CONTENT MODERATION AS SURVEILLANCE

Hannah Bloch-Wehba†

ABSTRACT

Technology platforms are the new governments, and content moderation is the new law, or so goes a common refrain. As platforms increasingly turn toward new, automated mechanisms of enforcing their rules, the apparent power of the private sector seems only to grow. Yet beneath the surface lies a web of complex relationships between public and private authorities that call into question whether platforms truly possess such unilateral power. Law enforcement and police are exerting influence over platform content rules, giving governments a louder voice in supposedly "private" decisions. At the same time, law enforcement avails itself of the affordances of social media in detecting, investigating, and preventing crime.

This Article, prepared for a symposium dedicated to Joel Reidenberg's germinal article *Lex Informatica*, untangles the relationship between content moderation and surveillance. Building on Reidenberg's fundamental insights regarding the relationships between rules imposed by legal regimes and those imposed by technological design, the Article first traces how content moderation rules intersect with law enforcement, including through formal demands for information, informal relationships between platforms and law enforcement agencies, and the impact of end-to-end encryption. Second, it critically assesses the degree to which government involvement in content moderation actually tempers platform power. Rather than effective oversight and checking of private power, it contends, the emergent arrangements between platforms and law enforcement institutions foster mutual embeddedness and the entrenchment of private authority within public governance.

DOI: https://doi.org/10.15779/Z389C6S202

† Associate Professor, Texas A&M School of Law. My thanks to Kendra Albert, Julie Cohen, Ignacio Cofone, Caroline Mala Corbin, Rebecca Crootof, Angel Diaz, evelyn douek, Miriam Estrin, Nik Guggenberger, Thomas Kadri, Christina Koningisor, Rachel Levinson-Waldman, Przemek Palka, Christopher Reed, Alicia Solow-Niederman, Jennifer Urban, and participants at the Freedom of Expression Scholars Conference, WIPIP and Platgov for helpful comments on this project. I am very grateful to Joshua Frechette for excellent research assistance. Finally, the terrific editors at the *Berkeley Technology Law Journal*, including Loc Ho, Justine McCarthy Potter, Barbara Rówińska, and Dakota Sneed, who helped get this Article across the finish line. All mistakes are, of course, my own.

TABLE OF CONTENTS

I. INTRODUCTION

In September 2020, after a summer of uprisings against police violence, a series of wildfires broke out in the Northwest. It didn't take long for rumors that the fires had been started by antifa activists, or by the Proud Boys, to start spreading on social media. Soon, vigilantes set up roadblocks, searching for the responsible parties and, in the process, obstructing traffic and heightening tensions. Law enforcement agencies, tasked with enforcing evacuation orders, grew increasingly concerned about viral misinformation making their jobs even harder.[1]

1. Dennis Romero, *Facebook to Take Down False Reports of Antifa Arson in Oregon*, NBC NEWS (Sept. 13, 2020, 2:58 AM), https://www.nbcnews.com/tech/tech-news/facebook-take-down-false-reports-antifa-arson-oregon-n1239966.

After first working to attach misinformation "warning labels" to the posts, Facebook ultimately announced that it would delete the posts altogether.[2] Facebook's action was welcome, but puzzling to some. After a year of epic failures in addressing misinformation about public health, elections, and social movements, why did Facebook act so quickly—and so aggressively—in shutting down misinformation about the Oregon wildfires? This Article proposes a potential answer: law enforcement's assertion of its own demands and needs shaped Facebook's content moderation rules and affected Facebook's response to crisis.

This Article suggests that law enforcement's impact on content governance is not sporadic or fleeting. Policing is, instead, a durable influence on the rules, standards, and technical processes by which platforms govern their communities. Nor is this influence limited to high profile examples of unlawful speech, such as terrorism, incitement of violence, or sex trafficking. In more quotidian contexts, platforms also play a crucial role as intermediaries in evidence-gathering processes.[3] As police increasingly depend upon digital evidence in investigating and prosecuting crime, content governance strategies also shape the kinds of data that are germane to investigations and affect how law enforcement does its job.[4]

This commingling of public and private authority raises conceptual questions about the nature of content and data governance. While a robust literature considers how and why platforms have developed "community standards" by which they govern user behavior in online spaces, the "private" character of these standards and rules is often taken for granted.[5] Indeed, platforms are often described as governments in their own right, equally powerful and sovereign as the states in which they are headquartered.[6] The structure of intermediary liability law reaffirms this conception of platform

2. Reuters Staff, *Facebook Removes Posts Linking Oregon Wildfires to Activist Groups*, REUTERS (Sept. 13, 2020, 3:19 AM), https://www.reuters.com/article/us-usa-wildfires-facebook-idUSKBN264013.

3. *See infra* Part III.A.

4. *See infra* Part III.B.

5. *See, e.g.*, Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018); Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353 (2017); NICOLAS P. SUZOR, LAWLESS: THE SECRET RULES THAT GOVERN OUR DIGITAL LIVES 11 (2019) ("The legal reality is that social media platforms belong to the companies that create them, and they have almost absolute power over how they are run.").

6. Kristen E. Eichensehr, *Digital Switzerlands*, 167 U. PA. L. REV. 665, 672 (2019); *see also* JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM 122 (2019) (noting that platforms "have worked to position themselves as both essential partners and competing sovereigns in the quest to instantiate states of exception algorithmically").

governance as "private." At least under U.S. law (for now), platforms are largely immune from liability for hosting even unlawful user-generated content, leading scholars to describe the voluntary mechanisms they enforce as a category of private regulation adopted without legal obligation.

In fact, however, the purportedly private rules of content moderation emerge and operate within a political context in which law enforcement acts as a particularly powerful stakeholder. For example, law enforcement has encouraged platforms to adopt more stringent rules on certain categories of harmful content, such as child sexual abuse imagery (CSAM) or violent rap music in the UK.[7] In Europe, police agencies have formed special "internet referral units" to report and flag violations of platforms' content rules for takedown.[8] Platforms' private decision-making thus provides a new avenue for law enforcement to regulate the public sphere.[9] As platform firms turn to automation and artificial intelligence to scale up their efforts to address harmful online content, the technical infrastructures of content moderation increasingly reflect government influence.[10]

Just as law enforcement seeks expanded influence over platforms' private decision-making, the processes and technical affordances of content governance also affect and shape law enforcement investigations in more mundane contexts. Police rely on social media to identify purported gang members, generate investigative leads, map networks and associations, and monitor activity by the public.[11] The prevalence of social media as an

7. *See infra* Part II.A.

8. 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 (2017); Hannah Bloch-Wehba, *Global Platform Governance: Private Power in the Shadow of the State*, 72 SMU L. REV. 27, 45–46 (2019).

9. Susan Benesch, *But Facebook's Not a Country: How to Interpret Human Rights Law for Social Media Companies*, 38 YALE J. ON REGUL. BULL. 86, 99 (2020) ("Company content moderation is also used as a means for states to carry out silent and invisible censorship."); *see also* Bloch-Wehba, *supra* note 8, at 45–46 (distinguishing between legal takedown orders and the expanded global sweep of takedowns under platforms' internal terms of service).

10. Hannah Bloch-Wehba, *Automation in Moderation*, 53 CORNELL INT'L L.J. 41, 69–70 (2020).

11. *See* Katherine J. Strandburg, *Freedom of Association in a Networked World: First Amendment Regulation of Relational Surveillance*, 49 B.C. L. REV. 741, 743–44 (2008) (describing how law enforcement began to use communications traffic data to map social relationships and group memberships, and naming these strategies "relational surveillance"); Desmond Upton Patton, Douglas-Wade Brunton, Andrea Dixon, Reuben Jonathan Miller, Patrick Leonard & Rose Hackman, *Stop and Frisk Online: Theorizing Everyday Racism in Digital Policing in the Use of Social Media for Identification of Criminal Conduct and Associations*, 3 SOCIAL MEDIA + SOCIETY 1 (2017) (describing the use of social media information in gang databases); Megan Behrman, *When Gangs Go Viral: Using Social Media and Surveillance Cameras to Enhance Gang Databases*, 29 HARV. J. L. & TECH. 315 (2015); Keegan Stephan, *Conspiracy: Contemporary Gang*

investigative tool also makes investigations, to some degree, reliant on platforms' own decisions about what content-related behaviors to permit or forbid. For instance, users' ability to delete posts, photos, videos, emails, and messages has prompted law enforcement agencies to procure new tools to scrape and retain user data.[12] Ironically, as platforms have cracked down on certain types of unlawful content, they have arguably made law enforcement's jobs in ferreting out unlawful activity that much more difficult.[13]

The chief goal of this Article is to illuminate the close relationship between platforms and police by examining how content-related decision-making within private platforms can advance or inhibit law enforcement surveillance practices. In so doing, I bring together two distinct bodies of scholarship. The first emphasizes platforms' roles as private guarantors of free expression and views government pressures on content-related rules as a toxic form of "jawboning" or collateral censorship through which the government seeks to regulate the public sphere indirectly when it could not do so directly.[14] The second examines how government can compel disclosure or otherwise extract information about users from social media and the role of internet platforms in accommodating, facilitating, and resisting those demands.[15] As Joel

Policing and Prosecutions, 40 CARDOZO L. REV. 991, 1021 (2018) ("The NYPD has admitted that communicating with the wrong person on social media is enough to get someone placed on a gang database . . . ."); Sam Biddle, *Police Surveilled George Floyd Protests with Help from Twitter-Affiliated Startup Dataminr*, THE INTERCEPT (July 9, 2020, 8:00 PM), https://theintercept.com/2020/07/09/twitter-dataminr-police-spy-surveillance-black-lives-matter-protests/; Bill Dries, *Police Documents Show Protest Spreadsheet and Fear of 'Radical'*, MEMPHIS DAILY NEWS (July 31, 2018), https://www.memphisdailynews.com/news/2018/jul/31/police-documents-show-protest-spreadsheet-and-fear-of-radicals//print.

12. Kate Knibbs, *The Race to Preserve the DC Mob's Digital Traces*, WIRED (Jan. 7, 2021, 5:40 PM), https://www.wired.com/story/archive-social-media-footage-pro-trump/; *see infra* text accompanying notes 122–129.

13. *See, e.g.*, Mike Masnick, *More Police Admitting That FOSTA/SESTA Has Made It Much More Difficult to Catch Pimps and Traffickers*, TECHDIRT, https://www.techdirt.com/articles/20180705/01033440176/more-police-admitting-that-fosta-sesta-has-made-it-much-more-difficult-to-catch-pimps-traffickers.shtml (last visited Mar. 25, 2021).

14. *See, e.g.*, Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. PA. L. REV. 11 (2006); Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296 (2013); Derek E. Bambauer, *Against Jawboning*, 100 MINN. L. REV. 51 (2015); Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035 (2017).

15. Daniel J. Solove, *The First Amendment as Criminal Procedure*, 82 N.Y.U. L. REV. 112 (2007); Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change Special Feature: Cyberlaw*, 70 MD. L. REV. 614 (2010–11); Jonathan Manes, *Online Service Providers and Surveillance Law Transparency*, 125 YALE L.J. F. 343 (2015–16); Alan Z. Rozenshtein, *Surveillance Intermediaries*, 70 STAN. L. REV. 99 (2018); Rebecca Wexler, *Privacy as Privilege: The Stored Communications Act and Internet Evidence*, 134 HARV. L. REV. 2721 (2021).

Reidenberg suggested in *Lex Informatica*, these two conceptions address two siloed visions of the role of platforms in constituting and governing the public sphere.[16]

This Article makes three contributions. First, it complicates existing narratives about the respective roles of social media platforms and law enforcement agencies regarding effective policing. Second, the Article maps how law enforcement both influences and relies upon platform content governance. Although law enforcement seeks to influence *lex informatica*, the substantive, procedural, and technical rules of platforms also shape law enforcement itself. Finally, the Article examines the implications of this public-private cooperation for the law of criminal procedure. Understanding how (public) law enforcement and (private) platform rules mutually inform and co-constitute each other complicates the existing division in U.S. law between state and private action.[17]

The rest of the Article proceeds in three parts. Part II reviews how protections from intermediary liability encouraged the development of private platform governance through technology, even as law enforcement needs remained a powerful influence on firms. Today, contemporary debates over changes to intermediary liability rules highlight the risk that new regulations might promote state censorship laundered through private actors.[18] Even without legal change, however, recent decisions by payment processors and social media platforms reflect the continuing influence of law enforcement even in "private" domains, as Part II.B recounts.

Part III develops the idea that the emergence of online commerce and communication has fundamentally reshaped law enforcement investigative practices. Part III further illustrates that the technological affordances of platforms drive policing's appetite for more data.[19] And while platforms sometimes constrain policing through privately developed policy, the mechanisms of private governance also advance law enforcement strategies.[20] The result is that the technological modalities of governing online content also

16. Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553, 554 (1997) (distinguishing between "[t]he treatment of content" and "the treatment of personal information").

17. *Cf.* Kiel Brennan-Marquez, *The Constitutional Limits of Private Surveillance*, 66 U. KAN. L. REV. 485, 490 (2018) (arguing that whether surveillance is conducted by state or private actors may not matter if it "threaten[s] the integrity of social life").

18. *See infra* Part II.C.

19. *See infra* Part III.A.

20. *See infra* Part III.B.

have increasing resonance for law enforcement investigations despite their "private" character.[21]

Part IV considers the implications of the increasing enmeshment of private platforms and law enforcement for the law of criminal procedure. The turn toward automated modalities of content governance will create new types and sources of information relevant to new kinds of investigations.[22] Yet more extensive collaboration between law enforcement and platforms will raise difficult questions about how best to vindicate important accountability and transparency values when private firms play an increasingly significant role in facilitating public functions.

II. POLICING'S INFLUENCE ON PLATFORMS

People use social media to keep up with their friends and family, watch music and cooking videos, and consume news and political commentary. But social media is also home to a slew of unlawful content. For example, YouTube hosts videos that infringe copyright,[23] Facebook Marketplace features posts advertising drugs, sex, and guns,[24] and Twitter is home to coded posts advertising child sexual abuse imagery.[25] Yet under Section 230 of the Communications Decency Act of 1996, none of these sites can be held liable for hosting content that violates the law, with only a few exceptions.[26] This Part explores how, in spite of existing protections insulating them from liability, platforms have developed many formal and informal mechanisms for advancing law enforcement interests. Although not uniform, these mechanisms illustrate that platforms frequently accommodate law

21. *Id.*

22. *See infra* Part III.

23. *See* Kristelia García, *Monetizing Infringement*, 54 U.C. DAVIS L. REV. 265, 286 (2020) (describing how the scale of copyright infringement on YouTube led the platform to develop Content ID, an automated content screening tool).

