"Outside Experts": Expertise and the Counterterrorism Industry in Social Media Content Moderation

Amre Metwally*

INTRODUCTION

Political violence is as old as civilization itself. Wars—local, civil, regional, and global—are, by their very natures, both political and violent. The

* J.D. Candidate, Harvard Law School. The author was previously a Policy and Enforcement Manager covering political extremism, counterterrorism, and graphic violence for YouTube. The author would like to thank the expert advice from Professor Naz Modirzadeh and Cecil Yongo Abungu in helping to bring this endeavor from thought to finished product. A special thanks as well must be extended to my fellow participants (Alev Erhan, Shaiba Rather, Carla Yoon, Kathryn Reed, Stephanie Gullo, Molly Richmond, and Marta Canneri) in Professor Modirzadeh's International Law writing group for their incisive feedback and generosity. My deepest gratitude to Adam Silow and the team at the *Journal of National Security Law and Policy* for their assistance throughout the editing process. All errors are strictly my own.

phenomenon of political violence can be reported, categorized, labeled, and distilled into the distinct acts that comprise it: a killing, a bombing, an execution, and the list goes on. During the late twentieth and early twenty-first centuries, particularly from the mid-1900s onward, our lexicon complicated this understanding of political violence to introduce a new term: terrorism.

News reports of politically motivated violent acts began advancing this new term which quickly blossomed. "The year 1972 marked a major transition in the framing of the media's treatment of political violence. Events that previously were covered under the rubrics of assassination, bombing, torture, repression, massacre, etc., were now classified as 'terrorism.' The word (and hence the concept) was catching on."[1] This fascination with political, violent, and politically violent drama spread immeasurably with the September 2001 attacks. Since then, America's, and more broadly the West's, "war on terror"[2] spurred the justification for new invasions,[3] creation of new government agencies,[4] and establishment of new intelligence and information-sharing efforts[5] all designed to "protect" the United States from its purported enemy, which President George W. Bush described as "a radical network of terrorists and every government that supports them."[6]

We know how this story unfolds: invasions into Iraq and Afghanistan, countless civilian lives lost, covert torture programs, surveillance programs targeting Muslims in America, and countries that remain deeply entrenched in political dysfunction and violence that American intervention either introduced or further exacerbated.[7]

While this war on terrorism has already profoundly disrupted the twenty-first century, there has been at the same time another force unfolding: the technology industry, and specifically, social media companies. In 2004, the ambitious but

---

1. JOSEBA ZULAIKA & WILLIAM A. DOUGLASS, TERROR AND TABOO: THE FOLLIES, FABLES, AND FACES OF TERRORISM 46 (2006).

2. *Text: President Bush Addresses the Nation*, WASH. POST (Sept. 20, 2001), https://perma.cc/D8CF-TT7V.

3. *See*, *e.g.*, *Full Text: Bush's Speech*, GUARDIAN (Mar. 17, 2003, 9:22 PM), https://perma.cc/YFJ8-M7UR.

4. *See*, *e.g.*, Homeland Security Act of 2002, Pub. L. No. 107–296, 116 Stat. 2135 (2002) (creating the United States Department of Homeland Security).

5. *See*, *e.g.*, Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act (USA PATRIOT Act) Act of 2001, Pub. L. No. 107–56, 115 Stat. 272 (2001). The PATRIOT Act has ultimately been amended and reauthorized since the initial 2001 law, extending its powers, *see, e.g.*, The USA PATRIOT Improvement and Reauthorization Act of 2005, Pub. L. No. 109-177 (2005). The USA FREEDOM Act ultimately made permanent many provisions of the PATROIT Act while also limiting the National Security Agency's bulk collection of communications material belonging to US citizens. *See* Uniting and Strengthening America by Fulfilling Rights and Ensuring Effective Discipline Over Monitoring Act of 2015, Pub. L. No. 114-23, Stat. 268 (2015).

6. *Text: President Bush Addresses the Nation*, *supra* note 2.

7. This does not even include the American military's effort to expand targeted killing programs through the use of drone aircraft. For more on the drone program, see Christopher J. Fuller, *The Origins of the Drone Program*, LAWFARE BLOG (Feb. 18, 2018, 10:00 AM), https://perma.cc/MZS7-8BHA.

naïve Mark Zuckerberg launched Meta.[8] Steve Chen, Chad Hurley, and Jawed Karim created YouTube in 2005, now the world's largest video platform.[9] And just one year later, in 2006, Jack Dorsey (who was until recently the company's Chief Executive Officer), Evan Williams, and their fellow co-founders changed the way we communicate, 140 characters at a time.[10]

The war on terrorism and technology began to collide in 2014 when social media intersected with the West's newest "terrorists" to emerge from Iraq, the Islamic State in Iraq and Syria (ISIS). As ISIS erupted onto the world stage:

> Far from keeping their operation a secret, though, these [ISIS] fighters made sure everyone knew about it. [. . .] To maximize the chances that the internet's own algorithms would propel it to virality, the effort was organized under one telling hashtag: #AllEyesOnISIS. [. . .] [The hashtag] took on the power of an invisible artillery bombardment, its thousands of messages spiraling out in front of the advancing force. Their detonation would sow terror, disunion, and defection.[11]

