts because he or she had previously defamed a person. Courts might regard this as a prior restraint and no more constitutional than denying future access to a public park to a person who had previously defamed someone in the park.[74] Moreover, removing all of an end user's content as punishment, even when significant parts of that content constitute protected speech, would seem to raise serious First Amendment problems.

The result is that—at least until the courts begin to treat cyberspaces differently from other public fora—applying First Amendment law would cripple social media sites' abilities to impose civility norms. When spaces seem unsafe and are riddled with racist speech and personal abuse, many people will avoid them.

Second, under a First Amendment regime, social media sites would be unable to curate content in order to provide personalized feeds. The creation of personalized feeds is inevitably content-based—social media sites have to decide what content is likely to be most interesting to their end users.[75] As Professor Tarleton Gillespie has pointed out, social media sites thrive on content-based moderation, even if the moderation is invisible to most users.[76] The same is true of search engines; ideally, their purpose is to help end users reach information that is relevant to their search engine queries. Furthering this task requires multiple content-based distinctions about the relevance and arrangement of links.[77]

---

73. See, e.g., Jack Nicas, Alex Jones and Infowars Content Is Removed from Apple, Facebook and YouTube, N.Y. Times (Aug. 6, 2018), https://www.nytimes.com/2018/08/06/technology/infowars-alex-jones-apple-facebook-spotify.html (on file with the *Columbia Law Review*) (noting that Apple removed the majority of Alex Jones's podcasts as hate speech, Facebook removed four Facebook pages for glorifying violence and dehumanizing speech, and YouTube removed Jones's entire channel for repeated violations of its terms of service).

74. Cf. Near v. Minnesota, 283 U.S. 697, 712–13 (1931) (striking down a Minnesota law that banned future publication of newspapers that had previously published defamatory material).

75. Tarleton Gillespie, Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media 5 (2018) (arguing that all platforms must make content-based decisions in order to govern their spaces).

76. Id. at 6–7.

77. See James Grimmelmann, Speech Engines, 98 Minn. L. Rev. 868, 874, 893–95 (2014) (offering a theory about the proper behavior of search engines, which argues that a search engine acts as a "trusted advisor").

The free speech values characteristic of this environment—and that promise to work best within it—do not necessarily conform to First Amendment doctrine. To be sure, the degree of free speech protection that exists on these sites is due in no small part to the fact that they were originally created by American-led companies and have been deeply influenced by American free speech values.[78] But because social media companies operate around the world, they cannot realistically apply American First Amendment doctrines everywhere, in part because American free speech law requires Americans to tolerate all sorts of things that people in other countries simply will not put up with.[79]

Given its global reach, Facebook has decided to have a worldwide policy, which applies more or less uniformly in every country.[80] According to this policy, the company takes down hate speech and disrespectful speech that would almost certainly be protected by the American First Amendment.[81] This policy is hardly surprising from a company that seeks to do business around the world. A global company that governs a global online community requires global rules on freedom of expression that are not necessarily American free speech rules, much less doctrines originally designed for public streets and parks.

C. *Privatized Bureaucracy*

So far, this Essay has introduced the ideas of old-school and new-school speech regulation and the crucial connection between new-school speech regulation and private governance. The last piece of the puzzle emerges out of this connection. It is the emerging system of *privatized bureaucracy*.[82]

For the reasons described above, new-school speech regulation depends on the expansion and promulgation of private governance.[83] Indeed, new-school speech regulation and private governance egg each other on. As digital-infrastructure companies become increasingly powerful in governing their spaces and collecting and analyzing content from their end users, nation-states may demand more from them through new-

78. See Klonick, New Governors, supra note 24, at 1621–22 (describing the influence of American free speech values on the moderation policies of Facebook, Twitter, and YouTube).

79. See id. at 1623 (describing problems of applying American free speech norms around the world).

80. Id. at 1642; Julia Angwin & Hannes Grassegger, Facebook's Secret Censorship Rules Protect White Men from Hate Speech but Not Black Children, ProPublica (June 28, 2017), https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms [https://perma.cc/RDF8-JEQB] (describing Facebook's attempts to enforce its hate speech rules worldwide and the arbitrariness of its categories).

81. See supra notes 67–72 and accompanying text.

82. See Balkin, Algorithmic Society, supra note 6, at 1226–28 (describing how nation-states co-opt private governance to create a new kind of bureaucracy).

83. See supra section II.A.

school speech regulation. As nation-states attempt to co-opt and coerce private-infrastructure operators, they increasingly attempt to get private companies to take on state functions of speech regulation and surveillance. As social media and search engine companies develop governing bureaucracies and algorithms, nation-states seek to harness that capacity to accomplish the nation-state's governance goals. These processes lead to a new phenomenon: privatized bureaucracy. Bureaucracies within private-infrastructure companies (including their contractors) serve as the front line for the nation-state's governance of online speech and conduct.

Two examples demonstrate how this phenomenon works: the right to be forgotten, which applies in the European Union generally, and Germany's recent NetzDG law.[84]

Consider how the European Union protects the right to be forgotten. Suppose that a petitioner objects to the presence of an embarrassing article on a search engine such as Google. The European Court of Justice (ECJ) has ordered Google to develop a bureaucratic system for deciding in the first instance whether a particular article should be delinked from its search engines. If the petitioner disagrees with Google's decision, he or she can sue in the courts.[85] This is, in essence, a system of administrative law, requiring an exhaustion of administrative remedies before one can use the court system. But the administrative agency in this case is a private company.

The ECJ chose this solution because the European Union and its member states lack the technical capacity to monitor the internet and protect the right to be forgotten on their own.[86] The number of complaints is likely to be very large and processing these complaints would require the creation of a sizeable new bureaucracy in each member state.[87] In order to protect those rights that the ECJ asserts should exist under European law, the European Union has essentially deputized a private company to serve as its bureaucracy.[88] This deputizing of privately owned infrastructure companies is the culmination of the logic of new-school speech regulation.[89]

---

84. See supra notes 1–2 and accompanying text.

85. See generally European Comm'n, Factsheet on the "Right to Be Forgotten" Ruling (C-131/12), https://www.inforights.im/media/1186/cl_eu_commission_factsheet_right_to_be-forgotten.pdf [https://perma.cc/ERG6-D3EP] (last visited Aug. 1, 2018) (describing Google's obligations under European law to make initial determinations about the right to be forgotten); see also Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, 2014 E.C.R. 317, ¶¶ 93–94 (explaining that parties should first apply for relief to Google, which has the initial obligation to investigate).

86. Balkin, Algorithmic Society, supra note 6, at 1206 (describing the reasons why nation-states privatize bureaucratic governance).

87. Id. at 1206–07.

88. Id. at 1207.

89. Hate speech regulation offers another example. In May 2016, the European Commission reached an agreement with Facebook, Microsoft, Twitter, and YouTube to create a "code of conduct on countering illegal hate speech online" that requires them to work with local authorities and NGOs to identify and take down hate speech in social media. Code of Conduct on Countering Illegal Hate Speech Online 1–3, https://www.statewatch.

Similarly, Germany's new NetzDG law was designed to co-opt social media companies into monitoring and taking down prohibited content in Germany, including hate speech.[90] Although some internet companies, such as Facebook, already have hate speech policies, Germany demanded stricter enforcement and prompt takedown—essentially within 24 hours of notice for "manifestly unlawful" speech.[91] Failure to comply with the state's requirements leads to sanctions against the company.[92]

From one perspective, NetzDG is just collateral censorship—a nation-state puts pressure on digital-infrastructure companies to block, take down, and censor content by end users. But from another perspective, NetzDG involves an agreement between the German state and various private companies in which the companies act as a private bureaucracy that implements the state's speech policies. Because Germany currently lacks the technical capability to perform this task on its own, it coerces or co-opts Facebook, Google, and Twitter to do it instead. Once again, this is the logical outcome of new-school speech regulation.

One might make four different objections to a government program like NetzDG, and it is important to distinguish them because they represent four different objections to new-school speech regulation. Three of these concern speech, while the last concerns surveillance.

First, one might object to Germany's substantive hate speech doctrines as insufficiently speech protective. This objection is really not about internet regulation at all, for Germany would presumably enforce the same restrictions on speech that did not appear on social media sites.

Second, one might object that Germany will attempt to impose its content regulation outside of its geographical boundaries. Because German citizens may access the internet everywhere (or use internet proxies to simulate being outside the country), Germany may eventually

---

org/news/2017/sep/eu-com-illegal-content-online-code-of-conduct.pdf [https://perma.cc/CG3N-2YX7] (last visited Aug. 1, 2018); Press Release, European Comm'n, Countering Illegal Hate Speech Online #NoPlace4Hate (July 11, 2018), https://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=54300 [http://perma.cc/L29F-3YGP].

90. Jenny Gesley, Germany: Social Media Platforms to Be Held Accountable for Hosted Content Under "Facebook Act," Library of Cong.: Glob. Legal Monitor (July 11, 2017), http://www.loc.gov/law/foreign-news/article/germany-social-media-platforms-to-be-held-accountable-for-hosted-content-under-facebook-act/ [https://perma.cc/XKT3-PFG9] ("[T]he so-called Facebook Act . . . aims to combat hate speech and fake news in social networks."); Overview of the NetzDG Network Enforcement Law, Ctr. for Democracy & Tech. (July 17, 2017), https://cdt.org/insight/overview-of-the-netzdg-network-enforcement-law/ [https://perma.cc/Z6WS-Q9A3] (summarizing the new law).

