Masthead Logo

SMU Law Review

Volume 72 | Issue 1

Article 9

2019

Global Platform Governance: Private Power in the Shadow of the State

Hannah Bloch-Wehba
Drexel University Thomas R. Kline School of Law, hcb38@drexel.edu

GLOBAL PLATFORM GOVERNANCE: PRIVATE POWER IN THE SHADOW OF THE STATE

*Hannah Bloch-Wehba**

ABSTRACT

Online intermediaries—search engines, social media platforms, even e-commerce businesses—are increasingly required to make critical decisions about free expression, individual privacy, and property rights under domestic law. These requirements arise in contexts that include the right to be forgotten, hate speech, "terrorist" speech, and copyright and intellectual property. At the same time, these disputes about online speech are increasingly borderless. Many laws targeting online speech and privacy are explicitly extraterritorial in scope. Even when not, some courts have ruled that they have jurisdiction to enforce compliance on a global scale. And governments are also demanding that platforms remove content—on a global scale—that violates platforms' terms of service, leading to the deletion of information that is legal in one jurisdiction and illegal in the next.

Existing accounts of platforms' governance role are incomplete and unsatisfying. These accounts tend to neglect the impact of cross-border and transnational pressures upon company policies that affect user rights. Observers have also tended to mischaracterize or underspecify the kinds of action that platforms take as governments outsource certain decision-making functions and attempt to extend domestic law and norms beyond territorial limits.

The Article contends that platforms are operating as privately owned bureaucracies charged with overseeing and implementing complex statutory and constitutional schemes. Platforms are engaged in both rulemaking and adjudication: they develop and promulgate regulations, statements of policy, and guidance that govern free expression and privacy online, and adjudicate disputes concerning those fundamental rights.

---

* Assistant Professor of Law, Drexel University Kline School of Law; Affiliated Fellow, Yale Law School Information Society Project. My thanks to Jack Balkin, Emily Berman, Rebecca Crootof, Jen Daskal, Kristen Eichensehr, Miriam Estrin, Claudia Haupt, Daphne Keller, Jane Kirtley, Heidi Kitrosser, Kate Klonick, Christopher Reed, and Rory Van Loo for helpful conversations and generous feedback. I am also grateful for comments from participants at the AALS New Voices in National Security Law panel, Amsterdam Privacy Conference, Freedom of Expression Scholars Conference, Internet Law Scholars Works in Progress Conference, and at All Things In Moderation: The People, Practices and Politics of Online Content Review. Finally, I am grateful to the editors of the *SMU Law Review* for their careful work on this Article. This Article reflects developments through February 2019, when it was finalized for publication. Any errors are my own.

That these governance mechanisms rely on private actors to be carried out does not, by itself, suggest that they are not legitimate. But basic principles of administrative law—transparency, participation, reason-giving, and review—remain essential to ensure that platform governance is accountable to the public. These protections are largely, if not entirely, absent from the status quo, due in part to longstanding industry practice—and in part to legal obstacles that prevent platforms from instituting the kinds of rigorous safeguards that are urgently needed.

TABLE OF CONTENTS

I. INTRODUCTION

In the last several years, providers of Internet communication services have come under increasing pressures from governments across the world to police speech and privacy online. Companies large and small alike have been enlisted by governments—both autocratic and demo-

cratic—to aid in efforts to stem the tide of nasty content, to assist in law enforcement investigations, to protect and vindicate privacy and intellectual property rights, and to police propaganda, misinformation, and "fake news." At times, companies have resisted these efforts, pushing back on government demands that go too far.[1] Frequently, however, they cooperate with governments and become crucial participants in efforts to make the Internet a better, more pleasant, and safer place.

These demands for cooperation are both geographically and ideologically diverse, and reflect varying degrees of government-imposed obligations on industry to "moderate content"—in other words, to remove unwanted speech. Increasingly, they also reflect mounting pressures on platforms not just to block or filter content in a particular jurisdiction, but rather to delete it on a global scale. Indeed, today's battles to control online speech are, in many cases, explicitly global in nature. At times, local regulators overtly attempt to extend their influence over the global Internet by fighting legal battles over jurisdiction.[2] But often, platforms play a global role in governing speech simply by enforcing their own terms of service (ToS)—applicable to all users, no matter where they are located.[3]

States have, accordingly, begun to exert a new kind of pressure on platforms. Rather than simply seeking to enforce domestic law online—whether globally or locally—states are leveraging the infrastructure of private ordering that has long characterized the global web in order to carry out their own policy preferences. In essence, platforms are acting as regulators or bureaucracies—implementing the favored policy solutions of executive and legislative branch actors—on a global scale. Indeed, platforms are performing quintessentially administrative functions. Like administrative agencies, platforms set rules by promulgating internal regulations about user speech and privacy.[4] And as the number of disputes about online rights grows beyond that which domestic courts can handle,

---

1. *See* Alan Z. Rozenshtein, *Surveillance Intermediaries*, 70 STAN. L. REV. 99, 115 (2018) (arguing that technology companies have "powerful incentives to resist government surveillance," particularly after the Snowden revelations); Julie Cohen, *Law for the Platform Economy*, 51 U.C. DAVIS L. REV. 133, 173 (2017).

2. *See, e.g.*, Request for a Preliminary Ruling from the Conseil d'État (France), Case C-507/17, Google Inc. v. Comm'n Nationale de l'Informatique et des Libertés (CNIL), 2017 O.J. (C 347) 23 (presenting the question of whether search results must be delisted on all of the domains used by Google globally).

3. *See* Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035, 1055–57 (2018) (observing, as a general matter, that ToS are applicable across the globe); *see also* Rebecca MacKinnon, Andi Wilson, & Liz Woolery, *Internet Freedom at a Crossroads: Recommendations for the 45th President's Internet Freedom Agenda*, OPEN TECH. INST. 12 (Dec. 2016), https://www.newamerica.org/oti/policy-papers/internet-freedom-crossroads/ (describing how Silicon Valley has responded to government pressure by "amending their terms of service and strengthening enforcement measures against a wide range of content that advocates or is associated with violent extremism").

4. *See, e.g.*, Rory Van Loo, *Rise of the Digital Regulator*, 66 DUKE L.J. 1267, 1269 (2017) (arguing, in the context of domestic commercial interactions, that digital intermediaries regulate "by influencing behavior in ways similar to public actors.").

platforms increasingly create adjudicatory and appellate mechanisms to absorb this responsibility as well.[5] While platforms have acknowledged that their rules and policies in some ways serve as the "law" of online content,[6] a closer look at the relationship between platforms and governments around the world is warranted in light of this increasing entanglement and overlap.

These developments are noteworthy because this relationship—what this Article calls "global platform governance"—exhibits several accountability deficits. States are increasingly coercing online platforms and intermediaries to instantiate and enforce *public* policy preferences regarding online speech and privacy through *private* regulation—including not only ToS but also hash-sharing and other purportedly cooperative arrangements—that lacks critical accountability mechanisms. These coercive measures convert what might otherwise be private action into heterodox, hybrid public-private governance arrangements in which state and private power are commingled. In short, governments can avoid responsibility for their policy preferences if they force platforms to carry their water.

This dynamic is of particular concern because the Internet's global reach heightens the likelihood of jurisdictional conflict concerning privacy and speech rights.[7] There is little agreement about either the substance or the process by which platforms and governments should make critical decisions about civil liberties online. Fundamental divisions exist among governments, industry, and civil society about the appropriate role that communications platforms and online intermediaries ought to play in governing online speech. On the one hand, the companies that have built and maintained platforms for communication have also long created and applied policies and ToS that govern what users can—and cannot—say online. Equipped with this experience, companies themselves may be the actors best suited to make decisions about governing online speech. On the other hand, many of the most urgent decisions about online speech and privacy implicate fundamental human and constitutional rights, sug-

5. *See, e.g.*, Rory Van Loo, *The Corporation as Courthouse*, 33 YALE J. ON REG. 547, 555–59 (2016) (describing how Amazon, eBay, and other platforms act as adjudicators for consumer purchases).

6. Alexis C. Madrigal, *Inside Facebook's Fast-Growing Content-Moderation Effort*, ATLANTIC (Feb. 7, 2018), https://www.theatlantic.com/technology/archive/2018/02/what-facebook-told-insiders-about-how-it-moderates-posts/552632/ (describing how Facebook employees "emphasized the similarity between what Facebook is doing and what government does"); *see also* Ingrid Burrington, *Could Facebook Be Tried for Human-Rights Abuses?*, ATLANTIC (Dec. 20, 2017), https://www.theatlantic.com/technology/archive/2017/12/could-facebook-be-tried-for-war-crimes/548639/ (quoting Susan Benesch, who called terms of service the "Constitution" of online rights).

7. *See* Paul Schiff Berman, *The Globalization of Jurisdiction*, 151 U. PA. L. REV. 311, 319 (2002) ("Cross-border interaction obviously is not a new phenomenon, but in an electronically connected world the effects of any given action may immediately be felt elsewhere with no relationship to physical geography at all."); *see also* Andrew Keane Woods, *Against Data Exceptionalism*, 68 STAN. L. REV. 729, 748 (2016) (noting that, although there is "very little precedent" regarding cross-border demands for the production of user data, these demands "are soon to be commonplace").

gesting that they should, perhaps, be adjudicated by the government itself through methods that enshrine values of due process, transparency, and accountability.

Take, for example, the "right to be forgotten" under European law. In the 2014 *Google Spain* case, the Court of Justice of the European Union (CJEU) held that Europe's Data Protection Directive protected an individual's right to demand that a search engine delete information that is "inadequate, irrelevant or excessive."[8] In response to the ruling, Google created a portal to receive and manage requests to delete information, which it considers on a case-specific basis. Google has developed internal guidelines for compliance with the *Google Spain* requirements, adjudicates each request to delete information itself, and employs a specialized staff that considers the requests. To date, Google has removed nearly 900,000 links.[9]

Or take the requirements that the European Commission and some European member states are beginning to impose on platforms in order to curb hate speech online. In May 2016, Facebook, Microsoft, Twitter, and YouTube implemented a "Code of Conduct" by which they committed to "have in place clear and effective processes to review notifications regarding illegal hate speech on their services so they can remove or disable access to such content."[10] Hate speech is disallowed by all four of the companies' ToS. Under the Code of Conduct, the companies promised to "review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such content, if necessary."[11] Compared to other new laws, the Code of Conduct has a soft touch: last year, Germany enacted a law requiring social media companies to delete illegal comments within twenty-four hours, or face heavy fines.[12]

The right to be forgotten and hate speech examples illustrate an emerging form of government pressure on platforms: rather than simply compelling intermediaries to delete specific content, governments are foisting upon platforms increasing responsibility for making *legal determinations* regarding speech—a task that might previously have belonged to a court, administrative agency, or other government body accountable to the pub-

8. Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos (AEPD), 2014 E.C.R. 317, ¶¶ 92, 94 [hereinafter *Google Spain*]; *see also* Ignacio Cofone, Google v. Spain*: A Right to Be Forgotten*, 15 CHI.-KENT J. INT'L & COMP. L. 1, 4–5 (2015).

9. *Search removals under European privacy law*, GOOGLE, https://transparencyreport.google.com/eu-privacy/overview?hl=en [https://perma.cc/H479-RCWK] (last visited Feb. 9, 2019).

10. European Comm'n, Code of Conduct on Countering Illegal Hate Speech Online (May 31, 2016) [hereinafter *Code of Conduct*], http://ec.europa.eu/newsroom/document.cfm?doc_id=42985 [https://perma.cc/E7QA-EJBX].

11. *Id.*

12. Melissa Eddy & Mark Scott, *Delete Hate Speech or Pay Up, Germany Tells Social Media Companies*, N.Y. TIMES (June 30, 2017), https://www.nytimes.com/2017/06/30/business/germany-facebook-google-twitter.html [https://perma.cc/G4WB-6C28].

lic. As a result, intermediaries bear increasing duties to make important decisions regarding sensitive civil liberties issues.

This Article contributes to a growing literature that explores how technology companies are functioning alongside states to create and enforce transnational policy on user speech, privacy, and security.[13] Specifically, the Article takes up two noteworthy developments resulting from this shift toward private governance. First, by enlisting intermediaries to carry out their desired policy preferences through ToS or other corporate standards, governments can give local policy global effect relatively cheaply. Nonetheless, most scholarly accounts have examined the substantive impact of these private governance arrangements on user speech through a domestic, not international, lens.[14] Second, although states are increasingly compelling platforms to adjudicate users' speech and privacy rights, they have stymied platforms' ability to provide procedural protections comparable to those that would typically accompany legal adjudications. The Article illuminates how the application of global administrative law both casts doubt on the legal sufficiency of existing procedural mechanisms to protect user rights, and points toward some solutions to enhance the accountability and transparency of platform decision-making.

Part I situates platforms' new governance role in the context of cyberlaw's embrace of private ordering. The premise that private actors rather than states should set many, if not all, of the rules of engagement online was fundamental to early cyber-enthusiasts, and even their harshest critics did not seriously dispute that large swaths of the Web should be self-governed. Indeed, self-governance was posited as a way of resolving jurisdictional tensions online. But as the Internet grew and became commercialized, platforms became increasingly susceptible to government control and pressure to extend the reach of local law. The Article examines two examples from Europe—hate speech and the right to be forgotten—and illustrates that, today, governments are in fact instrumentalizing platforms' own internal corporate architecture to instantiate their preferred policy outcomes.

Platform decision-making on the sensitive human rights and civil liberties issues implicated here has given rise to a number of complaints about the substance of platform governance and its accountability failures. In Part II, the Article offers a new way to conceptualize platforms' participation in *transnational* speech and privacy regulation, and a path forward

13. *See, e.g.*, Citron, *supra* note 3; Eyal Benvenisti, *Upholding Democracy Amid the Challenges of New Technology: What Role for Global Governance?*, 29 EUR. J. INT'L L. 9 (2018); Julie Cohen, *supra* note 1; Jennifer Daskal, *Borders and Bits*, 71 VAND. L. REV. 179, 219 (2018); Daphne Keller, *Who Do You Sue? State and Platform Hybrid Power Over Online Speech* (Hoover Inst., Stanford Univ., Aegis Series Paper No. 1902, 2019); Andrew Keane Woods, *Litigating Data Sovereignty*, 128 YALE L.J. 328 (2018); Evelyn Mary Aswad, *The Future of Freedom of Expression Online*, 17 DUKE L. & TECH. REV. 26 (2018); Kristen Eichensehr, *Digital Switzerlands*, 167 U. PA. L. REV. (forthcoming 2019).

14. *See, e.g.*, Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1599 (2018); Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353, 1358 (2018).

to hold them accountable in this role. When Internet platforms engage in policing and moderating content online, they engage in a form of private governance that reaches across geographic borders. These governance arrangements exhibit a complex web of international relationships between industry, government, and civil society organizations. And, like many global governance arrangements, they suffer from a "democratic deficit."[15] Viewed from this perspective, common critiques of platform decision-making are rooted in concerns—familiar to administrative law scholars—about the quandary of ensuring democratic accountability and legitimacy when private organizations serve as regulators.[16] These concerns are equally reflected in the body of scholarship centered on democratic accountability for institutions of global governance.[17]

Recognizing platforms' important regulatory role suggests its own solution to the problems of legitimacy and accountability: the application of administrative law principles and values to hold platforms to account. Part III explains how principles of transparency, participation, reasoned decision-making, and judicial review could be deployed to render private speech governance both legitimate and accountable. What's more, looking to procedural values like these would help to ensure that all the parties affected by content removal have a voice in the process, and would avoid fruitless attempts to create substantive rules for resolving a bevy of online disputes. Examining what it would take to implement these principles also makes clear that current frameworks governing content deletion are insufficient because they stymie the ability to achieve any of the four. In short, the imposition of these standards would benefit all parties—the public, the government, and the companies themselves.

II. GLOBAL PLATFORM GOVERNANCE—BACKGROUND

A. Private Ordering in the Shadow of the State

In order to understand why platforms' decisions on individual privacy and speech rights are increasingly contested, it is helpful to understand how the framework for online self-governance has shifted over time.

15. Andrew Moravcsik, *Is there a 'Democratic Deficit' in World Politics? A Framework for Analysis*, 39 Gov't & Opposition 336, 337 (2004).

16. *See, e.g.*, Government By Contract: Outsourcing and American Democracy (Jody Freeman & Martha Minow eds., 2009) (offering many different explanations of the accountability problems posed by widespread privatization); Jody Freeman, *Private Parties, Public Functions and the New Administrative Law*, 52 Admin. L. Rev. 813 (2000); Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. Rev. 543 (2000); Gillian E. Metzger, *Privatization as Delegation*, 103 Colum. L. Rev. 1367 (2003); Paul Starr, *The Meaning of Privatization*, 6 Yale L. & Pol'y Rev. 6 (1988); Daniel Guttiman, *Public Purpose and Private Service: The Twentieth Century Culture of Contracting Out and the Evolving Law of Diffused Sovereignty*, 52 Admin. L. Rev. 859 (2000); *see also* Van Loo, *supra* note 5.

17. *See, e.g.*, Robert O. Keohane, *Global Governance and Democratic Accountability*, *in* Taming Globalization: Frontiers of Governance 130–32 (David Held & Mathias Koenig-Archibugi eds., 2002); Patrizia Nanz & Jens Steffek, *Global Governance, Participation and the Public Sphere*, 39 Gov't & Opposition 314, 314 (2004); Moravcsik, *supra* note 15, at 336–37.

Early Internet thinkers embraced a radically individualistic vision of cyber self-governance oriented around autonomous communities, freedom of movement, and free expression online. While many scholars dismissed this view as idealistic or misguided—particularly the belief that the Internet should exist beyond the reach of national laws—elements of it persist to this day. In particular, at critical junctures, governments supported the emergence of online self-governance and self-regulation in the belief that it would stimulate innovation.[18] Those policy choices gave rise to the framework that governs social media today: ToS, privacy policies, and "community standards" constitute the most visible rules and regulations governing online speech and privacy.[19] Yet private governance of online speech has not insulated the Internet from pressure by national governments seeking to regulate. Today, numerous territories lay claim to the ability to regulate online activity. Increasingly, their weapon of choice is not the crude cudgel of national regulation, but *private ordering itself*.

Early Internet acolytes embraced the idea that cyberspace would be a new "place" or territory, beyond the jurisdiction of any "territorially based sovereign."[20] In 1996, John Perry Barlow, in a famous articulation of this principle, rejected the imposition of "legal concepts of property, expression, identity, movement, and context" on cyberspace as a "hostile and colonial measure."[21] Barlow's utopian narrative of Internet exceptionalism[22] suggested not only that the digital world should not be held to the same rules that apply to "meatspace"[23] or "wetware"[24] users, but also that the Internet would be *better* regulated by itself than by government, celebrating "the unwritten codes that already provide our society more order than could be obtained by any of your impositions."[25]

Although Barlow's quirky narrative was dismissed by some, it also, as Neil Weinstock Netanel noted, "resounded in thoughtful scholarship."[26] Most significantly, David Post and David R. Johnson had wrestled with Barlow's vision and concurred with his assessment that cyberspace posed

---

18. *See, e.g.*, White House, *The Framework for Global Electronic Commerce: Read the Framework* (July 1997), https://clintonwhitehouse4.archives.gov/WH/New/Commerce/read.html [https://perma.cc/3D2X-KGMQ].