24. Parmy Olson & Zusha Elinson, *Gun Sellers are Sneaking Onto Facebook's Booming Secondhand Marketplace*, WALL ST. J. (Aug. 20, 2019, 5:57 PM), https://www.wsj.com/articles/gun-sellers-are-sneaking-onto-facebooks-booming-secondhand-marketplace-11566315198; Ananya Bhattacharya, *Facebook's New Marketplace is Already Flooded with Illegal Guns, Drugs, Sex, and Wildlife*, QUARTZ, https://qz.com/799943/facebooks-fb-new-marketplace-is-already-flooded-with-illegal-guns-drugs-sex-and-wildlife/ (last visited July 15, 2021).

25. Olivia Solon, *Child Sexual Abuse Images and Online Exploitation Surge During Pandemic*, NBC NEWS (Apr. 23, 2020, 9:01 PM EST), https://www.nbcnews.com/tech/tech-news/child-sexual-abuse-images-online-exploitation-surge-during-pandemic-n1190506.

26. 47 U.S.C. § 230; *see also* Danielle Keats Citron & Benjamin Wittes, *The Internet will Not Break: Denying Bad Samaritans Sec. 230 Immunity*, 86 FORDHAM L. REV. 401, 403 (2017) (describing how Section 230 immunity has been extended to "immunize platforms dedicated to abuse and others that deliberately host users' illegal activities").

enforcement needs (although, at times, they also resist law enforcement demands).

A. INTERMEDIARY PROTECTION AND PRIVATE GOVERNANCE

While Section 230(c)(1) immunizes platforms from liability for most content posted by users,[27] Section 230(c)(2)'s "Good Samaritan" provision also protects providers that restrict "obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable" content.[28] The Good Samaritan provision explicitly immunizes online platforms that choose to edit or curate content in ways that would violate the First Amendment if done by the government itself.[29]

The result is that Section 230 ranks "among the most important protections of free expression in the United States in the digital age."[30] It also set the stage for the emergence and growth of what Joel Reidenberg called "lex informatica."[31] Section 230's Good Samaritan provision created breathing room within which self-regulation and private standard setting became the norm.[32] Without the obligation to monitor, filter, or block content, intermediaries nonetheless began to do so, developing both new rules to shape their communities and new enforcement technologies.[33] The example of spam filtering is illustrative: facing a flood of unsolicited commercial advertising,

27. 47 U.S.C. § 230(c)(1) ("No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.").

28. 47 U.S.C. § 230(c).

29. *See, e.g.*, United States v. Stevens, 559 U.S. 460, 468 (2010) (concluding that depictions of animal cruelty do not fall into a category of speech that is unprotected by the First Amendment); FACEBOOK COMMUNITY STANDARDS, *Coordinating Harm and Publicizing Crime*, https://www.facebook.com/communitystandards/coordinating_harm_publicizing_crime/ (banning content "depicting, admitting to or promoting[,] [a]cts of physical harm against animals"); s*ee also* Domen v. Vimeo, Inc., 991 F.3d 66, 68 (2021) (reasoning that Section 230(c)(2) protects online video hosting service from liability when it deletes a user account that violates its policy against the promotion of conversion therapy).

30. Balkin, *supra* note 14, at 2313.

31. Reidenberg, *supra* note 16, at 555; *see also* Jack L. Goldsmith, *Against Cyberanarchy*, 65 U. CHI. L. REV. 1199, 1215–16 (1998) ("Private legal ordering thus has the potential to resolve many, but not all, of the challenges posed by multijurisdictional cyberspace activity.").

32. Reidenberg, *supra* note 16, at 583 ("Law may encourage the development of Lex Informatica by imposing liability on various network actors, and law may provide immunity or safe harbors for implementation of technical rules."); *see also* Klonick, *supra* note 5, at 1603–04 (linking private governance to legal immunity).

33. *See, e.g.*, Bloch-Wehba, *supra* note 10, at 52 (describing the development of spam filtering).

platforms developed anti-spam rules, protocols, and software to filter out unwanted ads.[34]

But while protections for intermediaries allowed "private" regulation to flourish, formal immunity from liability does not equate to immunity from government pressure.[35] Even with Section 230's liability shield intact, government agencies often engage in efforts to coerce, compel, or convince intermediaries to take down harmful content or provide information about the users who posted it.[36] These dynamics may transform online intermediaries into engines of unaccountable private censorship. Scholars of free speech worry that in controversial cases, the government might pressure online intermediaries to go along with the state's own preferences for online speech, a form of "soft censorship" or "jawboning."[37]

Take the example of drill music, a genre of rap pioneered on Chicago's South Side and popular in its own right in the United Kingdom.[38] To earn a living, drill artists rely on social media to distribute music videos that contain "morally charged caricatures of themselves," replete with guns, violent lyrics, and drugs.[39] But drill music's violent content and links to offline crime have also earned it the attention of law enforcement.[40] During a rise in violent crime

34. *Id.* at 55 (describing how platforms turned to automated technology to scale the fight against spam but adopted different definitions of prohibited spam activity).

35. Balkin, *supra* note 14, at 2314 ("What a system of intermediary immunities and safe harbors does not protect, however, constitutes a system of intermediary liability and, hence, of potential collateral censorship."); *see also* Chris Montgomery, *Can Brandenburg v. Ohio Survive the Internet and the Age of Terrorism?: The Secret Weakening of a Venerable Doctrine*, 70 OHIO ST. L.J. 141, 168–78 (2009) (describing how law enforcement has encouraged voluntary action by ISPs and communications service providers).

36. Jonathan Zittrain, *Internet Points of Control*, 44 B.C. L. REV. 653, 674 (2003) (describing Pennsylvania law that required internet service providers to remove or block access to child pornography within five business days); Bambauer, *supra* note 14, at 67–68 (recounting how, under pressure from law enforcement institutions, states adopted laws meant to hold Backpage.com liable for posts submitted by users, knowing that those laws were likely unenforceable).

37. Derek E. Bambauer, *Orwell's Armchair*, 79 U. CHI. L. REV. 863, 905 (2012) (describing process-oriented problems with "soft censorship"); Bambauer, *supra* note 14, at 61 (defining "jawboning" as "enforcement through informal channels, where the underlying authority is in doubt").

38. Lambros Fatsis, *Policing the Beats: The Criminalisation of UK drill and Grime Music by the London Metropolitan Police*, 67 SOCIO. REV. 1300, 1302 (2019); Ben Beaumont-Thomas, *Is UK Drill Music Really Behind London's Wave of Violent Crime?* (Apr. 9, 2018), http://www.theguardian.com/music/2018/apr/09/uk-drill-music-london-wave-violent-crime.

39. FORREST STUART, BALLAD OF THE BULLET: GANGS, DRILL MUSIC, AND THE POWER OF ONLINE INFAMY 6 (2020).

40. *YouTube Must Crack Down on Videos Pushing Violence & Knife Crime*, MAYOR OF LONDON (Aug. 7, 2017), https://www.london.gov.uk//city-hall-blog/youtube-must-crack-down-videos-pushing-violence-knife-crime.

in London, Cressida Dick, the Commissioner of the Metropolitan Police, began to pressure social media platforms to take down UK drill videos, citing their relationship to knife crime.[41] In response, YouTube began aggressively taking down drill videos pursuant to police requests and developed special policies "specifically to help tackle videos related to knife crime in the UK."[42] YouTube also embraced close relationships with the police, publicizing its "dedicated process for the police to flag videos directly to [YouTube's] teams."[43] From the police perspective, stemming the dissemination of drill videos was only one part of a multiprong strategy. Police also obtained a "criminal behavior order" enjoining five people from "mentioning death or injury" in their online videos.[44] The Metropolitan Police also announced that it was indexing and tracking an extensive list of drill videos.[45]

Sometimes, however, platforms push back against government demands. Consider, for example, the infamous "Innocence of Muslims" video, an Islamophobic "film" that sparked violent protests across the world and reportedly led to the attack on the U.S. consulate in Benghazi, Libya.[46] As violence spread, the White House reportedly called YouTube to ask the firm

41. *Met Police Chief Calls on YouTube to Take Down Drill Music to Curb Gang Crime*, LBC (May 18, 2018, 9:17 AM), https://www.lbc.co.uk/radio/presenters/nick-ferrari/met-police-chief-calls-on-youtube-drill-music/.

42. Lizzie Dearden, *Police Targeting Drill Music Videos in Controversial Crackdown on Social Media That 'Incites Violence'*, THE INDEPENDENT (May 29, 2018, 12:04 AM), https://www.independent.co.uk/news/uk/crime/drill-music-stabbings-london-youtube-violence-police-knife-crime-gangs-a8373241.html; Jim Connolly, *Home Secretary: 'Sweep the Net, Take Down Knife-crime Posts'*, BBC NEWS: NEWSBEAT (Feb. 13, 2019), https://www.bbc.com/news/newsbeat-47211631; Ed Clowes, *For British Drill Stars, the Police are Listening Closely*, N.Y. TIMES (Jan. 11, 2021), https://www.nytimes.com/2021/01/11/arts/music/digga-d-drill-music.html (charting rise in YouTube's takedown numbers).

43. Dearden, *supra* note 42.

44. *Ladbroke Grove Banned From Making 'Violent Drill Music'*, BBC NEWS (June 15, 2018), https://www.bbc.com/news/uk-england-london-44498231; Lanre Bakare, *'New Stop and Search': Rappers Condemn Police Over Drill Bans*, THE GUARDIAN (June 14, 2019), http://www.theguardian.com/music/2019/jun/14/rappers-konan-krept-condemn-police-criminalisation-of-drill (describing how two rappers were sentenced to prison for breaching a gang injunction prohibiting them from performing violent lyrics).

45. Jim Edwards, *YouTube Deleted 130 Rap Videos to Help Police Fight Street Gangs Responsible for Thousands of Stabbings*, BUS. INSIDER (June 29, 2019, 12:52 PM), https://www.businessinsider.com/uk-drill-rap-videos-banned-by-police-2019–6 (quoting Met police as saying that their database contained over 2,000 music videos, while they had filed only 154 takedown requests with YouTube).

46. Michael Joseph Gross, *The Making of The Innocence of Muslims: Cast Members Discuss the Film That Set Fire to the Arab World*, VANITY FAIR (Dec. 27, 2012), https://www.vanityfair.com/culture/2012/12/making-of-innocence-of-muslims.

to review whether the video complied with its terms of service.[47] President Obama told 60 Minutes that while "we believe in the First Amendment," the film "is not representative of who we are and our values."[48] Civil liberties advocates chafed at the White House's use of quasi-official channels to pressure YouTube to take down the offensive but lawful video, and YouTube ultimately resisted the calls to take the video down for a U.S. audience.[49]

These two illustrations demonstrate YouTube's power to either facilitate or obstruct law enforcement priorities. In the "Innocence of Muslims" case, YouTube's own content-related policies led it to resist the White House's encouragement to take down the video. Across the pond, however, YouTube created new content-related rules at law enforcement's behest, offering itself as a vital partner to police. In both cases, YouTube's decisions were formally voluntary, free of government coercion.[50]

B. FORMAL INDEPENDENCE, INFORMAL ENTANGLEMENT

Notwithstanding platforms' status as private actors, government preferences continue to shape firms' internal content moderation systems, rules, and practices in a more general sense. Yet these kinds of pressures rarely amount to the kind of government coercion extensive enough to amount to a plausible First Amendment claim.[51] It can be difficult to draw a line between changes to content-related decisions that occur because of jawboning and those that occur because of reputational or business risk.[52] Using the examples of terrorist content and sex work, this subpart shows that firms' behavior might be attributed as much to political climate as to unambiguous legal obligations.

1. Terrorist Content

Platforms have touted their ability to use artificial intelligence, automation, and hash matching to detect and prevent the dissemination of online terrorist content, advertising their abilities to proactively remove ISIS and al-Qaeda

47. Josh Gerstein, *Activists Troubled by White House Call to YouTube*, POLITICO (Sept. 14, 2012, 4:42), https://www.politico.com/blogs/under-the-radar/2012/09/activists-troubled-by-white-house-call-to-youtube-135618.

48. *Id.*

49. *Id.*

50. However, as I have previously argued elsewhere, threats of regulation can also generate "voluntary" proactive measures by platforms. Bloch-Wehba, *supra* note 10, at 58.

51. *See* Montgomery, *supra* note 35, at 172–73.

52. Kreimer, *supra* note 14, at 50; Ass'n of Am. Physicians & Surgeons v. Schiff, CV 20-106 (RC), 2021 WL 354174, at *6 (D.D.C. Feb. 2, 2021) (concluding that plaintiffs could not demonstrate that congressional statements led to private action by social media companies that lessened traffic to plaintiffs' website).

terrorist content.[53] Notwithstanding claims of technical sophistication, however, critics have observed that platforms continued to allow designated foreign terrorist organizations such as Hamas, Hezbollah, and the FARC to maintain profiles and post content online, long after becoming aware of their activities.[54] Still, the numerous attempts to hold social media companies liable for the proliferation of online terrorist content have been unsuccessful.[55]

From one perspective, in the absence of liability, social media firms have allowed themselves to be used as conduits for terrorist speech.[56] At the same time, however, firms have continued to engage in what Alexander Tsesis calls "corporate self-policing."[57] For example, Zoom cancelled a 2020 San Francisco State University event with Leila Khaled, a Palestinian activist and member of the Popular Front for the Liberation of Palestine, a designated foreign terrorist organization.[58] Zoom argued that providing the platform for the talk would have violated federal laws prohibiting providing material support to terrorist organizations.[59] Similarly, in 2021, Google reportedly terminated the account of an activist sharing materials regarding Palestine on Google Drive, also citing violations of terrorism laws.[60]

In many instances, these decisions go above and beyond what the law appears to require. In numerous cases, courts have held that platforms are not civilly liable when their services are used by terrorists.[61] While regulators have pressured platforms to take more proactive steps to address terrorist content, platforms' blunt approaches to removing terrorist content also risk over

53. Bloch-Wehba, *supra* note 10, at 59.

54. Citron & Wittes, *supra* note 26, at 403; Luis Jaime Acosta, *Social Networks Clamp Down on Colombian FARC Dissident Accounts*, REUTERS (Jan. 15, 2021, 11:40 AM), https://www.reuters.com/N/us-twitter-colombia-idUSKBN29K2HI.

55. *See, e.g.*, Fields v. Twitter, 881 F.3d 739 (9th Cir. 2018); Gonzalez v. Google, 282 F. Supp. 3d 1150 (N.D. Cal. 2017); Crosby v. Twitter, 303 F. Supp. 3d 564 (E.D. Mich. 2018); Force v. Facebook, 934 F.3d 53 (2d Cir. 2019).

56. Alexander Tsesis, *Social Media Accountability for Terrorist Propaganda*, 86 FORDHAM L. REV. 605, 611 (2017–18).

57. *Id.* at 613.

58. Alice Speri & Sam Biddle, *Zoom Censorship of Palestine Seminars Sparks Fight Over Academic Freedom*, THE INTERCEPT (Nov. 14, 2020, 4:00 AM), https://theintercept.com/2020/11/14/zoom-censorship-leila-khaled-palestine/.