91. Network Enforcement Act, Sept. 1, 2017, BGBl I, at 3352, § 3(1) (Ger.), https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG_engl.pdf?__blob=publicationFile&v=2 [https://perma.cc/W2B8-JWHT] ("The provider of a social network shall maintain an effective and transparent procedure for handling complaints about unlawful content . . . ."); id at § 3(2).2 ("The procedure shall ensure that the provider of the social network . . . removes or blocks access to content that is manifestly unlawful within 24 hours of receiving the complaint . . . .").

92. See id. §§ 4(1).3, 4(1).4, 4(2).

demand global jurisdiction for its speech regulations. Similarly, the European Union and its member states may try to enforce the right to be forgotten around the world.[93]

In some cases, it may not even be necessary to formally assert jurisdiction over a company's operations around the world. Governments can achieve similar effects through indirection. They may encourage companies to alter their terms of service to better conform to the state's substantive law.[94] Then governments and associated law enforcement agencies can inform companies that content violates the company's terms of service and request removal, achieving worldwide enforcement by other means.[95]

This objection is distinct from the question of whether Germany or the European Union have adopted the correct substantive understanding of the right of freedom of speech. Instead, this objection is related to new-school speech regulation because enforcing universal jurisdiction involves coercing or co-opting infrastructure providers to enforce particular speech norms universally. The better infrastructure providers are at locating and enforcing speech regulations around the world, the more nation-states may be tempted to harness these technical capacities for their own ends.

Third, quite apart from concerns about substantive law and global jurisdiction, speakers get no judicial determination of whether their speech is legally protected or unprotected when private companies block, censor, or take down their speech. Instead, some nontransparent form of private governance or bureaucracy serves as prosecutor, judge, jury, and executioner. Speakers thus get no due process protections of notice or hearing. This is a problem of collateral censorship, which, as noted above in section I.A, has aspects of administrative prior restraint.

Fourth, when nation-states co-opt private infrastructure to regulate speech, they may also co-opt private infrastructures' capacities for surveillance, data collection, and analysis to solve their own problems of governance and control.[96]

One must pay attention to all four of these issues when considering any question of online speech today. The first issue—the question of

---

93. See, e.g., Miquel Peguera, Right to Be Forgotten and Global Delisting: Some News from Spain, Stanford Ctr. for Internet & Soc'y (Dec. 17, 2017), http://cyberlaw.stanford.edu/blog/2017/12/right-be-forgotten-and-global-delisting-some-news-spain [https://perma.cc/A6P9-LRUP] (describing the ongoing controversy over global delisting to protect the right to be forgotten).

94. See Danielle Citron, Extremist Speech, Compelled Conformity, and Censorship Creep, 93 Notre Dame L. Rev. 1035, 1038 (2018) ("By insisting upon changes to platforms' speech rules and practices, EU regulators have exerted their will across the globe. Unlike national laws that apply only within a country's borders, terms of service apply wherever platforms are accessed.").

95. Id.

96. Balkin, Old-School/New-School, supra note 6, at 2298–99, 2304–05, 2329–30.

substantive standards—is often the most salient to Americans because Americans generally have a much more libertarian free speech policy than the rest of the world. The second is the looming possibility of global jurisdiction, with countries vying with each other to impose their parochial speech regulations on the entire world—a race to the top, which, in American eyes, would be a race to the bottom. The third and fourth problems—privatized bureaucracy and surveillance—arise from the combination of new-school speech regulation and ever more technically effective private governance.

The result is a complicated array of relationships of power, control, and surveillance. End users, citizens, legacy media, and civil-society organizations are now targets of both old-school and new-school speech regulation by nation-states, as well as the subjects of private governance by digital-infrastructure companies.

End users are by no means powerless in this new environment—for example, coordinated campaigns by end users and mass media may pressure companies to change their policies.[97] The larger point, however, is that speakers face multiple threats from public *and* private governance and power, instead of merely the traditional threats of old-school speech regulation.

III. PROTECTING FREE SPEECH VALUES IN A PLURALIST SYSTEM OF REGULATION

If the characteristic feature of free speech regulation in our time is a triangle that combines new-school speech regulation with private governance, then the best way to protect free speech values today is to combat and compensate for that triangle's evolving logic of public and private regulation. The solution will not necessarily—or even primarily—involve enforcing the doctrines of the First Amendment jot for jot against private-infrastructure providers. To be sure, it *will* concern the free speech *values* that animate the First Amendment. But the best way to protect those values is not to apply doctrines developed for states as rules for private actors. Instead, protecting free speech in a digital age often involves technical, regulatory, and administrative solutions that apply in contexts where the First Amendment does not reach.[98] Judge-made

97. See, e.g., Yuki Noguchi, Facebook Changing Privacy Controls as Criticism Escalates, NPR: The Two-Way (Mar. 28, 2018), https://www.npr.org/sections/thetwo-way/2018/03/28/597587830/criticism-prompts-facebook-to-change-privacy-controls (on file with the *Columbia Law Review*).

98. See Balkin, Digital Speech, supra note 51, at 2, 50–52; Jack M. Balkin, The Future of Free Expression in a Digital Age, 36 Pepp. L. Rev. 427, 432–33 (2009) [hereinafter Balkin, Future of Free Expression] (arguing that policies that promote the development of the infrastructure of free expression "better serve the interests of freedom of speech in the long run, even though such innovation policies do not, at least on their face, seem to be about government censorship").

doctrines of First Amendment law can do only limited work, and sometimes they will actually hinder necessary reforms.[99]

Protecting free speech values in a pluralist model has two basic goals. The first goal is to prevent or ameliorate, as much as possible, collateral censorship and new forms of digital prior restraint. The second goal is to protect people from new methods of digital surveillance and manipulation—new methods that emerged from the rise of large multinational companies that depend on the collection, surveillance, analysis, control, and distribution of personal data.[100]

The four problems mentioned above—collateral censorship, privatized prior restraint, surveillance, and manipulation—are predictable features of private governance, of new-school speech regulation, and of public–private cooperation and co-optation. To protect free speech values in the new era, one must aim at all of them.

A. *Permissible Government Regulations: Structural Reform and Procompetition Policies*

Protecting free speech values does not mean rejecting all government regulation of digital infrastructure. Some regulations do not produce the problems of new-school speech regulation—collateral censorship and digital prior restraint.[101] To the contrary, they actually protect free speech values. These kinds of regulations may not only be perfectly appropriate, they may actually be necessary to provide practical freedom for end users.

One example of permissible regulation is the structural regulation of telecommunications facilities. Examples include municipal broadband projects; open-access rules, which make it possible to have many different kinds of internet service providers;[102] and network neutrality rules, which

---

99. See Balkin, Digital Speech, supra note 51, at 50–52; Balkin, Future of Free Expression, supra note 98, at 437–39, 443–44.

100. The success of social media companies, for example, depends on increasing advertising revenues, which depends on garnering ever-larger shares of scarce attention. Grabbing scarce attention, in turn, depends on discovering ever-new ways to attract and manipulate end users and collect and analyze the data that emerges from their behavior and interactions. See Tufekci, Facebook's Surveillance Machine, supra note 53 (describing Facebook's surveillance of its users to increase advertising revenue). Social media companies are hardly unique in this respect. Many other businesses—and government programs—also depend on the collection, analysis, and use of data to predict behavior and control populations. See Zuboff, supra note 53, at 75–76 (describing surveillance capitalism).

101. See supra section I.A.

102. See Jonathan E. Nuechterlein & Philip J. Weiser, Digital Crossroads: Telecommunications Law and Policy in the Internet Age 192–96 (2d ed. 2013) (explaining the idea of open access as a method for ISP competition and how it was eventually displaced in the United States).

prevent discrimination against content and applications.[103] These structural regulations of infrastructure are not new-school speech regulation: They do not encourage collateral censorship or digital prior restraints, and they do not raise the same problems of due process. In fact, network neutrality helps avoid many of the problems of collateral censorship because broadband providers may not block or slow down traffic based on its content or viewpoint in a network neutrality regime.[104]

A second and very important kind of regulation is procompetition regulation, which might include both antitrust law and media concentration rules.[105] Procompetition policies are important not only because of the potential power of privately owned bureaucracies but also because of their potential vulnerabilities.[106] The hacking of the 2016 election might have been different, and possibly less effective, if there were multiple Facebooks with different affordances, technologies, and advertising models.[107] There is only one Facebook today not simply because of economies of scale and network effects but because Facebook strategically

---

103. See Barbara van Schewick, Internet Architecture and Innovation 72–73, 220 (2010) (arguing that a key feature of all versions of network neutrality is nondiscrimination against content and applications).

104. See id.

105. I use the more general term "procompetition policy" instead of "antitrust" because, at least as it has developed in the United States in the past forty years, antitrust law has tended to focus on consumer welfare, and some scholars, following Robert Bork, have argued that this should be its only focus. See, e.g., NCAA v. Bd. of Regents of the Univ. of Okla., 468 U.S. 85, 107–08 (1984) ("'Congress designed the Sherman Act as a "consumer welfare prescription."' . . . Restrictions on price and output are the paradigmatic examples of restraints of trade that the Sherman Act was intended to prohibit." (citation omitted) (quoting Reiter v. Sonotone Corp., 442 U.S. 330, 343 (1979))); Robert H. Bork, The Antitrust Paradox: A Policy at War with Itself 7 (1978) ("[T]he only legitimate goal of antitrust is the maximization of consumer welfare."). Because social media and search engines offer their services for free, demonstrating that their business practices harm consumer welfare takes some ingenuity—for example, one might focus on control of digital advertising networks. Perhaps more important, an exclusive focus on consumer welfare may miss the point of what is most troubling about these business practices. A larger class of procompetition policies, by contrast, might focus on the effects of anticompetitive behavior on democracy and free expression. Rather than rehash debates about the "true" purposes of current antitrust law, I simply employ the more general term to describe possible reforms.