19. *See* Klonick, *supra* note 14, at 1599.

20. David R. Johnson & David Post, *Law and Borders—the Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367, 1375 (1996).

21. John Perry Barlow, *A Declaration of the Independence of Cyberspace*, ELEC. FRONTIER FOUND. (Feb. 8, 1996), https://www.eff.org/cyberspace-independence [https://perma.cc/M98D-KBR6].

22. *See* Eric Goldman, *The Third Wave of Internet Exceptionalism*, TECH. & MKTG. L. BLOG (Mar. 11, 2009), https://blog.ericgoldman.org/archives/2009/03/the_third_wave.htm.

23. Dan Hunter, *Cyberspace as Place and the Tragedy of the Digital Anticommons*, 91 CALIF. L. REV. 439, 447 (2003).

24. DIANA SACO, CYBERING DEMOCRACY: PUBLIC SPACE AND THE INTERNET 77 (2002) ("Finally, the concept of *wetware*, a common hacker slang for human beings . . . refers to the actual users of cyberspace") (citation omitted).

25. Barlow, *supra* note 21 ("Where there are real conflicts, where there are wrongs, we will identify them and address them by our means. We are forming our own Social Contract. This governance will arise according to the conditions of our world, not yours.").

26. Neil Weinstock Netanel, *Cyberspace Self-Governance: A Skeptical View from Liberal Democratic Theory*, 88 CALIF. L. REV. 395, 398 (2000).

intrinsic challenges to the application of national law, concluding that "no physical jurisdiction has a more compelling claim than any other to subject these [online] events exclusively to its laws."[27] Post and Johnson pointed out that many users would be readily capable of understanding which online "rules" governed their behavior—whether in the form of ToS, norms, or guidelines—but unable to determine "which territorial-based authority might apply its laws to your conduct."[28]

In a sense, Barlow's Internet exceptionalism has not aged particularly well. As Tim Wu pointed out only a few years later, "[t]he metaphor of place did not exactly stand the test of time."[29] Dan Hunter described Barlow's rhetoric as "amusing and intentionally overblown."[30] Taken seriously, Hunter pointed out, Barlow and his fellow travelers should be understood as celebrating the promise of self-regulation as "the only appropriate governance structure" for the Internet.[31] Accordingly, alternative accounts soon emerged to challenge this vision for the online world. First, scholars took on the general assertion that the legal challenges posed by "cyberspace" were substantially distinct from those that courts had been grappling with for years.[32] For instance, Jack Goldsmith rejected the legal distinctions between cyberspace and real space relied upon by "regulation skeptics" as overstated or unsupported.[33] And Tim Wu brushed off the "imaginative" suggestions that digital space would develop its own rules, norms, and institutions that would substitute for government: "It's time to move on."[34] Second, Orin Kerr reframed the rhetoric of "cyberspace as place" as a matter of "perspective," casting doubt on claims that the framework was uniquely suited to explain what made the Internet special.[35] In addition, scholars such as Dan Hunter also pointed out that the reification of the Internet as a distinct "place" had unwanted effects, resulting in "an undesirable policy outcome: the staking out of private claims in cyberspace and concomitant reductions in the public 'ownership' of the space."[36] Meanwhile, Julie Cohen observed that

27. Johnson & Post, *supra* note 20, at 1376.

28. *Id.* at 1380.

29. Timothy Wu, *When Law & the Internet First Met*, 3 GREEN BAG 2D 171, 172 (2000).

30. Hunter, *supra* note 23, at 447.

31. *Id.*

32. Most famously, Judge Frank Easterbrook denigrated cyberlaw as "just the law of the horse," suggesting that lawyers and scholars would do better to "develop a sound law"—of intellectual property, in this instance—and then apply it to any unique circumstances presented in digital space. Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207, 208 (1996).

33. Jack L. Goldsmith, *Against Cyberanarchy*, 65 U. CHI. L. REV. 1199, 1200–01 (1998).

34. Wu, *supra* note 29, at 176–77.

35. Orin S. Kerr, *The Problem of Perspective in Internet Law*, 91 GEO. L.J. 357, 361–62 (2003).

36. Hunter, *supra* note 23, at 499; *see also* Mark A. Lemley, *Place and Cyberspace*, 91 CALIF. L. REV. 521, 529 (2003) (arguing that courts have recognized the differences between cyber and real space in rejecting the "[r]ote application of personal jurisdiction rules").

all of these theories—whether supportive of or opposed to cyberspace as its own "place"—failed to account for "both the embodied, situated experience of cyberspace users and the complex interplay between real and digital geographies."[37]

But while Barlow's utopian vision of a space beyond the state was criticized and rejected, few American observers challenged the foundational belief that private ordering, rather than public regulation, should govern much online activity. Even the most trenchant critics of the view that the Internet should operate free from terrestrial legal constraints observed that "private legal ordering" would fill many, albeit not all, legal gaps in cyberspace.[38] Yet there was also substantial uncertainty—rarely articulated in the legal scholarship—about what that private ordering would consist of. On the one hand, some full-throated arguments in favor of cyberspace exceptionalism appeared to envision semi-autonomous communities, comprised of individuals, that were largely self-governed by consensus.[39] Although these arguments had different emphases—sometimes, for example, stressing the Internet's potential as a site of direct democracy, and sometimes stressing individual freedom and choice[40]—they were based on a fundamentally common premise that each community would control the formulation of its own rules and norms, leading to substantial diversity in approaches across the Internet. Moreover, that vision of community control was a democratic one. Post and Johnson suggested that, while system operators were capable of unilaterally exercising the power of "banishment" to enforce online rules and norms, online communities increasingly had started "explicitly to recognize that formulating and enforcing such rules should be a matter for principled discussion, not an act of will by whoever has control of the power switch."[41] As Anne Wells Branscomb noted in 1995, in some online contexts, "[the] law of the wild reigns, with each individual sysop [system operator] acting as the Lord Chancellor and High Executioner;" but in other "communities," rules and regulations were made and enforced, often with clarity and purpose.[42]

Far from being outright rejected, elements of this highly individualistic vision for online self-governance not only became reality, but were embraced by governments. In 1997, the Clinton Administration explicitly ap-

37. Julie Cohen, *Cyberspace As/And Space*, 107 COLUM. L. REV. 210, 213 (2007).

38. Goldsmith, *supra* note 33, at 1215–16 ("Private legal ordering thus has the potential to resolve many, but not all, of the challenges posed by multijurisdictional cyberspace activity.").

39. Johnson & Post, *supra* note 20, at 1393 ("If the sysops and users who collectively inhabit and control a particular area of the Net want to establish special rules to govern conduct there, and if that rule set does not fundamentally impinge upon the vital interests of others who never visit this new space, then the law of sovereigns in the physical world should defer to this new form of self-government.").

40. *See* Netanel, *supra* note 26, at 404 (describing "cyberpopulist," "cybersyndicalist," and "cyberanarchist" arguments in favor of online self-governance).

41. Johnson & Post, *supra* note 20, at 1388.

42. Anne Wells Branscomb, *Anonymity, Autonomy, and Accountability: Challenges to the First Amendment in Cyberspaces*, 104 YALE L.J. 1639, 1668–69 (1995).

plauded self-regulation as the primary mechanism for regulating the Internet in its "Framework for Global Electronic Commerce," writing that, as a matter of policy, "governments should encourage industry self-regulation wherever appropriate and support the efforts of private sector organizations to develop mechanisms to facilitate the successful operation of the Internet. Even where collective agreements or standards are necessary, private entities should, where possible, take the lead in organizing them."[43] In the interest of promoting innovation, governments also opted to create immunity provisions shielding platforms from civil or criminal liability for content posted by others—and permitting companies to make their own rules to govern which speakers and what content would be allowed on their platforms.

Intermediary protections from liability for content posted by others became prevalent across the globe.[44] In the United States, § 230 of the Communications Decency Act cemented this approach by immunizing platforms and hosts from liability for information posted by third parties and protecting their efforts to block or filter "offensive" material from their services.[45] Even the European Union, which initially appeared to support strong regulation of the Internet, turned toward self-regulation in the late 1990s,[46] and toward "co-regulation" thereafter.[47] Europe also embraced the need to "promote international standards" to protect intermediaries "from the obligation of blocking Internet content without prior due process."[48] Under the European Commission's E-Commerce Directive, intermediary service providers were shielded from liability in Member States if they lacked "actual knowledge of illegal activity or information," or "act[ ] expeditiously to remove or to disable access" once they gain knowledge.[49] The E-Commerce Directive likewise bars Member States from imposing general obligations on intermediaries "to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity."[50] In addition, Article 10 of the European Convention on Human Rights places

43. White House, *supra* note 18.

44. *See About the World Intermediary Liability Map*, STAN. CTR. FOR INTERNET & SOC'Y, http://wilmap.law.stanford.edu/about [https://perma.cc/XMG8-J328] (last visited Feb. 2, 2019).

45. 47 U.S.C. § 230(c) (2012); *see also* Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1169 (9th Cir. 2008) (holding that web site was not immune under § 230 where its connection to illegal content was "direct and palpable"); Chicago Lawyers' Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc., 519 F.3d 666, 670 (7th Cir. 2008); Zeran v. Am. Online, Inc., 129 F.3d 327, 330 (4th Cir. 1997).

46. Matthew J. Feeley, *EU Internet Regulation Policy: The Rise of Self-Regulation*, 22 B.C. INT'L & COMP. L. REV. 159, 168 (1999).

47. *See, e.g.*, CHRISTOPHER T. MARSDEN, INTERNET CO-REGULATION: EUROPEAN LAW, REGULATORY GOVERNANCE AND LEGITIMACY IN CYBERSPACE (2011).

48. Council of the European Union, *EU Human Rights Guidelines on Freedom of Expression Online and Offline*, ¶ 34 (May 12, 2014) [hereinafter *EU Human Rights Guidelines*], https://eeas.europa.eu/sites/eeas/files/eu_human_rights_guidelines_on_freedom_of_expression_online_and_offline_en.pdf [https://perma.cc/9DC6-QCK6].

49. Council Directive 2000/31, art. 14, 2000 O.J. (L 178) (EC).

50. *Id.* art. 15.

some limits on third-party liability for user-generated online content.[51]

Although online self-regulation and self-governance carried the day, however, their current form looks distinct from John Perry Barlow's original vision for cyberspace autonomy. The forms of online rules and regulations that are most visible to the Internet-using public—ToS, privacy policies, and "community standards," particularly on social media—scarcely resemble a set of rules and norms that were negotiated and agreed upon through direct democratic participation and a user-generated process. While companies may occasionally solicit public feedback on draft policy changes, doing so is more of a publicity stunt than a negotiation tactic; online platforms need not take user criticisms into account.[52]

In other words, the world of online "self-governance" in 2018 is neither the libertarian utopia envisioned by Barlow nor the anarchist dystopia envisioned by critics. The Internet is not a space apart, but rather one closely governed by corporate entities. At the same time, Barlow was not wrong to anticipate that the most important governors online would not be "territorial sovereigns."[53] But although states have taken a back seat to private enterprise in regulating online speech and privacy, the Internet hardly resembles Barlow's "world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity."[54] Rather, online life is continually—and increasingly aggressively—governed by rules made and enforced by corporate entities like Google, Facebook, and Twitter.[55]

In a sense, this development is not a departure from Barlow's vision, but its logical conclusion. Because the ethos of cyberspace exceptionalism was, at its core, neoliberal and market-driven,[56] it should come as no surprise that some platforms have gained increasing market dominance and concomitant control. Yet this asymmetric relationship between consumers and corporate platforms is a far cry from a universe in which an online community would agree upon collective norms, and a system operator—accountable to that community—would enforce them.

---

51. MTE v. Hungary, App. No. 22947/13, Eur. Ct. H.R. (2016) (finding that Article 10 of the ECHR limits third-party liability for online content). *But see* Delfi AS v. Estonia, App. No. 64569/09, Eur. Ct. H.R. (2015) (finding that imposition of liability for unlawful online hate speech does not offend Article 10).

52. *See* Josh Constine, *Facebook rewrites Terms of Service, clarifying device data collection*, TECHCRUNCH (Apr. 4, 2018), https://techcrunch.com/2018/04/04/facebook-terms-of-service/; Lizzie O'Shea, *Time to cut ties with the digital oligarchs and rewire the web*, GUARDIAN (Mar. 20, 2018), https://www.theguardian.com/commentisfree/2018/mar/20/digital-oligarchs-rewire-web-facebook-scandal; Margot E. Kaminski & Kate Klonick, *Facebook, Free Expression and the Power of a Leak*, N.Y. TIMES (June 27, 2017), https://www.nytimes.com/2017/06/27/opinion/facebook-first-amendment-leaks-free-speech.html.

53. *See* Johnson & Post, *supra* note 20, at 1392.

54. Barlow, *supra* note 21.

55. *See, e.g.*, Klonick, *supra* note 14, at 1601–03.

56. *See* Netanel, *supra* note 26, at 410 ("[F]or the most part, the cyberian project is a neoliberal one. They view liberal democracy as a second-best alternative to private agreement.").

But perhaps the most significant error the cyber-exceptionalists made was in thinking that private ordering would resolve issues of legitimacy and accountability for online governance. As the early cyberlaw scholars recognized, the Internet's global reach heightens substantive disagreements among nations about the scope of speech, privacy, and property protections. In suggesting that the Internet would be a new "place" or territory, beyond the jurisdiction of any "territorially based sovereign,"[57] cyber-exceptionalists predicted that the Internet could escape these disagreements by permitting online communities to create rules and norms to govern themselves.

This prediction could not have been more wrong. To the contrary, the development of the platform ecosystem has replicated jurisdictional conflict through the lens of private ordering.[58] Internet self-governance resulted in the privatization of surveillance and speech regulation and the emergence of "new-school" methods of speech regulation.[59] As Jack Balkin has demonstrated, the growth of platforms' own power to control speech and privacy has rendered them vulnerable to state control and pressure through a new set of techniques: "collateral censorship," "public-private cooperation and cooptation," and "digital prior restraint."[60] Moreover, the companies that have gained competitive control of the "infrastructure of free expression" provide only weak protections when a government "uses that infrastructure, or its limitations, as leverage for regulation or surveillance."[61] These features do not insulate the Internet from global pressure, but rather make it more vulnerable.

Indeed, platforms' weaknesses are compounded when multiple national governments pressure them to adopt new policy changes that are often in conflict with other national policies. Although some have suggested that online intermediaries are increasingly resistant to government pressures to assist in surveillance,[62] the evidence demonstrates that platforms are also increasingly willing to modulate their own rules on online speech and privacy to better fit a range of government demands. At the same time, governments seeking to control online speech are increasingly leveraging the infrastructure of private ordering itself—in particular, through "voluntary" agreements among enterprise and through the application and enforcement of ToS—to achieve global effects for their policy preferences.

---

57. Johnson & Post, *supra* note 20, at 1375.

58. *See* Daskal, *supra* note 13, at 219.

59. Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2298–99 (2014).

60. *Id.*

61. *Id.* at 2303.

62. *Chapter One: Cooperation or Resistance?: The Role of Tech Companies in Government Surveillance*, 131 HARV. L. REV. 1722, 1727 (2018) (pointing to an uptick in "intermediary-initiated litigation") [hereinafter *Cooperation or Resistance?*].

B. Private Ordering and Online Speech: Intellectual Property

As a general matter, efforts to enlist intermediaries to curb unwanted speech on a global basis are nothing new, although they are constantly evolving. Rightsholders have enlisted a variety of online intermediaries in efforts to apply U.S. intellectual property laws on a global scale. Section 512 of the Digital Millennium Copyright Act of 1998 (DMCA) provides online service providers with statutory immunity from secondary liability when infringing material is transmitted through their platforms.[63] In order to maintain their immunity under § 512's safe harbor, service providers implement "notice and takedown" procedures: a rightsholder can alert an online service provider of allegedly infringing content. The intermediary will typically remove that content and notify the alleged infringer, who can file a "counter notice" challenging the removal.[64]

The DMCA is a domestic law, but platforms' notice and takedown procedures give it global effect: when an intermediary deletes content under the DMCA, it typically does so across the entire platform.[65] This practice, which has largely gone unstudied, raises a number of jurisdictional and choice of law issues. As Alex Feerst, head of legal at Medium, has pointed out, the legal status of DMCA takedown requests originating from rightsholders outside of the United States is particularly unclear.[66] Moreover, some platforms have resisted demands to take down infringing content when compelled to do so by foreign courts.[67] Global deletion only compounds the significant problems with notice-and-takedown identified by several scholars, who have argued that the DMCA framework leads to over-deletion of lawful content. For instance, a recent quantitative analysis of a random sample of over 1,800 takedown requests found a significant number of requests either incorrectly identified or insufficiently specified the allegedly infringing work.[68]

---

63. 17 U.S.C. § 512 (2012).

64. *See* Jennifer M. Urban et al., *Notice and Takedown: Online Service Provider and Rightsholder Accounts of Everyday Practice*, 64 J. Copyright Soc'y U.S. 371, 373 (2017).

65. Annemarie Bridy, *Intellectual Property*, *in* Law, Borders, and Speech: Proceedings and Materials, Stan. Ctr. for Internet & Soc'y 13 (Daphne Keller ed., 2017), https://cyberlaw.stanford.edu/publications/proceedings-volume [https://perma.cc/SP9A-NE7Y] ("Has notice and takedown for copyright become de facto harmonized through platforms' global application of the DMCA?").

66. *Id.* ("Do such notices effectively signal acceptance of US legal jurisdiction over the dispute, and potentially even waiver of remedies under other countries' laws? Must right holders assert only copyright claims that are valid under US law, or can they use the DMCA to assert claims under the law of the sender's country?").

67. *See, e.g.*, Equustek Sols. Inc. v. Jack, 2014 BCSC 1063, 63 B.C.L.R. 5th 145 (Can. B.C. S.C.) (issuing injunction compelling Google to remove links to hardware distributor that had infringed plaintiff's trademark on a global basis).

68. Jennifer M. Urban et al., *Notice and Takedown in Everyday Practice* 2 (U.C. Berkeley, Pub. Law & Legal Theory Research, Paper No. 2755628, 2017); *see also* Daphne Keller, *Empirical Evidence of "Over-Removal" by Internet Companies Under Intermediary Liability Laws*, Stan. Ctr. for Internet & Soc'y (Oct. 12, 2015, 8:23 AM), http://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-companies-

Copyright is far from the only context, however, in which the law creates incentives for intermediaries to participate in private online governance that reaches across national borders. For instance, Annemarie Bridy has demonstrated that some stakeholders are increasingly seeking to broaden the role of ICANN—the Internet Corporation for Assigned Names and Numbers—in resolving disputes about online content.[69] ICANN, which is an independent corporation responsible for administering the Domain Name System (DNS),[70] has long been involved in resolving conflicts that concerned "cybersquatting"—bad-faith use of another's trademark in a domain name. Rather than adjudicating these disputes itself, ICANN adopted the Uniform Domain Name Dispute Resolution Policy (UDRP), an "ICANN-administered alternative dispute resolution system in which ICANN-accredited arbitrators decide disputes via a streamlined, web-enabled process."[71] The UDRP was mandatory for adjudication of registrars' and registrants' cybersquatting disputes.