59. *Id.*

60. @hotgirlhala, TWITTER (Apr. 22, 2021, 2:41 PM), https://twitter.com/hotgirlhala/status/1385212069679702020.

61. *See, e.g.*, Fields v. Twitter, 881 F.3d 739 (9th Cir. 2018); Crosby v. Twitter, 303 F. Supp. 3d 564 (E.D. Mich. 2018), aff'g 921 F.3d 617 (6th Cir. 2019); Force v. Facebook, 934 F.3d 53 (2d Cir. 2019); Sinclair for Tucker v. Twitter, Inc., C 17-5710 SBA, 2019 WL 10252752 (N.D. Cal. Mar. 20, 2019); Clayborn v. Twitter, Inc., 17-CV-06894-LB, 2018 WL 6839754 (N.D. Cal. Dec. 31, 2018).

censorship.[62] At the same time, this focus on Islamic terrorism, and particularly on ISIS and al-Qaeda, led to a severely under inclusive approach to other threats. Like many law enforcement agencies, social media companies paid little attention to problems of White nationalism and White extremism until after the Christchurch attacks in 2019.[63] Even then, platforms' mechanisms to address White nationalism and White supremacy have been haphazard and incomplete.[64]

There are several potential explanations for platforms' voluntary actions to address terrorism. Most notable, perhaps, is the adoption of new regulations in Europe that require platforms to take down terrorist content within an hour, or else face liability.[65] In the United States, other pressures are in play. The threat that social media companies may face potential criminal liability under the material support statutes, as Tsesis and others have urged, may have encouraged platforms to address terrorism more aggressively.[66] Or perhaps the burgeoning calls to rethink Section 230's liability shield have led platforms to be more proactive, even in the absence of regulatory change. But government interests also provide powerful motivation for businesses to address harmful online content even when firms face no legal obligation to do so. Firms'

62. *See* ELEC. FRONTIER FOUN., SYRIAN ARCHIVE & WITNESS, *Caught in the Net: The Impact of Extremist Speech Regulations on Human Rights Content* (2019), https://syrianarchive.org/en/lost-found/impact-extremist-human-rights#content-moderation-and-extremist-content (last visited Aug. 9, 2021) [hereinafter *Caught in the Net*] (describing how reliance on automated tools to block and delete "terrorist content" also suppress human rights reporting, journalism, and other socially valuable posts).

63. Bloch-Wehba, *supra* note 10, at 60; Amna Akbar, *Policing Radicalization*, 3 UC IRVINE L. REV. 809, 827 (2013) (describing how indicators of Muslim religious observance were transmuted into signals of "radicalization").

64. *See, e.g.*, Alex Kaplan, *YouTube Removed Some Channels Affiliated with White Nationalism—But Not All*, MEDIA MATTERS FOR AMERICA, https://www.mediamatters.org/white-nationalism/youtube-removed-some-channels-affiliated-white-nationalism-not-all (last visited June 22, 2021); Julia Carrie Wong, *White Nationalists are Openly Operating on Facebook. The company Won't Act*, THE GUARDIAN (Nov 21, 2019, 11:00 GMT), http://www.theguardian.com/technology/2019/nov/21/facebook-white-nationalists-ban-vdare-red-ice.

65. Regulation 2021/784 of the European Parliament and of the Council of Apr. 29, 2021, On Addressing The Dissemination of Terrorist Content Online ("TERREG"), annex, 2021 O.J. (L 172). In prior work, I have explored how platforms reacted to the emergence of new obligations in Europe, which have since been codified in the TERREG. *See* Bloch-Wehba, *supra* note 8, at 43–48 (detailing the evolution of European rules and platform responses on terrorist content).

66. Tsesis, *supra* note 56, at 625–26 (arguing that the material-support statute could support charges against recalcitrant social media service providers); Benjamin Wittes & Zoe Bedell, *Tweeting Terrorists, Part I: Don't Look Now but a Lot of Terrorist Groups are Using Twitter*, LAWFARE (Feb. 14, 2016, 5:05 PM), https://www.lawfareblog.com/tweeting-terrorists-part-i-dont-look-now-lot-terrorist-groups-are-using-twitter.

takedown priorities appeared to align with the government's law enforcement interests: in the context of a now decades-long war on (Islamic) terror, platforms likewise prioritized takedowns of ISIS and al-Qaeda content.[67]

2. Sex Work

The experience of adult service businesses offers another illustration. In 2013, the Department of Justice initiated what it called "Operation Choke Point," a program meant to encourage financial institutions to take a more active role in curtailing fraudulent businesses' access to the banking system.[68] Critics of the program soon began to worry that banks were also cutting off legitimate businesses that they simply found distasteful, like pornographers, gun dealers, and payday lenders.[69] Although an audit later found that the Federal Deposit Insurance Corporation (FDIC) had not wrongly pressured banks to drop "high-risk" clients, it acknowledged that the agency's regulatory activities "created a perception among some bank executives. . .that the FDIC discouraged institutions" from pursuing or maintaining business relationships with high-risk merchants.[70] In 2017, the Trump administration announced that it would put a stop to Operation Choke Point.[71]

Even in the absence of any legal requirements, many banks and payment processors have chosen to avoid providing services to adult businesses, perhaps because of social pressure or perceptions of other business risks.[72] However, although online payment processors have no legal obligation to deny service to "high-risk" adult services clients, they nevertheless continue to keep

67. Bloch-Wehba, *supra* note 10, at 76–77; Caroline Mala Corbin, *Terrorists are Always Muslim but Never White: At the Intersection of Critical Race Theory and Propaganda*, 86 FORDHAM L. REV. 455, 458–60 (2017) (describing how popular culture, media narratives, and government priorities link "terrorism" to Muslim identity).

68. Richard P. Eckman, Richard J. Zack, Christina O. Hud, Jonathan N. Ledsky & Scott J. Helfand, *Update on the Short-Term Lending Industry: Government Investigations and Enforcement Actions*, 70 BUS. LAW. 657 (2014–2015).

69. Elizabeth Nolan Brown, *DOJ's 'Operation Choke Point' may be Root of Porn Star Bank Account Closings*, REASON.COM (Apr 29, 2014, 8:40 PM), https://reason.com/2014/04/28/doj-operation-chokepoint-and-porn-stars/.

70. FED. DEPOSIT INS. CORP. OFF. OF INSPECTOR GEN., *The FDIC's Role in Operation Choke Point and Supervisory Approach to Institutions that Conducted Business with Merchants Associated with High-Risk Activities*, at 11 (2015), https://www.fdicoig.gov/publications/fdics-role-operation-choke-point-and-supervisory-approach-institutions-conducted (last visited Dec. 29, 2021).

71. Victoria Guida, *Justice Department to End Obama-era 'Operation Choke Point'*, POLITICO (Aug. 17, 2017, 10:41 PM), https://politi.co/2lObBHh.

72. *See* E. Christopher Johnson, Jr., *The Important Role for Socially Responsible Businesses in the Fight Against Human Trafficking and Child Labor in Supply Chains*, BUSINESS LAW TODAY (Jan. 22, 2015), https://www.americanbar.org/groups/business_law/publications/blt/2015/01/02_johnson/.

adult services providers at arm's length, reflecting the perception of legal or business risk caused by providing such services to "high-risk" clients.[73] And payment processors are powerful intermediaries, critical to "people's practical ability to speak"—and in this case, to post photos and videos or to maintain an online presence at all.[74]

New intermediary obligations have heightened the sense that businesses must do more to address adult content. In 2018, Congress enacted the Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA), a statute designed to promote platform accountability for sex trafficking.[75] As Eric Goldman has written, FOSTA responded to an apparent accountability gap that had allowed Backpage, a website primarily used for commercial sex advertising, to profit from advertisements of trafficking victims.[76] FOSTA expanded federal criminal liability for sex trafficking and for intentionally promoting or facilitating prostitution through interactive computer services.[77] Yet recent reporting suggests that FOSTA has hardly changed prosecutors' ability to charge and convict sex traffickers. A 2021 Government Accountability Office report indicates that, in the past three years, prosecutors had only brought one case under FOSTA's criminal provision.[78] In addition, as of June 2021, civil damages have never been awarded under FOSTA.[79]

Despite its apparently sparse impact on criminal and civil liability, FOSTA clearly discouraged online platforms from hosting sexual content. In the wake of FOSTA's passage, as Goldman recounts, several online service providers determined that they could no longer bear the risk of hosting *any* adult content at all. In 2018, online marketplace Craigslist stopped hosting personal ads entirely, citing the risk of criminal liability under FOSTA if adult content was

73. Sarah Manavis, *The PayPal ASMR banning Shows Us that Tech Companies Don't Understand Their Users*, NEWSTATESMAN (Sept. 20, 2018), https://www.newstatesman.com/science-tech/technology/2018/09/paypal-asmr-ban-youtube-monetise-patreon (describing broad application of PayPal's sexual content policy); Margot Cleveland, *How Mastercard's Rules Could be Used to Ban Conservatives from Banking*, FEDERALIST (Ap. 19, 2021), https://thefederalist.com/2021/04/19/how-mastercards-rules-against-child-pornographers-could-be-used-to-ban-conservatives-from-banking/.

74. Jack M. Balkin, *Free Speech is a Triangle*, 118 COLUM. L. REV. 2011, 2014–15 (2018).

75. *See* Aja Romano, *A new law Intended to Curb Sex Trafficking Threatens the Future of the Internet as we Know It*, VOX (updated July 2, 2018, 1:08 PM EDT), https://www.vox.com/culture/2018/4/13/17172762/fosta-sesta-backpage-230-internet-freedom.

76. Eric Goldman, *The Complicated Story of Fosta and Section 230*, 17 FIRST AMEND. L. REV. 279, 281 (2018).

77. *Id.* at 284; 18 U.S.C. § 2421A.

78. *Sex Trafficking: Online Platforms and Federal Prosecutions* 25–26, Gov't Accountability Office, GAO-21-385 (June 21, 2021), https://www.gao.gov/products/gao-21-385.

79. *Id.*

posted.[80] Even OnlyFans, the pay-per-view website known for risqué content, has a strict policy against escorts that has dramatically affected the livelihoods of sex workers.[81] While OnlyFans hosts adult content by amateurs and celebrities, sex workers report being shunned by the platform, perhaps because of its assessment of the risk of liability under FOSTA.[82] More broadly, sex workers have reported that FOSTA's enactment has "increased their exposure to violence and left those who rely on sex work as their primary form of income without many of the tools they had used to keep themselves safe."[83]

Broadly speaking, then, both the examples of terrorist content and sex work illustrate that, even without an obvious enforcement mechanism, laws can encourage platforms to take aggressive private action against certain forms of speech, in alignment with government's own priorities. Some scholars might view this as a form of coercion or "jawboning," as Derek Bambauer and others have argued.[84] But others might describe platforms' actions here as the result of a more subtle form of government influence rather than a clear result of ham-fisted proxy censorship.[85] And when private incentives align with public policy, private governance provides a powerful new mechanism by which government can obtain its desired results without costly inconveniences such as accountability or oversight.

---

80. Merrit Kennedy, *Craigslist Shuts Down Personals Section After Congress Passes Bill on Trafficking*, NPR (Mar. 23, 2018, 3:52 PM), https://www.npr.org/sections/thetwo-way/2018/03/23/596460672/craigslist-shuts-down-personals-section-after-congress-passes-bill-on-traffickin; *see also* Heidi Tripp, *All Sex Workers Deserve Protection: How FOSTA/SESTA Overlooks Consensual Sex Workers in an Attempt to Protect Sex Trafficking Victims*, 124 PENN. ST. L. REV. 219 (2019) ("[M]any ISPs completely shut down certain services on their websites or began over-censoring content beyond what was necessary to comply with FOSTA/SESTA.").

81. Shae Ashbury, *How OnlyFans Steals from Sex Workers and Fans*, (Aug. 13, 2019), https://www.shae-ashbury.com/shae-ashburys-blog/2019/8/13/how-onlyfans-steals-from-sex-workers; Mark Serrels, *Thanks to US laws, Sex Workers are Fighting to Stay Online,* CNET (Feb. 26, 2021), https://www.cnet.com/features/thanks-to-us-laws-sex-workers-are-fighting-to-stay-online/.

82. Natalie Jarvey, *How OnlyFans Has Become Hollywood's Risque Pandemic Side Hustle*, HOLLYWOOD REP. (Dec. 11, 2020, 7:00 AM), https://www.hollywoodreporter.com/news/how-onlyfans-has-become-hollywoods-risque-pandemic-side-hustle; *see also* Alexis Okeowo, *The Fragile Existence of Sex Workers During the Pandemic*, NEW YORKER (May 21, 2021), https://www.newyorker.com/news/news-desk/the-fragile-existence-of-sex-workers-during-the-pandemic (describing how sex workers began to post on OnlyFans after SESTA-FOSTA); Serrels, *supra* note 81.

83. Danielle Blunt & Ariel Wolf, *Erased: The Impact of FOSTA-SESTA & the Removal of Backpage*, HACKING//HUSTLING 1 (2020), https://hackinghustling.org/erased-the-impact-of-fosta-sesta-2020/.

84. *See* Bambauer, *supra* note 37, at 891–99, 943.

85. *See, e.g.*, Janice Nadler, *Expressive Law, Social Norms, and Social Groups*, 42 L. & SOC. INQUIRY 60, 64 (2017).

C. NEW INCENTIVES FOR PLATFORMS?

At one level, platforms have seemed eager to demonstrate their willingness and capacity to carry out government priorities through private policing. Yet governments (particularly outside of the United States) have also struggled to incentivize platforms to address unlawful content more aggressively. In the aftermath of the March 2019 massacre at two Christchurch mosques, governments proposed and adopted new legislation imposing penalties on online platforms that fail to remove unlawful content.[86]

Both law enforcement and platforms see the potential for artificial intelligence and other automated techniques to enhance compliance with these measures and speed up takedowns. In Australia, for example, the law now imposes criminal penalties on providers of online services that do not remove "abhorrent violent material" "expeditiously."[87] In Germany, the Network Enforcement Act of 2018 similarly requires platforms to quickly remove unlawful content, sometimes within 24 hours, or pay large fines.[88] The European Union recently finalized its regulation on terrorist content online, which will not only require platforms to take down terrorist content more quickly, but also require them to adopt more proactive measures to prevent the spread of terrorist content in the first place.[89]

These kinds of pressures have led free speech advocates and scholars to see in government regulatory proposals the clear threat of proxy censorship. The dominant accounts of law enforcement interests in this space describe governments as seeking more extensive takedowns, more limits on speech, and more aggressive enforcement of private and public rules, while platforms resist the imposition of these and similar obligations.[90] Faced with platforms' independence and immunity from liability, governments seek to require them to behave more aggressively in filtering out unlawful content, ideally through adopting new technologies of decision-making.