106. See Sally Hubbard, Fake News Is a Real Antitrust Problem, CPI Antitrust Chron., Dec. 2017, at 1, 1, https://www.competitionpolicyinternational.com/wp-content/uploads/2017/12/CPI-Hubbard.pdf [https://perma.cc/9PME-JHX2] ("Facebook and Google['s] . . . algorithms have an outsized impact on the flow of information, and fake news purveyors can deceive hundreds of millions of users simply by gaming a single algorithm.").

107. See id. at 5 ("Having two dominant algorithms controlling the flow of information enables deception on a massive scale, meaning that the concentration of the search and social markets is directly related to the scope of fake news' damage."); Sean Illing, Why "Fake News" Is an Antitrust Problem, Vox (July 18, 2018), https://www.vox.com/technology/2017/9/22/16330008/eu-fines-google-amazon-monopoly-antitrust-regulation [https://perma.cc/Z6L5-N3ZY] (quoting Sally Hubbard for the proposition that multiplying social media would make it harder for fake news sites to manipulate digital companies' algorithms).

bought up a number of potential competitors, incorporating some of their innovations and blocking others.[108] In this way, it forestalled the development of a wide range of potential competitors and innovators in social media.

More social media competitors, each with differing approaches and goals, would provide more platforms for innovation, more software features, more types of security measures that hackers would have to circumvent, more models for social spaces and communities, and a wider variety of speech policies.[109]

With stronger enforcement of antitrust and procompetition laws, innovations might have proliferated more widely, and we might have a healthier competition among social media companies and their sorting algorithms. Although we cannot be certain that this would have made it harder for foreign propaganda and fake news to proliferate and disrupt our democracy, it is generally harder to attack and compromise twelve targets than to attack and compromise one.

One might object that this degree of fragmentation—we might even call it balkanization—is undesirable.[110] But procompetition policies serve democratic values in a second way. Modern democracies increasingly rely on social media to facilitate public conversation, organize public discussion, and enforce civility norms.[111] Therefore, it is especially important to make sure that there are many such organizations, in order to prevent a small number of powerful for-profit companies from dominating how public opinion is organized and governed. Social media enforcement of civility norms is imperfect and often arbitrary,[112] and some organizations, like Facebook, attempt to impose the same standards around the world.[113] Thus, when people expect and even demand that a multinational corporation like Facebook ban hate speech, it is important to have many Facebooks, not a single one. The flip side of the expectation

---

108. See Erin Griffith, Will Facebook Kill All Future Facebooks?, Wired (Oct. 25, 2017), http://www.wired.com/story/facebooks-aggressive-moves-on-startups-threaten-innovation/ [http://perma.cc/W2L3-TP8E] (describing Facebook's strategy of preempting competition by purchasing startups and rival companies, potentially inhibiting innovation).

109. Even in the current landscape, it is easy to see that YouTube has different affordances and functions than Twitter, which has different affordances and functions than Snapchat, which has different affordances and functions than Facebook.

110. Of course, I myself have no objection to Balkinization!

111. See infra section IV.A.

112. See Gillespie, supra note 75, at 76–77, 107–08 (explaining the inherent difficulties of moderating content on a vast scale and the imperfections of algorithmic tools used to deal with the problem); Molly Roberts, Opinion, Alex Jones Does Not Compute, Wash. Post (Aug. 17, 2018), https://www.washingtonpost.com/opinions/twitter-infowars-and-techs-existential-crisis/2018/08/17/7c4c84bc-a232-11e8-8e87-c869fe70a721_story.html (on file with the *Columbia Law Review*) (observing the divergent, and widely criticized, responses by Facebook, Google, and Twitter to posts of inflammatory content by Alex Jones).

113. See Angwin & Grassegger, supra note 80.

that social media sites should enforce civility norms is the need for multiple social media sites serving different values and different publics.

One might also object that network effects will prevent broad diversity in social media because users will flock to the platforms with the largest user base.[114] Yet network effects will not necessarily prevent the growth of multiple social media sites.[115]

First, Facebook, like MySpace before it, will not be dominant forever. Often people—and especially generations—migrate from application to application without completely abandoning any of them.[116] How much time people spend on different sites may be fluid and may change over time as people age or have new experiences; moreover, the sites themselves may add new features as they compete for scarce attention.

Second, Professor Klonick has pointed out that people may see social media sites like Facebook, Twitter, and YouTube as complementary goods rather than rival goods.[117] People might use all three services for different purposes.

Third, social media sites are not like countries—one can both inhabit and be a "citizen" of many of them at the same time. The best model for the new digital public sphere is not the familiar model of competition between geographically distinct states.[118] Rather, it is one of diaspora, in which immigrants have connections both to their country of origin and to their current country and may also have relatives in many different countries. Digital diaspora may be a better model for thinking about the ecology of social media than the model of exclusionary network effects.

---

114. See, e.g., Michael A. Cusumano, Platform Wars Come to Social Media, Comms. ACM, Apr. 2011, at 31, 32–33 ("Because of the power of network effects and positive feedback, a relatively small number of sites will probably draw most of the user traffic and advertising dollars.").

115. See David S. Evans & Richard Schmalensee, Why Winner-Takes-All Thinking Doesn't Apply to the Platform Economy, Harv. Bus. Rev. (May 4, 2016), https://hbr.org/2016/05/why-winner-takes-all-thinking-doesnt-apply-to-silicon-valley [https://perma.cc/FF7Z-4FXQ] ("With low entry costs, trivial sunk capital, easy switching by consumers, and disruptive innovation showing no signs of tapering off, every internet-based business faces risk, even if it has temporarily achieved winner-takes-all status.").

116. See, e.g., Monica Anderson & Jingjing Jiang, Pew Research Ctr., Teens, Social Media & Technology 2018, at 2 (2018), http://assets.pewresearch.org/wp-content/uploads/sites/14/2018/05/31102617/PI_2018.05.31_TeensTech_FINAL.pdf [https://perma.cc/7S4N-BNM5] (describing the shift among younger Americans to use Facebook less and Instagram, YouTube, or Snapchat more).

117. See Klonick, New Governors, supra note 24, at 1630 (arguing that end users may employ multiple platforms because "[t]he commodity is not just the user, but rather it is the content created and engaged with by a user culture").

118. See Charles M. Tiebout, A Pure Theory of Local Expenditures, 64 J. Pol. Econ. 416, 423–24 (1956) (arguing that citizens would exit states in search of the most desirable combination of goods and services).

B. *The Responsibilities of Private Infrastructure*

Different parts of the internet infrastructure should have different responsibilities to protect freedom of speech online. For convenience, one might divide the digital infrastructure of free expression into three basic groups.[119]

---

119. The division of internet services offered in the text is related to two other sets of distinctions, although it is not, strictly speaking, identical with either.

The first approach analyzes internet traffic in terms of layers: for example, the hardware, protocol, applications, and content layers. See generally Lawrence B. Solum & Minn Chung, The Layers Principle: Internet Architecture and the Law, 79 Notre Dame L. Rev. 815 (2004) (arguing that legal regulation of the internet should recognize and respect the different layers of internet architecture).

The point of thinking in terms of layers is that different layers of internet traffic may require different regulatory norms. For example, one might contend that the hardware and protocol layers should remain neutral with respect to the carriage of content and applications, but actors in the applications and content layers should be free to curate, edit, and therefore discriminate on the basis of content. Governments should respect this basic division of labor between the various layers. They should not attempt to interfere with the efficiency of the hardware and protocol layers, for example, by requiring broadband companies or DNS servers to filter or block content.

The second approach argues that government regulation should respect the end-to-end principle of internet design. This principle distinguishes between decisions made at the edge of the internet (for example, by end users and applications companies) and decisions made in the core of the internet (for example, by internet service providers). See van Schewick, supra note 103, at 57–69 (explaining the different versions of the end-to-end principle).

Using this approach, one might argue that decisionmaking about content and applications should be located at the edge of the network and not in the middle. As a result, content regulation and discrimination should occur, if at all, at the edges of the network rather than in the center. It follows that broadband companies, which are located in the center of the network, should respect network neutrality—that is, they should not discriminate in content or applications. Likewise, governments should attempt to regulate content, if at all, by aiming only at the edges of the network rather than requiring players in the middle of the network to regulate or filter content. See Annemarie Bridy, Remediating Social Media: A Layer-Conscious Approach, 24 B.U. J. Sci. & Tech. L. (forthcoming 2018) (manuscript at 199–213), https://ssrn.com/abstract=3154117 (on file with the *Columbia Law Review*) (distinguishing between regulation of basic internet services and social media sites based on layers analysis and the end-to-end principle).

The approach in this Essay differs slightly from these two approaches for three reasons. First, payment systems are not, strictly speaking, layers of internet traffic; although they are edge services, I argue in this Essay that nondiscrimination norms should apply to them. See infra notes 127–128 and accompanying text.

Second, governments can reasonably require some services in the application layer—for example, email services—to be nondiscriminatory and open to all in much the same way as other basic internet services, while other services in the application layer—for example, social media services like Facebook and YouTube—should be treated as curators of content that are entitled to control access and make content-based decisions.

Third, the end-to-end design principle makes the most sense if we think of Facebook, Google, and other social media companies as located only at the "edge" of the internet—rather than squarely in the middle of it—because they provide, among other things, telecommunications, DNS, and hosting services. These companies' investments in telecommunications infrastructure have made them central governors and gatekeepers within the internet, straining the metaphor of "edge services." See supra note 40 and accompanying text.