The UDRP was never intended to adjudicate disputes about whether site content infringed intellectual property rights. Indeed, that appeared to be far beyond ICANN's purview.[72] However, Bridy documents how rightsholders and registry operators are entering into private, voluntary arrangements to facilitate the worldwide blocking of entire domains on which infringing content is hosted. In 2016, the Motion Picture Association of America (MPAA)—one of the biggest organizations representing rightsholders—entered into a voluntary "trusted notifier" arrangement with Donuts, a major registry operator that controls domains such as .MOVIE, .THEATER and .COMPANY.[73] Under the arrangement, MPAA can directly refer complaints of "pervasive copyright infringement" on websites to Donuts, which can then work with registrars to determine whether the websites violate Donuts' ToS.[74]

---

under-intermediary-liability-laws [https://perma.cc/383C-6VAH] (summarizing several studies).

69. Annemarie Bridy, *Notice and Takedown in the Domain Name System: ICANN's Ambivalent Drift into Online Content Regulation*, 74 WASH. & LEE. L. REV. 1345, 1347 (2017).

70. *About ICANN*, ICANN, https://www.icann.org/resources/pages/welcome-2012-02-25-en [https://perma.cc/5838-T9KR] (last visited Feb. 9, 2019).

71. Bridy, *supra* note 69, at 1356.

72. *See, e.g.*, David Post, *ICANN, copyright infringement, and "the public interest"*, WASH. POST (Mar. 9, 2015), https://www.washingtonpost.com/news/volokh-conspiracy/wp/2015/03/09/icann-copyright-infringement-and-the-public-interest/?utm_term=.994762d0df be [https://perma.cc/85QE-8T8H] ("Who authorized them to do that? What does that have to do with ICANN's fundamental mission (as stated in its own Charter): to 'coordinate . . . the global Internet's system of unique identifiers . . . to ensure the stable and secure operation' of that system?").

73. *One Year Later: Trusted Notifier Program Proves Effective*, MOTION PICTURE ASS'N AM. (Mar. 6, 2017), https://www.mpaa.org/press/one-year-later-trusted-notifier-program-proves-effective/ [https://perma.cc/K4ZR-ZRLB].

74. *Donuts and the MPAA Establish New Partnership to Reduce Online Piracy*, MOTION PICTURE ASS'N AM. (Feb. 9, 2016), https://www.mpaa.org/wp-content/uploads/2016/02/Donuts-and-MPAA-Establish-New-Partnership-2.9.16.pdf [https://perma.cc/5REX-DRGJ]. The MPAA has a similar partnership with Radix, a Dubai-based registry operator. *See Radix and the MPAA Establish New Partnership to Reduce Online Piracy*, MOTION

The MPAA's position is that voluntary cooperative measures to curb clearly illegal copyright infringement raise no concerns about chilling permissible speech.[75] But as Bridy points out, voluntary enforcement arrangements between rights holders and registries simply aggravate the accountability and transparency issues presented by the UDRP and other alternative dispute resolution methods.[76] When adjudication occurs in private, it raises questions about transparency, accountability, and the quality of decision making, none of which are visible to the public.[77] While the UDRP at least publishes its opinions,[78] the same cannot be said of many other ADR methods, in which neither the proceedings nor the outcomes are public.[79] As Bridy points out, rightsholders and registry operators might find private governance valuable and efficient because it allows them to avoid "[t]horny questions of sovereignty, jurisdiction, and choice of law."[80]

III. GLOBAL PLATFORM GOVERNANCE— TWO STORIES FROM EUROPE

In some contexts, too, private governance allows decision-making to occur out of the public eye—an outcome that might be particularly desirable in politically charged contexts. These features are not unique to the context of intellectual property. Indeed, new efforts to police and remove

---

PICTURE ASS'N AM. (May 13, 2016), https://www.mpaa.org/wp-content/uploads/2016/05/Radix-and-the-MPAA-Establish-New-Partnership-to-Reduce-Online-Piracy.pdf [https://perma.cc/8S8B-VYQP].

75. *One Year Later: Trusted Notifier Program Proves Effective*, *supra* note 73 ("Speech interests are not implicated nor 'content regulation' concerns triggered by cooperative efforts geared against wholesale piracy.").

76. Bridy, *supra* note 69, at 1359 ("Empirical study of the quality of UDRP decision-making is hampered by the fact that opinions alone are published without any of the parties' submissions. Lack of access to a full, public record makes it impossible to evaluate the provider's reasoning in light of the facts and competing arguments presented to it.").

77. *See, e.g.*, *ADR Advantages*, WORLD INTELL. PROP. ORG., http://www.wipo.int/amc/en/center/advantages.html [https://perma.cc/94Y4-XJKJ] (last visited Feb. 9, 2019) ("ADR proceedings are private."); Judith Resnik, *The Privatization of Process: Requiem For and Celebration of the Federal Rules of Civil Procedure at 75*, 162 U. PA. L. REV. 1793, 1821 (2014) ("[T]he promise of confidentiality is a linchpin of ADR's appeal, and the rules of the leading purveyors of ADR require it.").

78. *Rules for Uniform Domain Name Dispute Resolution Policy*, ICANN (Sept. 28, 2013), https://www.icann.org/resources/pages/udrp-rules-2015-03-11-en [https://perma.cc/7J9R-QFTM] ("Except if the Panel determines otherwise (see Paragraph 4(j) of the Policy), the Provider shall publish the full decision and the date of its implementation on a publicly accessible web site. In any event, the portion of any decision determining a complaint to have been brought in bad faith (see Paragraph 15(e) of these Rules) shall be published.").

79. Phillip Landolt & Alejandro García, *Commentary on WIPO Arbitration Rules*, WIPO ARB. & MEDIATION CTR. (2017), http://www.wipo.int/export/sites/www/amc/en/docs/2017commentrulesarb.pdf [https://perma.cc/5NXD-7F3R] ("Unique amongst the leading institutional rules, Article 75 of the WIPO provides that the parties to a WIPO arbitration are not able to disclose to third parties the existence of the arbitration. This obligation logically encompasses more specific information about the arbitration, for example, the cause of action, remedies sought, IP Rights in issue (where applicable) and the composition of the arbitral Tribunal.").

80. Bridy, *supra* note 69, at 1376.

illegal content online double down on trends toward private governance of online content, drawing on the same kinds of voluntary, cooperative arrangements and notice-and-takedown procedures. The result is that governments are increasingly able to leverage platforms' own procedures to make global policy—without public accountability.

A. Hate Speech, Terrorist Speech, and the Hash Database

Rising pressures on platforms to assist in monitoring and deleting illegal hate speech vividly illustrate the increasingly entwined, and often fraught, interests of states and platforms. Despite Europe's historically strong intermediary protections,[81] in recent years, both Member States and the European Commission have moved toward imposing additional obligations—through mechanisms both soft and hard—for intermediaries to monitor their own platforms and to delete illegal hate speech and terrorist speech. Over time, intermediaries' obligations have shifted and become less clear, partly because platforms have participated in voluntary self-regulatory and co-regulatory initiatives, through which they increasingly instantiate state speech preferences through private ordering.[82]

The European Union's initial effort to address online hate speech, the 2008 Framework Decision on combating certain forms and expressions of racism, was directed at Member States, not online service providers. Under the Framework Decision, Member States were bound to criminalize certain "expressions of racism and xenophobia."[83] These included "publicly inciting to violence or hatred directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin," or "publicly condoning, denying or grossly trivialising crimes of genocide, crimes against humanity and war crimes . . . when the conduct is carried out in a manner likely to incite to violence or hatred against such a group or a member of such a group."[84]

In May 2014, the Council of the European Union adopted the EU Human Rights Guidelines on Freedom of Expression Online and Offline.[85] Broadly speaking, the guidelines recognized the import of online communications for the freedoms of opinion, expression, and privacy protected in human rights instruments, noting that "[a]ll human rights that exist offline must also be protected online."[86] In invoking the language of human rights, the EU also gestured to the UN Guiding Princi-

81. *See* Daphne Keller, *The Right Tools: Europe's Intermediary Liability Laws and the EU 2016 General Data Protection Regulation*, 33 Berkeley Tech. L.J. 289, 295, 305–06 (2018) [hereinafter Keller, *The Right Tools*].

82. *See* Citron, *supra* note 3, at 1047 ("No matter how often EU lawmakers describe the recent changes to private speech practices as 'voluntary' decisions, they can only be reasonably understood as the product of government coercion.").

83. Council Framework Decision 2008/913/JHA of 28 November 2008, 2008 O.J. (L 328) 55, 55.

84. *Id.* at 55–56.

85. *EU Human Rights Guidelines*, *supra* note 48.

86. *Id.* ¶ 6.

ples on Business and Human Rights, recognizing that although human rights law binds only states, corporations too may have obligations to respect human rights.[87] The EU vowed to "promote" corporate responsibility for human rights violations online, but also recognized "the need to promote international standards, including standards protecting intermediaries from the obligation of blocking Internet content without prior due process."[88]

In adopting the guidelines, the EU highlighted the importance of procedural safeguards, noting that legislation constraining expressive rights "must be applied by a body which is independent of any political, commercial or other unwarranted influence in a manner that is neither arbitrary nor discriminatory, and with adequate safeguards against abuse, including the possibility of challenge and remedy against its abusive application."[89] Finally, the guidelines recognized the difficult balance between regulating unlawful speech and freedom of expression, noting that "[h]ate speech legislation should not be abused by governments to discourage citizens from engaging in legitimate democratic debate on matters of general interest."[90]

In 2015, France saw a slew of attacks. In January, the Kouachi brothers carried out a mass shooting at the offices of Charlie Hebdo, a satirical magazine. Days later, an associate carried out a second attack at a kosher supermarket in Paris. Then, in November, coordinated attacks across Paris—including at the Bataclan, a soccer stadium, and restaurants—killed 130 people and wounded hundreds more.[91] In the wake of the attacks in Paris and Brussels, European governments became increasingly vigilant about domestic threats. Europol, the European Union's law enforcement agency, pointed in part to social media, writing that terrorists relied on the Internet and social media for "dissemination of propaganda material but also for recruitment and fundraising," and that they had "adapt[ed] to efforts made by social media platforms and authorities to contain their online activities."[92]

Since 2015, European states and regional actors have taken a number of approaches to remediating hateful and violent speech online. First, platforms have engaged in partnerships with government entities to find a resolution to the problem of harmful online speech. In December 2015,

87. U.N. Office of the High Comm'r, *Guiding Principles on Business and Human Rights*, 13, U.N. Doc. HR/PUB/11/04 (2011), https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf [https://perma.cc/H24P-MZ2P] (Principle 11: "Business enterprises should respect human rights. This means that they should avoid infringing on the human rights of others and should address adverse human rights impacts with which they are involved.").

88. *EU Human Rights Guidelines*, *supra* note 48, ¶ 34.

89. *Id.* ¶ 22.

90. *Id.* annex I.

91. EUROPOL, EUROPEAN UNION TERRORISM SITUATION AND TREND REPORT 22 (2016), https://www.europol.europa.eu/sites/default/files/documents/europol_tesat_2016.pdf [https://perma.cc/DSZ6-UAPR].

92. *Id.* at 16.

the European Commission launched the EU Internet Forum, a multi-stakeholder forum intended "to reach a joint, voluntary approach based on a public-private partnership to detect and address harmful material online."[93] Ve?ra Jourová, Commissioner for Justice, Consumer and Gender Equality, said: "Freedom of speech is a human right, but this right does not protect illegal hate speech inciting violence and hatred."[94] The forum included high-level representatives from Facebook, Microsoft, Twitter, and YouTube, as well as Europe, the interior ministers of several EU Member States, and representatives of the European Parliament.

Following the launch of the forum, in May 2016, Facebook, Microsoft, Twitter, and YouTube agreed to adhere to a voluntary "Code of Conduct" negotiated with the European Commission, under which the companies agreed to expedite the review of notifications to remove hate speech and implement new rules, community guidelines, and processes to bar hate speech from their platforms.[95] The Code of Conduct required the companies to establish "clear and effective processes" for reviewing removal requests, to "prohibit the promotion of incitement to violence and hateful conduct" on their platforms, and to review the removal demands against both their own internal rules as well as national laws.[96]

The companies also agreed to partner with "trusted reporters" or "trusted flaggers" who would notify platforms of the existence of content that violated their ToS.[97] The Code of Conduct envisioned that the trusted reporters would include non-government civil society organizations,[98] and that platforms would "review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such content, if necessary."[99] Some platforms have also given trusted flagger status to specialized law enforcement departments, known as "internet referral units," that monitor the web for illicit material, including terrorist content.[100] Two civil society partners withdrew from the Forum in protest, writing that the process of using ToS to delete content—rather than legal demands from law enforcement—was "established outside an accountable democratic framework" and "ex-

93. European Commission Press Release IP/15/6243, EU Internet Forum: Bringing Together Governments, Europol and Technology Companies to Counter Terrorist Content and Hate Speech Online (Dec. 3, 2015), http://europa.eu/rapid/press-release_IP-15-6243_en.htm [https://perma.cc/4BDD-GW3C].

94. *Id.*

95. *Code of Conduct*, *supra* note 10.

96. *Id.*

97. European Commission Press Release IP/16/1937, European Commission and IT Companies announce Code of Conduct on illegal online hate speech (May 31, 2016), http://europa.eu/rapid/press-release_IP-16-1937_en.htm [https://perma.cc/FAW5-YKRR].

98. *Code of Conduct*, *supra* note 10.

99. *Id.*

100. *See* Brian Chang, *From Internet Referral Units to International Agreements: Censorship of the Internet by the UK and EU*, 49 COLUM. HUM. RTS. L. REV. 114, 120–22 (2018) ("[S]ome ICT companies, such as Google and YouTube, give IRUs a 'trusted flagger' status.").

ploits unclear liability rules for companies."[101]

Platforms have also taken a voluntary approach to harmful online content. In December 2016, several companies announced that they had "commit[ed] to the creation of a shared industry database of 'hashes' — unique digital 'fingerprints' — for violent terrorist imagery or terrorist recruitment videos or images that we have removed from our services."[102] Hash sharing is not entirely new—a similar collaborative arrangement began in 2012 to "improve and accelerate the identification, removal and reporting of child abuse images across different digital networks."[103] By December 2017, a year later, the terrorism hash database included more than 40,000 hashes, according to *The Guardian*.[104]

In March 2017, the European Union enacted the Terrorism Directive, which broadened the list of terrorism offenses to include

> the distribution, or otherwise making available by any means, whether online or offline, of a message to the public, with the intent to incite the commission of [a terrorist offense], where such conduct, directly or indirectly, such as by the glorification of terrorist acts, advocates the commission of terrorist offences, thereby causing a danger that one or more such offences may be committed.[105]

The Terrorism Directive recognized that "[a]n effective means of combating terrorism on the internet is to remove online content constituting a public provocation to commit a terrorist offence at its source," and enumerated a number of options for doing so: legislative, non-legislative, and judicial.[106] However, the directive also appeared to preserve the shield against intermediary liability in the e-commerce directive, providing that "no general obligation should be imposed on service providers to monitor the information which they transmit or store, nor to actively seek out facts or circumstances indicating illegal activity."[107]

---

101. *EDRi and Access Now withdraw from the EU Commission IT Forum discussions*, EUROPEAN DIGITAL RIGHTS INITIATIVE (May 31, 2016), https://edri.org/edri-access-now-withdraw-eu-commission-forum-discussions/.

102. *Partnering to help curb the spread of terrorist content online*, GOOGLE (Dec. 5, 2016), https://blog.google/topics/google-europe/partnering-help-curb-spread-terrorist-content-online/ [https://perma.cc/G574-DJGM].

103. *Sharing hashes across industry*, THORN, https://www.thorn.org/reporting-child-sexual-abuse-content-shared-hash/ [https://perma.cc/QM4D-3R9J] (last visited Feb. 9, 2019); *see also* Tomas Rudl, *EU-Kommission: Immer mehr Plattformen sollen Uploads filtern*, NETZPOLITIK (Feb. 23, 2018), https://netzpolitik.org/2018/eu-kommission-immer-mehr-plattformen-sollen-uploads-filtern/ [https://perma.cc/LL8V-3D79] ("Andererseits hält sie immer mehr Plattformen dazu an, sich an der Datenbank zu beteiligen, in der die digitalen Fingerabdrücke der inkriminierten Inhalte abgelegt sind. Landet ein solcher „Hash" einmal in der Datenbank, ist ein erneutes Hochladen der jeweiligen Datei nicht mehr möglich.").

104. Samuel Gibbs, *EU warns tech firms: remove extremist content faster or be regulated*, GUARDIAN (Dec. 7, 2017), https://www.theguardian.com/technology/2017/dec/07/eu-warns-tech-firms-facebook-google-youtube-twitter-remove-extremist-content-regulated-european-commission [https://perma.cc/L5VY-ZUCB].

105. Council Directive (EU) 2017/541, art. 5, 2017 O.J. (L 88) 6, 14.

106. *Id.* pmbl., ¶ 22.

107. *Id.* ¶ 23. (" Furthermore, hosting service providers should not be held liable as long as they do not have actual knowledge of illegal activity or information and are not

Throughout 2017, the "voluntary" cooperation arrangements between government and industry took on a harder edge as governments called on industry to develop new tools to stamp out illegal content faster and more accurately. In June 2017, the European Council adopted a conclusion that "[i]ndustry has its own responsibility to help combat terrorism and crime online."[108] The Council wrote that it expected industry "to develop new technology and tools to improve the automatic detection and removal of content that incites to terrorist acts."[109] The G7 had likewise called on online platforms "to act urgently in developing and sharing new technology and tools to improve the automatic detection of content promoting incitement to violence."[110]

The threat of legislation was not far behind.[111] In September 2017, the European Commission issued a communication entitled "Tackling Illegal Content Online: Towards an enhanced responsibility of online platforms."[112] As one commentator noted, "the thrust of the Communication is apparent from its sub-title."[113] The central new development of the communication, however, is that the Commission encouraged platforms to "adopt effective *proactive* measures to detect and remove illegal content online and not only limit themselves to reacting to notices which they receive."[114] The communication also encouraged platforms to rely on "trusted flaggers, . . . specialised entities with specific expertise in identifying illegal content, and dedicated structures for detecting and identifying such content online."[115] Trusted flaggers, according to the communication, should be "expected to bring their expertise and work with high quality standards, which should result in higher quality notices and faster take-downs."[116]

By December 2017, when the third "ministerial meeting" of the forum took place, industry representatives had expanded considerably: Justpaste.it, Snap, Wordpress, and Yellow all joined the forum meeting.