---

86. Evelyn Douek, *Australia's New Social Media Law is a Mess*, LAWFARE (Apr. 10, 2019, 8:28 AM), https://www.lawfareblog.com/australias-new-social-media-law-mess.

87. Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019 (Cth) § 474.34 (Austl.).

88. Evelyn Douek, *Germany's Bold Gambit to Prevent Online Hate Crimes and Fake News Takes Effect*, LAWFARE (Oct. 31, 2017, 11:30 AM), https://www.lawfareblog.com/germanys-bold-gambit-prevent-online-hate-crimes-and-fake-news-takes-effect.

89. Regulation 2021/784 of the European Parliament and of the Council of Apr. 29, 2021, On Addressing The Dissemination of Terrorist Content Online ("TERREG"), annex, 2021 O.J. (L 172).

90. *See generally* Evelyn Douek, *Australia's "Abhorrent Violent Material" Law: Shouting "Nerd Harder" and Drowning Out Speech*, 94 AUSTL. L. REV. 41 (2020).

Yet powerful interests also cut in the opposite direction, encouraging platforms to keep unlawful content online as a form of intelligence for law enforcement to mine. The enactment of FOSTA has reportedly made it much more difficult for law enforcement to investigate and detect sex trafficking victims and perpetrators.[91] Similarly, mechanisms for removing terrorist content have diminished the availability of human rights reporting online.[92] By contrast, law enforcement has a strong interest in maintaining access to social media's trove of online evidence. The more aggressively social media platforms enforce their private rules, whether through automated technology or through manual review, the harder it becomes for law enforcement to conduct this kind of surveillance.[93]

III. PLATFORMS' INFLUENCE ON POLICING

The communicative and data-generating affordances of online platforms change user behavior and create legal challenges.[94] In turn, they also drive investigative strategy.[95] In the previous Part, I demonstrated that law enforcement sometimes seeks to control or influence the affordances of social media platforms, especially when it comes to dangerous or violent speech. But as this Part shows, the content-related decision-making of platforms also benefits law enforcement, creating new sources of information with new affordances for investigating online speech. As a result, police increasingly depend upon purportedly private content moderation rules, strategies, and techniques, and platforms have a growing role in facilitating law enforcement surveillance. The aim here is to complicate what has become a binary distinction between platform and government and illustrate the mutual entanglements of the two.

91. *See* Appellant's Br. at 54, Woodhull v. DOJ, No. 18-5298 (D.C. Cir. filed Feb. 13, 2019); Masnick, *supra* note 13.

92. *Caught in the Net*, *supra* note 62.

93. *See infra* text accompanying notes 110–111.

94. For example, the ability to upload user-generated content to YouTube has facilitated widespread copyright infringement. García, *supra* note 23, at 285–86. Surprisingly, rightsholders have sometimes encouraged infringement because they benefit from the free publicity and "Internet buzz." *Id.* at 298–99.

95. *Cf.* Orin S. Kerr, *An Equilibrium-Adjustment Theory of the Fourth Amendment*, 125 HARV. L. REV. 476, 482 (2011) (describing how "changing technology" and "social practice" might impel courts to respond by ratcheting up or down the rules that constrain police power).

A. SHAPING LAW ENFORCEMENT THROUGH TECHNOLOGY

It is widely appreciated that private governance plays an increasing role in state policy.[96] As Jack Balkin has observed, the growing capacity of internet firms to surveil and control content has also made them "more valuable targets" for regulation.[97] Perhaps less appreciated, however, is the degree to which the affordances of networked technologies increasingly shape law enforcement practices themselves. Design choices dictate what information is available to law enforcement—and thus what information law enforcement can demand and use in investigative contexts.[98] Law enforcement is engaged in a form of what Marion Fourcade and Jeffrey Gordon call "dataist statecraft," in which the availability of data minted by both public and private actors drives policy.[99]

1. Compelled Disclosure

Historically, many of the fights about law enforcement access to user information have been about compelled disclosure of customer records and communications.[100] The law of compelled disclosure governs the standards by which law enforcement can obtain access to different categories of user information in the possession of firms. For example, the Stored Communications Act (SCA) imposes a warrant requirement for communications that have been stored in an electronic communications system for 180 days or less.[101] If communications have been stored for greater than 180 days, then the government can seek access using a subpoena, court order, or a search warrant, accompanied by different indicia of suspicion and different notice obligations.[102] This legal structure has generated numerous

96. Balkin, *supra* note 74, at 2028 ("[N]ew-school speech regulation depends on the expansion and promulgation of private governance."); Robert Gorwa, *The Platform Governance Triangle: Conceptualising the Informal Regulation of Online Content*, 8 INTERNET POL'Y REV. 1, 7 (2019) ("[I]nformal regulatory arrangements have formed a key tool through which governance stakeholders—especially EU governments—have sought to shape the behaviour of firms on content issues.").

97. Balkin, *supra* note 74, at 2020.

98. Andrew Guthrie Ferguson, *Structural Sensor Surveillance*, 106 IOWA L. REV. 47, 54 (2020) ("[D]esign choices can directly impact the usefulness of the data collected.").

99. Marion Fourcade & Jeffrey Gordon, *Learning Like a State: Statecraft in the Digital Age*, 1 J. L. & POL. ECON. 78, 78 (2020).

100. *See generally* Kerr, *supra* note 15, at 1209–12 (describing how ambiguities in the Fourth Amendment's application to the Internet fostered legal uncertainty about compelled disclosure of user communications); *see also In re* 381 Search Warrants Directed to Facebook, Inc., 29 N.Y.3d 231 (2017).

101. 18 U.S.C. § 2703(a).

102. 18 U.S.C. § 2703(b); *see also* Hannah Bloch-Wehba, *Exposing Secret Searches: A First Amendment Right of Access to Electronic Surveillance Orders*, 93 WASH. L. REV. 145, 160 (2018)

legal battles regarding whether disclosure of a range of information—including historic cell site location data, web browsing histories, and the contents of emails—implicates the Fourth Amendment's definition of a search or seizure.[103]

Focusing on the appropriate standard for defining a government search, though, threatens to miss the degree to which the availability of networked technologies itself drives law enforcement strategy. Digital searches and seizures have vastly grown in number, reflecting the increased relevance of digital communications in investigations, the growing scale of networked technology applications and services, and the proliferation of different forms of information.[104] As the number of requests for user information has increased, the role of electronic communications service providers in facilitating, obstructing, and enabling surveillance has also grown apace.[105] Providers such as Google, Facebook, and Microsoft have large in-house compliance teams in order to process a growing number of law enforcement requests for customer data.[106]

Networked technologies are not only driving an increase in the degree of law enforcement surveillance and control; they are also fundamentally transforming the work of law enforcement. Consider the surveillance of cell site location information. In the last decade, significant ink has been spilled regarding law enforcement's acquisition of cell phone location information through real-time tracking, historic location data, cell tower dumps, cell site simulators, and data purchases.[107] In 2018, the Supreme Court decided in

(describing the different notice provisions for different forms of legal process to compel disclosure of customer records and communications).

103. Carpenter v. United States, 138 S. Ct. 2206, 2217 (2018) (holding that a disclosure of a week of cell site location information is a search); United States v. Forrester, 512 F.3d 500, 511 (9th Cir. 2008) (holding that acquiring IP addresses of websites user visited is not a search); United States v. Warshak, 631 F.3d 266, 288 (6th Cir. 2010) (holding that a disclosure of email contents is a search).

104. Rozenshtein, *supra* note 15, at 109 (arguing that digital intermediaries are "more central than ever to government surveillance").

105. *See id.* at 114; Manes, *supra* note 15, at 348.

106. *See, e.g.*, *Facebook Transparency Report*, https://transparency.facebook.com/government-data-requests/country/US (last visited June 21, 2021) (documenting a rise in the number of requests from 11,000 in the first half of 2013 to over 61,000 in the first half of 2020); *Google Transparency Report*, https://transparencyreport.google.com/user-data/overview (last visited July 15, 2021) (documenting a rise in the number of requests from 25,000 in the first half of 2013 to over 100,000 in the first half of 2020).

107. Brian L. Owsley, *The Fourth Amendment Implications of the Government's Use of Cell Tower Dumps in its Electronic Surveillance*, 16 UNIV. PA. J. CONST. L. 1 (2013); Stephen Wm. Smith, *Gagged, Sealed & Delivered: Reforming ECPA's Secret Docket*, 6 HARV. L. & POL'Y REV. 601 (2012); Andrew Guthrie Ferguson, *The Internet of Things and the Fourth Amendment of Effects*, 104 CALIF.

Carpenter v. United States that obtaining over six days of historical cell site location from a cell phone service provider constitutes a search for Fourth Amendment purposes.[108] However, the *Carpenter* Court expressly declined to decide whether cell tower dumps, which collect all the phone numbers that connected to a given cell tower during a given time period, held the same Fourth Amendment implications.[109] Soon afterward, law enforcement began to seek so-called geofence or reverse location information—information pertaining to every user in a given geographical radius during a given time period—from Google.[110] Google's location tracking—infamously difficult to turn off or opt out of—becomes the new equivalent of the cell tower.[111] Networked technologies, by design, collect and retain information from large numbers of users, in turn driving law enforcement to seek more data from these sources.[112]

Likewise, the emergence of the so-called Internet of Things and omnipresent embedded sensors are equally responsible for novel transformations in investigative strategy.[113] Law enforcement can now acquire data from connected speakers, fitness trackers, doorbell cameras, and smart streetlights.[114] As a result, consumer technology may drive not only self-

---

L. REV. 805 (2016); Byron Tau, *House Investigating Company Selling Phone Location Data to Government Agencies*, WALL ST. J: POLITICS (June 24, 2020, 3:19 PM), https://www.wsj.com/articles/house-investigating-company-selling-phone-location-data-to-government-agencies-11593026382.

108. Carpenter v. United States, 138 S. Ct. 2206 (2018).

109. *Id.* at 2220; *see also* Owsley, *supra* note 107, at 16–17 (arguing that cell tower dumps are more intrusive than simple pen registers).

110. *See, e.g.*, *In re* Search Warrant Application for Geofence Location, 497 F. Supp. 3d 345 (N.D. Ill. 2020).

111. Ryan Nakashima, *Google Tracks Your Movements, Like it or Not*, AP NEWS (Aug. 14, 2018), https://apnews.com/article/north-america-science-technology-business-ap-top-news-828aefab64d4411bac257a07c1af0ecb.

112. *See, e.g.*, *Carpenter*, 138 S. Ct. at 2218 ("[T]he retrospective quality of the data here gives police access to a category of information otherwise unknowable.").

113. Paul Ohm, *The Many Revolutions of Carpenter*, 32 HARV. J. L. & TECH. 357, 364–65 (2019).

114. Kayla Epstein, *Police Think Amazon's Alexa may have Information on a Fatal Stabbing Case*, WASH. POST (Nov. 2, 2019), https://www.washingtonpost.com/technology/2019/11/02/police-think-amazons-alexa-may-have-information-fatal-stabbing-case/; Christine Hauser, *Police Use Fitbit Data to Charge 90-Year-Old Man in Stepdaughter's Killing*, N. Y. TIMES (Oct. 3, 2018), https://www.nytimes.com/2018/10/03/us/fitbit-murder-arrest.html; John Herrman, *Who's Watching Your Porch?*, N. Y. TIMES (Jan. 19, 2020), https://www.nytimes.com/2020/01/19/style/ring-video-doorbell-home-security.html; Jesse Marx, *Police Used Smart Streetlight Footage to Investigate Protesters*, VOICE SAN DIEGO (June 29, 2020), https://www.voiceofsandiego.org/topics/government/police-used-smart-streetlight-footage-to-investigate-protesters/.

tracking, but law enforcement tracking as well.[115] As discussed in Part IV, networked, sensory technologies do not just create goldmines of information for law enforcement, but also fundamentally alter the legal mechanisms through which policing can be made transparent and accountable to the public.

2. "Open Source" Investigations

The growing role of compelled disclosure in law enforcement investigations illustrates the centrality of networked technology as a mechanism of surveillance, but it is just the tip of the iceberg. Although law enforcement can influence platform rules and practices through either takedown requirements (as in Part II) or compelled disclosure requirements (as in Part III.A), social media can also influence law enforcement by serving as a ready source of open-source information and evidence.

For instance, law enforcement regularly monitors public social media activity in both targeted investigations and as a source of dragnet intelligence.[116] As advocates at the Brennan Center have explained, social media surveillance often occurs when officers "view[] publicly available posts by searching for an individual, group, hashtag, or another search vector."[117] The extent, scope, and manner in which these results might be displayed depends on the affordances of the platform at issue. For example, the U.S. Department of Homeland Security has monitored Black Lives Matter groups and events using Twitter hashtags and location information.[118]

Law enforcement has also used surveillance services such as Geofeedia, Snaptrends, and others to access social media data in an automated fashion.[119]

115. GINA NEFF & DAWN NAFUS, SELF-TRACKING 178 (Mass. Inst. Tech., 2016) (describing the potential legal questions around self-tracking data).

116. Rachel Levinson-Waldman, *Private Eyes, They're Watching You: Law Enforcement's Monitoring of Social Media*, 71 OKLA. L. REV. 997, 999–1000 (2019); Rachel Levinson-Waldman, *Government Access to and Manipulation of Social Media: Legal and Policy Challenges*, 61 HOW. L.J. 523, 541–42 (2018); *see also* Daphna Renan, *The Fourth Amendment as Administrative Governance*, 68 STAN. L. REV. 1039, 1053 (2016) (noting that "generalized collection" can lead to targeted surveillance).

117. Rachel Levinson-Waldman & Ángel Díaz, *How to Reform Police Monitoring of Social Media*, BROOKINGS (July 9, 2020), https://www.brookings.edu/techstream/how-to-reform-police-monitoring-of-social-media/.

118. George Joseph, *Exclusive: Feds Regularly Monitored Black Lives Matter Since Ferguson*, THE INTERCEPT (July 24, 2015, 11:50 AM), https://theintercept.com/2015/07/24/documents-show-department-homeland-security-monitoring-black-lives-matter-since-ferguson/.