The first group is *basic internet services*. It consists of four types of companies:

(a) Hosting services (such as Amazon Web Services or Gmail), which host websites, software applications, and platforms;
(b) Telecommunications services, which include internet backbone operators, ISPs, transit providers, and certificate authorities (which issue SSL certificates to websites and other applications);
(c) Domain name services, which include registrars that register domain names (such as GoDaddy), registries that run top-level domains (such as Verisign), and DNS providers that resolve domain names (such as Cloudflare and Google); and
(d) Caching and defense services (such as Akamai and Cloudflare), which smooth and speed up internet traffic and may also provide cybersecurity and defense against DNS attacks.[120]

The second group consists of *payment services* that allow people to conduct business and make payments online (such as Mastercard, Visa, and PayPal). Although payment services do not regulate traffic flows, many online enterprises would be effectively impossible without them.[121]

The third group consists of *content curators*. These companies include both platforms (such as Facebook, YouTube, Twitter, and Instagram) and search engines (such as Google). These companies make regular and pervasive content-based decisions as part of their business models using human bureaucracies, algorithms, or some combination of the two.[122]

Generally speaking, basic internet services should adopt policies of nondiscrimination with respect to the content and viewpoint of traffic that flows through them or that is stored on them. Network neutrality rules attempt to enforce this principle against broadband providers and ISPs, but analogous principles should apply to the rest of the delivery system.

For example, caching and defense services like Cloudflare should not, as a general rule, discriminate on the basis of content or viewpoint. There are three reasons for this. First, as a practical matter, these services

---

120. For a list of key players and functions in basic internet services, see Matthew Prince, Why We Terminated Daily Stormer, Cloudflare (Aug. 16, 2017), https://blog.cloudflare.com/why-we-terminated-daily-stormer [https://perma.cc/8QKY-GK8H]; see also James Grimmelmann, Internet Law: Cases & Problems 27–35 (8th ed. 2018) (describing elements of the internet "stack"); Balkin, Old-School/New-School, supra note 6, at 2303–04 (listing elements of the digital infrastructure of free expression); Free Speech: Only as Strong as the Weakest Link, Elec. Frontier Found., https://www.eff.org/free-speech-weak-link [https://perma.cc/X2KB-WXB5] (last visited Aug. 1, 2018) (describing elements of digital infrastructure).

121. For example, when the U.S. government sought to shut down WikiLeaks in 2011, it pressured payment services to stop doing business with the organization. See supra note 35 and accompanying text; see also infra note 127 and accompanying text.

122. See infra Part IV.

are increasingly important for unpopular speakers or for speakers likely to be targeted by others.[123] Second, withholding caching or defense services for particular disfavored sites will likely have significant collateral effects for other content on those sites that deserves protection. Third, withholding services will have many features of an administrative prior restraint: The decisions will be nontransparent and lack due process.

The analysis of DNS services is a little different. The initial grant of a domain name is usually content-based, if only because two applicants cannot have the same domain name; moreover, permissible top-level domain names are regulated by the Internet Corporation for Assigned Names and Numbers (ICANN).[124] The point, rather, is that once a domain name has been granted, the DNS system should not refuse to resolve a domain name because DNS service providers disapprove of the content appearing on a site that employs a given domain name. A fortiori, governments should not try to leverage the domain name system to block or censor content.[125] Suspension of domain names, refusal to resolve domain names, and blocking content by domain name offer extreme examples of collateral censorship.[126] As before, these decisions will also likely lack transparency, notice, and due process.

The second group, payment systems, presents still another set of problems. We might best analogize payment systems to public accommodations. They should be open to all people, groups, and businesses that do not use the service to engage in illegal activities. Public accommodations usually protect people against discrimination based on their identities—race, religion, sex, and so forth. In this context, however, the goal is to prevent discrimination based on the content of what people lawfully publish online. Even so, payment systems should be able to refuse to do business with those who seek to use their systems to facilitate illegal enterprises, which may include the sale of content whose distribution is illegal in a particular jurisdiction. But where the publication of content is not illegal, payment systems should not discriminate among their customers. In 2011, for example, the United States put pressure on payment systems to refuse to do business with WikiLeaks.[127]

---

123. See Prince, supra note 120 ("[I]f you don't have a network like Cloudflare in front of your content, and you upset anyone, you will be knocked offline.").

124. About ICANN, ICANN, https://www.icann.org/resources/pages/welcome-2012-02-25-en [https://perma.cc/55YL-4EDY] (last visited Aug. 1, 2018).

125. See, e.g., Mark Lemley, David S. Levine & David G. Post, Don't Break the Internet, 64 Stan. L. Rev. Online 34, 34–38 (2011), http://www.stanfordlawreview.org/wp-content/uploads/sites/3/2011/12/64-SLRO-34_0.pdf [https://perma.cc/6LV8-5KLV] (explaining that congressional attempts to protect intellectual property by commandeering the DNS system would have disastrous policy consequences).

126. See Balkin, Old-School/New-School, supra note 6, at 2318 (arguing that using the DNS system for content regulation is especially overbroad).

127. Id. at 2327–29; see also Yochai Benkler, WikiLeaks and the PROTECT-IP Act: A New Public-Private Threat to the Internet Commons, Daedalus, Fall 2011, at 154, 156–57

What made this episode especially worrisome was that, at that point, WikiLeaks was in much the same position as an American newspaper, which could not be prosecuted for publishing the very same information.[128]

Because basic internet services and payment systems should not engage in content regulation—with certain exceptions for the DNS system noted above—government regulation that enforces nondiscrimination obligations should ordinarily not be objectionable on free speech grounds. These companies should not exercise editorial control in the first place; hence government regulations that enforce obligations similar or analogous to common carriage, nondiscrimination, or public accommodation should normally be appropriate, both from a First Amendment and a more general free speech perspective.[129]

The problem, of course, is that nation-states may be tempted to do precisely the opposite—not to forestall content discrimination but to demand it through new-school speech regulation. Once the telecommunications system, the DNS system, and the system of electronic payments begin blocking, censoring, or discriminating against certain speakers, nation-states will attempt to piggyback on their technical capabilities. As we have seen, state pressure on infrastructure owners to surveil, block, and filter content creates predictable problems of collateral censorship and privatized prior restraint.

IV. THE OBLIGATIONS OF DIGITAL CURATORS—CURATIONAL DUE PROCESS AND INFORMATION FIDUCIARIES

Curators are in a different position than either payment systems or basic internet services because they curate and personalize information for end users. They also facilitate communication through curation. For example, an end user's Facebook feed does not offer every possible

---

(noting that "American political figures widely denounced the disclosures" and that a number of private parties severed ties to WikiLeaks).

128. Jack M. Balkin, The First Amendment Is an Information Policy, 41 Hofstra L. Rev. 1, 22 (2012) (arguing that WikiLeaks and the *New York Times* are essentially in the same position with respect to First Amendment doctrine); Yochai Benkler, A Free Irresponsible Press: Wikileaks and the Battle over the Soul of the Networked Fourth Estate, 46 Harv. C.R.-C.L. L. Rev. 311, 314 (2011) ("[I]t is not, as a matter of law, sustainable to treat Wikileaks or Assange any differently than the *New York Times* and its reporters . . . ."); Jack Goldsmith, Seven Thoughts on WikiLeaks, Lawfare (Dec. 10, 2010), https://www.lawfareblog.com/seven-thoughts-wikileaks [https://perma.cc/Y7HQ-X8LY] ("I do not understand why so much ire is directed at Assange and so little at the *New York Times*.").

129. See U.S. Telecom Ass'n v. FCC, 825 F.3d 674, 740–44 (D.C. Cir. 2016) (rejecting a First Amendment challenge to the FCC's network neutrality rules); Stuart Minor Benjamin, Transmitting, Editing, and Communicating: Determining What "the Freedom of Speech" Encompasses, 60 Duke L.J. 1673, 1696–712 (2011) (arguing that network neutrality rules and common-carriage obligations in telecommunications law do not violate the First Amendment); Susan Crawford, First Amendment Common Sense, 127 Harv. L. Rev. 2343, 2375–78 (2014) (same).

posting from the user's Facebook friends in the order they were posted; instead, Facebook decides which posts are most relevant and in what order to display them.[130] Google tries to provide the links that are most helpful to end users who make search requests,[131] and it tries to deter companies that seek to game the system of search engine results.[132]

Because companies like Facebook and Google act as curators and personalizers, they cannot really avoid making decisions about content. We should therefore think about their obligations differently than payment systems and basic internet services. Familiar concepts like content and viewpoint neutrality are simply unhelpful in describing their responsibilities in the emerging global system of free expression. Above all, these curators need to be *trustworthy* providers of search and communications services and *nonarbitrary* in their governance of communities.

A. *Digital Curators as Professionals and the Successors of Twentieth-Century Mass Media*

The obligation of trustworthiness makes digital curators both similar to and different from twentieth-century (or "legacy") mass media such as newspapers, broadcasters, and cable channels. Like twentieth-century mass media, digital curators have become important custodians of the public sphere and of democratic self-government. Hence, whether they like it or not, digital curators have social and moral obligations to the public—as opposed to legal obligations. With those obligations comes a corresponding duty to develop and abide by professional norms of curation and governance.

Social media companies and search engines have social and moral obligations to the public because they perform three connected public services. First, they *facilitate public participation* in art, politics, and culture. Second, they *organize public conversation* so that people can easily find and communicate with each other. Third, they *curate public opinion* by providing individualized feeds and search results, and by enforcing civility norms through their terms-of-service obligations and community guidelines.

---

130. See, e.g., Victor Luckerson, Here's How Facebook's News Feed Actually Works, Time (July 9, 2015), http://time.com/collection-post/3950525/facebook-news-feed-algorithm/ [https://perma.cc/LLT8-U2DQ] ("[M]ost users see only a sliver of the potential posts in their network each day.").