---

aware of the facts or circumstances from which the illegal activity or information is apparent.").

108. Presidency Conclusions, Brussels European Council ¶ 2 (June 23, 2017), http://www.consilium.europa.eu/media/23985/22-23-euco-final-conclusions.pdf [https://perma.cc/A2B6-RW9C].

109. *Id.*

110. *G7 Taormina Statement on the Fight Against Terrorism and Violent Extremism*, G7 (May 26, 2017), http://www.consilium.europa.eu/en/press/press-releases/2017/05/26/statement-fight-against-terrorism/ [https://perma.cc/TE3W-NLVH].

111. Presidency Conclusions, *supra* note 108 ("This should be complemented by the relevant legislative measures at EU level, if necessary.").

112. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Tackling Illegal Content Online: Towards an enhanced responsibility of online platforms*, at 1, COM (2017) 555 final (Sept. 28, 2017) [hereinafter *Communication*], https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017DC0555 [https://perma.cc/96FK-C8F6].

113. Nick Aries, *Is Europe really moving away from protecting Online platforms*, IPWATCHDOG (Oct. 11, 2017), http://www.ipwatchdog.com/2017/10/11/europe-moving-away-protecting-online-platforms/id=88737/.

114. *Communication*, *supra* note 112, at 10 (emphasis added).

115. *Id.* at 8.

116. *Id.*

The Facebook representative spoke in support of automated means of deleting content, writing that "[t]oday, 99% of the ISIS and Al Qaeda-related terror content we remove from Facebook is content we detect before anyone in our community has flagged it to us, and in some cases, before it goes live on the site."[117] But European Commission representatives suggested that "our favoured cooperative approach with the industry" may not go "far enough and fast enough," and that legislation might be necessary.[118]

In March 2018, the European Commission released a recommendation on "measures to effectively tackle illegal content online."[119] The recommendation called on platforms to provide "fast-track procedures" to take down content referred by "competent authorities," "internet referral units," and "trusted flaggers," whether the content was illegal or was a violation of the platform's ToS.[120] The Commission also called on platforms to take "proportionate and specific proactive measures, including by using automated means," to find, remove, and prevent the reposting of terrorist content.[121] Finally, the recommendation called on government institutions to "encourage" platforms to "cooperate" in sharing technological tools to curb terrorist content.[122]

In the summer of 2018, the Commission announced that it would take "stronger action" by drafting a new regulation on preventing the dissemination of terrorist content online.[123] The Commission intended the draft regulation, released in September 2018, to strike an appropriate, "proportional" balance with respect to fundamental rights of free expression.[124] Nonetheless, the draft regulation adopts a relatively broad definition of "terrorist content," including not only direct incitement but also "glorifying" terrorist crimes or "promoting the activities" of terrorist groups.[125]

The draft regulation also imposes a variety of new obligations on platforms that may raise concerns about accountability. As an initial matter, the regulation requires platforms to remove or disable access to "terrorist content" within one hour.[126] Although the one-hour rule has gained sig-

---

117. European Commission Press Release IP/17/5105, Fighting Terrorism Online: Internet Forum pushes for automatic detection of terrorist propaganda (Dec. 6, 2017) (statement of Monika Bickert, Director of Global Policy Management, Facebook), http://europa.eu/rapid/press-release_IP-17-5105_en.htm [https://perma.cc/3GFB-G4CE].

118. *Id.* (statement of Julian King, Comm'r for the Security Union).

119. *Commission Recommendation on Measures to Effectively Tackle Illegal Content Online*, at ¶ 1, COM (2018) 1177 final (Mar. 1, 2018).

120. *Id.* at ¶¶ 23, 25, 32.

121. *Id.* at ¶¶ 36–37.

122. *Id.* at ¶ 38.

123. *Social media faces EU fine if terror lingers for an hour*, BBC (Aug. 20, 2018), https://www.bbc.com/news/technology-45247169 [https://perma.cc/7QXF-SD4F].

124. *Proposal for a Regulation of the European Parliament and of the Council on Preventing the Dissemination of Terrorist Content Online*, COM (2018) 640 final (Sept. 12, 2018) [hereinafter *Draft Terrorism Regulation*], https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-preventing-terrorist-content-online-regulation-640_en.pdf [https://perma.cc/JW9W-XRLH].

125. *Id.* art. 2.

126. *Id.* art. 4.

nificant media attention, other provisions are equally significant.[127] Article 6 of the draft regulation requires service providers to "take proactive measures" to prevent the dissemination of terrorist content.[128] These measures may include automated removal of content, automated prevention of re-uploading, or "detecting, identifying and expeditiously removing" new terrorist content.[129] If a "competent authority" of a member state finds that the platform's proactive measures do not suffice, it may impose "specific additional necessary and proportionate proactive measures" upon the platform.[130] The draft regulation instructs that these measures should account for the "economic capacity" of the platform and the effects of the measures on free expression and the free flow of information.[131] Nonetheless, it provides no information about how this balance ought to be struck. Moreover, the draft regulation provides no mechanism for platforms to appeal decisions that would require them to adopt new technological means of deleting or preventing the publication of speech.

In defense of the draft regulation, Julian King, the EU Security Commissioner, has observed that "[e]very attack over the last 18 months or two years or so has got an online dimension. Either inciting or in some cases instructing, providing instruction, or glorifying."[132] At least in theory, the information targeted by the draft regulation is equally illegal offline as it is online.[133]

At the same time, the draft regulation makes a sincere effort to bring accountability and transparency to platforms' practices. For example, recognizing that automated removal might present serious risks to free expression, the draft regulation also requires that platforms provide "effective and appropriate safeguards"—including "human oversight and verifications *where appropriate*"—to ensure that automated takedowns are "well-founded."[134] Article 8 requires platforms to publish transparency reports that include information about their measures to take down and prevent the re-upload of terrorist content.[135] The regulation also requires platforms to notify individuals whose content is deleted and to create "effective and accessible" complaint mechanisms so that users

127. *See, e.g.*, Natasha Lomas, *Europe to push for one-hour takedown law for terrorist content*, TechCrunch (Sept. 12, 2018), https://techcrunch.com/2018/09/12/europe-to-push-for-one-hour-takedown-law-for-terrorist-content/ [https://perma.cc/UAD4-8MVP]; Daniel Boffey, *Remove terror content or be fined millions, EU tells social media firms*, Guardian (Sept. 13, 2018), https://www.theguardian.com/media/2018/sep/13/social-media-firms-could-face-huge-fines-over-terrorist-content [https://perma.cc/2SQY-VZT2].

128. *Draft Terrorism Regulation*, *supra* note 124, art. 6.

129. *Id.*

130. *Id.*

131. *Id.*

132. Boffey, *supra* note 127.

133. *See* Patrick Smyth, *Crackdown on online terror content not censorship, says EU*, Irish Times (Sept. 13, 2018), https://www.irishtimes.com/news/world/europe/crackdown-on-online-terror-content-not-censorship-says-eu-1.3628352 [https://perma.cc/J2TA-WNSF].

134. *Draft Terrorism Regulation*, *supra* note 124, art. 9 (emphasis added).

135. *Id.* art. 8.

may appeal those decisions.[136]

While the Commission has taken an aggressive tack, the Council of Europe (CoE)—the international organization dedicated to promoting and protecting human rights in Europe—has been more moderate in its outlook. In 2008, the CoE issued the Human Rights guidelines for Internet service providers, in which the CoE recognized that platforms' adjudication "decisions and actions with regard to the accessibility of services" affected human rights and expressive freedoms.[137] The CoE guidelines instructed platforms that they should filter, block, or delete content "only after a verification of the illegality of the content;" deleting content without cause could violate rights of free expression and access to information.

The forum and the hash sharing arrangement reflect several strains in apparently voluntary efforts to curb hateful and violent content online. First, each reflects extensive cooperation among industry, civil society organizations, and government, raising questions about whether the decision-making on online takedowns is truly private or might be more accurately described as "co-regulatory."[138] This collaboration between government and industry is widespread, especially when it comes to law enforcement surveillance.[139] The use of special law enforcement units to track and request deletion of content is not unique to Europe—a recent article noted that Vietnam had developed a special Internet monitoring unit with over 10,000 workers, "tasked with fighting 'wrongful views.'"[140]

Second, both the forum and the hash sharing reflect increasing pressures on platforms to act "proactively" to develop new tools that will facilitate automatic content deletion. Despite the tension with longstanding intermediary immunity principles that imposed no requirements for platforms to monitor content, the European Commission has determined that platforms "should remove illegal content as fast as possible,"[141] and that doing so is somehow consistent with the E-Commerce Directive's immu-

136. *Id.* arts. 10, 11. Notably, however, the regulation provides that individuals shall *not* be notified of content deletion if a "competent authority decides that there should be no disclosure" because, for example, it is relevant to an investigation of terrorist offenses. *Id.* art. 11. For reasons discussed *infra*, this exception will often apply.

137. *Human Rights Guidelines for Internet Service Providers*, COUNCIL EUR. ¶ 6 (2008), https://rm.coe.int/16805a39d5 [https://perma.cc/XTC3-VC74].

138. Hui Zhen Gan, Note, *Corporations: The Regulated or the Regulators? The Role of IT Companies in Tackling Online Hate Speech in the EU*, 24 COLUM. J. EUR. L. 111, 113, 117 (2017); *see also* Fabrizio Cafaggi et al., *Private Regulation and Freedom of Expression*, 11 INT'L J. COMM. 1998, 2002 (2017) ("[T]he degree of state involvement determines whether private regulation becomes classifiable as co-regulation. When this is so, the state may also have direct responsibility to verify that the private regime complies with the principles enshrined in the ECHR.").

139. Hannah Bloch-Wehba, *Process Without Procedure: National Security Letters and First Amendment Rights*, 49 SUFFOLK U. L. REV. 367, 368 (2016).

140. Glyn Moody, *Revealed: Vietnam's 10,000-Strong Internet Monitoring Force, Tasked With Stamping Out 'Wrongful Views'*, TECHDIRT (Jan. 3, 2018), http://techdirt.com/articles/20180102/03180438906/revealed-vietnams-10000-strong-internet-monitoring-force-tasked-withstamping-out-wrongful-views.shtml [https://perma.cc/6DUH-ZD9W].

141. *Communication*, *supra* note 112, at 13.

nity provision.[142] This recommendation is also in tension with the Council of Europe's instruction that "[s]tate authorities should not directly or indirectly impose a general obligation on intermediaries to monitor content to which they give access, or which they transmit or store, be it by automated means or not."[143]

Finally, both the hate speech and terrorist speech arrangements suffer from serious transparency deficits. In the context of the hash sharing arrangement, it is unclear whether content in the hash database is shared with law enforcement, whether users are notified that content is in the hash database, or whether there are any penalties for users who post or repost hashed content. Moreover, it remains unclear whether counter-notice or an appeals process is available.

In the context of the Code of Conduct, while the communication urges platforms to provide users the ability to contest wrongful removal decisions through counter-notice, it explicitly notes that counter-notice would be inappropriate in criminal cases—which include, in many European jurisdictions, hate speech. Finally, the code encourages platforms to use "out-of-court dispute settlement bodies to resolve disputes" about content removal, but many of those bodies reflect a strong presumption that proceedings and awards shall not be made public.[144]

B. The Right to Be Forgotten

A second context in which platforms are playing a major governance role in adjudicating disputes about free expression and privacy online has arisen in the context of European data protection law. In May, 2014, the European Court of Justice (ECJ) recognized the "right to be forgotten," an individual right under the Data Protection Directive and the Charter of Fundamental Rights of the European Union.[145] In *Google Spain*, Mario Costeja Gonzalez brought suit against La Vanguardia, Google Inc., and Google Spain, seeking an order that would result in the deletion of

142. *See Draft Terrorism Regulation*, *supra* note 124, § 1.2 ("The present proposal is consistent with the acquis related to the Digital Single Market and in particular the E-Commerce Directive.").

143. *Recommendation CM/Rec(2017x)xx of the Committee of Ministers to Member States on the Roles and Responsibilities of Internet Intermediaries*, Council Eur. ¶ 1.3.3 (2016) [hereinafter *Council of Europe Draft Recommendation*], https://rm.coe.int/recommendation-cm-rec-2017x-xx-of-the-committee-of-ministers-to-member/1680731980 [https://perma.cc/3U55-S5WV].

144. *See, e.g.*, *International Dispute Resolution Procedures*, Int'l Ctr. Disp. Resol. arts. 30, 37 (June 1, 2014), https://www.icdr.org/sites/default/files/document_repository/ICDR_Rules.pdf [https://perma.cc/JMM8-L8MC] (Article 30: "An award may be made public only with the consent of all parties or as required by law") (Article 37: "Except as provided in Article 30, unless otherwise agreed by the parties or required by applicable law, the members of the arbitral tribunal and the Administrator shall keep confidential all matters relating to the arbitration or the award."); *see also EU Privacy Shield Procedure Rules*, Better Bus. Bureau, https://www.bbb.org/EU-privacy-shield/rules-and-policies/ [https://perma.cc/WFG8-SUBA] (Rule 8.2: "All deliberations, meetings, proceedings and writings of the Procedure other than the Settlement Agreement or the Decision shall be sealed from public access and shall be treated as confidential by CBBB.").

145. *Google Spain*, *supra* note 8, ¶ 96.

search results concerning 1998 "attachment proceedings for the recovery of social security debts."[146] Costeja Gonzalez argued that the search results were totally irrelevant—the attachment proceedings had long been resolved. Under Article 6 of the 1995 Data Protection Directive, moreover, data controllers were required to ensure that

> personal data are processed "fairly and lawfully," that they are "collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes," that they are "adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed," that they are "accurate and, where necessary, kept up to date" and, finally, that they are "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed."[147]

The ECJ found that, even where the processing of data was "initially lawful," it may, over time,

> become incompatible with the directive where those data are no longer necessary in the light of the purposes for which they were collected or processed. That is so in particular where they appear to be inadequate, irrelevant or no longer relevant, or excessive in relation to those purposes and in the light of the time that has elapsed.[148]

In other words, Costeja Gonzalez had a "right to be forgotten"—a right he could assert against Google to force the company to take down search results that were inadequate, irrelevant, or excessive.

While Costeja Gonzalez could prevail against Google, however, the Spanish data protection authority had dismissed his claim against La Vanguardia. Under the Data Protection Directive, EU Member States had to provide "exemptions or derogations" from data protection mandates for the "processing of personal data carried out solely for journalistic purposes or the purpose of artistic or literary expression"—albeit only in situations in which the right to privacy conflicted with free expression.[149] Accordingly, the ECJ distinguished between the "activity of a search engine" and "publishers of websites,"[150] and determined that search engines could not benefit from these exemptions.[151] For that reason, the court held that Costeja Gonzalez need not successfully obtain relief from La Vanguardia before approaching Google.[152]

Google Spain suggested a partial list of factors that search engines would have to weigh in considering right to be forgotten requests. As-

146. *Id.* ¶ 14.
147. *Id.* ¶ 72.
148. *Id.* ¶ 93.
149. Council Directive 95/46, art. 9, 1995 O.J. (L 281) 31, 43 (EC).
150. *Google Spain*, *supra* note 8, ¶ 35.
151. *Id.* ¶ 85.
152. *Id.* ("It cannot therefore be ruled out that in certain circumstances the data subject is capable of exercising the rights referred to in Article 12(b) and subparagraph (a) of the first paragraph of Article 14 of Directive 95/46 against that operator but not against the publisher of the web page.").

suming a data subject established that the information to be deleted was "inadequate, irrelevant or no longer relevant, or excessive in relation to the purposes of the processing at issue carried out by the operator of the search engine,"[153] the court suggested that, as a rule, those rights would "override" the public interest in the information.[154] But the ECJ also anticipated that search engines would have to undertake a careful factual balancing that would

> depend, in specific cases, on the nature of the information in question and its sensitivity for the data subject's private life and on the interest of the public in having that information, an interest which may vary, in particular, according to the role played by the data subject in public life.[155]

The court insinuated that there might be other reasons for a heightened public interest in the information, but did not enumerate them.[156]

As Edward Lee has put it, the *Google Spain* opinion was largely silent on "how to operationalize or put into practice, in the EU, a procedure and a set of criteria for determining claims invoking the right to be forgotten in search engine results."[157] Instead, the burden fell on Google to do so itself. Google convened an "Advisory Council" to assist it in formulating "criteria that Google should use in striking a balance, such as what role the data subject plays in public life, or whether the information is outdated or no longer relevant."[158] In addition, the Article 29 Working Party published guidelines and criteria for implementing the *Google Spain* ruling.[159]

By July 2014, Google reported that it had received more than 70,000 take-down demands pertaining to 250,000 websites.[160] The company complained of the "difficult value judgments" it was having to make about content including negative reviews, coverage of past crimes, and criticism of politicians and policy choices: "in each case, someone wants the information hidden, while others might argue it should be out in the open."[161] Google listed some of the criteria it used to weigh the deletion demands, including whether

---

153. *Id.* ¶ 94.
154. *Id.* ¶ 97.
155. *Id.* ¶ 81.
156. *Id.* ("such as").
157. Edward Lee, *Recognizing Rights in Real Time: The Role of Google in the EU Right to Be Forgotten*, 49 U.C. DAVIS L. REV. 1017, 1035 (2016).
158. *The Advisory Council to Google on the Right to be Forgotten*, GOOGLE (Feb. 6, 2015), https://static.googleusercontent.com/media/archive.google.com/en//advisorycouncil/advisement/advisory-report.pdf [https://perma.cc/EL4X-HH4Q].
159. Article 29 Data Prot. Working Party, *Guidelines on the Implementation of the Court of Justice of the European Union Judgment on "Google Spain and Inc v. Agencia Espanola de Proteccion de Datos (AEPD) and Mario Costeja Gonzalez" C-131/12* (Nov. 26, 2014) [hereinafter *Guidelines on the Implementation of Google Spain*], https://www.pdpjournals.com/docs/88502.pdf [https://perma.cc/MLM5-YE5Q].
160. *Searching for the right balance*, GOOGLE (July 11, 2014), https://googleblog.blogspot.be/2014/07/searching-for-right-balance.html [https://perma.cc/YQG8-JB4M].
161. *Id.*

> the information relates to a politician, celebrity, or other public figure; if the material comes from a reputable news source, and how recent it is; whether it involves political speech; questions of professional conduct that might be relevant to consumers; the involvement of criminal convictions that are not yet "spent"; and if the information is being published by a government.[162]

The company also noted that, while it strived to be transparent about the takedowns, it struggled to do so in a way that would not further infringe on the privacy right of the data subject.[163] Indeed, Google's decision-making on the right to be forgotten raised hackles almost immediately. In 2015, the BBC published a list of web pages that Google had delisted from search results.[164] The network reasoned that "the integrity of the BBC's online archive is important and, although the pages concerned remain published on BBC Online, removal from Google searches makes parts of that archive harder to find."[165] Julia Powles critiqued the republication, arguing that it "only accentuates the demand for data protection rights."[166]

In addition to deciding whether a user's request to delete a link was valid, the company also had to determine whether the link should be deleted in the user's home country, in Europe, or worldwide. In May 2015, the CNIL, the French data protection authority, issued an order requiring Google to delist links across *all* of its geographic extensions, including .com, and not only across .fr, .eu, and other European domains.[167] Google refused to comply, but agreed to restrict access to the deleted content within Europe by using geographic indicators, including IP addresses, to limit any access to the delisted links.[168] In March 2016, the Restricted Committee of the CNIL rebuffed this proposal, reasoning that only global deletion, "regardless of the extension used or the geographic

---

162. *Id.*

163. *Id.* In 2017, the Higher Regional Court of Munich enjoined Google from forwarding takedown notices to the Lumen database, which documents efforts to remove online content. Kieren McCarthy, *When we said don't link to the article, Google, we meant DON'T LINK TO THE ARTICLE!*, REGISTER (June 15, 2017), https://www.theregister.co.uk/2017/06/15/google_germany_right_to_be_forgotten_court_case/ [https://perma.cc/JZ8P-BSFA].