ISIS was not the only group to masterfully wield YouTube, Twitter, and Meta to broadcast hostage executions, threats, recruitment messages, propaganda, and battlefield victories.[12] Other actors and organizations flocked to these platforms as well. For example, Rudaw, a Kurdish news agency, set up a live stream to capture the fighting; contractors at the US State Department would engage directly with individuals who seemed likely to join ISIS, and the Iraqi military would broadcast their wins against the Islamic State.[13]

While ISIS may have been one of the first actors to exploit social media's potential for its bloody aims in a highly visible manner, they were not the last. In 2019, Brenton Tarrant live-streamed himself entering multiple mosques in Christchurch, New Zealand, killing 51 worshippers.[14] He teased his murders on Twitter and then broadcasted the terrorist attack on Meta Live; his manifesto flourished in YouTube comments and 8chan boards.[15] While the graphic footage of terrorist attacks circulating online is not new, in many ways, this terrorist attack was "a first — an internet-native mass shooting, conceived and produced entirely within the irony-soaked discourse of modern extremism."[16]

---

8. *Our History*, ABOUT META, https://perma.cc/4929-N3Z6. The Article uses Facebook's new name "Meta" after the company recently changed it.

9. Laura Fitzpatrick, *Brief History YouTube*, TIME (May 31, 2010), https://perma.cc/37B8-7ZH4.

10. *Our Leadership*, TWITTER, https://perma.cc/N7NQ-DY8B.

11. P.W. SINGER & EMERSON T. BROOKING, LIKEWAR: THE WEAPONIZATION OF SOCIAL MEDIA 4-5 (2018).

12. *See*, *e.g.*, Ahmed Shehabat & Teodor Mitew, *Black-boxing the Black Flag: Anonymous Sharing Platforms and ISIS Content Distribution Tactics*, 12 PERSPECTIVES ON TERRORISM 81, 83-85 (2018).

13. SINGER & BROOKING, *supra* note 11, at 10.

14. *Christchurch Shootings: 49 Dead in New Zealand Mosque Attacks*, BBC (Mar. 15, 2019), https://www.bbc.com/news/world-asia-47578798.

15. Kevin Roose, *A Mass Murder of, and for, the Internet*, N.Y. TIMES (Mar. 15, 2019), https://perma.cc/PT3M-Y3YQ.

16. *Id.*

During the rapid spread of the Islamic State and after the Christchurch terrorist attack, many have criticized social media platforms for the role they play in helping extremist content multiply in the digital realm and for their role in radicalizing people to commit acts of violence.[17] The Royal Commission of Inquiry into the Terrorist Attack on Christchurch Mosques on 15 March 2019 notes in their almost-800 page report that:

> We [the Commission] have no doubt that the individual's internet activity was considerably greater than we have been able to reconstruct [. . .] He [the terrorist] also visited other sites and discussion boards where there was discussion promoting extreme right-wing and ethno-nationalist views similar to his own and sometimes supporting violence. He also spent much time accessing broadly similar material on YouTube. His exposure to such content may have contributed to his actions on 15 March 2019 – indeed, it is plausible to conclude that it did.[18]

The Commission also writes that "[t]he individual [i.e., the shooter] claimed that . . . YouTube was, for him, a far more significant source of information and inspiration."[19]

While ISIS, the Christchurch Shooter, and other violent assailants flocked to social media, there was another story that intersected with all of theirs: my own. For nearly three years, I worked at YouTube as the company's policy manager for counterterrorism, political extremism, and graphic violence. I wrote YouTube's counterterrorism policy, which laid out what signals and legal designations the company will rely on to consider an actor a terrorist entity, a label that, once conferred, invited much harsher action compared to other actors in the extremism umbrella.

As an "expert" on terrorism content policy, it is frightening to consider how much power I had in shaping YouTube's position. For a video platform that receives over 500 hours of uploads *every minute*,[20] I possessed an inordinate level of influence: a notion that becomes particularly clear after YouTube's stance shifted right at the end of my tenure to become "tougher" on Hamas and Hezbollah material.[21] Why? In part, of course, because YouTube was under

17. *See*, *e.g.*, Alan Travis, *MPs say Meta, Twitter, and YouTube 'Consciously Failing' to tackle Extremism*, GUARDIAN (Aug. 24, 2016, 7:01 PM), https://perma.cc/7WH2-L7WR;; Kevin Roose, *The Making of a YouTube Radical*, N.Y. TIMES (June 8, 2019), https://perma.cc/365J-KXT4; Tanya Basu, *YouTube's Algorithm Seems to be Funneling People to Alt-Right Videos*, MIT TECH. REV. (Jan. 29, 2020), https://perma.cc/2SR4-LBC4.

18. ROYAL COMM'N OF INQUIRY INTO THE TERRORIST ATTACK ON CHRISTCHURCH MOSQUES ON 15 MARCH 2019, REPORT OF THE ROYAL COMMISSION OF INQUIRY INTO THE TERRORIST ATTACK ON CHRISTCHURCH MASJIDAIN ON 15 MARCH 2019 234 (2020), https://perma.cc/PBS8-2NSU [hereinafter ROYAL COMMISSION]. For more commentary on the Commission's reports, see Cecilia D'Anastasio, *The Christchurch Shooter and YouTube's Radicalization Trap*, WIRED (Dec. 8, 2020, 7:51 PM), https://www.wired.com/story/christchurch-shooter-youtube-radicalization-extremism/.

19. ROYAL COMMISSION, *supra* note 18, at 193.

20. *YouTube for Press*, OFFICIAL YOUTUBE BLOG, https://perma.cc/D5KD-GMEX.