131. See How Search Algorithms Work, Google, https://www.google.com/search/howsearchworks/algorithms/ [https://perma.cc/MEP7-GMJ9] (last visited Aug. 1, 2018) ("Google ranking systems . . . are made up of a series of algorithms that analyze what it is you are looking for and what information to return to you.").

132. See Kaspar Szymanski, Google Penalties and Messages Explained—Search Engine Land's Ultimate Guide, Search Engine Land, https://searchengineland.com/guide/google-penalties [https://perma.cc/TL6C-NGVG] (last visited Aug. 1, 2018) (explaining how and why Google demotes links to penalize firms that attempt to manipulate its search rankings system).

Social media companies and search engines present themselves as more than ordinary profit-making enterprises.[133] They explain that they use their special technological expertise to promote public-spirited goals like access to knowledge, freedom of expression, and community building.[134] In this way, they encourage the idea that they do act and should act according to public-regarding, professional norms. Moreover, social media companies and search engines invoke these professional and public-regarding norms to justify their decisions to organize search-engine results, to curate public discourse, and to enforce (or sometimes refrain from enforcing) civility norms.[135] The public, politicians, and civil-society organizations repeatedly push back, claiming that digital

---

133. For example, as Google's founders explain: "Google is not a conventional company. We do not intend to become one." 2004 Founders' IPO Letter, Alphabet, http://abc.xyz/investor/founders-letters/2004/ipo-letter.html [https://perma.cc/NW3Q-BN82] (last visited Aug. 30, 2018). Google's stated purpose is to "organize the world's information and make it universally accessible and usable." Our Company, Google, https://www.google.com/about/our-company/ [https://perma.cc/UFS8-PSJN] (last visited Aug. 21, 2018).

134. In addition to its goal of "organiz[ing] the world's information and mak[ing] it universally accessible and usable," Google also aims "to develop services that significantly improve the lives of as many people as possible." 2004 Founders' IPO Letter, supra note 133.

Twitter explains that it "offer[s] Twitter and other services in order to give everyone the power to create and share ideas and information instantly, without barriers." Our Services and Corporate Affiliates, Twitter, https://help.twitter.com/en/rules-and-policies/twitter-services-and-corporate-affiliates [https://perma.cc/AE8Z-TJFR] (last visited Aug. 21, 2018).

In 2017, Facebook announced a new mission statement: "To give people the power to build community and bring the world closer together." Heather Kelly, Mark Zuckerberg Explains Why He Just Changed Facebook's Mission, CNN (June 22, 2017), https://money.cnn.com/2017/06/22/technology/facebook-zuckerberg-interview [https://perma.cc/CNX2-VYC5] ("It's important to give people a voice, to get a diversity of opinions out there, but on top of that, you also need to do this work of building common ground so that way we can all move forward together." (internal quotation marks omitted) (quoting Mark Zuckerberg)).

135. See Community Standards, Facebook, https://www.facebook.com/communitystandards [https://perma.cc/L4US-ZB7Z] (last visited Aug. 21, 2018) ("The goal of our Community Standards is to encourage expression and create a safe environment. We base our policies on input from our community and from experts in fields such as technology and public safety."); How Search Works: Our Mission, Google, https://www.google.com/search/howsearchworks/mission/ [https://perma.cc/SFP4-SL3P] (last visited Aug. 21, 2018) ("From innovations like the Knowledge Graph to updates to our ranking algorithms that ensure we're continuing to highlight relevant and authoritative content, our goal is always to improve the usefulness of your results."); The Twitter Rules, Twitter, https://help.twitter.com/en/rules-and-policies/twitter-rules [https://perma.cc/DZ2P-E62A] (last visited Aug. 21, 2018) ("We believe that everyone should have the power to create and share ideas and information instantly, without barriers. In order to protect the experience and safety of people who use Twitter, there are some limitations on the type of content and behavior that we allow.").

companies have fallen short of these professions of public-spiritedness and demanding that companies act according to the public interest.[136]

A similar process happened to journalism in the early twentieth century. Newspapers were confronted with the rise of propaganda, advertising, and public relations. Seeking to differentiate themselves from these practices, newspapers gradually accepted that they had distinctive professional obligations to the public in how they covered and reported the news.[137] This growing sense of responsibility to the public developed into what we now know as the professional norms of modern journalism.[138] The twentieth-century vision of objective journalism in the public interest did not arise overnight—it was shaped by economic, social, and technological developments.

Just as in the case of twentieth-century mass media, however, the state constitutionally cannot force these professional norms—or their twenty-first-century equivalents—on digital curators. But this does not mean that the public cannot or should not demand these norms. We are beginning to see a slow and halting evolution of platforms' self-understanding precisely along these lines. This learning process is the result of wave after wave of public pressure on companies like Google, Facebook, and Twitter, often facilitated by journalists who themselves apply professional norms of news reporting developed in the previous century.[139] Companies that once viewed themselves purely as technology

---

136. See, e.g., Carole Cadwalladr, Google, Democracy and the Truth About Internet Search, Guardian (Dec. 4, 2016), https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook [https://perma.cc/WS2B-MFEF] (criticizing Google's search results for promoting anti-Semitism); Cecilia Kang & Kate Conger, Inside Twitter's Struggle over What Gets Banned, N.Y. Times (Aug. 10, 2018), https://www.nytimes.com/2018/08/10/technology/twitter-free-speech-infowars.html (on file with the *Columbia Law Review*) (describing internal debates in response to mounting public criticism of Twitter for its failure to discipline Alex Jones and Infowars); Alyssa Newcomb, A Timeline of Facebook's Privacy Issues—And Its Responses, NBC News (Mar. 24, 2018), https://www.nbcnews.com/tech/social-media/timeline-facebook-s-privacy-issues-its-responses-n859651 [https://perma.cc/H9A2-9UHQ] (describing successive episodes of public criticism of Facebook for its privacy policies).

137. See Michael Schudson, The Objectivity Norm in American Journalism, 2 Journalism 149, 159–63 (2001).

138. Id.

139. See, e.g., Angwin & Grassegger, supra note 80 (describing public objections to Facebook's policies for removing content and sanctioning end users); Vindu Goel, Some Privacy, Please? Facebook, Under Pressure, Gets the Message, N.Y. Times (May 22, 2014), https://www.nytimes.com/2014/05/23/technology/facebook-offers-privacy-checkup-to-all-1-28-billion-users.html (on file with the *Columbia Law Review*); Dave Lee, Facebook Amends 'Real Name' Policy After Protests, BBC (Dec. 15, 2015), http://www.bbc.com/news/technology-35109045 [https://perma.cc/9WF3-MJFS]; Joel Schectman, Facebook Releases New Privacy Safeguards After Ceding to Pressure from Advertisers, Reuters (June 13, 2018), https://www.reuters.com/article/us-facebook-privacy-broker/facebook-releases-new-privacy-safeguards-after-ceding-to-pressure-from-advertisers-idUSKBN1J924P [https://perma.cc/6UV4-GUWN]. Investigative journalists like Carole Cadwalladr and Emma Graham-Harrison of the *Guardian* uncovered the Cambridge Analytica scandal, discussed infra at

companies are beginning to understand their public responsibilities as twenty-first-century media companies.

In saying this, it is important not to wax nostalgic over twentieth-century mass media or to assume that the twentieth century represents a lost, golden age of media responsibility.[140] Nor should we assume that twenty-first-century media will adopt professional norms identical to those of twentieth-century journalism. The point is rather that twentieth-century media, with all of its faults, served as a countervailing force to government power in a democracy. In the same way, twenty-first-century media companies, at best, may provide platforms for democratic organization and protest and act as checks on the power of territorial governments, even as these governments are necessary checks on technology companies' burgeoning economic and political power.

B. *Legal Obligations—Curational Due Process*

If digital curators were just like twentieth-century mass media, that would be the end of the story. These companies would have public obligations—that is, moral duties—to develop professional norms in the public interest, and the public (and legacy media) might pressure them into adopting and adhering to those norms. The state, however, would be forbidden from enforcing these norms by law. But that is not the end of the story. The new digital curators differ from twentieth-century mass media in two important respects.

The first difference is that digital media companies have curatorial obligations of due process. These obligations made little sense in a world in which very few people had access to mass media but are central in a world in which everyone is a broadcaster. For this reason, to the extent that digital curators block, censor, or take down content from their end users, they have obligations of due process toward their end users.[141] That is especially so if their content regulation is at the behest of nation-states employing new-school speech regulation.

To see how these due process obligations might operate in practice, a good place to start is the Manila Principles on Intermediary Liability, a

---

notes 168–177 and accompanying text, which led to increased public pressure for the reform of social media.

140. See Morgan N. Weiland, The Paradox of Platforms-as-Press: Unwinding This Analogy to Solve the Platform Accountability Problem 2–3 (Apr. 10, 2018) (unpublished manuscript) (on file with the *Columbia Law Review*). The famous Hutchins Commission Report of 1947 argued that the press had a social responsibility to the public in a democracy and warned against commercialization, tendencies toward monopoly, ownership conflicts of interest, and sensationalism. See A Free and Responsible Press: A General Report on Mass Communication (Robert D. Leigh ed., 1947) (Hutchins Commission Report). Many of these concerns are still with us today. Although highly influential, the report hardly quelled concerns about whether the press was living up to its social responsibilities in a democracy.

141. Balkin, Algorithmic Society, supra note 6, at 1197.

series of reform proposals developed by civil society organizations in 2015.[142] The Manila Principles require, among other things: (1) clear and public notice of the content-regulation policies companies actually employ; (2) an explanation and an effective right to be heard before content is removed; and (3) when this is impractical, an obligation to provide a post facto explanation and review of a decision to remove content as soon as practically possible.[143] One might call these and similar norms the obligations of *curational due process*.