164. Neil McIntosh, *List of BBC web pages which have been removed from Google's search results*, BBC (June 25, 2015), www.bbc.co.uk/blogs/internet/entries/1d765aa8-600b-4f32-b110-d02fbf7fd379 [https://perma.cc/8DSC-AJTT].

165. *Id.*

166. Julia Powles, *Why the BBC is wrong to republish 'right to be forgotten' links*, GUARDIAN (July 1, 2015), https://www.theguardian.com/technology/2015/jul/01/bbc-wrong-right-to-be-forgotten.

167. *Right to be delisted: the CNIL Restricted Committee imposes a _100,000 fine on Google*, CNIL (Mar. 24 2016) [hereinafter CNIL, *Right to be Delisted*], https://www.cnil.fr/en/right-be-delisted-cnil-restricted-committee-imposes-eu100000-fine-google [https://perma.cc/3F4U-LWQN].

168. *Id.*; Natasha Lomas, *Google's right to be forgotten appeal heading to Europe's top court*, TECHCRUNCH (July 19, 2017), https://beta.techcrunch.com/2017/07/19/googles-right-to-be-forgotten-appeal-heading-to-europes-top-court/ [https://perma.cc/T9D4-7DCE] ("We're doing this because we want to ensure that people have access to content that is legal in their country.") (statement of Peter Fleischer, Global Privacy Counsel, Google).

origin of the person performing the search," could effectively protect privacy and data protection rights.[169] The CNIL reasoned that all of Google's search engines constituted a unified "data processing" operation, rendering each and every top level domain subject to French data protection law.[170] Google appealed the decision to the Conseil d'État, which referred the question to the European Court of Justice; the case is still pending.

Google's role in adjudicating right to be forgotten requests ruffled some feathers. In August 2014, Johannes Masing, a judge on Germany's Federal Constitutional Court, wrote an essay predicting that the *Google Spain* judgment would elevate search engine providers to "private arbitrators with far-reaching decision-making authority" over online communication.[171] As Edward Lee has pointed out, it is easy to "envision a different procedure," in which domestic data protection authorities—not Google—would serve as the gateway to adjudicate right to be forgotten requests, "[b]ut that's not what happened."[172]

In 2018, Google published a significant report, *Three Years of the Right to Be Forgotten*, detailing how the company has responded to the hundreds of thousands of removal requests it has received.[173] "In broad terms," the company relies on four criteria to balance the public interest in the information against the requester's privacy interest: the "validity" of the request, the requester's identity, the content of the information requested to be delisted, and the source of the content.[174] The company identified two main categories of requests for delisting: content that revealed "personal information" on social media or directory websites, and content that revealed "legal history and professional information," often on news websites.[175]

Like the Code of Conduct, the implementation of the right to be forgotten raises a number of concerns about the absence of evenhanded mechanisms for affected parties to vindicate their rights. Under *Google Spain* and the Article 29 Working Party guidelines, only right to be forgotten *requesters* have standing to contest Google's decisions not to take down content. In other words, nobody has standing to challenge a decision by Google to *delete* content.

---

169. CNIL, *Right to be Delisted*, *supra* note 167.

170. *Id.*

171. Johannes Masing, *RiBVerfG Masing: Vorläufige Einschätzung der, Google-Entscheidung" des EuGH*, VERFASSUNGSBLOG (Aug. 14, 2014), https://verfassungsblog.de/ribverfg-masing-vorlaeufige-einschaetzung-der-google-entscheidung-des-eugh/ [https://perma.cc/CES9-DLKH] ("Durch die Entscheidung des EuGH werden Suchmaschinenbetreiber als für Löschungsanträge Verantwortliche zu einer privaten Schiedsinstanz mit weitreichenden Entscheidungsbefugnissen über die Kommunikation im Netz erhoben.").

172. Lee, *supra* note 157, at 1036.

173. THEO BERTRAM ET AL., GOOGLE, THREE YEARS OF THE RIGHT TO BE FORGOTTEN (Sept. 7, 2018), https://drive.google.com/file/d/1H4MKNwf5MgeztG7OnJRnl3ym3gIT3HUK/view.

174. *Id.* at 2–3.

175. *Id.* at 15.

In fact, although the European Court of Justice concluded in the *Telekabel* case that users have standing to contest over-removal that results from judicial action, there is no clear path to do so when the removal appears to result from private action.[176] Although the Article 29 Working Party urged search engines to balance publishers' rights against the data protection interests of right to be forgotten requesters,[177] only the latter have the legal ability to enforce their rights. Indeed, the Article 29 Working Party's guidelines concluded that search engines should avoid notifying web publishers about requests to de-list links to their content, reasoning that publishers "have no control or influence" over data protection decisions and that "search engines do not recognize a legal right of publishers to have their contents indexed and displayed, or displayed in a particular order."[178]

This lopsidedness chafes against free expression norms and values recognized in Europe and beyond. As the EU Human Rights Guidelines suggest, "Any restriction that prevents the flow of information offline or online must be in line with permissible limitations as set out in international human rights law."[179] Successful efforts to extend the right to be forgotten beyond search engines to newspapers underscore the dangers to free expression. In 2016, the Belgian Court of Cassation upheld a decision requiring a Belgian newspaper to anonymize the online version of an article regarding an individual's involvement in a fatal car accident in 1994.[180] The Belgian court determined that anonymizing the newspaper's archived article was less burdensome to free expression than de-indexing or unpublishing the article. In the summer of 2018, the Spanish Constitutional Court required the newspaper *El País* to de-index an article concerning allegations of drug trafficking, concluding that anonymization would pose a "greater interference" with free expression than de-indexing.[181] It is hard to imagine that limitations left to the discretion of platforms and enforceable only by those who request deletion could comport with human rights law's requirements.[182]

---

176. *See* Keller, *The Right Tools*, *supra* note 81, at 347.

177. *Guidelines on the Implementation of Google Spain*, *supra* note 159, at 10 ("In any case, that [publisher's] interest should be balanced with the rights, freedoms and interests of the affected data subject.").

178. *Id.*

179. *EU Human Rights Guidelines*, *supra* note 48.

180. *Belgian Court of Cassation Rules on Right to Be Forgotten*, PRIVACY & INFO. SECURITY L. BLOG, HUNTON ANDREWS KURTH (June 1, 2016), https://www.huntonprivacyblog.com/2016/06/01/belgian-court-of-cassation-rules-on-right-to-be-forgotten/ [https://perma.cc/Q2BE-ECZG].

181. Raul Rubio et al., *Spain — Landmark decision regarding the implementation of the "right to be forgotten"*, LEXOLOGY (Oct. 4, 2018), https://www.lexology.com/library/detail.aspx?g=9f7db79b-d75c-4a7f-b6fe-91825ba27392 [https://perma.cc/W9N4-8S82].

182. *See* CHRISTINA ANGELOPOULOS ET AL., STUDY OF FUNDAMENTAL RIGHTS LIMITATIONS FOR ONLINE ENFORCEMENT THROUGH SELF-REGULATION 14 (2016) ("The European Court of Human Rights has developed a standard test to determine whether Article 10, ECHR, has been violated. Put simply, whenever it has been established that there has been an interference with the right to freedom of expression, that interference must first of all be prescribed by law. In other words, it must be adequately accessible and reasonably

IV. CRITIQUES OF PLATFORM GOVERNANCE

Each of the foregoing examples of the regimes surrounding online content implicates user privacy, expressive freedoms, the free flow of information, and countervailing interests. Platforms play varying roles in governance in each of these settings, and each reflects a different approach to adopting public law norms or including public institutions within these governance arrangements.

In the context of the effort to tamp down hate speech and terrorist speech, for example, takedown requests come from across the spectrum of private individuals, civil society organizations, and state institutions such as internet referral units. By participating in the voluntary "hash database," platforms proactively attempt to identify terrorist content for deletion themselves, without the prodding of outside actors. In this context, speakers' interests are underrepresented in comparison with the government's interest in monitoring, suppressing, and counteracting illegal speech online. While the European Commission's Communication on Tackling Illegal Content Online instructs that "[r]obust safeguards to limit the risk of removal of legal content also should be available, supported by a set of meaningful transparency obligations to increase accountability of the removal processes," the shape of these safeguards is far from clear.[183] This directive raises further questions about whether safeguards intended to protect free expression have been developed entirely by private entities,[184] or with the input or oversight of government actors.

Of course, it is also anyone's guess as to whether these mechanisms are effective to protect free expression.[185]

These kinds of questions have given rise to two major critiques of the global pressures on online platforms to govern speech. At the risk of oversimplifying, I cast these critiques as (1) substantive and (2) procedural. While both critiques have some merit, they are also both partially wrong—and incomplete.

---

foreseeable in its consequences. Second, it must pursue a legitimate aim (i.e., correspond to one of the aims set out in Article 10(2)). Third, it must be necessary in a democratic society, i.e., it must correspond to a 'pressing social need', and it must be proportionate to the legitimate aim(s) pursued."); *see also* Aleksandra Kuczerawy, *The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression*, 8 J. INTELL. PROP. INFO. TECHNOLOGY & E-COM. L. 226, 233 (2017) (observing that, under the European Convention on Human Rights, freedom of expression has positive as well as negative dimensions).

183. *Communication*, *supra* note 112, ¶ 4; *see also* Daphne Keller, *Counter-Notice Does Not Fix Over-Removal of Online Speech*, STAN. CTR. FOR INTERNET & SOC'Y (Oct. 5, 2017) [hereinafter Keller, *Counter-Notice*], http://cyberlaw.stanford.edu/blog/2017/10/counter-notice-does-not-fix-over-removal-online-speech [https://perma.cc/74J5-UZ3D].

184. *See generally* Klonick, *supra* note 14.

185. *Communication*, *supra* note 112, ¶ 4; *see also* Keller, *Counter-Notice*, *supra* note 183.

A. Substantive Critiques: Platforms are Achieving the Wrong Balance

At its core, the substantive critique complains that when platforms weigh speech rights against other rights, they strike the wrong balance. This critique has emerged particularly sharply in the context of the right to be forgotten. It is difficult, if not impossible, for American observers to square the right to be forgotten with the First Amendment's strong protections for publishing truthful information.[186] As some media law practitioners have noted, European privacy law has a different perspective on what is newsworthy than the First Amendment.[187] Some are concerned that search engines and others will adopt a default presumption in favor of deletion in order to avoid financial penalties, thus chilling even lawful speech.[188] This concern is especially pronounced in light of Google's finding that a significant portion of links requested to be delisted are from news websites.[189]

In essence, many of these critics simply object on normative grounds to the European approach, which appears to prioritize individual privacy rights equally highly or perhaps even above press freedoms and the free flow of information. This balance simply would not pass muster under U.S. constitutional law, which privileges First Amendment freedoms above privacy rights.[190] These concerns have been exacerbated by the ongoing battles over whether the right to be forgotten should extend abroad. U.S.-based civil society organizations have explicitly rejected the idea that the right to be forgotten could apply globally because of concerns about its impact on First Amendment rights. In response to the CNIL's decision that Google was required to delist links across all of its domains in order to vindicate the right to be forgotten, the Reporters Committee for Freedom of the Press, a U.S.-based association of report-

186. Jasmine E. McNealy, *The Emerging Conflict Between Newsworthiness and the Right to Be Forgotten*, 39 N. Ky. L. Rev. 119, 128 (2012) (concluding that various approaches to defining "newsworthiness" "seem to offer great First Amendment protection for the publication of truthful, yet private, information").

187. Ashley Messenger, *What Would a "Right to Be Forgotten" Mean for Media in the United States*, 29 Comm. Law. 29, 35 (2012); *see also* Kurt Wimmer, *The Long Arm of the European Privacy Regulator: Does the New EU GDPR Reach U.S. Media Companies?*, 33 Comm. Law. 16, 18 (2017) (arguing that a successful challenge to the application of the GDPR to a U.S. media organization would require showing that the GDPR conflicts with the First Amendment, "and that the publisher's free expression interests outweigh the European Union's interest in safeguarding its citizens' privacy rights"); Charles D. Tobin & Christine N. Walz, *Right to Be Forgotten, Expungement Laws Raise New Challenges on the 40th Anniversary of* Cox Broadcasting v. Cohn, 31 Comm. Law. 4, 4 (2015) (contending that the right to be forgotten threatens the "black-and-white" protection for reporting of truthful facts).

188. Jeffrey Rosen, *The Right to Be Forgotten*, 64 Stan. L. Rev. Online 88, 90–91 (2012).

189. Bertram et al., *supra* note 173, at 15.

190. *See generally* Eugene Volokh, *Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People From Speaking About You*, 52 Stan. L. Rev. 1049 (2000); Sorrell v. IMS Health Inc., 564 U.S. 552, 580 (2011) (striking down Vermont's regulation of the sale and use of prescriber-identifiable information for drug marketing purposes on First Amendment grounds).

ers and editors, wrote a letter expressing concern about exporting censorship, arguing that "[i]nternational free expression cannot survive on the Internet if every nation's laws apply to every website."[191] The Electronic Frontier Foundation (EFF) published a report that stated in part that the right to be forgotten "fundamentally contradicts U.S. laws and rights, including those protected by the First Amendment."[192]

Although the right to be forgotten has drawn particularly potent objections—perhaps because its balance between the public interest in the free flow of information and the private interest in delisting explicitly subverts the balance under existing First Amendment jurisprudence—platforms have also been accused of censoring user speech in other contexts. Numerous civil society organizations in the United States as well as in Europe have opposed the European Commission's Code of Conduct on the basis that it is over broad and may chill lawful expression.[193] As a general matter, it is well recognized that "broad secondary liability for intermediaries may also chill speech or stop legal conduct as a result of over-enforcement."[194] Jennifer Urban and others have documented how Section 512 of the Digital Millennium Copyright Act has resulted in over-enforcement of copyright claims and under-enforcement of speech protections.[195]

---

191. Letter from The Reporters Comm. for Freedom of the Press to Isabelle Falque-Pierrotin, President, Commission nationale de l'informatique et des libertés (Sep. 14, 2015), https://www.rcfp.org/sites/default/files/RCFP_CNIL_Sept14-English.pdf [https://perma.cc/L4UB-L5VB].

192. David Greene et al., *Rights at Odds: Europe's Right to Be Forgotten Clashes with U.S. Law*, ELEC. FRONTIER FOUND. (Nov. 2016), https://www.eff.org/files/2016/11/29/rtbf-us_law_legal_background.pdf [https://perma.cc/UB53-S7EE].

193. *See, e.g.*, Jillian C. York, *European Commission's Hate Speech Deal With Companies Will Chill Speech*, ELEC. FRONTIER FOUND. (June 3, 2016), https://www.eff.org/deeplinks/2016/06/european-commissions-hate-speech-deal-companies-will-chill-speech [https://perma.cc/9KNV-PFYS]; ARTICLE 19, EU: EUROPEAN COMMISSION'S CODE OF CONDUCT FOR COUNTERING ILLEGAL HATE SPEECH ONLINE AND THE FRAMEWORK DECISION 15–16 (June 2016), https://www.article19.org/data/files/medialibrary/38430/EU-Code-of-conduct-analysis-FINAL.pdf [https://perma.cc/V7NX-FB5F] (expressing "serious concerns"); Joe McNamee, *Guide to the Code of Conduct on Hate Speech*, EUR. DIGITAL RIGHTS (June 3, 2016), https://edri.org/guide-code-conduct-hate-speech/ [https://perma.cc/VFX3-7DUV].

194. Graeme B. Dinwoodie, *A Comparative Analysis of the Secondary Liability of Online Service Providers*, *in* 25 IUS COMPARATUM – GLOBAL STUDIES IN COMPARATIVE LAW 18 (2017); *see also Council of Europe Draft Recommendation*, *supra* note 143, ¶ 1.3.3 ("The imposition of sanctions for non-compliance may prompt over-regulation and speedy takedown of all dubious content, which may result in an overall chilling effect for the freedom of expression online.").

195. *See, e.g.*, Jennifer M. Urban & Laura Quilter, *Efficient Process or "Chilling Effects"? Takedown Notices under Section 512 of the Digital Millennium Copyright Act*, 22 SANTA CLARA COMPUTER & HIGH TECH. L.J. 621, 682 (2006) ("[T]he implications for expression on the Internet of this extrajudicial process appear, from our limited data, significant. Removal of speech from the Internet, with very little or no process, is a strong remedy for allegations of infringement, especially where there are so few recourses available to the targeted speaker."). *See also* Urban et al., *supra* note 68, at 10 ("[R]elying on machines to make decisions about sometimes-nuanced copyright law raises questions about the effect on expression."); Wendy Seltzer, *Free Speech Unmoored in Copyright's Safe Harbor: Chilling Effects of the DMCA on the First Amendment*, 24 HARV. J.L. & TECH. 171, 176 (2010).

These critiques of platform "censorship" of speech are in substantial tension with the concept of platforms as "editors" or "speakers" in their own right.[196] As Heather Whitney notes, in domestic litigation, "Google and others have argued that their decisions concerning their platforms—for example, what sites to list (or delist) and in what order, who can buy ads and where to place them, and what users to block or permanently ban—are analogous to the editorial decisions of publishers."[197] This account of platforms serving as "editors," curating content and making choices about what information is newsworthy, suggests—in an American perspective—that those choices are worthy of protection from government meddling.

To the extent platforms are truly functioning as editors, it can hardly be argued that their decisions about what content to display are illegitimate. Yet it is equally clear that when platforms are required or coerced to delete content in order to avoid legal liability, a different paradigm ought to apply. When platforms act as "proxy censors"[198] or "collateral censors,"[199] they are distinct from neutral or beneficent "curators" or "editors." Indeed, implicit in calls to require additional safeguards for content deletion is a rejection of the editorial analogy in the context of coercive censorship, and with it a rejection of the platforms' power over private speech.

B. Procedural Critiques: Platforms Shouldn't Make The Rules

Scholars and civil society organizations have also criticized the procedural deficits of global speech governance, arguing that private companies such as Google, Twitter, and Facebook are ill-suited to the task of making decisions about what kinds of speech are—and are not—protected.