119. Matt Cagle, *Facebook, Instagram, and Twitter Provided Data Access for a Surveillance Product Marketed to Target Activists of Color*, ACLU: N. CAL. (Oct. 11, 2016), https://www.aclunc.org/blog/facebook-instagram-and-twitter-provided-data-access-surveillance-product-marketed-target; *see also* Ryan Devereaux, *Homeland Security Used a Private Intelligence Firm to Monitor Family Separation Protests*, THE INTERCEPT (Apr. 19, 2019, 8:25 AM), https://theintercept.com/2019/

Social media surveillance can allow law enforcement agencies to assess social media information for potential risks and threats and map connections between investigative targets and other subjects.[120] On the other hand, police do not always recognize the gravity of online threats or "chatter." On January 5, 2021, Dataminr reached out to police at the U.S. Capitol to notify them of an uptick in chatter regarding the upcoming riots, but law enforcement reportedly took no preparatory action.[121]

To some extent, the emergence of third-party social media surveillance tools like Dataminr and Geofeedia is a direct response to platform firms' user affordances and content policies. In April 2021, the New York Police Department (NYPD) published a draft "impact and use policy" for public comment on NYPD's social media surveillance systems.[122] The policy stressed that NYPD only accesses "publicly available information, or information that is viewable as a result of user privacy settings or practices." However, the policy also explained that third party surveillance tools help to fill critical investigative gaps that result when users or platforms delete content relevant to an investigation.[123]

04/29/family-separation-protests-surveillance/; Colin Daileda, *Twitter Cuts Ties with Another Social Media Surveillance Company*, MASHABLE, (Oct. 20, 2016) https://mashable.com/article/twitter-social-media-surveillance-snaptrends; Colin Daileda, *Geofeedia isn't the Only Social Media Surveillance Company Giving Data to Police*, MASHABLE, (Oct. 12, 2016) https://mashable.com/article/geofeedia-social-media-surveillance-police.

120. JOHN HOLLYWOOD, MICHAEL JOHN DEVRIES VERMEER, DULANI WOODS, SEAN GOODISON & BRIAN JACKSON, USING SOCIAL MEDIA AND SOCIAL NETWORK ANALYSIS IN LAW ENFORCEMENT: CREATING A RESEARCH AGENDA, INCLUDING BUSINESS CASES, PROTECTIONS, AND TECHNOLOGY NEEDS 8–9 (Rand Corporation, 2018) (describing social media monitoring for "worrisome activity" and in order to identify individuals "at high risk of being involved in violence").

121. Zachary Cohen & Whitney Wild, *Internal Emails Reveal Capitol Security Officials Dismissed Warnings About Troubling Social Media Posts Before January 6 Riot*, CNN (Apr. 28, 2021, 6:16 AM), https://www.cnn.com/2021/04/28/politics/capitol-security-emails-social-media-riot/index.html.

122. Police Department City Of New York, *Social Network Analysis Tools: Impact and Use Policy* (Apr. 11, 2021), https://www1.nyc.gov/assets/nypd/downloads/pdf/public_information/post-final/social-network-analysis-tools-nypd-impact-and-use-policy_4.9.21_final.pdf. The policy was published pursuant to the Public Oversight of Surveillance Technology Act, a law enforcement reform bill that requires the New York Police Department to publish reports about its "surveillance technology."; Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CALIF. L. REV. 917, 955 (2021).

123. POLICE DEP'T N.Y.C., *Social Network Analysis Tools: Impact and Use Policy* 3 (Apr. 11, 2021), https://www1.nyc.gov/assets/nypd/downloads/pdf/public_information/post-final/social-network-analysis-tools-nypd-impact-and-use-policy_4.9.21_final.pdf. ("NYPD may miss information critical to investigations because users can easily remove information posted on social media and social media platforms routinely delete content and deactivate accounts for violations of terms of service. Accordingly, social network analysis tools allow the NYPD

In addition to dragnet surveillance of events and people of interest, law enforcement uses social media in more targeted ways, often in contexts in which police rely upon undercover operations and confidential informants.[124] For instance, local police and the Federal Bureau of Investigation (FBI) have reportedly used undercover social media accounts to surveil groups and individuals and to develop probable cause to arrest suspected lawbreakers.[125] Similarly, law enforcement routinely uses fake social media accounts to engage in investigations.[126] For example, sex trafficking investigators often create fake social media accounts to "befriend, identify, and monitor people suspected of engaging in criminal activities, as well as those who are presumed to be victims."[127] Although creating a fake social media account often violates a platform's terms of service and other content-related rules, this practice appears prevalent.

3. Deputizing Users

The public-facing character of social media itself can feed into law enforcement strategies. Although law enforcement frequently monitors social media content and demands access to the wealth of data that online firms collect and retain, police also engage with users much as other ordinary users

to retain information on social networking platforms relevant to investigations and alert investigators to new activity on queried social media accounts.").

124. *See* Cyrus Farivar & Olivia Solon, *FBI Trawled Facebook to Arrest Protestors for Inciting Riots, Court Records Show*, NBC NEWS (June 19, 2020, 1:26 PM), https://www.nbcnews.com/tech/social-media/federal-agents-monitored-facebook-arrest-protesters-inciting-riots-court-records-n1231531 (describing FBI's use of social media to "infiltrate activist groups").

125. Betsy Woodruff Swan, *Feds Comb Facebook to Hunt down Alleged Rioters and Looters*, POLITICO (June 12, 2020, 4:30 AM), https://www.politico.com/news/2020/06/12/facebook-riot-loot-justice-department-314567.

126. *See, e.g.*, Dave Maass, *Facebook Warns Memphis Police: No More Fake "Bob Smith" Accounts*, ELECTR. FRONTIER FOUND. (Sept. 24, 2018), https://www.eff.org/deeplinks/2018/09/facebook-warns-memphis-police-no-more-fake-bob-smith-accounts; Dave Maass, *Four Steps Facebook Should Take to Counter Police Sock Puppets*, ELECTR. FRONTIER FOUND. (Apr. 14, 2019), https://www.eff.org/deeplinks/2019/04/facebook-must-take-these-four-steps-counter-police-sock-puppets; Jon Schuppe, *Undercover Cops Break Facebook Rules to Track Protestors, Ensnare Criminals*, NBC NEWS, (Oct. 5, 2018, 12:08 PM), https://www.nbcnews.com/news/us-news/undercover-cops-break-facebook-rules-track-protesters-ensnare-criminals-n916796; Tami Abdollah, *U.S. Plan to Use Fake Social Media Profiles for Surveillance is Against Facebook Rules*, PBS (Sept. 3, 2019, 5:19 PM), https://www.pbs.org/newshour/nation/u-s-plan-to-use-fake-social-media-profiles-for-surveillance-is-against-facebook-rules.

127. JENNIFER MUSTO, CONTROL AND PROTECT: COLLABORATION, CARCERAL PROTECTION, AND DOMESTIC SEX TRAFFICKING IN THE UNITED STATES 57–58 (1st ed. 2016).

do.[128] And, of course, law enforcement uses social media to disseminate routine information to a mass audience.[129]

Law enforcement also relies on social media to generate tips and investigative leads. For instance, police sometimes post videos to social media to solicit the public's help in identifying a suspect.[130] This form of crowdsourced public assistance can be crucial for investigating crimes but raises complex questions about online vigilantism, anonymity, and accountability. In the wake of the January 6 putsch at the U.S. Capitol, the FBI called for "the public's assistance in identifying individuals who made unlawful entry into the U.S. Capitol building and committed various other alleged criminal violations."[131] Though the vast majority of the insurrectionists walked away from the scene at the Capitol, social media users, private investigators, and the press identified dozens of individuals who were later charged.[132] This is not the first time in which a group of self-appointed internet users have tried to identify and hold accountable lawbreakers. In 2017, after the Unite the Right Rally in Charlottesville, online sleuths identified and outed, or "doxxed," several right-wing and White supremacist protestors.[133] But online vigilantes sometimes identify the wrong people, leading to harassment of innocent

128. *See* Levinson-Waldman, *Private Eyes*, *supra* note 116, at 999 ("[I]f a targeted user has a public Twitter account, police can go on the site to check the user's recent posts and interactions with other users without needing any special third-party software.").

129. Benesch, *supra* note 9, at 93 ("The very functions of routine governance are also carried out, increasingly, on social media platforms."); *see also* Knight First Amendment Inst. at Colum. Univ. v. Trump, 928 F.3d 226, 235–36 (2019) (describing how President Trump used his Twitter account "as an important tool of governance and executive outreach").

130. *Aggravated Assault 1 South Broad St. Dc 21 06 015773*, YOUTUBE (May 6, 2021), https://www.youtube.com/watch?v=FCeCu8WT3es; *Severe Injury Hit and Run Traffic Collision in Northeast Area NR21122wc*, YOUTUBE (May 5, 2021), https://www.youtube.com/watch?v=zNoDjrw5W-M.

131. FBI, *U.S. Capitol Violence*, https://www.fbi.gov/wanted/capitol-violence (last visited June 21, 2021); *see also* FBI Washington Field (@FBIWFO), TWITTER https://twitter.com/FBIWFO/status/1347407275300954112 (last visited Dec. 29, 2021).

132. Jaclyn Peiser, *Internet Detectives are Identifying Scores of Pro-Trump Rioters at the Capitol. Some have Already been Fired.*, WASH. POST (Jan. 8, 2021, 6:54 AM), https://www.washingtonpost.com/nation/2021/01/08/capitol-rioters-fired-doxed-online/; Sara Morrison, *The Capitol Rioters Put Themselves All Over Social Media. Now they're Getting Arrested.*, VOX, https://www.vox.com/recode/22218963/capitol-photos-legal-charges-fbi-police-facebook-twitter (last updated: Jan 19, 2021, 6:52 PM); Greg Myre, *How Online Sleuths Identified Rioters at the Capitol*, NPR (Jan. 11, 2021, 9:45 AM) https://www.npr.org/2021/01/11/955513539/how-online-sleuths-identified-rioters-at-the-capitol.

133. Vegas Tenold, *To Doxx a Racist: How a Dead White Supremacist Sparked the Debate About the Tactics Used Against the Extreme Right*, THE NEW REPUBLIC (July 26, 2018), https://newrepublic.com/article/150159/doxx-racist; Emma Grey Ellis, *Whatever Your Side, Doxing is a Perilous Form of Justice—Even When it's Outing Nazis*, WIRED (Aug. 17, 2017, 8:00 AM), https://www.wired.com/story/doxing-charlottesville/.

individuals. For example, in 2013, users of the subreddit Find Boston Bombers misidentified several people as suspects in the Boston Marathon attack.[134]

Scholars such as Mary Anne Franks and Danielle Citron have warned that doxing can be part of a campaign of online harassment and abuse, with particularly devastating results for women.[135] But unlike these earlier episodes, the hundreds of Capitol putsch arrests and prosecutions appear to rely heavily on identifications made using information gleaned from social media, whether crowdsourced or obtained directly from platforms.[136] Indeed, in light of major social media platforms' decisions to take down much of the evidence related to the Capitol putsch, crowdsourcing may have been particularly essential to identifying individuals.[137]

The public also uses social media to alert law enforcement to suspicious activity in more mundane settings. Consider Nextdoor, a social media platform designed for "neighbors" to exchange information with each other.[138] It is a unique surveillance tool because it facilitates voluntary, private surveillance by those who choose to join the platform.[139] As Sam Levin has documented, White Nextdoor users have deployed the platform to report unsubstantiated claims of suspicious activity and to organize noise complaints against Black

134. Dave Lee, *Boston Bombing: How Internet Detectives Got it Very Wrong*, BBC NEWS: TECHNOLOGY (Apr. 19, 2013), https://www.bbc.com/news/technology-22214511.

135. Mary Anne Franks, *Sexual Harassment 2.0*, MD. L. REV. 655, 678–79 (2012) (describing episode of sexual harassment in which online forum users "posted personal information of their targets" and encouraged forum participants to contact victims directly); DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE 53–54 (2016); *see also* David M. Douglas, *Doxing: A Conceptual Analysis*, 18 ETHICS INFO. TECH. 199, 200 (2016) ("In cases where exposing wrongdoing is in the public interest, deanonymizing and delegitimizing doxing is permissible only to the extent necessary to reveal that wrongdoing has occurred.").

136. Craig Timberg, Drew Harwell & Spencer S. Hsu, *Police Let Most Capitol Rioters Walk Away. But Cellphone Data and Videos Could Now Lead to More Arrests.*, WASH. POST (Jan. 8, 2021), https://www.washingtonpost.com/technology/2021/01/08/trump-mob-tech-arrests/ ("The countless hours of video—much of it taken by the rioters themselves and uploaded to social media—also offers an ideal data set for facial recognition."); Drew Harwell & Craig Timberg, *How America's Surveillance Networks Helped the FBI Catch the Capitol Mob*, WASH. POST (Apr. 2, 2021), https://www.washingtonpost.com/technology/2021/04/02/capitol-siege-arrests-technology-fbi-privacy/.

137. Knibbs, *supra* note 12 (describing how efforts to preserve online documentation of the Capitol putsch related to FBI's efforts to seek evidence for use in criminal proceedings); *How Facebook is Responding to the Violence at the US Capitol*, FACEBOOK (Jan. 11, 2021, 1:00 PM), https://www.facebook.com/business/news/facebooks-actions-in-response-to-washington-dc-violence (describing Facebook's actions to remove content that "incites, praises, or encourages violence or harm," including support for the Capitol putsch).

138. *About Nextdoor*, https://about.nextdoor.com/ (last visited June 21, 2021).

139. Rahim Kurwa, *Building the Digitally Gated Community: The Case of Nextdoor*, 17 SURVEILLANCE & SOC'Y 111, 113 (2019) (describing the "co-production of community through participation in surveillance").

residents.[140] In response to criticism that its platform was amplifying patterns of racial profiling and harassment, Nextdoor adopted changes to its service to discourage users from posting unsubstantiated, racialized accusations.[141] In 2020, during nationwide uprisings against police violence, Nextdoor announced that it was removing its "Forward to Police" feature, which permitted users to forward posts directly to law enforcement partners.[142] Although Nextdoor has tried to nudge users away from using crime reporting to perpetuate racial harassment, crime prevention is still a core part of Nextdoor's offerings and appeal.[143]

4. Resistance Through Design

While this Article focuses on how platforms can enable and facilitate law enforcement surveillance, in recent years, firms have also made design choices that can obstruct policing, generating substantial legal and political backlash. Although these choices can take many forms, I highlight two here.

First, firms can choose to collect and store data about user communications in ways that are more or less vulnerable to law enforcement demands.[144] For example, Signal, a secure messaging provider, simply "does

140. Sam Levin, *Racial Profiling via Nextdoor.Com*, EAST BAY EXPRESS (Oct. 7, 2015), https://eastbayexpress.com/racial-profiling-via-nextdoorcom-2-1/.

141. Sam Levin, *What Happens when Tech Firms End Up at the Center of Racism Scandals?*, THE GUARDIAN (Aug. 30, 2016), http://www.theguardian.com/technology/2016/aug/30/tech-companies-racial-discrimination-nextdoor-airbnb (describing Nextdoor's adoption of a new system that warns users about racial profiling before they post a crime and safety message); *see also* Tatyana Mamut, *Announcing Our New Feature to Promote Kindness in Neighborhoods*, NEXTDOOR: BLOG (Sept. 18, 2019), https://blog.nextdoor.com/2019/09/18/announcing-our-new-feature-to-promote-kindness-in-neighborhoods/ (describing Nextdoor's "Kindness Reminder" feature, which nudges users to reconsider offensive or hurtful posts before publishing); Team Nextdoor, *Standing in Solidarity with Black Neighbors—Nextdoor*, (Mar. 25, 2021), https://blog.nextdoor.com/2021/03/25/standing-in-solidarity-with-black-neighbors/ (prohibiting All Lives Matter and Blue Lives Matter content "when used to undermine racial equality or the Black Lives Matter movement").