The Manila Principles focus on removing content that is illegal in a given country,[144] but the same principles could also apply to content that violates the company's terms of service or end-user license agreement. In fact, instead of passing new speech regulations, nation-states may find it more convenient to press curators to enforce their existing terms of service. This has the additional advantage of leveraging private speech codes that operate globally to serve the nation-state's parochial ends.[145]

Curational due process made little sense for twentieth-century mass media because twentieth-century mass media largely published their own content or the content of a relatively small number of people and businesses. Most people had no access to mass-media publication, and few people expected an explanation or a right to be heard either before or after the *New York Times* rejected their proposed letters to the editor. Twenty-first-century media companies, by contrast, primarily publish content by the general public. Digital curators exist to facilitate mass cultural participation, and their end users expect and depend on the fact that curators will help them in this process. Therefore, curators need to provide assurances that when they block or limit participation, they are not being overbroad or arbitrary.

Legislation that requires digital curators to provide due process would not necessarily violate the First Amendment. Curators would probably argue that such rules would interfere with their editorial functions.[146] But one can avoid constitutional problems by making due process obligations part of a safe harbor from intermediary liability.

---

142. See Manila Principles on Intermediary Liability, https://www.manilaprinciples.org/ [https://perma.cc/U7SD-VCUW] (last visited Aug. 2, 2018).

143. See Manila Principles on Intermediary Liability 2– 5 (2015), https://www.eff.org/files/2015/10/31/manila_principles_1.0.pdf [https://perma.cc/W693-TSUW].

144. See id.

145. See supra text accompanying notes 94–95. Thus, when the European Union pressed four major curators to help it combat hate speech, it essentially asked for more prompt and efficient enforcement of the curators' own hate speech policies. See, e.g., Code of Conduct on Countering Illegal Hate Speech Online, supra note 89, at 1.

146. See Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 636 (1994) (holding that cable broadcasters exercise editorial functions protected by the First Amendment "[t]hrough 'original programming or by exercising editorial discretion over which stations or programs to include in [their] repertoire'" (quoting Los Angeles v. Preferred Commc'ns, Inc., 476 U.S. 488, 494 (1986))).

Section 230 of the 1996 Telecommunications Act protects digital curators from liability for content appearing on their sites.[147] Some aspects of intermediary immunity are probably required by the Constitution, so that if Congress repealed § 230, certain constitutional protections would still be in force.[148] For example, it might be unconstitutional to hold digital curators strictly liable for any defamatory or obscene content that appears on their sites.[149] But the boundaries of constitutional protection are uncertain. Would a negligence standard be sufficient? What about other kinds of unlawful content?[150] What if digital curators are notified that the material is defamatory, tortious, or otherwise unlawful?[151] These and similar questions remain unsettled. Moreover, § 230(c)(2)—which holds digital curators harmless for editing or deleting content of end users—is probably not required by the First Amendment.[152]

The best way to guarantee curatorial due process, therefore, is to resolve these uncertainties by creating a safe harbor provision that would

---

147. See 47 U.S.C. § 230 (2012); see also Zeran v. Am. Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997) (holding that § 230 immunized an online service provider from liability for content appearing on its site created by another party).

148. See Note, Section 230 as First Amendment Rule, 131 Harv. L. Rev. 2027, 2028, 2030 (2018) (arguing that certain aspects of § 230 immunity for defamation are required by the Constitution despite the fact that "[j]udges and academics are nearly in consensus in assuming that the First Amendment does not require § 230").

149. See Smith v. California, 361 U.S. 147, 152–55 (1959) (striking down a California law that held booksellers strictly liable for possession of obscene books with no requirement of knowledge of the contents of the books).

150. Section 230 excludes content that violates intellectual property law, federal privacy law, and federal criminal law from its immunity. See 47 U.S.C. § 230(e) (listing exemptions). In 2018, Congress passed the Allow States and Victims to Fight Online Sex Trafficking Act of 2017, Pub. L. No. 115-164, 132 Stat. 1253 (2018) (to be codified at 18 U.S.C. § 2421A; 47 U.S.C. § 230(e)(5)), which removes § 230 immunities for sex-trafficking and prostitution-related offenses.

151. Section 512(g) of the Digital Millennium Copyright Act (DMCA), 17 U.S.C. § 512(g) (2012), provides less protection for hosting materials that infringe copyright than § 230 does for hosting defamatory materials. Section 512(g) creates a safe harbor from copyright liability if an online service provider removes content upon notice. The notice-and-takedown rules create incentives for collateral censorship. See Balkin, Old-School/New-School, supra note 6, at 2314; Mulligan, supra note 16, at 181–84; Rebecca Tushnet, Power Without Responsibility: Intermediaries and the First Amendment, 76 Geo. Wash. L. Rev. 986, 1003 (2008) ("Because DMCA notice requirements are minimal and ISPs have no incentive to investigate, the notice-and-takedown process can be used to suppress critical speech as well as copyright infringement.").

152. See Balkin, Old-School/New-School, supra note 6, at 2318 & n.82 (arguing that if online service providers are not state actors, then their filtering of internet content, which is protected from liability under § 230(c)(2), does not violate the First Amendment). Although it is certainly possible to imagine scenarios under which a grant of legal immunity to one private party for censoring or editing the expression of another private party might violate the second party's First Amendment rights, § 230(c)(2) is probably facially constitutional.

amend § 230.[153] If digital curators agree to adopt something similar to the Manila Principles, they will retain their intermediary immunity. If not, they will have to be content with constitutional limitations on intermediary liability, which are both uncertain and likely to be far less than the statutory guarantee.

Note that this solution would employ a new-school technique because it would reduce the intermediary immunity of digital curators with respect to the status quo. But it would use that technique for the opposite goal of most new-school speech regulation: It would attempt to protect the free speech interests of end users.

C. *Legal Obligations—Information Fiduciaries*

The second difference between twentieth-century mass-media companies and twenty-first-century digital curators is that twenty-first-century companies have developed elaborate technologies and techniques for individualized surveillance, manipulation, and control that were not really possible for twentieth-century mass media.[154] To be sure, twentieth-century mass media also hoped to appeal to certain demographics in order to attract advertisers. Yet their abilities to surveil, target, manipulate, and even addict their audiences could not be so effective or so precise as those of twenty-first-century companies.

Indeed, the characteristic feature of twenty-first-century digital media companies is not merely that they enable mass participation. It is that in doing so they also engage in mass data collection and surveillance, and that they develop ever more effective means for influencing (and thus potentially manipulating) their audiences in order to gain their scarce attention.[155] The flip side of mass cultural participation is mass personal surveillance, and the danger of widespread digital participation is widespread digital manipulation.

Digital curation is not simply the selection of content for end users; it also involves using knowledge about end users to control, shape, and govern their behavior.[156] Digital curators are private governors not only

---

153. Such a safe harbor might be modeled along the lines of the DMCA's § 512(g). See supra note 151.

154. See Wu, Attention Merchants, supra note 49, at 323–25; Louise Matsakis, Facebook's Targeted Ads Are More Complex than It Lets On, Wired (Apr. 25, 2018), https://www.wired.com/story/facebooks-targeted-ads-are-more-complex-than-it-lets-on/ [https://perma.cc/6BWV-JG7H] (explaining that in comparison with twentieth-century mass-media companies, advertisers "who use Facebook have a near-endless number of data points with which to target their ads, and can show them to much narrower slices of the population").

155. See Wu, Attention Merchants, supra note 49, at 323–25.

156. See, e.g., Zuboff, supra note 53, at 85 ("[S]urveillance capitalism . . . produces the possibility of modifying the behaviors of persons and things for profit and control."); Samuel Gibbs, Facebook Apologises for Psychological Experiments on Users, Guardian (July 2, 2014), https://www.theguardian.com/technology/2014/jul/02/facebook-apologises-psychological-experiments-on-users [https://perma.cc/56UY-S372] ("[Facebook's] researchers decided after

in establishing and enforcing community norms but also in their attempts to govern and direct end users through surveillance. The problem of digital curators, which makes them different in kind from twentieth-century mass-media companies, is the far greater danger that they will engage in acts of manipulation and breach of trust through the use of personal data.

In the algorithmic age, many digital companies—and not merely digital curators—take on new kinds of obligations. These new obligations arise from people's increasing dependence on and vulnerability to digital services that collect data about them but whose operations are not transparent to them. Companies that create and maintain these relations of digital dependence and vulnerability should be considered *information fiduciaries* toward their end users.[157]

We rely on digital companies to perform many different tasks for us. In the process, these companies learn a great deal about us, but we do not know very much about their operations.[158] As a result, we are especially vulnerable to them, and we have to trust that they will not betray us or manipulate us for their own ends.

The law has long recognized that clients or patients of professionals like doctors and lawyers are in a similar situation: We need to trust these professionals with sensitive personal informational about ourselves, but they could injure us as a result. Therefore the law treats them as fiduciaries.[159] Fiduciary relationships are relationships of good faith and loyalty toward people who are in special positions of vulnerability. Fiduciaries have special duties of care, confidentiality, and loyalty toward their clients and beneficiaries.[160]

---

tweaking the content of people's 'news feeds' that there was 'emotional contagion' across the social network."); Luckerson, supra note 130 (describing how Facebook attempts to manage end users' behavior through personalized news feeds); Tufekci, Facebook's Surveillance Machine, supra note 53 (describing the use of Facebook data by Cambridge Analytica as "an all-too-natural consequence of Facebook's business model, which involves having people go to the site for social interaction, only to be quietly subjected to an enormous level of surveillance").