In that vein, scholars and civil society actors have criticized the role that private companies are playing in adjudicating disputes and making rules about online speech.[200] For example, Eldar Haber has described the role of search engines in administering the right to be forgotten as akin to

196. *See, e.g.*, Heather Whitney, *Search Engines, Social Media, and the Editorial Analogy*, Knight First Amend. Inst. 9–10 (2018), https://knightcolumbia.org/sites/default/files/content/Heather_Whitney_Search_Engines_Editorial_Analogy.pdf [https://perma.cc/F7Y9-VKTF]; *see also* Klonick, *supra* note 14, at 1609 ("Depending on the type of intermediary involved, courts have analogized platforms to established doctrinal areas in First Amendment law—company towns, broadcasters, editors—and the rights and obligations of a platform shift depending on which analogy is applied.").

197. Whitney, *supra* note 196, at 3.

198. Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. Pa. L. Rev. 11, 14 (2006).

199. Balkin, *supra* note 59, at 2309.

200. *See On the "Right to Be Forgotten": Challenges and Suggested Changes to the Data Protection Regulation*, Ctr. for Democracy & Tech. (May 2, 2013), https://cdt.org/files/pdfs/CDT-Free-Expression-and-the-RTBF.pdf [https://perma.cc/F8RE-E2UB] ("Private companies are ill-equipped to take responsibility for decisions that balance the right to privacy with the right to free expression."); Aleksandra Kuczerawy & Jef Ausloos, *From*

a private judiciary, arguing that existing mechanisms are insufficient to render Google's decision-making accountable.[201] Kyle Langvardt has likewise argued that speech governance by online platforms is inappropriate because the companies are not chosen through a democratic process, but rather are "politically unaccountable technology oligarchs [who] exercise state-like censorship powers."[202] A number of international organizations have also addressed the proper role of private companies in the context of intermediary liability—the Organization for Security and Cooperation in Europe, for example, has recommended that "excessive and disproportionate" takedown obligations "create a clear risk of transferring regulation and adjudication of Internet freedom rights to private actors and should be avoided."[203]

Like the substantive critique of speech governance by platforms, the procedural critique dismisses the "editorial analogy" that compares platforms to newspapers in their own right.[204] Seeing platforms, instead, as (more or less) willing collaborators with or stand-ins for government censors—as arms of the state—critics have sought to apply the same standards to private platforms as to states themselves.[205]

Some new developments suggest that platforms have taken this charge seriously and are responding in ways intended to bolster their own legitimacy. For instance, when Google published its *Three Years of the Right to Be Forgotten* report, it explicitly recognized the vital importance of its own decision-making role, writing, "Understanding how we make these types of decisions—and how people are using new rights like those

---

Notice-and-Takedown to Notice-and-Delist: Implementing Google Spain, 14 COLO. TECH. L.J. 219, 232 (describing concerns with "search engines as arbitrators").

201. Eldar Haber, *Privatization of the Judiciary*, 40 SEATTLE U. L. REV. 115, 152–53 (2017).

202. Langvardt, *supra* note 14, at 1358; *see also* Lee, *supra* note 157, at 1066–67 ("[T]his delegation of power raises serious concerns for democratic accountability and due process.").

203. Press Release, Dunja Mijatoviæ, OSCE Rep. on Freedom of the Media, Org. for Security & Cooperation in Europe, Communiqué No.1/2016 (Jan. 29, 2016), https://www.osce.org/fom/219391?download=true; *see also Council of Europe Draft Recommendation*, *supra* note 143, ¶ 1.3.2 ("State authorities shall seek to obtain an order by a judicial authority or other fully independent and impartial state entity when demanding intermediaries to restrict access to unlawful content. State authorities should not require internet intermediaries to restrict access to third-party content based on their own assessment of its lawfulness. They should further ensure that internet intermediaries do not restrict access to third-party content based on their own assessment of its lawfulness without ensuring proper redress mechanisms and adherence to due process guarantees.").

204. Whitney, *supra* note 196, at 5.

205. *See, e.g.*, Plaintiff's Memorandum in Support of Motion for Preliminary Injunction at 14, Prager Univ. v. Google, No. 5:17-cv-06064-LHK, 2018 WL 1471939 (N.D. Cal. Dec. 29, 2017) ("Although Defendants are not public entities, they are 'state actors' for the purposes of both the First Amendment and Liberty of Speech claims because they operate and hold YouTube out to the public as the functional equivalent of a public forum, where the general public is invited and solicited to use the platform as place for 'freedom of expression' and 'freedom of information' where 'everyone's voice can be heard.'").

granted by the European Court—is important."[206] Facebook also plans to create an "independent body" to render "transparent and binding" decisions on content-related decisions, writing that independence would "create accountability and oversight."[207]

Platforms' recognition of their own governance role—and their steps to enhance their legitimacy and accountability—may not satisfy critics who also object to the extensive cooperation between platforms and governments on procedural grounds. For example, the Center for Democracy and Technology (CDT) has taken issue with law enforcement agencies' role as "trusted flaggers" that notify platforms of content that violates their ToS.[208] In the recommendation on measures to effectively tackle illegal content online, the European Commission applauded the use of Internet referral units—law enforcement units intended to ferret out unlawful Internet activities—in a "trusted flagger" role.

The use of public police agencies as enforcers of private companies' ToS exhibits at least two of the features of what Jack Balkin calls "new school" speech regulation.[209] First, the European Commission's Code of Conduct and the other "co-regulatory" initiatives to curb hate speech have taken place against the background of explicit threats to regulate and penalize online platforms for illicit content.[210] This is precisely what Balkin calls "collateral censorship" and what Seth Kreimer terms "proxy censorship": the state's technique of coercing private companies to censor speech that the government could not itself lawfully sanction.[211]

---

206. Michee Smith, *Updating our "right to be forgotten" Transparency Report*, GOOGLE (Feb. 26, 2018), https://www.blog.google/around-the-globe/google-europe/updating-our-right-be-forgotten-transparency-report/.

207. Mark Zuckerberg, *A Blueprint for Content Governance and Enforcement*, FACEBOOK (Nov. 15, 2018), https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/; *see also* Nick Clegg, *Charting a Course for an Oversight Board for Content Decisions*, FACEBOOK NEWSROOM (Jan. 28, 2019), https://newsroom.fb.com/news/2019/01/oversight-board/ (publicizing plan to "build a board that creates accountability and oversight of our content policy and enforcement decisions"); *Draft Charter: An Oversight Board for Content Decisions*, FACEBOOK NEWSROOM, https://fbnewsroomus.files.wordpress.com/2019/01/draft-charter-oversight-board-for-content-decisions-2.pdf (describing plans to create an independent "oversight board" for Facebook).

208. Emma Llansó, *Who Needs Courts? A Deeper Look At the European Commission's Plans to Speed Up Content Takedowns*, CTR. FOR DEMOCRACY & TECH. (Mar. 1, 2018), https://cdt.org/blog/who-needs-courts-a-deeper-look-at-the-european-commissions-plans-to-speed-up-content-takedowns/ [https://perma.cc/ZD7W-4CCE].

209. Balkin, *supra* note 59, at 2306.

210. *See, e.g.*, Gibbs, *supra* note 104; Jessica Elgot, *May and Macron plan joint crackdown on online terror*, GUARDIAN (June 12, 2017), https://www.theguardian.com/politics/2017/jun/12/may-macron-online-terror-radicalisation [https://perma.cc/34QN-CPHK] ("Theresa May will attempt to reassert control over the political agenda by agreeing a new counter-terror strategy with the French president, vowing to fine tech companies such as Facebook and Google if they do not step up efforts to combat online radicalisation.").

211. Balkin, *supra* note 59, at 2309 ("Collateral censorship occurs when the state holds one private party A liable for the speech of another private party B, and A has the power to block, censor, or otherwise control access to B's speech."); Kreimer, *supra* note 198, at 28 (noting that proxy censorship "provides a mechanism for the exercise of authority over otherwise ungovernable conduct").

The second aspect of government participation in enforcing platforms' ToS that has troubled scholars and civil society organizations is the blurring of distinctions between the power wielded by public agencies and that held by private companies. As Balkin notes, "Public/private cooperation and co-optation are hallmarks of new-school speech regulation."[212] Cooperation and co-optation are particularly concerning where they slide into "soft censorship" or "jawboning" by persuading or pressuring platforms to adopt government's favored limitations on speech, because they limit the ability of the public to hold the government accountable for those limitations.[213] Just so, in the context of the trusted flagger controversy, CDT objected that "government actors must be held accountable for the action they take to censor speech, even if another party is the ultimate implementer of the decision."[214]

The "procedural" critique of platform governance is thus twofold. First, the procedures that platforms themselves use to determine whether content should be deleted are insufficient to safeguard user rights. Some platforms, including Google and Facebook, appear to be attempting to solve this problem through more transparent and accountable mechanisms. But the ways in which platforms cooperate with governments—especially in informal arrangements—pose particular obstacles to procedural fairness and accountability. This challenge remains daunting even in light of the steps platforms are taking to enhance their own accountability.

C. Looking for Laws in All the Wrong Places

At bottom, both the procedural and substantive critiques of current modes of online speech governance are primarily concerned with democratic accountability and legitimacy in the context of decision-making by private actors. Procedural critics have decried the lack of oversight, the opacity, and the lack of effective notice or appeal mechanisms that characterize both the Code of Conduct and the right to be forgotten.[215] Substantive critics appear particularly concerned that foreign laws might unduly influence the information to which American Internet users have

212. Balkin, *supra* note 59, at 2324.

213. *See* Derek E. Bambauer, *Orwell's Armchair*, 79 U. Chi. L. Rev. 863, 905 (2012) ("Soft censorship is deeply problematic from the perspective of the process-oriented legitimacy methodology."); Derek E. Bambauer, *Against Jawboning*, 100 Minn. L. Rev. 51 (2015) (arguing that informal government pressure on platforms, or "jawboning," suffers from legitimacy deficits); *see also* Jon D. Michaels, *All the President's Spies*, 96 Calif L. Rev. 901, 923–24 (2008) (describing how "informal" relationships between the government and private sector in the intelligence context might facilitate avoidance of public scrutiny, oversight, and accountability).

214. Llansó, *supra* note 208.

215. Langvardt, *supra* note 14; *see also* Lee, *supra* note 157; Haber, *supra* note 201; Daphne Keller, *Notice and Takedown under the GDPR: An Operational Overview*, Stan. Ctr. for Internet & Soc'y (Oct. 29, 2015) [hereinafter Keller, *Notice and Takedown*], http://cyberlaw.stanford.edu/blog/2015/10/notice-and-takedown-under-gdpr-operational-overview#_ftnref11 [https://perma.cc/BL8F-D4WA].

access online.[216]

Some have thus called for additional government oversight of corporate speech governance. For instance, Kyle Langvardt has suggested that online content moderation should be subject to administrative and judicial review to ensure that it comports with First Amendment principles and values.[217] Edward Lee has called for the creation of a "hybrid agency" that would oversee right to be forgotten requests and takedowns in Europe and "bridge" the divide between private enterprise and government.[218] Eldar Haber has argued in favor of leaving the determination of right to be forgotten requests in the judicial system rather than in private adjudicators' hands.[219] But although these proposals appear motivated by the laudable goal of clarifying the respective roles of states and platforms in governance, that blurring of influence and power is in fact useful to the state, which may prefer "soft" methods of censorship to overt mechanisms precisely because they are less visible.[220]

In a sense, these concerns are similar to those that have animated an extensive body of scholarship examining legitimacy and accountability within the administrative state, particularly as government roles become increasingly privatized.[221] In the domestic context, these concerns have prompted scholars to explore whether, and when, private governance might constitute "state action" and thus be subject to constitutional limitations.[222] This effort reflects a growing sense that the blurred boundaries between public and private power raise significant questions for individual rights and liberties.[223] But, as Jody Freeman has pointed out, the state

---

216. *See, e.g.*, Jeffrey Rosen, *The Right to Be Forgotten*, ATLANTIC (July–Aug. 2012), https://www.theatlantic.com/magazine/archive/2012/07/the-right-to-be-forgotten/309044/ ("This would transform Facebook and Google from neutral platforms into global censors and would clash directly with the principle, embedded in U.S. free-speech law, that people can't be restricted from publishing embarrassing but truthful information.").

217. Langvardt, *supra* note 14, at 1377.

218. Lee, *supra* note 157, at 1083.

219. Haber, *supra* note 201, at 149–50, 172.

220. *See* Bambauer, *Orwell's Armchair*, *supra* note 213, at 930 ("Should the government censor the Net, however, it should do so directly—using legislation that is tailored to the problem, that incorporates safeguards informed by the history of prior restraint, and that creates a system that is open, transparent, narrow, and accountable. Hard censorship is superior to soft censorship in achieving legitimacy."); *see also* Michaels, *supra* note 213, at 924 (describing how "informal collaboration with minimal-to-no-reporting to Congress or the courts" allows intelligence agencies and contractors to benefit from secrecy).

221. *See, e.g.*, Alfred C. Aman, Jr., *Administrative Law for a New Century*, *in* THE PROVINCE OF ADMINISTRATIVE LAW 91–92 (Michael Taggart ed., 1997); David Mullan, *Administrative Law at the Margins*, *in* THE PROVINCE OF ADMINISTRATIVE LAW 135 (Michael Taggart ed., 1997); Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461, 515 (2003) ("A model fixated on accountability cannot adequately address the concerns for arbitrariness necessary for a truly legitimate administrative state."); Freeman, *supra* note 16, at 546.

222. *See, e.g.*, Freeman, *supra* note 16, at 576 ("[M]any scholars have argued that, in certain contexts, private actors ought to submit to oversight by agencies, courts, and the legislature, and to be constrained by the Constitution in the same manner as traditional public agencies are.").

223. Mullan, *supra* note 221, at 153 (suggesting that "in an era of deregulation, privatization, and corporatization," the number of "inherent state domestic roles" may be dwindling as the responsibilities of private actors grow).

action doctrine is not the only option for promoting accountability of private companies' decision-making processes. As Freeman demonstrates, private and soft law can also become a fruitful source of accountability-enhancing mechanisms by "importing" public law norms and enshrining accountability mechanisms in contract, relying on private standard-setting groups, and other soft-law tools designed to enhance accountability and legitimacy.[224] Imposing various post hoc safeguards may "render agencies indirectly accountable to the electorate," even in the absence of direct government participation or involvement.[225]

In seeking to boost public accountability through government oversight, critics of platform governance largely overlook the broad range of alternative methods that exist to hold private entities accountable for regulatory choices and determinations that affect individuals.[226] They also tend to underplay the dangers of promoting a robust role for government in determining the legality of online content. Classical free expression thinking suggests that private decision-making is to be preferred to government interference: government involvement transforms editorial selection into censorship. Indeed, protections for editorial choices are premised on the view that selectivity itself implicates free expression rights. In a U.S. framework, the principle that government interventions limiting speech must be justified stems from First Amendment doctrine; private actors, of course, can choose to speak (or not to) arbitrarily or for any reason whatsoever.

More troubling, neither the substantive nor the procedural critiques of platform governance pay sufficient heed to the global context in which platforms are operating or to jurisdictional conflicts regarding free expression and privacy values. Arguments that the First Amendment provides the appropriate benchmark wrongly assume that the U.S. domestic context is the most relevant one. For example, Langvardt's suggestion that First Amendment standards should guide platforms when they take down content that violates terms of service ignores the global reach of platforms' community standards and assumes that the U.S. Constitution is, and should be, the operative legal constraint on international businesses.[227] But falling back on the First Amendment as the appropriate legal standard essentially doubles down on American unilateralism online.

Far from offering an achievable solution to governments' increasingly overlapping and conflicting demands in the areas of speech and privacy governance, overreliance on U.S. legal standards replicates the worst features of American exceptionalism; it uncritically assumes not only that

224. Freeman, *supra* note 16, at 588–90.

225. *Id.* at 546.

226. An exception is Edward Lee, *supra* note 157, whose work is exceptionally thoughtful about the range of constraints to private power that might apply in the context of the right to be forgotten.

227. Langvardt, *supra* note 14, at 1385 (arguing for content moderation that "aligns . . . with the doctrine and the priorities of First Amendment law").

American law *does* govern, but also that it is normatively preferable and should supply the baseline standard for a de facto global regulation.[228] First Amendment doctrine is not value-neutral, and in the words of Yochai Benkler, the incorporation of one nation's values "into the technology of communication shared by many displaces those of other nations."[229] When critics assume that the imposition of First Amendment standards on global platform governance is value-neutral, they embrace a unilateral approach to speech and privacy on the global web.[230]

In essence, the harshest substantive and procedural critics of platform governance are preoccupied with the domestic effects of global platform governance. But because of the global reach of internal platform governance structures, ToS, and other rules, platform governance has become an intractable *global* problem not susceptible to an easy fix either by applying domestic substantive law or procedural protections.[231] Nor is a domestic focus sufficient to understand platforms' roles in governing speech and privacy online. Rather, a new vocabulary is necessary to understand the legitimacy and accountability problems that confront platform governance.

V. DEMOCRATIC ACCOUNTABILITY, LEGITIMACY, AND GLOBAL GOVERNANCE

Platform governance muddles state and private power. At a minimum, the examples of the Code of Conduct and the right to be forgotten illustrate that in numerous circumstances, platforms' decisions about online content are not theirs alone, but are informed—if not compelled—by government actors as well. In short, platform governance is characterized by "a deep interdependence among public and private actors in accomplishing the business of governance."[232] This mixed responsibility for decision-making on online speech and privacy makes ensuring accountability more complicated, but it is not a hopeless task.

Private governance across international borders does raise special concerns about legitimacy and accountability because it is two steps removed from a democratic process: platform governance occurs outside traditional state-based accountability mechanisms and is instantiated by

228. *See* Frederick Schauer, *The Exceptional First Amendment*, *in* AMERICAN EXCEPTIONALISM AND HUMAN RIGHTS 30 (Michael Ignatieff ed., 2005) (positing that the First Amendment is "generally stronger" than its international peers, "but stronger in ways that may also reflect an exceptional though not necessarily correct understanding of the relationship between freedom of expression and other goals, other interests, and other rights").

229. Yochai Benkler, *Internet Regulation: A Case Study in the Problem of Unilateralism*, 11 EUR. J. INT'L L. 171, 174 (2000).

230. *See* Franz Mayer, *Europe and the Internet: The Old World and the New Medium*, 11 EUR. J. INT'L L. 149, 161 (2000) (arguing that, with regard to internet governance, "Europeans suspect that public and private interests in the US are aiming at structuring the use of and the behaviour in the digital networks along American lines").

231. Daskal, *supra* note 13, at 219; Eichensehr, *supra* note 13; Citron, *supra* note 3.

232. Freeman, *supra* note 16, at 547.

unelected, non-state actors.[233] As in many settings in which unelected organizations create and implement rules bearing on national sovereignty and individual rights—to name just two interests—instead of the utopian direct democracy we might have expected in the heady days of the 1990s, we have a not insignificant democratic deficit instead. These concerns are heightened by substantive distinctions between national legal systems in this area. But these concerns are not unique to the digital context. Indeed, the same issues have long been at the core of efforts to demonstrate that global governance by a variety of private, public, and hybrid actors can be legitimated and made accountable to the public.