142. Team Nextdoor, *Nextdoor Removes "Forward to Police" Feature*, NEXTDOOR: BLOG (June 18, 2020), https://blog.nextdoor.com/2020/06/18/nextdoor-removes-forward-to-police-feature/.

143. Joseph Porcelli, *Nextdoor for Public Agencies Crime Prevention Engagement Plan*, MEDIUM (May 22, 2019), https://medium.com/nextdooragencyresources/nextdoor-for-public-agencies-crime-prevention-engagement-plan-1bf92c34b360; *see, e.g.*, Timothy Hayden, Arlington Policy Department, *Requesting Assistance in Identifying Suspect that Broke into Vehicles in Your Neighborhood.*, (Apr. 19, 2021), https://nextdoor.com/agency-post/tx/arlington/arlington-police-department/requesting-assistance-in-identifying-suspect-that-broke-into-vehicles-in-your-neighborhood-184037858/.

144. SHOSHANNA ZUBOFF, SURVEILLANCE CAPITALISM 385 (2019) (describing how government officials "must work, at least in part, through the [private] surveillance capitalists" to access and make use of consumer data).

not collect user metadata," the kind of to- and from- information that accompanies a text message or phone call.[145] In contrast, ordinary telephone and online service providers collect and store this information for months, if not years.[146] The divergence between these two poles has, not surprisingly, fostered legal controversy. For example, traceability obligations introduced in India in 2021 would require platforms to be able to identify individuals who "originate information" online and would appear to require firms like Signal to redesign their services or to exit the market entirely.[147] Firms also choose where to store user data. This has led some states to require that multinational firms ensure that data about domestic users is stored domestically, in part so it is readily available to law enforcement.[148]

Second, firms can choose to implement encryption for communications that are in storage or in transit. In October 2014, then-FBI Director James B. Comey spoke at the Brookings Institution about the threat that important sources of evidence and intelligence were rapidly "going dark."[149] As major communications providers increasingly adopted end-to-end encryption, he argued, "the bad guys" could communicate without being intercepted by law enforcement.[150] In the ensuing years, major communications platforms have further committed to embracing end-to-end encryption.[151]

145. Anna Wiener, *Taking Back Our Privacy*, THE NEW YORKER (Oct. 26, 2020), https://www.newyorker.com/magazine/2020/10/26/taking-back-our-privacy.

146. Suzanne Choney, *How Long Do Wireless Carriers Keep Your Data?*, NBC NEWS (Sept. 29, 2011), https://www.nbcnews.com/technolog/how-long-do-wireless-carriers-keep-your-data-120367; Lily Hay Newman, *How to Limit How Long Google Keeps Your Data*, WIRED (May 7, 2019, 6:02 AM), https://www.wired.com/story/google-auto-delete-data-privacy-setting/ ("By default, Google will continue to indefinitely retain the Web & Activity data you've set it to collect—everything by default.").

147. Times Now Digital, *WhatsApp Sues Centre Over New Social Media Rules: Why Facebook-owned Company has Claimed they Endanger Privacy*, TIMES NOW NEWS (May 26, 2021, 17:23 IST), https://www.timesnownews.com/india/article/whatsapp-sues-centre-over-new-social-media-rules-why-facebook-owned-company-has-claimed-they-endanger-privacy/762133.

148. Anupam Chander & Uyen P. Le, *Data Nationalism*, 64 EMORY L.J. 677, 708–13 (2015) (summarizing data localization mandates).

149. James B. Comey, *Going Dark: Are Technology, Privacy, and Public Safety on a Collision Course?*, FEDERAL BUREAU OF INVESTIGATION (Oct. 16, 2014), https://www.fbi.gov/news/speeches/going-dark-are-technology-privacy-and-public-safety-on-a-collision-course.

150. *Id.*

151. Cory Doctorow, *Google Announces End-to-End Encryption for Gmail (a Big Deal!)*, BOING BOING (June 4, 2014, 6:00 PM), https://boingboing.net/2014/06/04/google-announces-end-to-end-en.html; Aatif Sulleyman, *This is why the Government is so Worried about WhatsApp Encryption*, THE INDEPENDENT (Mar. 27, 2017, 16:45), https://www.independent.co.uk/life-style/gadgets-and-tech/features/whatsapp-encryption-what-it-how-does-it-work-why-ban-it-backdoor-access-secret-messages-a7652396.html; Mark Zuckerberg, *A Privacy-Focused Vision for Social Networking*, https://m.facebook.com/nt/screen/?params=%7B"note_id"%

Encryption has become a major point of contention for law enforcement in the United States and elsewhere. Domestically, for example, the San Bernardino shootings generated legal controversy when Apple refused to unlock the shooter's iPhone.[152] In 2020, several Republican senators introduced the Lawful Access to Encrypted Data (LAED) Act, which "would bring an end to warrant-proof encryption in devices, platforms, and systems."[153] The LAED Act would "require device manufacturers and service providers to assist law enforcement with accessing encrypted data if assistance would aid in the execution of the warrant."[154] Similar approaches have been adopted elsewhere. In the United Kingdom, the Investigatory Powers Act permits the government to issue a "technical capability notice" that requires firms to be able to assist in executing lawful warrants.[155]

As outlined above, technological design choices like these have prompted substantial controversy and strife between regulators and platforms. Both encryption and data storage choices can make it more difficult for platforms or law enforcement to access information about user speech that is either harmful or unlawful.[156] However, some automated mechanisms for screening

---

3A2420600258234172%7D&path=%2Fnotes%2Fnote%2F&refsrc=http%3A%2F%2Ft.co%2F&_rdr (last visited June 22, 2021) (announcing plans to work on end-to-end encryption); *but see* Andy Greenberg, *Facebook Says Encrypting Messenger by Default Will Take Years*, WIRED (Jan. 10, 2020, 4:54 PM), https://www.wired.com/story/facebook-messenger-end-to-end-encryption-default/.

152. Ellen Nakashima & Reed Albergotti, *The FBI Wanted to Unlock the San Bernardino Shooter's IPhone. It Turned to a Little-Known Australian Firm.*, WASH. POST (Apr. 14, 2021), https://www.washingtonpost.com/technology/2021/04/14/azimuth-san-bernardino-apple-iphone-fbi/.

153. United States Senate Committee on the Judiciary, *Graham, Cotton, Blackburn Introduce Balanced Solution to Bolster National Security, End Use of Warrant-Proof Encryption that Shields Criminal Activity*, https://www.judiciary.senate.gov/press/rep/releases/graham-cotton-blackburn-introduce-balanced-solution-to-bolster-national-security-end-use-of-warrant-proof-encryption-that-shields-criminal-activity (last visited Dec. 29, 2021).

154. *Id.* The EARN IT Act proposed in 2020 adopted a similar approach, requiring platforms to qualify for a statutory safe harbor under Section 230 of the Communications Decency Act by showing that they abided by "best practices" to fight child sexual exploitation. As several commentators noted, those "best practices" were likely incompatible with strong encryption.; Lily Hay Newman, *The EARN IT Act is a Sneak Attack on Encryption*, WIRED (Mar. 5, 2020, 8:22 PM), https://www.wired.com/story/earn-it-act-sneak-attack-on-encryption/; Riana Pfefferkorn, *The EARN IT ACT is a disaster amid the COVID-19 crisis*, BROOKINGS INST. (May 4 2020), https://www.brookings.edu/techstream/the-earn-it-act-is-a-disaster-amid-the-covid-19-crisis/.

155. Investigatory Powers Act 2016 § 253.

156. Michael H. Keller & Gabriel J. X. Dance, *The Internet is Overrun with Images of Child Sexual Abuse. What Went Wrong?*, N.Y. TIMES (Sept. 28, 2019), https://www.nytimes.com/interactive/2019/09/28/us/child-sex-abuse.html ("[W]hen tech companies cooperate fully, encryption and anonymization can create digital hiding places for perpetrators.").

content may be compatible with end-to-end encryption.[157] Whether firms choose to deploy them is an entirely different design question.

B. SHAPING LAW ENFORCEMENT THROUGH PLATFORM POLICY

Like technological design, firms' internal policies also shape law enforcement behavior by encouraging or discouraging certain kinds of demands for different types of data. Partly because of the First Amendment implications of compelled disclosure, social media platforms have sometimes resisted government demands, citing the implications for their users. For example, technology companies have, at times, moved to quash government search warrants, attempting to advance the Fourth Amendment interests of their users.[158] Electronic communications service providers have also invoked their own expressive rights in efforts to lift nondisclosure orders that prevent service providers from notifying users of demands for users' information.[159]

Outside of litigation, firms can also engage in private standard-setting to raise the standard that the government must meet when it demands user information.[160] Again, consider the example of government requests for historical location information. As Matthew Tokson has observed, the Supreme Court's *Carpenter* decision is "exceedingly vague and cautious" with regard to its application to new technologies and forms of surveillance.[161] Therefore, substantial ambiguity remains about whether government activity constitutes a search or a seizure.[162]

---

157. Jonathan Mayer, *Content Moderation for End-to-End Encrypted Messaging*, 5 (Oct. 6, 2019), https://www.cs.princeton.edu/~jrmayer/papers/Content_Moderation_for_End-to-End_Encrypted_Messaging.pdf.

158. *See, e.g.*, *In re* 381 Search Warrants Directed to Facebook, Inc., 29 N.Y.3d 231 (2017).

159. *See, e.g.*, Microsoft Corp. v. United States Dep't of Justice, 233 F. Supp. 3d 887, 908 (W.D. Wash. 2017) (concluding that Microsoft had adequately supported its argument that nondisclosure orders under the Electronic Communications Privacy Act violated its First Amendment rights); *In re* Nat'l Sec. Letter, 863 F.3d 1110 (9th Cir. 2017) (rejecting petitioner's First Amendment challenge to nondisclosure orders that accompanied National Security Letters); Twitter, Inc. v. Barr, 445 F. Supp. 3d 295 (N.D. Cal. 2020) (rejecting Twitter's First Amendment challenge to the government's prohibition on publishing certain types of data regarding legal process the platform had received under the Foreign Intelligence Surveillance Act).

160. *Cf.* Klonick, *supra* note 5, at 1615 (developing the idea of private governance in the context of platform content moderation standards).

161. Matthew Tokson, *The Next Wave of Fourth Amendment Challenges after Carpenter*, 59 WASHBURN L.J. 1, 1 (2020).

162. *Id.*; *see also* United States v. Hammond, 996 F.3d 374, 391–92 (7th Cir. 2021) (concluding that real-time collection of location information for several hours was not a "search" in the meaning of the Fourth Amendment).

In the wake of *Carpenter*, does the government's demand for location information from Google raise a Fourth Amendment issue? At times, private platform policymaking can preempt this inquiry. In one case, police investigating a string of fires sought information from Google regarding user devices near six different locations.[163] Although the government has argued that *Carpenter* does not extend to reverse location information, in practice, Google will only provide this information in response to a search warrant.[164] That creates a default practice in which the government must satisfy a higher modicum of suspicion to satisfy Google's policy notwithstanding the absence of controlling legal precedent.

Firms also use the mechanisms of private governance—particularly policies and terms of service—to limit government access to data in other ways. In 2016, the American Civil Liberties Union (ACLU) obtained information, through public records requests, showing that law enforcement agencies were procuring social media monitoring software from third-party vendors. In response, major social media firms such as Facebook, Twitter, and Instagram publicly announced that they would cut off access to their application programming interfaces (APIs) by firms that sold surveillance software to law enforcement.[165]

Yet it appears that these policy-based limitations on government access are often ineffective. During the nationwide uprisings against police violence after George Floyd's murder in 2020, for example, it became clear that law enforcement agencies were continuing to use social media monitoring services such as Dataminr to keep tabs on protest, activism, and dissent.[166] In order to circumvent Twitter's terms of service, which barred API users from "tracking, alerting, or monitoring sensitive events," Dataminr has rebranded itself as a breaking news service that takes advantage of access to Twitter's "firehose" to provide alerts to law enforcement clients.[167] The upshot is that while platform firms' formal content and privacy policies appear to bar use of their services

163. *In re* Search Warrant Application for Geofence Location, 497 F. Supp. 3d at 351.

164. *See In re* Search of Info. Stored at Premises Controlled by Google, 481 F. Supp. 3d 730, 736 (N.D. Ill. 2020) (finding that, because the government had sought a search warrant, it had "forfeited the argument" that the Fourth Amendment didn't apply); *see also id. In re* Search Warrant Application for Geofence Location, 497 F. Supp. 3d at 360 (noting that Google "will only produce the information upon presentation of a warrant").

165. Levinson-Waldman, *Government Access*, *supra* note 116, at 556–57.

166. Biddle, *supra* note 11; *see also* Lee Fang, *FBI Expands Ability to Collect Cellphone Location Data, Monitor Social Media, Recent Contracts Show*, THE INTERCEPT (June 24, 2020, 8:56 PM), .https://theintercept.com/2020/06/24/fbi-surveillance-social-media-cellphone-dataminr-venntel/; Sahar F. Aziz & Khaled A. Beydoun, *Fear of a Black and Brown Internet: Policing Online Activism*, 100 B.U. L. REV. 1151, 116869 (2020).

167. Biddle, *supra* note 11.

for law enforcement surveillance, a veritable cottage industry of surveillance and monitoring firms has sprung up to help police take full advantage of the wealth of intelligence that social media can provide.

C. VOLUNTARY PRIVATE-PUBLIC SURVEILLANCE ARRANGEMENTS

It is not just that private governance sometimes serves as an ineffective check on law enforcement; at times, private decision-making can in fact advance law enforcement goals. Indeed, voluntary private decision-making can give rise to a systematic relationship with law enforcement investigations, arrests, and prosecutions. Faced with competing pressures to both take down more harmful content and to facilitate law enforcement surveillance, the private sector has increasingly turned to voluntary, cross-platform arrangements that allow them to pool technical and policy resources across firms.[168]

Collaboration to eradicate child sexual abuse imagery provides one illustration. While technology firms are not required to proactively monitor user-uploaded or -generated content for unlawful child sexual abuse imagery, many do so voluntarily.[169] For example, Microsoft's PhotoDNA program, a hash-matching tool, scans images and videos against a database of unlawful images.[170] Thorn, a nonprofit organization, has developed a technical tool for the same purposes for smaller companies to use.[171] When a firm detects a match, federal law requires the firm to report it to the National Center for Missing and Exploited Children (NCMEC).[172] NCMEC, in turn, discloses the information to law enforcement and plays an essential coordinating role with law enforcement agencies investigating the crime.[173] While NCMEC itself is a private organization, it is funded through annual grants by the government and, pursuant to federal law, must coordinate several distinct public and private programs.[174] At least one federal court has concluded that NCMEC's

168. EVELYN DOUEK, THE RISE OF CONTENT CARTELS 5–6 (Knight First Amendment Inst. Colum. Univ., 2020) (describing voluntary cross-industry arrangements as "content cartels").