157. See Jack M. Balkin, Information Fiduciaries and the First Amendment, 49 U.C. Davis L. Rev. 1183, 1209 (2016) [hereinafter Balkin, Information Fiduciaries] ("An information fiduciary is a person or business who, because of their relationship with another, has taken on special duties with respect to the information they obtain in the course of the relationship."); Jack M. Balkin, The Three Laws of Robotics in the Age of Big Data, 78 Ohio St. L.J. 1217, 1228 (2018) [hereinafter Balkin, Three Laws of Robotics] ("When fiduciaries collect and process information about their clients, . . . [t]hey are information fiduciaries.").

158. See Frank Pasquale, The Black Box Society 3–4 (2015) (emphasizing the knowledge asymmetries between digital companies and end users).

159. Tamar Frankel, Fiduciary Law 42–45 (2011) [hereinafter Frankel, Fiduciary Law] (listing traditional fiduciaries, including professionals like doctors and lawyers).

160. See Balkin, Information Fiduciaries, supra note 157, at 1206–08 (describing duties of care, loyalty, and confidentiality); Deborah A. DeMott, Beyond Metaphor: An Analysis of Fiduciary Obligation, 1988 Duke L.J. 879, 882 (explaining that fiduciaries

Because digital companies collect enormous amounts of data about their end users, and use this data to predict and control what end users will do—including, among other things, matching them with advertisers—digital curators are perhaps the most important example of the new information fiduciaries of the digital age.[161] Even so, we should treat the analogy to doctors and lawyers with some care. The kinds of fiduciary duties that a company has depend on the nature of its social role and its business.[162] Digital companies are not trained professionals like doctors and lawyers. They offer a different set of services, consumers expect different things from them, and therefore we should expect that they will not have all of the same obligations as doctors and lawyers.[163]

For example, unlike doctors and lawyers, social media companies and search engines offer their services for free in return for the right to serve targeted ads to their end users. The practice of offering free or heavily subsidized services in return for surveillance and collection of data creates a potential conflict of interest between end users and digital companies. Companies will always be tempted to use the data in ways that sacrifice the interests of their end users to the company's economic or political interests. Nevertheless, unless governments outlaw the practice of financing free (or subsidized) digital products altogether, one must start with the assumption that the law can cure potential conflicts of interest through appropriate regulation; if so, this means that social media companies will be able to monetize personal data in some ways but not in others. Their fiduciary duties will constrain the ways they are allowed to collect, monetize, and employ end-user data. What constitutes a breach of trust depends on the nature of their business, and this, in turn, depends on what consumers would reasonably consider unexpected or abusive for digital companies to do.[164]

A good example of how information fiduciaries should *not* act—and how fiduciary duties would constrain their behavior—is the story of how

---

"must be loyal to the interests of the other person" and that "[t]he fiduciary's duties go beyond mere fairness and honesty; they oblige him to act to further the beneficiary's best interests").

161. See Balkin, Algorithmic Society, supra note 6, at 1162; Balkin, Information Fiduciaries, supra note 157, at 1221.

162. See Frankel, Fiduciary Law, supra note 159, at 53 (noting that "[t]he process of recognizing new fiduciary relationships is ongoing," depending on the nature of their services, the power relations and temptations they create, and the ability of institutions and markets to control them); Balkin, Information Fiduciaries, supra note 157, at 1225 ("The duties that we impose on traditional fiduciaries can be fairly extensive; but the duties we might justifiably impose on online service providers may be different and sometimes considerably narrower, especially if we want these duties to be consistent with the First Amendment.").

163. See Balkin, Information Fiduciaries, supra note 157, at 1227–29 (describing three differences between digital-information fiduciaries and traditional fiduciaries).

164. Id. at 1229; see also Tamar Frankel, Fiduciary Law, 71 Calif. L. Rev. 795, 810 (1983) ("Fiduciary relations vary by the extent to which each type of fiduciary can abuse his power to the detriment of the entrustor.").

Facebook allowed third parties to exploit its end users' data. This practice came to light in the Cambridge Analytica scandal in the spring of 2018.[165]

Until the middle of 2014, Facebook had a policy of sharing access to end-user data with third parties, including for-profit companies and academic researchers.[166] This practice offered Facebook additional ways to monetize consumer data.[167] In 2014, Facebook allowed a data scientist, Aleksandr Kogan, to conduct social-science experiments using end-user data.[168] Kogan used Amazon's Mechanical Turk and similar platforms to find people who were willing to take a personality quiz for a few dollars; the participants signed onto the test using their Facebook accounts.[169] This gave Kogan access to the data that Facebook associated with their personal accounts as well as the data of all of their Facebook friends.[170] In this way, Kogan was able to leverage the approximately 300,000 users who took the quiz to obtain access to some 87 million end users' data profiles.[171]

What Kogan did not tell Facebook, however, was that he was secretly working with Cambridge Analytica, a for-profit consulting company that

---

165. See Alvin Chang, The Facebook and Cambridge Analytica Scandal, Explained with a Simple Diagram, Vox (May 2, 2018), https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram [https://perma.cc/7QCM-5QPZ].

166. Paul Lewis, 'Utterly Horrifying': Ex-Facebook Insider Says Covert Data Harvesting Was Routine, Guardian (Mar. 20, 2018), https://www.theguardian.com/news/2018/mar/20/facebook-data-cambridge-analytica-sandy-parakilas [https://perma.cc/4GPN-NZNJ] [hereinafter Lewis, Covert Data Harvesting] (explaining that under the policy, "'a majority of Facebook users' could have had their data harvested by app developers without their knowledge" (quoting Sandy Parakilas, former platform operations manager at Facebook)).

167. See id. ("Facebook took a 30% cut of payments made through apps, but in return enabled their creators to have access to Facebook user data.").

168. See Carole Cadwalladr & Emma Graham-Harrison, How Cambridge Analytica Turned Facebook 'Likes' into a Lucrative Political Tool, Guardian (Mar. 17, 2018), https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm [https://perma.cc/VJR6-KPCK] [hereinafter Cadwalladr & Graham-Harrison, How Cambridge Analytica Turned Facebook 'Likes']

169. Id.

170. Id.

171. See Michael Riley et al., Understanding the Facebook-Cambridge Analytica Story: QuickTake, Wash. Post (Apr. 9, 2018), https://www.washingtonpost.com/business/understanding-the-facebook-cambridge-analytica-story-quicktake/2018/04/09/0f18d91c-3c1c-11e8-955b-7d2e19b79966_story.html [https://perma.cc/PKV5-9SGX] (estimating that 300,000 people participated and that 87 million users had their data harvested); Matthew Rosenberg et al., How Trump Consultants Exploited the Facebook Data of Millions, N.Y. Times (Mar. 17, 2018), https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html (on file with the *Columbia Law Review*) ("Only about 270,000 users—those who participated in the survey—had consented to having their data harvested."); Mike Schroepfer, An Update on Our Plans to Restrict Data Access on Facebook, Facebook Newsroom (Apr. 4, 2018), https://newsroom.fb.com/news/2018/04/restricting-data-access/ (on file with the *Columbia Law Review*) (offering an updated estimate of 87 million persons, including some 70 million in the United States, whose data was harvested by Cambridge Analytica).

uses personal data to serve targeted political ads based on psychological profiles constructed from the data.[172] Kogan violated Facebook's platform policy for researchers and scientists by turning over the data to a for-profit company.[173] Facebook learned about the arrangement in 2015 but did not reveal it to the public.[174] It asked Kogan and Cambridge Analytica to delete the data they had harvested but did not ensure that the data was actually erased, and Cambridge Analytica kept the data.[175] When the news of the arrangement leaked out in the spring of 2018, it caused a scandal, and Facebook's founder, Mark Zuckerberg, was asked to testify before Congress.[176] Following the disclosures, Zuckerberg admitted that the company had "made mistakes" and described the scandal as a "breach of trust" toward his end users.[177] That well describes the central issue.

As an information fiduciary, Facebook has three different kinds of duties toward its end users: a duty of confidentiality, a duty of care, and a duty of loyalty.[178] The duties of confidentiality and care mean that Facebook must keep its customers' data confidential and secure. It must make sure that fiduciary duties "run with the data": In other words, Facebook must ensure that anyone who shares or uses the data is equally trustworthy and is legally bound by the same legal requirements of confidentiality, care,

---

172. Riley et al., supra note 171 ("Facebook says Kogan 'lied to us' by saying he was gathering the data for research purposes and violated the company's policies by passing the data to Cambridge Analytica[,] . . . a company that 'uses data to change audience behavior,' both commercially and politically, according to its website.").

173. Cadwalladr & Graham-Harrison, How Cambridge Analytica Turned Facebook 'Likes,' supra note 168.

174. Carole Cadwalladr & Emma Graham-Harrison, Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach, Guardian (Mar. 17, 2018), https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election [https://perma.cc/J2BS-QCDW].

175. Id.; see also Paul Grewal, Suspending Cambridge Analytica and SCL Group from Facebook, Facebook Newsroom (Mar. 16, 2018), https://newsroom.fb.com/news/2018/03/suspending-cambridge-analytica/ (on file with the *Columbia Law Review*).

176. Mark Zuckerberg Testimony: Senators Question Facebook's Commitment to Privacy, N.Y. Times (Apr. 10, 2018), https://www.nytimes.com/2018/04/10/us/politics/mark-zuckerberg-testimony.html (on file with the *Columbia Law Review*).

177. Mark Zuckerberg, Facebook (Mar. 21, 2018), https://www.facebook.com/zuck/posts/10104712037900071 (on file with the *Columbia Law Review*) ("This was a breach of trust between Kogan, Cambridge Analytica and Facebook. But it was also a breach of trust between Facebook and the people who share their data with us and expect us to protect it. We need to fix that.").