21. *See* Sheera Frenkel & Ben Hubbard, *After Social Media Bans, Militant Groups Found Ways to Remain*, N.Y. TIMES (Apr. 19, 2019), https://perma.cc/SED4-F9Q6.

pressure to moderate content in lockstep with its peers, Meta and Twitter. However, a larger, more compelling factor to consider is because YouTube had an army of academics, research firms, and an "Intelligence Desk"[22] that saw the matter differently and outnumbered me. The "experts" won.

In nearly every crisis involving violent extremism—and the accompanying critiques that social media platforms have failed to do more to stop this problem—company executives, in return, have consistently responded with promises to be better. In their blog posts and interviews, however, an interesting word keeps appearing: experts. "[W]e need a lot more experts," Susan Wojcicki, YouTube's CEO, has said.[23] In a 60 Minutes interview, Wojcicki has said that "[f]or every area we work with experts."[24] Vijaya Gadde, Twitter's Chief Legal Officer and Head of Trust, said in an interview about the platform's policies around white supremacist content and the company's decision not to allow individuals with affiliations to the KKK to hold a Twitter account:

> I want to be very clear that that is our policy, we'll continue to enforce that. We do have work to do in terms of understanding what more we should be doing. That is the work that we're engaging in. I don't want to make that decision all by myself, because there are a lot of experts who work on radicalization on the ground in these communities. Engaging in these conversations. I want the benefit of their expertise and their opinions before I make further changes.[25]

Meta's content policy team holds bi-weekly meetings in which potential changes to their community guidelines are discussed.[26] Before these sessions, "members of our [Meta's] content policy team reach out to internal and external experts" as well.[27]

This expertise and the accompanying rhetoric of reliance on expertise neither starts nor stops with social media companies. The cult of the terrorism expert extends to multiple sectors that are increasingly intertwined with Silicon Valley, from think tanks to military and intelligence agencies, government bodies to the media, and new private companies offering services for the tech titans that create, monitor, and control our digital public spaces. The reality is that because of the constant preoccupation with terrorism in our society—physical and digital—the business of terrorism experts is booming.[28] More concerning is that with this

22. *See* discussion *infra* Part II.B for more on YouTube's "Intelligence Desk" team.

23. Julia Alexander, *YouTube Needs More Experts to Help Tackle Dangerous Content, Says CEO*, POLYGON (Jan. 29, 2018, 10:44 PM), https://perma.cc/A4ZH-75SS.

24. Connie Loizos, *In '60 Minutes' Appearance, YouTube's CEO Offers a Master Class in Moral Equivalency*, TECHCRUNCH (Dec. 1, 2019, 8:16 PM), https://perma.cc/ZYN8-AU7M.

25. Eric Johnson, *Twitter's Kayvon Beykpour and Vijaya Gadde: The Code Conference Interview (transcript)*, VOX (June 27, 2019, 6:20 AM), https://perma.cc/6KSJ-ZW8F.

26. *Writing Meta's Rulebook*, ABOUT META (Apr. 10, 2019), https://perma.cc/P56X-92KN. *See infra* Part III.A and Part III.B for more on Meta's community guidelines procedures.

27. *Writing Meta's Rulebook*, *supra* note 26.

28. *See infra* Part II.B for more information on these third-party companies offering services to social media platforms.

expertise comes the domination of certain types of discourses or ideology over others, raising genuine questions of bias or politically motivated agendas as experts consult with, or work at, social media platforms. Another concern is simply with the control these experts have over how companies think about, define, and enforce content policies.

This paper details, examines, and challenges the expertise industry in terrorism and social media content moderation efforts. Ultimately, its loftier goal is to offer a critical case study of how self-anointed experts create orthodox narratives of terrorism and cajole companies into embracing these viewpoints by collaborating with companies in several ways—and why resistance through the adoption of unorthodox points of view is critical. Part I provides an overview of the critiques and studies of expertise and also summarizes key debates on experts and expertise within the field of Critical Terrorism Studies. Part II documents how expertise is produced, legitimated, and communicated both inside social media companies, with experts from the public sector moving into the technology industry and the growth of "Intelligence Desks," and around the social media platforms, with third-parties increasingly meddling in companies' content policy and enforcement work. Part III discusses in more detail why these organizations and experts raise significant concerns as companies are under increasing pressure to be ever more aggressive in identifying and removing terrorist content online, a particularly lucrative point from the perspective of nation states as they engage in counterterrorism work and circumvent limitations of international counterterrorism law. Part IV examines potential solutions and recommendations that could be employed to address this unregulated revolving door.