Moreover, the legitimacy concerns surrounding "private governance"[234]—even in the context of the Internet—are hardly new, and are not limited to censorship or speech. In 2000, James Boyle predicted that Internet governance would be characterized by increased privatization: "unable to respond at Internet speed, and limited by pesky constitutional constraints, the state can use private surrogates to achieve its goals."[235] In the same symposium, Michael Froomkin and Jonathan Weinberg likewise raised questions about ICANN's unaccountable governance of the domain name system.[236] More recently, Jennifer Daskal has recognized that states' efforts to regulate cross-border data flows, particularly in the context of law enforcement efforts to compel production of data stored abroad, constitute "a new form of international rulemaking," one achieved by market forces rather than states, or international organizations.[237]

Yet although mechanisms exist for promoting the accountability and legitimacy of global speech governance, they have gone largely unexplored.[238] Against this background, the literature on global governance and global administrative law is particularly resonant.

A. Legitimacy and Accountability in Private Governance

At bottom, the two critiques of platform governance outlined above reflect concerns about the legitimacy and accountability of rights govern-

233. *See, e.g.*, *id.* at 574 ("Private actors exacerbate all of the concerns that make the exercise of agency discretion so problematic. They are one step further removed from direct accountability to the electorate.").

234. *See generally* Klonick, *supra* note 14.

235. James Boyle, *A Nondelegation Doctrine for the Digital Age?*, 50 Duke L.J. 5, 10 (2000).

236. A. Michael Froomkin, *Wrong Turn in Cyberspace: Using ICANN to Route Around the APA and the Constitution*, 50 Duke L.J. 17, 29 (2000) (analyzing accountability deficit resulting from ICANN's governance of the DNS system); Jonathan Weinberg, *ICANN and the Problem of Legitimacy*, 50 Duke L.J. 187, 191–92 (2000); *see also* Dawn Nunziato, *Freedom of Expression, Democratic Norms, and Internet Governance*, 52 Emory L.J. 187, 187–88 (2003) (analyzing ICANN's role in governing online speech).

237. Daskal, *supra* note 13, at 233 ("[T]hrough the de facto operation of the market and the multinational corporations that operate across borders—rather than the more formal and mutually agreed upon process of treaty making amongst states or international organizations setting standards that impose obligations on participating states.").

238. *Cf. id.* at 238.

ance beyond the state. Reframing these critiques cuts to the heart of the issue: rather than debating about the nature of platforms' power, the right question is whether the exercise of that power is legitimate, or worthy of recognition.[239] While questions of legitimacy in states and governments have long preoccupied political theorists, scholars have also observed that, as globalization diversified the sources of power in both domestic and global politics, new kinds of legitimacy questions emerged.[240] Many legitimacy-based critiques of international and supranational institutions focus on the redistribution of power from nation states to specific international organizations such as the World Trade Organization, International Monetary Fund, and World Bank,[241] or to supranational institutions such as the European Union.[242] These actors were increasingly capable of exercising "many of the powers of a state," but lacked the "typical processes and structures" of democratic accountability that were necessary to hold states to account.[243] Scholars have sought, too, to find ways of legitimating the exercise of political power by private actors.[244] As a result, the struggle to find democratic mechanisms to legitimize "governance beyond the nation-state" and on a global scale is familiar to scholars of global politics, international relations, and international law.[245]

---

239. Jürgen Habermas, *Legitimation Problems in the Modern State*, *in* THE HABERMAS READER 248–49 (William Outhwaite ed., 1996).

240. *See* Jürgen Habermas, *The Postnational Constellation and the Future of Democracy*, *in* THE POSTNATIONAL CONSTELLATION: POLITICAL ESSAYS 67 (Max Pensky ed., 1998) ("Which aspects of globalization could potentially degrade the capacity for democratic self-steering within a national society? Are there functional equivalents at the supranational level for deficits that emerge at the level of the nation-state?").

241. *See* Robert Howse & Kalypso Nicolaïdis, *Legitimacy and Global Governance: Why Constitutionalizing the WTO Is a Step Too Far*, *in* EFFICIENCY, EQUITY, AND LEGITIMACY: THE MULTILATERAL TRADING SYSTEM AT THE MILLENNIUM 247–48 (Roger B. Porter et al. eds., 2001); Robert O. Keohane & Joseph S. Nye, Jr., *The Club Model of International Cooperation and Problems of Democratic Legitimacy*, *in* EFFICIENCY, EQUITY, AND LEGITIMACY: THE MULTILATERAL TRADING SYSTEM AT THE MILLENNIUM 265 (Roger B. Porter et al. eds., 2001); Allen Buchanan & Robert O. Keohane, *The Legitimacy of Global Governance Institutions*, 20 ETHICS & INT'L AFF. 405, 406 (2006); Michael Zürn, *Global Governance and Legitimacy Problems*, 39 GOV'T & OPP. 260, 260–61 (2004).

242. *See, e.g.*, Francesca E. Bignami, *The Democratic Deficit in European Community Rulemaking: A Call for Notice and Comment in Comitology*, 40 HARV. INT'L L.J. 451, 451–52 (1999); Gráinne de Búrca, *The Quest for Legitimacy in the European Union*, 59 MOD. L. REV. 349, 352 (1996).

243. de Búrca, *supra* note 242, at 352.

244. *See* A. Claire Cutler et al., *Introduction*, *in* PRIVATE AUTHORITY AND INTERNATIONAL AFFAIRS 6 (A. Claire Cutler et al. eds., 1999) ("[C]ooperation among private sector actors can become authoritative or government-like, even in the international sphere, thus challenging our notions of the character of political authority itself."); Rodney Bruce Hall & Thomas J. Biersteker, *The Emergence of Private Authority in the International System*, *in* THE EMERGENCE OF PRIVATE AUTHORITY IN GLOBAL GOVERNANCE 4–5 (Rodney Bruce Hall & Thomas J. Biersteker eds., 2002); *see also* HABERMAS, *supra* note 239, at 248 ("Only political orders can have and lose legitimacy; only they need legitimation. Multinational corporations or the world market are not capable of legitimation.").

245. Michael Zürn, *Democratic Governance Beyond the Nation-State: The EU and Other International Institutions*, 6 EUR. J. INT'L RELATIONS 183 (2000); *see also* Gráinne de Búrca, *Developing Democracy Beyond the State*, 46 COLUM. J. TRANSNAT'L L. 221, 223 (2008).

The global governance literature also reflects two distinct concepts of legitimacy: the "normative" concept, which concerns whether an institution "has a right to rule," and the "descriptive or sociological" perspective, which concerns whether that authority is in fact accepted.[246] A normative perspective on legitimacy, which questions the "validity of political decisions and political orders and their claim to legitimacy," closely resembles the arguments lodged by the proceduralist critics of platform governance.[247] Are platforms the actors best equipped to govern disputes about online speech and privacy? Are they sufficiently accountable to the public, and do they operate within a democratic framework? These questions are different from the question of whether platforms are reaching the right outcomes when they participate in governance, a question which concerns the "societal acceptance" of those decisions.[248]

As a descriptive matter, scholars of global administrative law have also recognized that the sources of regulatory and adjudicative power are increasingly heterodox. As Nico Krisch and Benedict Kingsbury pointed out in 2006, new kinds of actors in the "global administrative space" perform "recognizably administrative and regulatory functions: the setting and application of rules by bodies that are not legislative or primarily adjudicative in character."[249] In *The Emergence of Global Administrative Law*, Kingsbury, Krisch and Stewart distinguish five types of "globalized administrative regulation" existing along a spectrum reflecting varying degrees of formality and participation by private institutions.[250] Kingsbury et al. observe that global governance is occurring within "hybrid intergovernmental–private arrangements" as well as "administration by private institutions with regulatory functions,"[251] and that, increasingly, "international bodies make decisions that have direct legal consequences for individuals or firms without any intervening role for national government action."[252] Most tellingly, the authors point out:

> international lawyers can no longer credibly argue that there are no real democracy or legitimacy deficits in global administrative governance because global regulatory bodies answer to states, and the governments of those states answer to their voters and courts. National administrative lawyers can no longer insist that adequate accountability for global regulatory governance can always be achieved

246. Daniel Bodansky, *Legitimacy in International Law and International Relations*, *in* INTERDISCIPLINARY PERSPECTIVES ON INTERNATIONAL LAW AND INTERNATIONAL RELATIONS: THE STATE OF THE ART 327 (Jeffrey L. Dunoff & Mark A. Pollack eds., 2013); *see also* Buchanan & Keohane, *supra* note 241, at 405 ("'Legitimacy' has both a normative and a sociological meaning."); Zürn, *supra* note 241, at 260–61; de Búrca, *supra* note 242, at 349 (proposing both normative and social views of legitimacy).

247. Zürn, *supra* note 241, at 260.

248. *Id.*

249. Nico Krisch & Benedict Kingsbury, *Global Governance and Global Administrative Law in the International Legal Order*, 17 EUR. J. INT'L L. 1, 3 (2006).

250. Benedict Kingsbury, Nico Krisch & Richard Stewart, *The Emergence of Global Administrative Law*, 68 L. & CONTEMP. PROBS. 15, 20 (2005).

251. *Id.*

252. *Id.* at 24.

> through the application of domestic administrative law requirements to domestic regulatory decisions.[253]

The standard critiques of platform governance reflect many of the same concerns as have long preoccupied scholars of global governance and global administrative law. In particular, the procedural critique of platform governance—which takes a dim view of platforms' capabilities to perform governance functions in a manner that is sufficiently deliberative, independent of undue pressure from interested parties, and with adequate procedural safeguards—mirrors many of the concerns expressed in the global governance literature on legitimacy and accountability.[254] The substantive critique, too, reflects a different sort of legitimacy question: whether the outcomes of platform decisions are correct or deserve social or political support.

The literature on global governance is particularly resonant now because so much of it emerged out of popular unrest and dissatisfaction with the performance of international institutions. The picture that emerges from the global governance scholarship is one of secretive, powerful institutions that purported to act in the public's name and whose substantive and procedural deficits led them to confront public resistance and protest. As Richard Stewart and Michelle Badin described the World Trade Organization in 2009, the organization had confronted "stringent criticism by civil society organizations and some members for closed decision-making, an unduly narrow trade focus, domination by powerful members and economic and financial interests, and disregard of social and environmental values and the interests of many developing countries and their citizens."[255]

These descriptions bear a striking similarity to the critiques of platforms today—from both the left and the right—as organizations out of touch with the public mood, unduly focused on the bottom line, dominated by the interests of Brussels and Washington, and oblivious to the privacy and dignity interests of their users. A fresh approach to questions of legitimacy and accountability should guide our understanding of the current moment of profound political and cultural backlash against platforms. Thinking about platform governance through the lens of global governance is doubly helpful because it both expands the unduly narrow view of the actors who are performing governance roles and enriches the

253. *Id.* at 26.

254. Although these two terms are separate, they are often conflated or addressed together. *See*, *e.g.*, Buchanan & Keohane, *supra* note 241, at 415 (arguing that cooperation by democratic governments provides a "democratic channel of accountability" but is insufficient to legitimate governance institutions); *id.* at 426–27 ("It is misleading to say that global governance institutions are illegitimate because they lack accountability and to suggest that the key to making them legitimate is to make them accountable . . . what might be called narrow accountability—accountability without provision for contestation of the terms of accountability—is insufficient for legitimacy.").

255. Richard B. Stewart & Michelle Ratton Sanchez Badin, *The World Trade Organization and Global Administrative Law* 1 (N.Y.U. Pub. Law & Legal Theory Working Papers, Paper 166, 2009), http://ssrn.com/abstract=1518606 [https://perma.cc/B4JJ-MQQ4].

vocabulary for considering potential solutions to legitimacy and accountability gaps.

B. Solving Legitimacy and Accountability Problems

The governance literature makes clear that state intervention is not the only way to promote accountability and legitimacy for private actors. Rather than shying away from their global governance role, platforms could willingly adopt new procedures and methods explicitly designed to increase their own legitimacy and accountability and to address their critics. Far from being unprecedented, scholars of global administrative law have observed that, to "bolster[ ] their legitimacy in the face of growing political challenges," some global regulators have turned to administrative law principles and values such as transparency, participation, reasoned decision-making, and judicial review in an effort to demonstrate their own accountability to the governed.[256] Indeed, as Facebook has begun to plan for the creation of its oversight board, it has often used the language of public law, expressing desire to "giv[e] people a voice" and to conduct effective oversight of decision-making.[257]

It is not difficult to imagine that online platforms, which face increasing pressures to create and implement rules and adjudicate disputes about online speech and privacy, could turn to the principles and values of administrative law as a way of enhancing their own legitimacy. As an initial matter, platforms are already beginning to recognize that they perform a quasi-regulatory role. This role is not out of place in the Internet governance arena, although most formal and informal international institutions of Internet governance focus on the Internet infrastructure rather than content and privacy issues.[258] Global, non-governmental regulatory bodies are already common in the Internet governance space—ICANN being perhaps the premier example. Other bodies might be seen as private "meta-regulators" that aim to coordinate among private entities and create standards that all can adhere to.[259] Meta-regulators are "private re-

256. Krisch & Kingsbury, *supra* note 249, at 4.

257. Zuckerberg, *supra* note 207. The board has been widely described in the popular press as a "Supreme Court." *See, e.g.*, Hanna Kozlowska, *Facebook will have a Supreme Court–like body within a year*, Quartz (Nov. 15, 2018), https://qz.com/1465898/mark-zuckerberg-facebook-to-have-a-supreme-court-within-a-year/; Evelyn Douek, *Facebook's New 'Supreme Court' Could Revolutionize Online Speech*, Lawfare (Nov. 19, 2018), https://www.lawfareblog.com/facebooks-new-supreme-court-could-revolutionize-online-speech; Kate Klonick & Thomas Kadri, *How to Make Facebook's 'Supreme Court' Work*, N.Y. Times (Nov. 17, 2018), https://www.nytimes.com/2018/11/17/opinion/facebook-supreme-court-speech.html.

258. *But see* Laura DeNardis, The Global War for Internet Governance 9 (2014) ("Traditional power structures increasingly view Internet governance technologies as mechanisms for controlling global information flows. One Internet governance theme is the escalating use of Internet governance technologies as a proxy for content control.").

259. *See, e.g.*, Fabrizio Cafaggi, *Transnational Private Regulation: Regulating Global Regulators*, *in* Research Handbook on Global Administrative Law 214 (Sabino Cassese ed., 2016) ("Private meta-regulation usually consists of general common principles to be applied to the regulators that subscribe to them on a voluntary basis. In this case, regulators become regulated entities by means of the meta-regulator.").

gimes in the public interest," as opposed to "those that aim to produce purely private benefits for the regulated."[260] For instance, the Global Network Initiative is a voluntary membership organization that "provides, among other things, a framework of best practices and an assessment mechanism for evaluating the human rights performance of companies in the information and communication technologies sector" and is active on issues related to content regulation.[261]

1. Transparency

What would it mean to apply the principles and values of administrative law in the context of global platform governance? To begin, platforms could commit to acting more transparently. Platforms already regularly publish transparency reports that document a range of statistics, including information about the number and types of requests to take down content that they receive.[262] For instance, Google reports that it removed almost ten million videos during the first quarter of 2018, over seven million of which were detected by automated means.[263] Google's transparency report informs readers that YouTube participates in the Global Internet Forum's hash sharing program, and that 98% of videos "removed for violent extremism" were identified by machine learning.[264] But the report provides no information about the actual number of videos that were removed for "violent extremism," nor does it document the number of videos flagged by Internet referral units. Facebook's transparency report is similarly spotty, revealing that the platform "took action on" 1.9 million pieces of content that contained "terrorist propaganda" in the first quarter of 2018.[265] But "taking action" could mean anything from "removing a piece of content," flagging content with

260. *Id.* at 213.

261. Molly Land, *Toward an International Law of the Internet*, 54 HARV. INT'L L.J. 393, 445 (2013); *see also* Luca Belli & Nicolo Zingales, *Preview of the 2017 DCPR Outcome: Platform Regulations (DC on Platform Responsibility)*, INTERNET GOVERNANCE F., https://www.intgovforum.org/multilingual/content/preview-of-the-2017-dcpr-outcome-platform-regulations-dc-on-platform-responsibility [https://perma.cc/CW4Y-9DUG] ("[A]t the 2014 Internet Governance Forum, the Dynamic Coalition on Platform Responsibility was created. The DCPR is a multistakeholder group established under the auspices of the United Nations Internet Governance Forum dedicated to the analysis of the role and responsibilities of online platforms from a technical, legal, social or economic perspective. Since its inception, DCPR has facilitated and nurtured a cross-disciplinary analysis of the challenges linked to the emergence of digital platforms and has promoted a participatory effort aimed at suggesting policy solutions.").

262. *See, e.g.*, *Facebook Transparency Report*, FACEBOOK, https://transparency.facebook.com/ [https://perma.cc/RC9X-LDKD] (last visited Feb. 9, 2019); *Transparency Report*, GOOGLE, https://transparencyreport.google.com/ [https://perma.cc/BSV4-4VN6] (last visited Feb. 9, 2019); *Twitter Transparency Report*, TWITTER, https://transparency.twitter.com/en.html [https://perma.cc/L6QX-4LWT] (last visited Feb. 9, 2019).

263. *YouTube Community Guidelines Enforcement*, GOOGLE, https://transparencyreport.google.com/youtube-policy/overview?hl=en [https://perma.cc/GT6P-H8DA] (last visited Feb. 9, 2019).

264. *Id.*

265. *Hard Questions: What Are We Doing to Stay Ahead of Terrorists?*, FACEBOOK (Nov. 11, 2018), https://newsroom.fb.com/news/2018/11/staying-ahead-of-terrorists/.

a "warning," or disabling an entire account.[266]

Although platform transparency reports are helpful to inform the public, the incomplete information they contain makes them of limited use in understanding how platform governance actually operates.[267] A fuller embrace of transparency principles would not only suggest that platforms should publish more granular information in their transparency reports, but also that they could coordinate with each other to make those reports more meaningful to the public.[268] A private organization such as the Global Network Initiative could easily serve as a "meta-regulator" to set standards for the kinds of information that transparency reports ought to convey. This is especially important because platforms are, indeed, already coordinating—through the Global Internet Forum, hash sharing, and other cooperative arrangements—both with each other and with government. In order to be truly informative, transparency reports should shed light on the ways that these arrangements impact platform governance.

Platforms could also embrace more robust transparency in ways extending beyond their ordinary reporting. Indeed, Google's experience implementing the right to be forgotten illustrates the limitations of transparency reporting in shedding light on platform governance. While Google does publish periodic reports about content deletion under the right to be forgotten,[269] the information is fairly limited in scope, showing the numbers of requests and URLs affected and the types of requesters. The public lacks access to the information most critical to understand the platform's implementation of the right to be forgotten: Google's criteria for deciding whether to delete information. The report maintains that Google "assess[es] each request on a case-by-case basis" and sets forth "[a] few common material factors" the platform uses, but the internal

266. *Community Standards Enforcement Preliminary Report*, FACEBOOK, https://transparency.facebook.com/community-standards-enforcement (last visited Nov. 11, 2018).