178. See Frankel, Fiduciary Law, supra note 159, at 106 (describing the duties of care and loyalty); see also Restatement (Third) of Agency § 8.05 (Am. Law Inst. 2006) ("An agent has a duty . . . (2) not to use or communicate confidential information of the principal for the agent's own purposes or those of a third party."); Restatement (Third) of the Law Governing Lawyers §§ 16, 49, 60 (Am. Law Inst. 2000) (stating lawyers' fiduciary duties to respect client confidences); Restatement (Second) of Torts § 874 reporter's note (Am. Law Inst. 1979) ("One breach of fiduciary duty that is more commonly regarded as giving rise to an action in tort is the disclosure of confidential information."); Mark A. Hall, Mary Anne Bobinski & David Orentlicher, Medical Liability and Treatment Relationships 171 (3d ed. 2013) (stating physicians' fiduciary duties of confidentiality).

and loyalty as Facebook is. The company must vet its potential partners to make sure that they are ethical and reliable, subject them to regular audits, and, if they violate the terms of their agreements, it must take steps to get back all the data that the company shared with them.

Facebook failed these duties in several ways. It did not properly limit who could use its data and for what purposes, it did not vet or audit its partners properly, and it did not claw back the data obtained in violation of its policies. In short, it breached its duties of confidentiality and care because it did not keep its data confidential and secure.

Next consider the duty of loyalty. The previous discussion assumed that Facebook would not breach its duty of loyalty simply by serving targeted ads for consumer products in return for free services. In part, that is because people more or less expect that Facebook will serve them ads based on the data it collects. But when Facebook departs from these consumer expectations to benefit itself to the disadvantage of its end users, it may breach its duty of loyalty. The problem arises when Facebook uses the data in unexpected ways that people would find offensive and a breach of trust. The problem is exacerbated when Facebook shares data with third parties without adequate controls over use and disclosure, or when Facebook allows third parties access to its end users by having end users sign in to third-party applications through their Facebook accounts. It is one thing for Facebook to serve you ads for shampoo; it is quite another for Facebook to hand your data off to third parties who have no qualms about manipulating you.

In fact, the Cambridge Analytica scandal appears to have been only the tip of the iceberg. In the hopes of increasing profit margins, Facebook granted access to end users' information to a host of for-profit companies without adequate safeguards as to whether companies were manipulating its end users.[179] This created a conflict of interest because it gave Facebook an incentive to look the other way, which it apparently did.[180] Facebook may have also breached its duty of loyalty by allowing third parties to perform social-science experiments on its end users without the equivalent of a human-subjects review board to minimize harm and to prevent overreaching and manipulation.[181] Finally, if, as critics

179. See Lewis, Covert Data Harvesting, supra note 166 (reporting an account of a former platform-operations manager at Facebook that Facebook deliberately avoided finding out whether and how data was being abused by third parties); Asher Schechter, Roger McNamee: "I Think You Can Make a Legitimate Case that Facebook Has Become Parasitic," ProMarket (Mar. 23, 2018), https://promarket.org/roger-mcnamee-think-can-make-legitimate-case-facebook-become-parasitic/ [https://perma.cc/WNZ3-AC2H] (describing an interview with Roger McNamee, an early Facebook investor, who argued that "Facebook's algorithms and business model essentially enable bad actors to harm innocent people" (internal quotation marks omitted)).

180. See Lewis, Covert Data Harvesting, supra note 166.

181. See Rey Junco, Why Facebook's User Manipulation Research Study Is Ethically Troubling, Venture Beat (July 6, 2014), https://venturebeat.com/2014/07/06/why-facebooks-user-manipulation-research-study-is-ethically-troubling/ [https://perma.cc/8BV8-LFPB] (arguing

charge, Facebook designed its applications and employed its end users' data to psychologically manipulate and addict them to the site,[182] it would also have breached its duty of loyalty, creating a conflict of interest between the company and its end users.

The duties of information fiduciaries depend in part on what is reasonable to expect from them given their business models. But the most general obligation of digital-information fiduciaries is that they may not act like con artists.[183] They may not induce trust in their end users to obtain personal information from them and then turn around and betray that trust by harming and manipulating them for the company's own benefit. Digital businesses may not hold themselves out as providing safe and welcoming digital communities that respect privacy and then manipulate their end users; nor should they be permitted to give access to end-user data to third parties who will not accept similar duties of care, confidentiality, and good faith.[184]

Although companies can violate these duties when they violate their privacy policies, fiduciary duties extend beyond the precise terms of those privacy policies to duties of good faith, respect, and nonmanipulation.[185] Social media companies engage in manipulation when—under conditions of extreme information asymmetry and vulnerability—end users must provide information about themselves in order to use the service, and companies use this information in ways that both benefit the

---

that some social media experiments should require an institutional review board and that "[r]esearchers are obliged not only to ensure they do no harm, but also to maximize the potential benefits[,] . . . minimize the potential harms of a study," and put checks and balances in place). But see Timothy J. Ryan, On the Ethics of Facebook Experiments, Wash. Post (July 3, 2014), https://www.washingtonpost.com/news/monkey-cage/wp/2014/07/03/on-the-ethics-of-facebook-experiments/ [https://perma.cc/9TFN-8DLF] (arguing that fears of manipulation are overblown and that many experiments do not require informed consent from human subjects).

182. See Mike Allen, Sean Parker Unloads on Facebook: "God Only Knows What It's Doing to Our Children's Brains," Axios (Nov. 9, 2017), https://www.axios.com/sean-parker-unloads-on-facebook-god-only-knows-what-its-doing-to-our-childrens-brains-1513306792-f855e7b4-4e99-4d60-8d51-2775559c2671.html [https://perma.cc/B8JV-FASH] (quoting a statement by the former president of Facebook that social media applications are designed to "exploit[] a vulnerability in human psychology" using psychological methods to "consume as much of your time and conscious attention as possible" and keep users locked into the site (internal quotation marks omitted)); Paul Lewis, 'Our Minds Can Be Hijacked': The Tech Insiders Who Fear a Smartphone Dystopia, Guardian (Oct. 6, 2017), https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia [https://perma.cc/P9AA-CU2X] (interviewing former employees at Google and Facebook who report that technologies are designed to addict users and monopolize their attention).

183. See Balkin, Algorithmic Society, supra note 6, at 1163; Balkin, Information Fiduciaries, supra note 157, at 1224.

184. See Balkin, Three Laws of Robotics, supra note 157, at 1229–30.

185. Balkin, Information Fiduciaries, supra note 157, at 1225–26 ("Digital information fiduciaries may be held to reasonable ethical standards of trust and confidentiality, even if they do not make specific representations, because of the nature and kind of business they are in.").

fiduciary and harm the end user. Governments may act to protect these obligations of good faith, respect, and nonmanipulation, which sound both in consumer protection and privacy.

Digital curators operating in the United States may object to any regulation of their operations on the ground that the First Amendment protects their right to collect, collate, analyze, use, and distribute data as they choose. Although the Supreme Court has suggested that data is speech,[186] the protection of fiduciary relationships between social media companies and their end users should be constitutional for two reasons.

First, information gathered by digital curators in the context of a fiduciary relationship is not part of public discourse any more than the information gathered in the course of other fiduciary relationships like those between clients and doctors, lawyers, and estate managers.[187] The First Amendment allows governments to regulate fiduciaries' collection, collation, use, and distribution of personal information to prevent overreaching and breach of trust.[188] In the same way, the First Amendment should not foreclose regulations designed to protect the relationships of trust between the new class of digital-information fiduciaries and their end users.

Second, Congress can avoid any potential constitutional difficulties under the First Amendment by creating safe harbors for digital companies as described above.[189] Professor Jonathan Zittrain and I have proposed a Digital Millennium Privacy Act under which the federal government would preempt state regulation if digital media companies accept the obligations of information fiduciaries toward their end users.[190] Offering digital media companies greater protections than the Constitution affords as part of a grand bargain to protect end users should be constitutional.

---

186. See Sorrell v. IMS Health Inc., 564 U.S. 552, 570 (2011) ("[T]he creation and dissemination of information are speech within the meaning of the First Amendment."); see also Jane Bambauer, Is Data Speech?, 66 Stan. L. Rev. 57, 71–72 (2014) (arguing that data should be treated as speech for purposes of the First Amendment).

187. Balkin, Information Fiduciaries, supra note 157, at 1209–10, 1215–20.

188. Id. at 1215–20; see also Lowe v. SEC, 472 U.S. 181, 210–11 (1985) (distinguishing between regulation of investment advisors addressing the general public and regulation of advisors counseling individual clients in order to prevent "fraud, deception, or overreaching"); Robert C. Post, Democracy, Expertise, and Academic Freedom: A First Amendment Jurisprudence for the Modern State 22–23 (2012) (arguing that a central distinction in First Amendment law is between public discourse, which the state can regulate only in limited ways, and professional speech, which the state can regulate broadly to protect the interests of clients and beneficiaries).

189. See supra notes 146–153 and accompanying text.

190. See 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/4QK5-SLFY] (proposing that if digital businesses "agree to a set of fair information practices, including security and privacy guarantees, . . . the federal government would preempt a wide range of state and local laws" affecting them).

CONCLUSION

Free speech today is a triangle. Its three corners are nation-states, private infrastructure, and speakers.

This triangle creates three problems: (1) new-school speech regulation that produces collateral censorship and digital prior restraint; (2) abuse by privatized bureaucracies that govern end users arbitrarily and without due process and transparency; and (3) digital surveillance that facilitates manipulation.

Three reforms will help address these problems: (1) structural regulation that promotes competition and prevents discrimination by payment systems and basic internet services; (2) guarantees of curatorial due process; and (3) the recognition of a new class of information fiduciaries with duties of trustworthiness and good faith toward their end users.