I. THE PHENOMENON OF EXPERTISE IN TERRORISM STUDIES

"Because of the complex nature of terrorism, we believe it is valuable to study it from a variety of perspectives," boasts King's College London (KCL)—home to the world-famous War Studies Department—on its page advertising a graduate program in Terrorism, Security, and Society.[29] Where exactly do terrorism studies lie? Not quite a field of its own, this discipline perhaps finds itself at home in the social sciences for its "interdisciplinary" focus. For example, KCL's master's degree pays homage to the different disciplines that touch on terrorism, writing that "[o]ur Terrorism, Security & Society MA is an interdisciplinary course that draws on history, political science, international relations, sociology, social psychology and risk studies to understand international security threats."[30]

The dilemma here is that just as it is unclear what constitutes "terrorism," it is also just as baffling to delineate what terrorism studies are, where it lies compared to other academic disciplines, and what constitutes expertise in it. Before diving deeper into this discussion on expertise and experts, it is worth clarifying the

29. *Terrorism, Security* & Society MA, KING'S COLL. LONDON, https://perma.cc/C3GP-SE4H.
30. *Id.*

ecosystem in which all these actors operate. The terrorism industry relies on multiple sectors. First, the public sector, composed of government officials and agencies that "establish policy and provide opinions and selected facts about official acts and plans on terrorist activity."[31] Second, the private sector that is comprised of research centers, security companies, and think tanks that "deal in risk analysis, personal and property protection, and training, and a body of terrorism 'experts.'"[32] The security firm officials are typically from intelligence and security agencies in the government—a reality that should be quite obvious when one considers the connections and reputational networks on which much of this industry is built.[33] The experts in this terrorism industry "are associated mainly with the institutes and think tanks, some of which are affiliated with academic institutions, but officials and analysts of security firms are also regarded as authorities on terrorism."[34]

So what makes someone an expert? A physician, for example, is an "expert" with a clearly defined credential followed by years of residency training. A lawyer, a professor, an educator—all examples of pathways and professions where a degree builds and signals expertise. Terrorism "experts," however, run the gamut. Some develop this expertise with no academic rigor but through years of work experience in different sectors or industries, whether through a government agency, the military, or as a media pundit. Others obtain graduate education in a discipline like political science or history, only to pivot and become a terrorism expert. In other words, there is "no set career path to becoming a terrorism expert, nor is there any recognized credentialing body."[35] As Lisa Stampnitzky notes:

> Even specialized research journals and conferences, which represent the most professionalized and internally regulated areas of the terrorism studies world, have been populated by a high proportion of one-time authors, those who enter with no significant background in the field, and then disappear. Of 1,796 individuals presenting at conferences on terrorism between 1972 and 2001, 1,505 (84%) made only one appearance. Similarly, a study of journal articles published on terrorism during the 1990s found more than 80% to be by one-time authors, while another study found that core journals in terrorism studies had significantly higher rates of contributions from non-academic authors than journals in political science or communications studies.[36]

---

31. EDWARD HERMAN & GARY O'SULLIVAN, THE "TERRORISM" INDUSTRY: THE EXPERTS AND INSTITUTIONS THAT SHAPE OUR VIEW OF TERROR 55 (1989).

32. *Id.*

33. *See id.* at 56.

34. *Id.* at 55. Herman and O'Sullivan's point that employees at firms are also seen as terrorism experts will hopefully become much more apparent in this paper when the third-party ecosystem that has emerged around Silicon Valley is discussed in Part II.B, *infra*.

35. Lisa Stampnitzky, *Disciplining an Unruly Field: Terrorism Experts and Theories of Scientific/Intellectual Production*, 34 QUALITATIVE SOC. 1, 8 (2011).

36. *Id.* at 8.

I illustrate these examples to show the difficulty behind understanding how expertise in this space is generated and how individual legitimacy is conferred. "[W]hile sociological studies of expertise . . . focus upon arenas in which the object of expertise is already 'formed,' . . . terrorism studies presents an example . . . in which the object of knowledge is not only not yet stabilized, . . . it is not clear that it will ever completely take settled form."[37] However, more important to consider is that this expertise is not necessarily meant to be in conversation with fellow terrorism experts (or, at the very least, not confined to conversations with peers). This poorly regulated field has, as its audience, "not an ideal-typical scientific community, but rather the public and the state."[38] This expertise exists in "multiple arenas of knowledge production, consumption, and legitimation, including academia, the media, and the state."[39] Terrorism studies, and those who claim to be terrorism experts, tend to be heavily reliant on the government, not only for funding but also for the production of knowledge in this field.[40]

This politicization of academia and academic output by the state is not new.[41] Because "the military seek expertise whilst the academic seeks funding and an outlet for research,"[42] an option for mutually beneficial knowledge production emerges. This relationship moves to a point where "the line between academic expert at a distance and functionary of the state has, at least, been blurred."[43]

As such, Critical Terrorism Studies (CTS) has emerged as a growing discourse meant to wield critical theory to challenge predominant narratives on terrorism, terrorism scholars, and knowledge production in this space. Writings critical of mainstream terrorism research and scholarship point to a "lack of conceptual clarity and theoretical sterility to political bias and a continuing dearth of primary research data."[44] Critical scholars have used their publications to point out how terrorism experts attempt to legitimize their self-described "expertise." For example, Richard Jackson discusses how many experts provide testimony in governmental hearings or commissions, only to cite their own testimony later to create the illusion and self-legitimization of their expertise.[45]

Other critical voices have examined how terrorism experts have created an "expert nexus" where status quo or orthodox points of view of terrorism are shared and presented as the dominant discourse, a significant danger when one