169. Bloch-Wehba, *supra* note 10, at 58.

170. *Id.* at 58.

171. Olivia Solon, *To Fight Online Child Sexual Abuse, Tech Companies Turn to a Nonprofit Startup*, NBC NEWS (July 22, 2020, 3:16 PM), https://www.nbcnews.com/tech/tech-news/fight-online-child-sexual-abuse-tech-companies-turn-nonprofit-startup-n1234569.

172. 18 U.S.C. § 2258A.

173. United States v. Ackerman, 831 F.3d 1292, 1296 (10th Cir. 2016).

174. 34 U.S.C. § 11293(b).

statutory obligations give rise to "special law enforcement duties and powers" that distinguish it from other private entities.[175]

Platforms also collaborate on efforts to filter and block terrorist content online. Consider the Global Internet Forum to Counter Terrorism (GIFCT), a private, voluntary consortium of technology firms that uses both hash-based and artificial intelligence-based filtering to detect unlawful terrorist content across platforms.[176] The GIFCT's database is limited to industry members and is not shared directly with law enforcement.[177] But new European regulations will require platforms to take proactive measures to remove terrorist content and to preserve it for law enforcement purposes for six months.[178] The result is that the GIFCT hash-matching database is likely to yield a substantial number of posts that platforms will be required to preserve for potential law enforcement use and possibly to report to the "competent authorities."[179]

These two examples illustrate a strikingly similar dynamic: technology firms have voluntarily adopted monitoring technology to enforce content-related rules, the use of which gives rise to an escalating set of legal obligations. In the context of child sexual abuse imagery, platforms must report information to NCMEC, a nominally private center, which then funnels it to law enforcement.[180] In the context of terrorist imagery, platforms are required to report certain kinds of terroristic threats to European authorities and likewise required to preserve a broader range of information for future law enforcement use.[181] The result is that the monitoring technology used to detect

175. *See Ackerman*, 831 F.3d at 1296–97. A second court has concluded that NCMEC can act as part of the "prosecution team" for purposes of discovery, and of obligations to disclose exculpatory evidence pursuant to Brady v. Maryland.; United States v. Rosenschein, CR 16-4571 JCH, 2019 WL 2298810, at *7 (D.N.M. May 30, 2019), clarified on denial of reconsideration, CR 16-4571 JCH, 2020 WL 2750247 (D.N.M. May 27, 2020) ("It was NCMEC's acts of investigating the location and providing CyberTipline information to the geographically appropriate law enforcement agency that effectively commenced the prosecution of this case.").

176. *Explainers*, GLOBAL INTERNET F. COUNTER TERRORISM,https://gifct.org/explainers/ (last visited June 21, 2021); Robert Gorwa, Reuben Binns & Christian Katzenbach, *Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance:* 7 BIG DATA & SOCIETY 1 (2020) (describing GIFCT's use of hash-based and machine learning techniques).

177. GLOB. INTERNET F. COUNTER TERRORISM, *supra* note 176.

178. Regulation 2021/784 of the European Parliament and of the Council of Apr. 29, 2021, On Addressing The Dissemination of Terrorist Content Online ("TERREG"), art. 6 sec. 3, art. 3.

179. *Id.* art. 6(1)–(2) (requiring hosting service providers to "preserve terrorist content" for six months); art. 14(5).

180. *Ackerman*, 831 F.3d at 1296–97.

181. TERREG arts. 6 and 14.

lawbreaking itself lies at the heart of investigations and prosecutions, yielding increasing entanglements between law enforcement and platform governance.[182]

The increased reliance on automated content moderation also creates new opportunities and frameworks within which platforms share user communications with law enforcement. Ordinarily, the Stored Communications Act (SCA) bars electronic communications service providers from voluntarily disclosing user communications to law enforcement, with a few exceptions.[183] Broadly speaking, the SCA is meant to limit the circumstances in which user communications are shared with law enforcement to those circumstances in which the government has met the appropriate standard.[184] But if the communications service provider obtains the contents of communications "inadvertently" and they "appear to pertain to the commission of a crime," then the provider may disclose the contents to a law enforcement agency.[185]

Firms that detect legal violations using technical moderation tools are arguably free to voluntarily disclose that information to law enforcement pursuant to the SCA because they learned of it "inadvertently" and the contents "appear to pertain to the commission of a crime."[186] Alternatively, law enforcement could use a search warrant, administrative subpoena, or 2703(d) order to compel a platform to disclose subscriber information for any user who has uploaded content that has been flagged as unlawful.[187] Law enforcement has pursued this dragnet approach before. In 2017, the government obtained a search warrant to compel DisruptJ20, a website that had been used to organize protests against Donald Trump's inauguration, to disclose records related to a huge number of people who had visited the site.[188] It is well within the realm of possibility that law enforcement may use a similar process to seek information about individuals who have been flagged through

182. *See, e.g.*, State v. Lizotte, 197 A.3d 362, 366 (Vt. 2018) (describing AOL's use of its "Image Detection Filtering Process" in the context of a defense motion to suppress).

183. 18 U.S.C. § 2702(b).

184. *See* 18 U.S.C § 2701(a)–(c) (making it a criminal offense to access stored communications, except if doing so is authorized); 18 U.S.C. § 2703(b)–(c) (setting forth procedural requirements that law enforcement must meet in order to access different types of communications information).

185. 18 U.S.C. § 2702(b)(7)(A).

186. *Id.*

187. 18 U.S.C. § 2703.

188. Coalition: Justice Department's demand for protest website data raises privacy and civil liberty concerns, OPENTHEGOVERNMENT.ORG (Aug. 24, 2017), https://www.openthegovernment.org/coalition-justice-departments-demand-for-protest-website-data-raises-privacy-and-civil-liberty-concerns/.

the GIFCT database or who have been suspended for posting terrorist-related content.

IV. IMPLICATIONS FOR CRIMINAL PROCEDURE

As Reidenberg rightly anticipated, today's public sphere is shaped as much by private technological and design choices as by formal law and regulation.[189] But the emergence of private platforms as regulatory forces in their own right has not uniformly diminished the role or power of the state. Certainly, platform intransigence on content-related issues has, at times, posed challenges for law enforcement.[190] But platforms can also expand and facilitate law enforcement power by encoding and enforcing law enforcement demands in the rules, norms, and technological infrastructures of online governance.

The current alignment between private technology firms and public law enforcement has expanded the authority and the power of both firms and states. At the same time, as governments seek to incentivize technology firms to prevent the dissemination of unlawful speech through private governance and technology, they also enlist technology firms to aid in digital surveillance, both directly and indirectly.[191] As outlined in Parts II and III, these efforts can sometimes occur at cross-purposes; increasing deletion of online content has, at times, created obstacles for law enforcement investigating criminal activity online. But here, too, technology itself can provide a workaround. New surveillance technology tools and practices emerge, taking advantage of the affordances of social media for law enforcement's gain.

A. THE EMERGENCE OF NEW FORMS OF DISCLOSURE

New legal and technological developments only underscore the mutual dependency between private firms and the public sector. As both governments and tech firms herald the growing capacity and use of artificial intelligence and automated content moderation systems, content- and data-related decision-making itself is increasingly becoming entwined with law enforcement objectives. Yet the more extensive public-private cooperation becomes, the weaker the opportunities for accountability appear.

189. Reidenberg, *supra* note 16, at 571 ("The political-governance process ordinarily establishes the substantive law of the land. For Lex Informatica, however, the primary source of default rule-making is the technology developer and the social process by which customary uses evolve.").

190. *See, e.g.*, Danielle Keats Citron & Benjamin Wittes, *The Problem isn't Just Backpage: Revising Section 230 Immunity*, 2 GEO. L. TECH. REV. 453, 466 (2018) (listing myriad bad actors who were protected from legal liability under Section 230 of the Communications Decency Act).

191. Balkin, *supra* note 74, at 2019–20.

Emerging forms of content and data governance generate new demands by law enforcement for consumer data.[192] Again, the example of the Internet of Things is illustrative: why conduct a search of a person's home in real time when a set of networked home technologies makes it possible to do so retrospectively? As online platforms turn increasingly toward automated moderation techniques to proactively filter user-generated content, the content moderation process itself becomes an increasingly appealing target for law enforcement. Platforms engaged in automated content moderation will obtain access to a huge amount of content that either violates or appears to violate the law. Indeed, the explicit goal of automated moderation is to scale enforcement of platform rules and practices to respond to the growing volume of online content.[193] But because the technology is not yet that sophisticated, automated techniques are often necessarily over inclusive.[194] This means that, at times, automated moderation will sweep in more content than it was intended to.

In the United States, the examples of NCMEC and Thorn already illustrate how voluntary content moderation processes can feed law enforcement demands.[195] But existing laws governing the sharing of data between private and public sector actors are ill-equipped to address these emerging practices. The Stored Communications Act (SCA) presumptively limits the sharing of private user information between communications firms and law enforcement to a defined set of circumstances governed by appropriate statutory limitations. For example, the SCA explicitly provides that platforms may voluntarily disclose user data to NCMEC in connection with a statutorily required report regarding child sexual abuse imagery.[196] As tech firms engage in more extensive collaboration, including with nonprofits and independent organizations such as Thorn and GIFCT, neither the SCA nor the Fourth Amendment are likely to promote private accountability. For its part, the SCA's voluntary disclosure limitations extend only to actors who provide a "remote computing service" or "electronic communication service" *to the public*, which coalitions such as Thorn and GIFCT do not.[197] Moreover, the SCA explicitly permits platforms to share data among themselves without constraint.[198] The lax attitude toward private data sharing will encourage more voluntary, private arrangements to

192. *See infra* Part III.

193. Gorwa et al., *supra* note 176, at 2.

194. Gorwa et al., *supra* note 176, at 5 (describing how machine learning systems "risk over-blocking in cases in which the word may be acceptable in context").

195. *See* Ohm, *supra* note 113.

196. 18 U.S.C. § 2702(c)(5).

197. 18 U.S.C. § 2702(a).

198. 18 U.S.C. § 2702(c)(6).

emerge, while ignoring how those arrangements feed law enforcement demands for data.[199]

Outside the United States, new statutory initiatives already make clear that law enforcement has a growing appetite to deputize the content moderation process in service of investigative needs. In Germany, legislators introduced a new version of the Network Enforcement Act (NetzDG) alongside a package of measures intended to strengthen criminal law enforcement.[200] The new initiatives would require social network providers to report content that violated certain criminal prohibitions directly to law enforcement, along with the user's IP address and passwords.[201] As platforms continue to ramp up their efforts to police harmful and unlawful content through technology and through policy, the data they collect will, itself, become a rich source of evidence for law enforcement.

B. NEW INVESTIGATIVE METHODS

Firms' private decisions regarding both design and policy do not only shape law enforcement practices. They also shape the law of criminal procedure itself.

First, the design of technical infrastructure that facilitates the simultaneous collection and retention of information from large numbers of users encourages a shift from individualized suspicion to larger scale "dragnets," often with unclear consequences for Fourth Amendment protections.[202] For

199. *Cf.* Hannah Bloch-Wehba, *Transparency after Carpenter*, 59 WASHBURN L.J. 23, 28 (2020) ("[P]rivate sector data collection has created a rich source of information for law enforcement, yet goes hand in hand with stringent limitations on government conduct.").

200. Evelyn Douek, *Germany's Bold Gambit to Prevent Online Hate Crimes and Fake News Takes Effect*, LAWFARE (Oct. 31, 2017, 11:30 AM), https://www.lawfareblog.com/ germanys-bold-gambit-prevent-online-hate-crimes-and-fake-news-takes-effect; Amelie Heldt, *Germany is Amending its Online Speech Act NetzDG. . .But Not Only That*, INTERNET POL'Y REV. (Apr. 6, 2020), https://policyreview.info/articles/news/germany-amending-its-online-speech-act-netzdg-not-only/1464.

201. Patrick Beuth, *Was Sie über das Gesetz gegen Hasskriminalität Wissen Müssen*, DER SPIEGEL, (Feb. 18, 2020, 5:10 PM) https://www.spiegel.de/netzwelt/netzpolitik/gesetz-gegen-hasskriminalitaet-was-sie-darueber-wissen-muessen-a-1f995e2b-80a9-4e11-aecc-75f3250c69b9 (reporting that social network providers would be required to report certain violent threats, neo-Nazi propaganda, and incitement of hatred along with the IP addresses and port numbers of the subscribers to the German federal criminal police; in addition, providers may also be required to share passwords with law enforcement or intelligence agencies).

202. Renan, *supra* note 116, at 1053; Barry Friedman & Cynthia Benin Stein, *Redefining what's Reasonable: The Protections for Policing*, 84 GEO. WASH. L. REV. 281, 303–04 (2016) (describing the turn toward "dragnet searches"); Christopher Slobogin, *Government Dragnets*, 73 LAW & CONTEMP. PROBS. 107, 110 (2010) (defining "dragnets" as "programmatic government efforts to investigate, detect, deter, or prevent crime or other significant harm by subjecting a

example, information gleaned from generalized social media surveillance might be included in law enforcement databases and in targeted investigations. Gang policing is illustrative: Police frequently use social media information in gang databases, which collect and maintain information about alleged gang members.[203] Posts in which a user "admits" to gang membership, photos that include gang signs or other alleged gang members, and "liking" other users' gang related posts are all reportedly sufficient to land a social media user in a gang database.[204] Data from social media also makes its way into predictive policing tools, immigration enforcement, and domestic terrorism investigations.[205]

Law enforcement often describes the scraping, analysis, and use of huge amounts of publicly available data to predict and control behavior as an essential tool for high-priority investigations.[206] But perhaps this is exactly backwards—perhaps it is the availability of the data itself, and the possibilities for analysis and interpretation, that drive law enforcement's turn toward new priorities.[207] And, of course, it is not just the ease of acquiring privately held data that incentivizes law enforcement to adopt new data-driven techniques,

group of people, most of whom are concededly innocent of wrongdoing or of plans to engage in it, to a deprivation of liberty or other significant intrusion.").

203. Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CALIF. L. REV. 917, 950–54 (2021).

204. Sara Robinson, *When a Facebook Like Lands You in Jail*, BRENNAN CTR. FOR JUST. (July 6, 2018), https://www.brennancenter.org/our-work/analysis-opinion/when-facebook-lands-you-jail; STUART, *supra* note 39, at 9; Meredith Broussard, *When Cops Check Facebook*, THE ATLANTIC (Apr. 19, 2015), https://www.theatlantic.com/politics/archive/2015/04/when-cops-check-facebook/390882/.