267. *See* Hannah Bloch-Wehba, *How much of your online life should police have access to? SCOTUS prepares to weigh in*, VOX (Nov. 29, 2017, 8:33 AM), https://www.vox.com/the-big-idea/2017/11/22/16687420/fourth-amendment-online-searches-constitution-facebook-gag-orders [https://perma.cc/VW2X-D38Z] (observing, in the context of electronic surveillance data, that transparency reports "provide the most useful and granular information" available to the public, despite their limitations).

268. *See* Alex Feerst, *Implementing Transparency About Content Moderation*, TECHDIRT (Feb. 1, 2018, 1:38 PM), https://www.techdirt.com/articles/20180131/22182339132/implementing-transparency-about-content-moderation.shtml [https://perma.cc/8UC8-6SNX]. Civil society groups have already called on platforms to provide more detailed information and more robust notice and appeal mechanisms for content moderation decisions. *The Santa Clara Principles on Transparency and Accountability in Content Moderation* (May 7, 2018), https://newamericadotorg.s3.amazonaws.com/documents/Santa_Clara_Principles.pdf; *see also* Spandana Singh & Kevin Bankston, *The Transparency Reporting Toolkit: Content Takedown Reporting*, OPEN TECH. INST. (Oct. 25, 2018), https://www.newamerica.org/oti/reports/transparency-reporting-toolkit-content-takedown-reporting/ ("Transparency reporting on content takedowns is critically important because it helps holds companies accountable in their role as gatekeepers of our online speech.").

269. *Search removals under European privacy law*, GOOGLE, https://transparencyreport.google.com/eu-privacy/overview?hl=en [https://perma.cc/6YMA-2VTY] (last visited Feb. 9, 2019).

criteria that platform employees actually rely upon are not public.[270] Even the company's more detailed report, *Three Years of the Right to Be Forgotten*, provides fairly minimal information about how the company actually considers the criteria it uses to process takedown requests.[271] The report notes that this "emphasis on privacy creates a fundamental tension with standards for reproducible science."[272]

In a recent case in the United Kingdom, in which two anonymous businessmen successfully sought to compel Google to delist links concerning prior convictions of criminal offenses, the High Court noted that "Google relies on a number of factors" not present in the Article 29 Working Party criteria: "the nature of the third party publishers; the public availability of the information; the nature of the information and the alleged inaccuracies; the claimant's business activities and their relationship with the information published at the URLs complained of; and the claimant's own online descriptions of himself."[273] The lack of clarity regarding these criteria, which only emerged at trial,[274] simply underscores the complexity of making right to be forgotten decisions—and the public's urgent need to understand how these decisions are made.[275]

Finally, platforms could operate more transparently by publicly acknowledging and justifying their decisions to take action to limit access to content online. For example, platforms could publish a brief summary of the facts raised by each request to delist search results, the analysis, and the result. Google already publishes sample requests in its transparency report, including a one-sentence summary of the request and a one-sentence summary of the outcome.[276] This information is presented at a sufficient level of generality to avoid re-identifying the requester, but illustrates the kinds of questions the company confronts when it determines whether to delist search results. Unfortunately, however, no corollary explanation exists for content against which platforms take action pursuant to their own ToS. Indeed, platforms have made very little information public to explain their own implementation of their rules and guidelines on ToS takedowns and community standards; what has become public is largely due to leaks in the press.[277]

---

270. *Id.*

271. BERTRAM ET AL., *supra* note 173, at 2.

272. *Id.* at 4.

273. NT1 & NT2 v. Google LLC [2018] EWHC 799 (QB) [¶ 131] (UK), https://www.judiciary.uk/wp-content/uploads/2018/04/nt1-Nnt2-v-google-2018-Eewhc-799-QB.pdf [https://perma.cc/V2A8-GMSP].

274. Gareth Corfield, *Here is how Google handles Right To Be Forgotten requests*, REGISTER (Mar. 19, 2018), https://www.theregister.co.uk/2018/03/19/google_right_to_be_forgotten_request_process/ [https://perma.cc/7FH9-VW3C].

275. It remains to be seen whether the court's decision in *NT1 & NT2* will result in Google embracing the Working Party guidelines and abandoning its own internal rules.

276. *Search removals under European privacy law*, *supra* note 269.

277. *See* Julia Angwin & Hannes Grassegger, *Facebook's Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children*, PROPUBLICA (June 28, 2017, 5:00 AM), https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms [https://perma.cc/N3Q9-E3X6].

2. Reasoned Decision-Making

Beyond simple transparency, additional disclosures are vital for platforms to demonstrate that they satisfy another core principle of administrative law: reasoned decision-making. Platforms' failure to publish information about how, when, and why they take action regarding online content simply raises suspicions that they are applying their own (secret) rules in ways that are arbitrary or baseless.[278] That platforms routinely reach opposite conclusions about specific instances of online content underscores the public's perception of these decisions as random and illegitimate.[279] As Sarah Roberts has observed, moreover, decision-making about what content "is acceptable and what is not is a complex process, beyond the capabilities of software or algorithms alone," yet commercial content moderation still largely operates behind the scenes.[280] Far from justifying their decisions to the public, platforms continue to treat their governance mechanisms as if they were proprietary tools that did not affect public rights.

3. Participation

Nor have platforms embraced mechanisms that would enhance the public's ability to participate in governance, either directly or through representatives. When they have done so, platforms have not demonstrated that they take participatory governance seriously. For instance, from 2009 through 2012 Facebook experimented with participatory governance by asking users to vote on a variety of policy changes. But the platform's rules for engagement required that 30% of active Facebook users had to participate for the results of the process to be effective.[281] By requiring such an astronomical turnout threshold—by 2012 Facebook had more than a billion users—the platform virtually ensured that its participatory governance mechanisms would fall short. Facebook's prior failures, however, appear not to have dissuaded the platform from invoking the value of "public participation" in support of its proposed oversight

---

278. *See, e.g.*, Kevin Bankston & Liz Woolery, *We Need To Shine A Light On Private Online Censorship*, TECHDIRT (Jan. 31, 2018, 1:29 PM), https://www.techdirt.com/articles/20180130/22212639127/we-need-to-shine-light-private-online-censorship.shtml [https://perma.cc/8SVC-88R3] (discussing how public scrutiny of platforms' "bizarrely idiosyncratic rule sets" could "guard against discriminatory impacts on oft-marginalized communities").

279. *See, e.g.*, Aja Romano, *Twitter's stance on Infowars' Alex Jones should be a moment of reckoning for users*, VOX (Aug. 8, 2018, 2:00 PM), https://www.vox.com/2018/8/8/17662774/twitter-alex-jones-jack-dorsey [https://perma.cc/PU3W-73YK] (reporting that, while Apple, Facebook, YouTube, and other platforms had banned far-right conspiracy theorist Alex Jones and InfoWars from their platforms, Twitter had not).

280. Sarah T. Roberts, Behind the Screen: The Hidden Digital Labor of Commercial Content Moderation 13–15 (2014) (unpublished Ph.D. dissertation, University of Illinois at Urbana-Champaign), http://hdl.handle.net/2142/50401 [https://perma.cc/V73M-Y9Q2].

281. Severin Engelmann et al., *A Democracy Called Facebook? Participation as a Privacy Strategy on Social Media*, *in* PRIVACY TECHNOLOGIES AND POLICY: 6TH ANNUAL PRIVACY FORUM, APF 2018, BARCELONA, SPAIN, JUNE 13-14, 2018, REVISED SELECTED PAPERS (2018), https://www.cybertrust.in.tum.de/fileadmin/w00bzf/www/papers/2018-APF-Engelmann.pdf [https://perma.cc/YX5J-VCR5].

board.[282]

Informed by the development of more participatory global governance mechanisms by international organizations, networks, and other hybrid arrangements, however, platforms could readily adapt a variety of methods of boosting public participation. For instance, platforms could adopt rigorous "notice-and-comment" procedures for changes to their privacy policies, ToS, and community standards, to give the public input concerning the kinds of information that are made available online and collected and used by advertisers and other third parties.

Platforms could also expand the opportunities for civil society to participate in platform governance. Currently, platforms appear to engage with some NGOs as civil liberties partners, while they partner with others as "trusted flaggers"—and the two groups have very little overlap or interaction. This method of engaging with civil society is incomplete; as Robert Keohane and Joseph Nye have written, NGOs and civil society networks will often engage in overlapping or conflicting "transnational-transgovernmental coalitions."[283] Platforms could, therefore, explore more robust avenues for participation by civil society in formulating and expounding policies and rules that will, by dint of practical operation, have global effects. Avenues for participation in adjudication could also be broadened: platforms could appoint an "amicus curiae" to help make more informed and robust decisions in difficult cases regarding content or privacy.

As with reasoned decision-making, however, platforms' ability to demonstrate that they take participation seriously requires additional transparency as well. Full participation by the public and civil society is, indeed, *contingent* on transparency: as Richard Stewart points out, "information about an organization's ongoing and proposed decisions and policies is essential for outsiders to know when and where to make submissions on proposed decisions and how to make such submissions effective."[284] Moreover, civil society participation in some global governance contexts has been "sharply contested," with advocates asserting that participation can remedy democratic deficits and detractors contending that civil society organizations tend to support the policies sought by developed countries.[285]

4. Judicial Review

Finally, platform governance could be made more accountable and democratically legitimate were it subjected to judicial review. Under the current framework, an individual who requests that a search result be

282. Zuckerberg, *supra* note 207 (describing how the board will "giv[e] people a voice").

283. Keohane & Nye, *supra* note 241, at 271.

284. Richard B. Stewart, *Remedying Disregard in Global Regulatory Governance: Accountability, Participation, and Responsiveness*, 108 AM. J. INT'L L. 211, 263 (2014).

285. Steven Kochevar, *Amici Curiae in Civil Law Jurisdictions*, 122 YALE L.J. 1653, 1658 (2013).

delisted pursuant to the right to be forgotten has standing to contest a platform's failure to comply with the data protection authority in their country of residence or in court. However, the author or webmaster of the delisted result has no avenue to seek administrative or judicial redress.[286] Perhaps more troubling, the *public* has no opportunity to be heard on the deletion of information from search results, although Article 10 of the European Convention on Human Rights stipulates that the public has the right to "receive and impart information and ideas without interference by public authority and regardless of frontiers."[287] The result is that judicial review of platforms' decisions on the right to be forgotten is available only for one of the affected parties.

Nor is judicial review available when platforms take action against content or activity that violates their community standards or ToS. Although some litigants are testing the limits of this obstacle in U.S. courts, seeking to transform this into an issue with a legal remedy, they have so far not prevailed.[288] When content is deleted pursuant to a decision that it violates a platform's ToS, it is not even clear that platforms always provide users with an opportunity to contest that decision within the platform itself.[289] Even assuming that platforms do notify posters that their content will be taken down, numerous commentators have pointed out that relying on counter-notices by users is not effective to protect lawful content.[290] Some new developments in foreign courts, however, suggest that users may have remedies against platforms that wrongfully delete content. In Germany, the courts have long applied the *Drittwirkung* doctrine, which recognizes that public law values influence private rights.[291] In at least three German cases in 2018, courts held that, under the *Drittwirkung* doctrine, Facebook must observe fundamental rights when it determines whether to delete content pursuant to its ToS.[292]

The German cases suggest that, at least under some national laws, individuals *will* have standing to contest removal decisions in court. In many cases, this may raise significant questions about intermediary immunity as well as about platforms' claims that they act as curators or editors of content themselves.[293] But remedying the one-sided nature of judicial review

286. *Guidelines on the Implementation of Google Spain*, *supra* note 159, at 10.

287. European Convention for the Protection of Human Rights and Fundamental Freedoms art. 10(1), Nov. 4, 1950, 213 U.N.T.S. 221.

288. *See, e.g.*, Prager Univ. v. Google LLC, No. 17-CV-06064-LHK, 2018 WL 1471939, at *14 (N.D. Cal. Mar. 26, 2018) (dismissing complaint with leave to amend).

289. *See* Daphne Keller, *Internet Platforms: Observations on Speech, Danger, and Money* 18 (Hoover Institution, Aegis Series Paper No. 1807, 2018), https://www.hoover.org/sites/default/files/research/docs/keller_webreadypdf_final.pdf (calling on platforms to implement "rigorous procedural safeguards" to protect against over-removal of speech).

290. *See, e.g.*, Keller, *Counter-Notice*, *supra* note 183.

291. *Lüth Case*, 7 BVerfGE 198 (1958), *in* THE CONSTITUTIONAL JURISPRUDENCE OF THE FEDERAL REPUBLIC OF GERMANY 444 (3d ed. 2012) (holding that "every provision of private law must be compatible with [the Basic Law's] system of values, and every such provision must be interpreted in its spirit").

292. *See* Keller, *supra* note 13, at 12.

293. *See* Whitney, *supra* note 196, at 3.

for platform governance is a good thing when it comes to checking government coercion. If platforms' content moderation decisions increasingly result from government pressure, enhancing standing to contest removal decisions will provide a vital check. Even where such standing does not exist—as in the United States—platforms could still work to remedy the one-sided nature of judicial review for platform governance by committing to litigate test cases in which they seek to assert the rights of their users. In the surveillance context, platforms have already seen the utility of initiating litigation on behalf of users, potentially in part because standing up for user privacy is "a valuable marketing tool."[294] Just so, platforms could seek to stand up for their users' *expressive* rights in court as well.

C. Obstacles to Accountability

The above Part articulated a range of options for platforms to bolster their accountability and legitimacy with the public. However, none of these mechanisms are band-aids. Instead, platforms should adopt a range of additional measures to promote legitimacy and accountability. For the reasons set forth above, moreover, platforms must begin by making more information public: robust transparency is a precondition for public participation in policy and rulemaking, and is doubly required for public confidence in the quality of platform decision-making.

Yet the laws that create the obligations for private governance actually in some ways prevent the development of systems that could be accountable. In other words, the applicable legal regimes obstruct the application of safeguards that would constrain discretion and enhance democratic accountability.

First, applicable frameworks for content takedowns favor *speed* over *accuracy*. This balance is at the core of calls for more "proactive," automated measures in the context of hate speech and terrorist speech.[295] It is equally clear, however, that "[a]utomated techniques will block some legitimate content."[296] At bottom, frameworks that favor speech and automation will make it difficult for platforms to engage in the kind of careful, reasoned decision-making that is necessary to make informed decisions about online speech, and the development of proprietary algorithmic blocking tools will further hamper public understanding.

Second, and relatedly, applicable legal frameworks presumptively favor decisions to *delete content* over decisions to *maintain content* in edge

294. *Cooperation or Resistance?*, *supra* note 62, at 1726; *see also* Rozenshtein, *supra* note 1; Hannah Bloch-Wehba, *Exposing Secret Searches: A First Amendment Right of Access to Electronic Surveillance Orders*, 93 Wash. L. Rev. 145, 160 (2018) (describing how Microsoft initiated litigation to challenge the secrecy of electronic surveillance demands).

295. *See Communication*, *supra* note 112, at 13 ("It is in the entire society's interest that platforms should remove illegal content as fast as possible.").

296. *Chaining giants: Internet firms face a global techlash*, Economist (Aug. 10, 2017), https://www.economist.com/news/international/21726072-though-big-tech-firms-are-thriving-they-are-facing-more-scrutiny-ever-internet-firms [https://perma.cc/XY7H-TY9P].

cases. In the context of the right to be forgotten, for instance, Google bears no responsibility for wrongfully delisting content. Google has no counter-notice mechanism to notify the author of a delisted page regarding a right to be forgotten decision. Moreover, neither that author nor the public has standing to contest a removal decision.[297] But sanctions for failing to delist content can be quite stringent: data protection authorities may fine intermediaries, and right to be forgotten requesters may bring suit. This creates incentives for companies to err on the side of deletion—precisely the risk that intermediary liability has always raised. Identical risks of over-deletion exist in the contexts of hate speech and terrorist speech. The lopsided incentives reflect that parties have wildly varying incentives and opportunities to challenge decisions to remove—or not remove—online content.

Third, the law *discourages transparency* and *encourages opacity* regarding how platforms reach decisions concerning online speech and privacy. For instance, Google has struggled to inform the public regarding right to be forgotten requests because to do so often would constitute a further infringement on the privacy right of the data subject.[298] It is particularly difficult to attribute responsibility for censorship in the contexts of hate speech and terrorist speech, in which governments exert considerable pressure on platforms to delete content pursuant to their own ToS rather than under public law. In the context of the right to be forgotten, where platforms administer takedown schemes and adjudicate thousands of individualized disputes behind closed doors, understanding rule making and dispute resolution is practically impossible because of corporate opacity and the large number of demands. Lack of transparency itself, then, can prompt legitimacy concerns for companies engaged in policing online speech.

VI. CONCLUSION

Far from making the Internet more accountable, more legitimate, and more responsive to users, online self-governance and self-regulation have in fact obscured the boundaries between state and private actors. This confusion makes it increasingly difficult to attribute responsibility for online censorship and to disentangle platform governance from coercive actions taken by states across the globe. Because the public lacks key information about how, when, and at whose direction platform governance is taking place, it is extremely difficult for the outside observer to discern what is going on, and to distinguish private action from government pressure.

This doesn't have to be the case. Platforms themselves can embrace accountability mechanisms that would make their governance decisions more accountable and more legitimate to their users and to the public.

297. Keller, *Notice and Takedown*, *supra* note 215.
298. *Searching for the right balance*, *supra* note 160; *see also* McCarthy, *supra* note 163.

Indeed, as platforms increasingly seek to subdue popular backlash and ingratiate themselves with consumers, there are good reasons for them to take measures to "boost their legitimacy and effectiveness."[299] Adherence to the principles and values of administrative law can subject even informal, private, global governance to rule of law principles.

Accountability may not come naturally to the tech industry. Although private governance has long transpired behind closed doors, platforms will need to resist the impulse to operate in the dark, instead opting for meaningful disclosure practices and other transparency mechanisms that expose their own rulemaking and adjudicatory procedures to scrutiny. Rather than ignoring user input, platforms will need to actively solicit participation by users and civil society organizations in public-facing and public-regarding ways. Most difficult of all, these measures will require platforms to resist legal requirements that prevent or block them from developing robust safeguards and accountability mechanisms.

But platforms must take on these tasks themselves rather than waiting for government to act, because to wait is to allow the structures of private ordering to be coopted by state censors. Obscurity is not simply the result of bureaucracy: it is a key characteristic of collateral censorship, which relies on secrecy and an uninformed public for its success. As we find that would-be government censors all over the globe are increasingly relying on platforms' own internal rules, procedures, and decision-making mechanisms, it should be clear: now is the time to act.

299. Kingsbury, Krisch & Stewart, *supra* note 250, at 16.