37. *Id.* at 3.
38. *Id.* at 7.
39. *Id.* at 3.
40. *See* Andrew Silke, *An Introduction to Terrorism Research*, *in* RESEARCH ON TERRORISM: TRENDS, ACHIEVEMENTS, AND FAILURES 1, 15 (Andrew Silke ed., 2004).
41. *See* David Miller & Tom Mills, *Counterinsurgency and Terror Expertise: The Integration of Social Scientists into the War Effort*, 23 CAMBRIDGE REV. OF INT'L AFF. 203, 203 (2010).
42. *Id.*
43. *Id.* at 205.
44. Jeroen Gunning, *A Case for Critical Terrorism Studies?*, 42 GOV'T AND OPPOSITION 363, 363 (2007).
45. RICHARD JACKSON, *Knowledge, Power and Politics in the Study of Political Terrorism*, *in* CRITICAL TERRORISM STUDIES – A NEW RESEARCH AGENDA 66, 81 (Richard Jackson, Marie Breen Smyth & Jeroen Gunning eds., 2009).

contextualizes this nexus in the broader industrial web of the state, the military, private companies, and the media.[46] David Miller and Tom Mills use the sociology of science concept of "invisible colleges," which are "informal communication networks of scientists who come to form an elite and to dominate a field."[47] And in these colleges "there exists a sort of commuting circuit of institutions, research centers, and summer schools giving them [the scholars] an opportunity to meet piecemeal, so that over an interval of a few years everybody who is anybody has worked with everybody else in the same category."[48] Miller and Mills tracked an invisible college of 100 terrorism experts that appear most frequently in the press and found:

> A significant number of the experts (42 out of 100) are currently or have previously been a member of state institutions such as government, security or intelligence services, policing or the military. The majority of the experts (67 of the 100) are currently or have previously been members of private think-tanks or research institutes. Of the remaining experts, 16 out of 33 are currently or have previously worked in private security or intelligence firms, or alternatively state institutions such as government security or intelligence, policing or military service.[49]

The invisible college mapped in this study was broken down into three categories of expertise—orthodox, alternative, critical—with the vast majority of the college (73 of 100) espousing orthodox ideology, 26 supporting alternative viewpoints, and only one that was critical of the narrative of terrorism studies and Western foreign policy.[50] The study also revealed that the terrorism experts that appeared most often in the media had strong ties to government agencies, the military, and corporations and were mostly supporters of orthodox views on terrorism.[51] In other words, "terror experts are not simply expressions of the ideological needs of state and corporate actors, but are actually a functional part of the governing nexus . . . [A]lthough the orthodox experts are mostly 'far from the killing fields' their 'spirit' is 'there, on the front lines and in the torture chambers.'"[52]

Orthodoxy, and the fact that the vast majority of this sample of experts espouse it, warrants further explication. The problem is not necessarily that there is something inherently wrong with orthodox views. The issue with the dominance of orthodoxy is that terrorism expertise engages, as Richard Jackson frames it, in

46. *See generally* David Miller & Tom Mills, *The Terror Experts and the Mainstream Media: The Expert Nexus and Its Dominance in the News*, 2 CRITICAL STUD. ON TERRORISM 414 (2009).

47. *Id*. at 417.

48. DEREK DE SOLLA PRICE, LITTLE SCIENCE, BIG SCIENCE . . . AND BEYOND 75-76 (1986).

49. Miller & Mills, *supra* note 46, at 419.

50. *Id*. at 422.

51. *Id*. at 431.

52. *Id*. Miller and Mills also discuss "media-source relation" theory, a field in communication studies that examines the interplay between journalists/media production and the sources they depend on.

Foucault's concept of "subjugated knowledge"[53] in which certain knowledge is known but subjugated.[54] Experts contribute to the production of a particular narrative and with it an accompanying body of knowledge. Terrorism expertise works by positioning itself close to power (i.e., governments and influential private institutions) and harmonizing its rhetoric so that "it accords with hegemonic commonsense and cultural narratives of political violence, legitimacy and security."[55] Terrorism experts maintain legitimacy through a "code of authorisation" which includes "such factors as social scientific credentials, personal military or counterterrorism experience, institutional proximity to power (the defence establishment[,] academic institutions or state-recognised security think-tanks) and evidence of the provision of advice to policymakers."[56] The ideologies present in particular institutions, careers, and degrees are legitimated and seen as expert at others' expense. With this prevalent, hegemonic discourse created, concentrated, and then propagated by experts, any contradictory knowledge is ultimately buried or suppressed.[57]

This echo chamber of ideology that is then espoused to the public, the state, and private companies, is not the only concerning trend with the growth of terrorism experts. Since 2001, funding has increased significantly for terrorism research "for every group, every company, every sector of society, and every lobbyist."[58] The National Science Foundation, for example, gave almost $50 million for terrorism research after 2001.[59] The vast majority of funding comes from other agencies like the Defense Department, with "much of it going to consulting

53. Richard Jackson quotes Michel Foucault who says "When I say 'subjugated knowledges' I mean two things. On the one hand, I am referring to historical contents that have been buried or masked in functional coherences or formal systemizations. [. . .] Subjugated knowledges are, then, blocks of historical knowledges that were present in the functional and systematic ensembles, but which were masked, and the critique was able to reveal their existence by using, obviously enough, the tools of scholarship. Second . . . when I say 'subjugated knowledges' I am also referring to a whole series of knowledges that have been disqualified as nonconceptual knowledges, as insufficiently elaborated knowledges: naïve knowledges, hierarchically inferior knowledges, knowledges that are below the required level of erudition or scientificity." Robert Jackson, *Unknown Knowns: The Subjugated Knowledge of Terrorism Studies*, 5 CRITICAL STUD. ON TERRORISM 11, 13 (2012).