205. Drew Harwell & Nick Miroff, *ICE Just Abandoned its Dream of "Extreme Vetting" Software that Could Predict whether a Foreign Visitor would become a Terrorist*, WASH. POST (May 17, 2018), https://www.washingtonpost.com/news/the-switch/wp/2018/05/17/ice-just-abandoned-its-dream-of-extreme-vetting-software-that-could-predict-whether-a-foreign-visitor-would-become-a-terrorist/; Will Carless, *Feds are Tracking Americans' Social Media to Identify Dangerous Conspiracies. Critics Worry for Civil Liberties.*, USA TODAY (2021), https://www.usatoday.com/story/news/nation/2021/05/14/terrorist-social-media-narratives-focus-new-dhs-effort/5075237001/.

206. *See, e.g.*, Wes Simmons, *Big Data Does Not Have to Mean Big Brother or be a Big Deal*, POLICE CHIEF MAGAZINE (May 3, 2017), https://www.policechiefmagazine.org/big-data-does-not-have-to-mean-big-brother/ (recounting hypothetical case of an illegally armed individual identified and prevented from committing a violent act because of big data); *see also* Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 362–63 (2015) (describing how "law enforcement and private companies have embraced the idea of networking and sharing personal information").

207. *Cf.* Fourcade & Gordon, *supra* note 99, at 79–80 (describing how technology generates "new possibilities," often controlled and marketed to government by private firms).

but also the emergence of privately developed networked technologies that are themselves purpose-built for law enforcement uses.[208]

Second, the essential role of networked technology firms in facilitating surveillance also limits opportunities for oversight agencies, the public, and defendants to understand how law enforcement is doing its job. Once upon a time, the search of a home was the "canonical fact pattern" of Fourth Amendment law.[209] But today, rather than conducting a physical search of one's home, law enforcement can issue a remote request to a Silicon Valley firm to compel disclosure of reams of intimate data from inside the same four walls.[210] As Jon Michaels has pointed out with regard to intelligence, public-private cooperation can enhance secrecy and impede oversight.[211] The web of sealing and secrecy orders that often surrounds electronic surveillance tends to obscure the role that social media platforms play in facilitating law enforcement investigations.[212] In prior work, I have also suggested that the emergence of more secretive forms of surveillance has diminished the opportunities for defendants to use the law of criminal procedure to hold law enforcement accountable.[213]

In theory, networked technology firms should reduce, rather than increase, secrecy. As David Pozen has put it, a secret is "deep" if its existence is concealed from the public; a secret is "shallow" if "ordinary citizens understand they are being denied relevant information and have some ability to estimate its content."[214] Private firms, which exist outside the executive branch, are arguably able to increase and facilitate public knowledge of surveillance practices.[215]

---

208. Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y REV. 15, 15 (2016) ("New technologies have altered surveillance discretion by lowering its costs and increasing the capabilities of the police to identify suspicious persons."); Andrew Guthrie Ferguson, *The Exclusionary Rule in the Age of Blue Data*, 72 VAND. L. REV. 561, 597 (2019).

209. Orin S. Kerr, *Searches and Seizures in a Digital World*, 119 HARV. L. REV. 531, 536 (2005).

210. Ohm, *supra* note 113, at 395–96 (arguing that *Carpenter* requires law enforcement to get a warrant before obtaining smart home data from third-party technology providers).

211. *See* Jon D. Michaels, *All the President's Spies: Private-Public Intelligence Partnerships in the War on Terror*, 96 CALIF. L. REV. 901, 922–26 (2008) (describing how informal public-private partnerships can minimize leaks, negative publicity, and legal risk).

212. Smith, *supra* note 107, at 602; Manes, *supra* note 15, at 351.

213. Bloch-Wehba, *supra* note 122, at 14 ("Diminished Fourth Amendment protections have also made it much more difficult for courts, defendants, and the public to get critical information necessary to check the police.").

214. David E. Pozen, *Deep Secrecy*, 62 STAN. L. REV. 257, 274 (2009–10).

215. Manes, *supra* note 15, at 344 ("If these companies could win the right to speak about the *kinds* of records the government is ordering them to disclose, they would be able to provide

Yet the increasing role of private sector actors in policing has not appreciably diminished law enforcement secrecy, but rather shifted the locus of claims of secrecy to private sector actors. Today, technology companies routinely invoke trade secrecy and other corporate protections to avoid transparency about the role they play in facilitating law enforcement investigations and prosecutions.[216] In an atmosphere of increased calls to defund and reform policing, secrecy plays an essential role in protecting law enforcement's private partners from reputational risk.[217] By shielding private firms from the reputational costs of partnering with police, however, law enforcement also reduces the public's ability to monitor and understand how the government conducts surveillance.

Third, technology firms are unconstrained by constitutional limitations. For Fourth Amendment purposes, constitutional constraints on searches and seizures are limited to "unreasonable searches undertaken by the government or its agents—not private parties."[218] As Kiel Brennan-Marquez has documented, the traditional approach has been to examine whether a private party was "deputized" by the state to investigate; if so, the party loses its "private" status, and the Fourth Amendment applies to the search.[219] Brennan-Marquez points out, however, that courts have uniformly held that *voluntary* private hashing such as that accomplished by PhotoDNA software or the GIFCT is not covered by the Fourth Amendment.[220] This creates a legal gap, inviting law enforcement to exploit private action and informal relationships to extend its own power. In contrast, Brennan-Marquez suggests that, where the government relies on private action to extend the infrastructure of surveillance, the Constitution ought to follow.[221]

Here, regulatory action and jawboning have created new incentives for platforms to engage in private policing that itself drives law enforcement

---

the public with crucial information about how the surveillance laws have been interpreted and applied in practice.").

216. *See generally* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343 (2018); *see also* Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 87 (2019) (describing how Northpointe, a firm that provides risk assessment tools used at sentencing, conceals the weight of its risk scores based on trade secrecy); Natalie Ram, *Innovating Criminal Justice*, 112 NW. U. L. REV. 659, 668–71 (2018) (describing how, pursuant to non-disclosure agreements with Harris Corp., police concealed the use of Stingray technology).

217. Michaels, *supra* note 211, at 926 (arguing that "handshake collaborations" with government agencies may generate litigation risk for firms).

218. *Ackerman*, 831 F.3d at 1295.

219. Brennan-Marquez, *supra* note 17, at 488.

220. *Id.* at 504.

221. *Id.* at 505.

action. Yet Fourth Amendment protections have not followed. Voluntary private searches for unlawful content trigger no Fourth Amendment protection. Neither, seemingly, does the use of new technologies of surveillance that reimagine online speech as a source of law enforcement intelligence.[222] Indeed, the lack of constitutional constraints likely encourages informal relationships between tech firms and law enforcement by avoiding the heavy costs of the warrant requirement, reasonableness limitations, and the exclusionary rule.[223] As Brennan-Marquez suggests, these circumstances may warrant a reimagining of the constitutional status of purportedly "private" searches.[224]

To be clear, I do not suggest that any time a platform takes action on content in a manner favored by the government, the Constitution ought to attach. Technology firms are powerful, wealthy actors; the fact that their interests sometimes (or even often) align with the government's is hardly surprising, nor is it cause for inherent suspicion. But the extensive alignment between platforms and states ought to prompt scholars and policymakers to reexamine the prevalent assumption that technology firms are engaged in forms of private governance accountable to nobody except, possibly, their shareholders. The truth is that, while platforms can provide a powerful counterweight to state action, they can just as easily buttress it.

C. DESIGN AND LEGAL IMMUNITY

Another way in which the government might limit private accountability for partnering with law enforcement is through granting immunity from suit for firms. The government might simply decide to require technology firms to design their services and products in ways that are more amenable to law enforcement. For example, the Communications Assistance for Law Enforcement Act (CALEA) requires telecommunications carriers to design their services to be capable of providing access for law enforcement.[225] As Gus

222. *See* Jeramie D. Scott, *Social Media and Government Surveillance: The Case for Better Privacy Protections for Our Newest Public Space*, 12 J. BUS. & TECH. L. 151, 158–59 (2017) (explaining that Fourth Amendment does not protect social media posts "knowingly expose[d]" to the public eye); *cf.* Ken Dilanian, *DHS Launches Warning System to Find Domestic Terrorism Threats on Public Social Media*, NBC NEWS (May 10, 2021, 10:30 AM), https://www.nbcnews.com/politics/national-security/dhs-launches-warning-system-find-domestic-terrorism-threats-public-social-n1266707 (describing a Department of Homeland Security proposal to use social media to detect domestic terrorism threats).

223. *Cf.* William J. Stuntz, *Distribution of Fourth Amendment Privacy*, 67 GEO. WASH. L. REV. 1265, 1274–77 (1999) (observing that Fourth Amendment law makes certain investigative activities more "costly" than others).

224. Brennan-Marquez, *supra* note 17, at 489.

225. 47 U.S.C. § 1002(a).

Hurwitz has pointed out, CALEA was "arguably the first time that Congress had imposed affirmative design requirements on firms in order to support law enforcement capabilities."[226] Outside the United States, regulators are already pressuring platforms to redesign their moderation techniques and rules to prioritize law enforcement, as the NetzDG and the EU Terrorist Regulation both demonstrate. As outlined above, a growing chorus of proposals would also require platforms to retain and disclose user data for law enforcement as well.

Congress may also grant statutory immunity to firms that partner with law enforcement. Consider the response after the New York Times published a story in 2005 detailing how telecommunications companies had partnered willingly with federal law enforcement and intelligence agencies after September 11th to collect the contents of communications under what was known inside the Bush Administration as the "Terrorist Surveillance Program," which became colloquially known as "warrantless wiretapping."[227] In the following months and years, dozens of lawsuits were filed against the telecommunications companies themselves, as well as the National Security Agency (NSA).[228] In response, Congress enacted the Federal Intelligence Surveillance Act of 1978 Amendments Act of 2008 ("FISA Amendments Act"), which codified provisions authorizing the bulk collection of some foreign communications.[229] The FISA Amendments Act also granted conditional statutory immunity to private firms that worked with the Terrorist Surveillance Program under assurances that the program was lawful.[230] News organizations have also reported that telecommunications firms have assisted law enforcement with wiretapping in legally dubious circumstances after

226. Justin Hurwitz, *Encryption^Congress Mod(Apple + CALEA)*, 30 HARV. J. L. & TECH. 355, 379 (2017).

227. James Risen & Eric Lichtblau, *Bush Lets U.S. Spy on Callers without Courts*, N. Y. TIMES (Dec. 16, 2005), https://www.nytimes.com/2005/12/16/politics/bush-lets-us-spy-on-callers-without-courts.html.

228. *See, e.g.*, *In re* Nat'l Sec. Agency Telecomm. Rec. Litig., 671 F.3d 881, 890 (2011) (analyzing statutory immunity for telecommunications companies that collaborated with NSA warrantless wiretapping); Jewel v. Nat'l Sec. Agency, 673 F.3d 902, 906 (9th Cir. 2011) (permitting AT&T subscriber to proceed in First and Fourth Amendment challenge to warrantless wiretapping that AT&T conducted "in collaboration with the [NSA]") ; American Civil Liberties Union v. Nat'l Sec. Agency, 493 F.3d 644 (6th Cir. 2007) (finding that plaintiffs lacked standing for First and Fourth Amendment claims against the NSA regarding the Terrorist Surveillance Program).

229. Pub. L. No. 110-261.

230. 50 U.S.C. § 1885a(4)(A); *In re* Nat'l Sec. Agency Telecomm. Rec. Litig., 633 F. Supp. 2d 949, 959 (2009) (describing the FISA Amendments Act immunity provision as "sui generis" because of its limitations on subject-matter, time period, and those who could invoke it).

having been granted immunity from prosecution in what are known as 2511 letters.[231]

Finally, amid growing calls to rethink or repeal Section 230, immunity for partnering with law enforcement may become more significant. While platforms' "right to exclude" user-generated content has increasingly been called into question, the obligation to comply with law enforcement demands remains at the core of many proposals to redesign Section 230.[232] If Congress were to enact legislation that renders platforms immune from suit when they take down "unlawful" content, they may also immunize platforms from liability for their collaboration with law enforcement in determining whether content is, in fact, unlawful.

V. CONCLUSION

Amid broadening recognition that social media platforms aggravate offline harms, like election tampering, communal violence, public health risks, and genocide, platforms' collaborations with law enforcement institutions bring both positive and negative effects. Synergies between public policy and private platform decision-making surely strengthen the government's ability to put its priorities into action. At the same time, however, the emergent ecosystem of information-sharing, collaboration, and public-private cooperation undermines the conventional wisdom that platform power necessarily comes at the expense of state authority, and vice versa. Law enforcement exerts both direct and indirect pressure on platform content rules, urging platforms to adopt more restrictive community standards, facilitate speedier takedowns, and share more information about harmful content with regulators. At the same time, platforms' affordances are reshaping law enforcement investigations and advancing surveillance.

Instead, as Reidenberg recognized, *lex informatica* can advance the goals of regulation as easily as inhibit them. The truth is that platforms in many contexts reflect the values of governments and specifically reflect the need for effective law enforcement. The same features that make social media so

231. Janus Kopfstein, *AT&T Getting Secret Immunity from Wiretapping Laws for Government Surveillance*, THE VERGE (Apr. 24, 2013, 2:42 PM), https://www.theverge.com/2013/4/24/4261410/att-getting-secret-wiretapping-immunity-government-surveillance.

232. Biden v. Knight Inst., 593 U.S. __, 8 (2021) (Thomas, J., concurring); *see* Stop the Censorship Act, H.R. 4027 (116th Cong.), Sec. 2 (eliminating platforms' immunity for moderating content that it deems objectionable but preserving immunity for taking down "unlawful content"); Protecting Constitutional Rights from Online Platform Censorship Act, H.R. 83 (117th Cong.), Sec. 2 (making it unlawful for platforms to moderate "protected" content, and by implication excluding illicit material from the definition of "protected").

compelling as a technology of mass communication—the ability to instantaneously reach a broad audience—make it equally compelling as a technology of surveillance that is easy and cheap to use in both targeted and dragnet investigative contexts.

Yet twenty-three years after Reidenberg's germinal observations, U.S. law has made little progress in ensuring that *lex informatica* is as democratically legitimate or accountable as its regulatory equivalents. As technology firms increasingly rely on automation and predictive technology to define the boundaries between lawful and unlawful speech, the private policies and techniques of platform governance are increasingly transmuted into public law enforcement institutions. Despite the blurry boundaries between firm and state, the laws of surveillance and information-sharing continue to recognize a sharp divide between public and private actors. Amid growing calls to fundamentally rethink, reshape, or abolish U.S. policing, we should reconsider how the law enables private sector firms to act as a force multiplier for law enforcement.