54. Jackson gives many examples of known but subjugated knowledge in terrorism studies, including, but not limited to, the fact that terrorism can be carried out by state actors (and not just non-state ones) and that terrorism as a whole is a "statistically minor threat" to security in general. *Id.*

55. *Id.* at 16. Jackson provides the example of Peace Studies—a discipline that has often been dismissed as weak or inferior—as a discourse that is "counter-hegemonic/countercultural." *Id.* at 17.

56. *Id.* at 18.

57. *Id.* at 20. Jackson gives an example of when Rudolph Guiliani and a former Homeland Security Advisor gave public remarks in support of the Mujaheddin-e Khalq (MEK), although it is an organization that is labelled as a terrorist organization by the State Department. Although it is "a federal crime to engage in public advocacy of the group" the "knowledge that such actions were supporting terrorism neither prevented them from speaking nor induced the counterterrorism structures, the media or the terrorism studies field to react in the prescribed manner. Instead, the spasmodic contradiction was ignored, suppressed and ultimately tolerated." *Id.*

58. IAN LUSTICK, TRAPPED IN THE WAR ON TERROR 71 (2006).

59. *Id.* at 91.

firms, think tanks, and private research institutes."[60] In the post-9/11 period, funding from state and private sources has increased significantly.[61]

In sum, as has been highlighted throughout this part, determining why someone is an expert on terrorism is a very difficult question to ask. Not only because the path to becoming an anointed expert—by one's own declaration, through an invisible college, or the media and state—is quite unclear but also because terrorism expertise is a space, and process, through which certain discourses are legitimated over others, and funding can influence the research agendas of terrorism experts. "[D]ebates on terrorism invariably develop into contests of judgment as to whether particular acts, and actors, are, or are not, terrorist."[62] Conclusions or answers to these debates are decreed by these experts, particularly ones that espouse traditional viewpoints that may also be suppressing alternative interpretations of terrorism. As Part II will discuss in more detail, the outcome of expert decision-making extends not only to governments and the media, but now also social media platforms. "[T]here is a constant circulation of experts from internal positions such as the intelligence agencies to external sites such as think tanks. Further, large amounts of government research funding go to 'outside' terrorism experts, and state agencies regularly sponsor conferences composed of 'outside' experts and bring such experts in for consultations."[63] The next section examines and details the understudied issue of the terrorism industry's deep forays into and surrounding Silicon Valley.

II. THE EXPERTISE CONSTELLATIONS IN AND AROUND SILICON VALLEY

As discussed in Part I, the development and legitimation of expertise threatens to perpetuate certain worldviews and theories regarding terrorism and violence over others. Additionally, the funding incentives that flow from the state and military to research centers, universities, and academics can create ethical concerns of providing information and expertise to the military.[64] The tense relationship between "the military and the academy undermines the latter's obligation to remain critical and independent . . . . [A]t best the research agenda of the academy is being weighted more in the interests of power, and . . . at worst, particular experts are violating ethical norms."[65]

This scrutiny of terrorism expertise, knowledge production, the military, the state, and the private terrorism industry requires more examination to account for the dominance and prominence of Silicon Valley in our daily lives. The rise of social media has also generated a discourse of the need for experts and expertise

60. LISA STAMPNITZKY, DISCIPLINING TERROR: HOW EXPERTS INVENTED "TERRORISM" 197 (2013).

61. *See* Richard Jackson, *The Study of Terrorism 10 Years after 9/11*, INT'L REL., Winter 2012, at 1, 4-5.

62. STAMPNITZKY, *supra* note 60, at 203.

63. Lisa Stampnitzky, *Experts, États et Théorie des Champs: Sociologie de L'expertise en Matière de Terrorisme*, 59 CRITIQUE INTERNATIONALE 89, 90 n.17 (2013).

64. *See* Miller & Mills, *supra* note 41, at 204. Social media companies are also beginning to fund terrorism research through initiatives of their own. *See* discussion *infra* Part IV.

65. *Id.*

in the platforms' quest to combat controversial and offensive content—particularly violent extremist and terrorist material online. Part II first details the "inside experts" that social media companies have hired and turned to for help before examining the "outside expert" ecosystem of companies and research centers that provide services to, and advise, social media networks.

A. Inside Experts

Companies have increasingly turned to directly hiring individuals to help address the proliferation of terrorist propaganda online. In 2017, CNN ran an article with the eye-catching title "There's a new in-demand job at Meta: counterterrorism specialist."[66] That year alone, Meta had "more than 150 people who are mainly focused on fighting terrorism on the social network, including a mix of academics, analysts and former law enforcement agents."[67] YouTube also faced an avalanche of criticism over the proliferation of extremist content on its platform. Following controversies of advertisements running on such material,[68] YouTube expanded its content moderation staff to 10,000 and hired full-time terrorism specialists.[69] In a public blog post written by YouTube, the company noted that they have "also hired full-time specialists with expertise in violent extremism, counterterrorism, and human rights" and that they have "expanded regional expert teams."[70] Companies have also sprinted to staff up the leadership of teams dedicated to counterterrorism work. For example, Meta's head of counterterrorism is Brian Fishman—a former professor at West Point, military veteran, legislative assistant for Congresswoman Lynn Woolsey, and author of *The Master Plan: ISIS, al-Qaeda, and the Jihadi Strategy for Final Victory*.[71]

Counterterrorism experts who have worked with Fishman praised his appointment to run Meta's Dangerous Criminal Organizations team. *Wired* writes, "Fishman has a deep understanding of the online strategies deployed by Al-Qaeda and ISIS, and has used that expertise to help governments and nonprofits combat extremism. His background in academia will help Meta apply policies and technologies backed by strong research."[72] The piece cites another counterterrorism expert at the RAND Corporation—Colin Clarke—who served in Afghanistan with Brian Fishman. The piece writes that "Clarke says the clearest

66. Seth Fiegerman, *Meta Grows its Counterterrorism Team*, CNN (June 15, 2017, 1:59 PM), https://perma.cc/2RBV-TJ3V.

67. *Id.*

68. Jason Murdock, *Selling to Extremists: YouTube Ran Ads for Major Brands on Channels Promoting Nazis, Pedophilia, Propaganda*, NEWSWEEK (Apr. 20, 2018, 5:32 AM), https://perma.cc/A7E5-AY2D.

69. Jason Murdock, *Google: YouTube Hires Counterterrorism Experts to Help Police Website's Videos*, NEWSWEEK (Apr. 24, 2018, 5:47 AM), https://perma.cc/H7LQ-GHBH.

70. The YouTube Team, *More Information, Faster Removals, More People – An Update on What We're Doing to Enforce YouTube's Community Guidelines*, YOUTUBE OFFICIAL BLOG (Apr. 23, 2018), https://perma.cc/X68F-UAKD.

71. *Brian Fishman*, NEW AMERICA, https://perma.cc/8BHU-Y4ZA.

72. Emily Dreyfuss, *Meta's Counterterrorism Playbook Comes Into Focus*, WIRED (June 17, 2017, 7:00 AM), https://perma.cc/3KXZ-FRDM.

indication that Meta wants to get this right came last year when it hired Fishman to lead its counterterrorism efforts."[73] We see Fishman's legitimacy solidified in several ways: he served in the military, written a book, was as a professor at the West Point Military Academy, and, quite tellingly in the piece, has his legitimacy affirmed by another terrorism expert, one who worked with Fishman and has many of the same expertise "markers" as well.

Of course, Meta needs to hire individuals who can steer these policy teams through difficult terrain. And whether the company chose Brian Fishman or a different expert, concerns arising from orthodox points of view and the interconnected nature of relationships between experts, military, and the state remain. In fact, a study by César Ross and Gonzalo Montaner finds Brian Fishman to be one of the 20 most influential authors in the period between 2002-2012. The study's authors also note the "bulk of the intellectual production has revolved around the 20 authors . . . and the application and reproduction of the ideas of these authors, which have become dominant."[74] One of the study's conclusions is that these authors have reframed the focus of security studies since 9/11 to place "the security-religion axis as a matter of high relevance in international relations. Proof of this is that the actions and role of non-state organizations, such as Al Qaeda, dominate the agenda and have forced academics and researchers to expand their areas of security studies to understand this phenomenon."[75] In other words, not only is Brian Fishman leading terrorism-related work at Meta, he's contributing to the creation of dominant discourses on terrorism itself, ones that Meta and other social media companies are measured against in their content moderation efforts. Orthodoxy has moved from outside the company to inside it.

The use of inside experts is not confined to Meta. YouTube supplements its counterterrorism staff expansion with the creation of an intelligence staff called the Intelligence Desk. In an interview with YouTube's Chief Product Officer Neal Mohan on how the platform responds to the COVID-19 pandemic, the discussion veered towards conspiracy theories, notably the idea that 5G networks cause the coronavirus. In outlining some of the policy and technical features that have been in place to respond to this misinformation, Mohan says:

> One of the other tools that we established that's come in handy here is what we call an Intelligence Desk. This is a team of professionals who actually try to look kind of just over the horizon, if you will, in terms of where a conspiracy might be coming from, where misinformation might be coming from, so that we can do our best to sort of stay ahead of something that might be emerging before it becomes a challenge on our platform.[76]

---

73. *Id.*

74. César Ross & Gonzalo Montaner, *La Agenda de Los Estudios de Seguridad Post 9/11: ¿De Qué Y Quiénes Hablan?*, 12 REVISTA 15, 38 (2017).

75. *Id.* at 39.

76. Protocol, *YouTube's Chief Product Officer Neal Mohan*, YOUTUBE (Apr. 23, 2020), https://perma.cc/DH78-7N56.

This Intelligence Desk is not just focused on coronavirus misinformation. Susan Wojcicki, YouTube's CEO, has adressed this function in the past. In a profile of the executive, the *New York Times* noted that she "created an 'intelligence desk' to identify percolating issues on the internet more quickly."[77] Formed in January 2018,[78] this Intelligence Desk is part of a "multipronged 'early detection' initiative intended to ferret out controversial content before it spirals into a bigger problem."[79] In a statement to BuzzFeed News, a YouTube spokesperson confirmed that "part of those efforts [combating abuse on the platform] will include assembling new teams dedicated to protecting our platform against emerging trends and threats."[80] To achieve its ultimate goal of identifying new risks to the platform, the Desk relies "on Google data, user reports, social media trends, and **third-party consultants** to detect inappropriate content early, and either remove it or prevent advertiser messages from appearing near it."[81]

Commentators have worried about the impact this Intelligence Desk may have on YouTube creators, raising valid critiques along the way.[82] However, it is important to note that while the company has discussed this Desk's creation, it is unclear who runs it, the backgrounds of the employees that form this team, and who are the third-party consultants that have the ears of this branch of the video-sharing platform. And it matters greatly. As YouTube's Vice President of Government Affairs & Public Policy noted in a blog post she penned, "as a result of the Intelligence Desk's work to detect the evolving online tactics and impending statements of terrorist organizations, we shared 100,000 digital fingerprints (also known as hashes) of terror content to the Global Internet Forum to Counter Terrorism's hash-sharing database."[83] The Global Internet Forum to Counter Terrorism (GIFCT) is an initiative created by YouTube, Google, Meta, and Microsoft that established an industry-wide database of removed "terrorist" material that other platforms could contribute to or use to identify copycat uploads on other sites.[84] The volume of hashes already contributed by YouTube—coupled with the reality that these hashes are used by dozens of companies to find and potentially remove videos on their platforms—raises not only fears of censorship but also questions about who guides the Desk's decision making.

In a job posting for the Intelligence Desk, it notes under qualifications that a candidate should have "[a]nalytical experience in a **government intelligence**

77. Daisuke Wakabayashi, *The Most Measured Person in Tech is Running the Most Chaotic Place on the Internet*, N.Y. TIMES (Apr. 17, 2019), https://perma.cc/4NBY-7AM2.

78. The YouTube Team, *The Four Rs of Responsibility, Part 1: Removing Harmful Content*, YOUTUBE OFFICIAL BLOG (Sept. 3, 2019), https://perma.cc/JH7L-6FWG.

79. Alex Kantrowitz, *YouTube is Assembling New Teams to Spot Inappropriate Content Early*, BUZZFEED NEWS (Jan. 19, 2018, 1:43 PM), https://perma.cc/G6K4-T7EH.

80. *Id.* (internal quotation marks omitted).

81. *Id.* (emphasis added).

82. *See*, *e.g.*, Rachel Kaser, *YouTube's 'Intelligence Desk' Could Screw Legitimate Creators*, NEXT WEB (Jan. 23, 2018), https://perma.cc/FBU5-QJ3Q.

83. Leslie Miller, *How YouTube Supports Elections*, YOUTUBE OFFICIAL BLOG (Feb. 3, 2020), https://perma.cc/BT9X-MC7V.

84. *About*, GLOB. INTERNET F. TO COUNTER TERRORISM, https://perma.cc/25PU-MQGU.

organization or business environment."[85] The candidate posting also calls for "3 years of **experience in intelligence collection** and analysis."[86] The fact that the platform has such a function is, of course, perfectly within its rights. However, the notion that these roles in the Intelligence Desk need—or at the very least highly value—experience in intelligence jobs suggests that dependency on the state, in the form of its vast intelligence and defense apparatus, may be more than minimal. The idea of "revolving doors" between the state, the academy, and the military is reflected in the abundance of material discussing the types of careers available to former government employees.[87]

B. Outside Experts

This phenomenon of experts has been growing inside tech companies. However, expertise has become a growing business outside of tech companies, with an army of firms and research centers that offer services for Silicon Valley. The connections between a government and the private sectors of the terrorism industry transform the latter into "a virtual arm of the former."[88] This section will first detail some of the organizations that have created a business (both financial and ideological) of terrorism expertise for social media platforms. By peddling in orthodox viewpoints, the private sector enhances "the credibility of the official view by presenting this view, with minor variants, through purportedly 'independent' agencies" with the sector's ultimate purpose to "satisfy this demand for independent but credible authorities."[89] This relationship between the state and private sector in the terrorism industry is a helpful analogy to carry in mind when discussing the interplay between social media companies and the firms and research centers.[90]

The reality is, unfortunately, that much of the information around experts' work with social media companies remains somewhat opaque. This is not a fluke;

---

85. *Regional Analyst, YouTube Intelligence Desk*, HUNTR, https://perma.cc/25WV-VE8N (emphasis added).

86. *Id.* (emphasis added).

87. *See*, *e.g.*, *Brass Parachutes: Defense Contractors' Capture of Pentagon Officials Through the Revolving Door*, PROJECT ON GOV'T OVERSIGHT (Nov. 5, 2018), https://perma.cc/8JVU-975R.

88. Herman & O'Sullivan, *supra* note 31, at 55-6.

89. *Id.*

90. This is a helpful metaphor and nothing more. The author is not contesting that platforms are, or should be, seen as, government sources. After all, the Court in Manhattan Cmty. Access Corp. v. Halleck, 139 S. Ct. 1921 (2019) stressed that "merely hosting speech by others is not a traditional, exclusive public function and does not alone transform private entities into state actors subject to First Amendment constraints." *Id.* at 1930. In 2020 as well, the Ninth Circuit Court of Appeals resoundingly dismissed an appeal from the United States District Court for the Northern District of California in Prager Univ. v. Google LLC, 951 F.3d 991 (9th Cir. 2020). Prager University is a conservative media outlet that sued YouTube on the grounds that it violated Prager's First Amendment rights when the video-sharing platform chose to moderate Prager's content. Prager contested, among many arguments, that the size of the platform must be taken into consideration. The Ninth Circuit included reference to *Halleck* of course but also stressed that the search engine giant's size is not a relevant dimension to this discussion, writing that "[d]espite YouTube's ubiquity and its role as a public-facing platform, it remains a private forum, not a public forum subject to judicial scrutiny under the First Amendment." *Id.* at 995.

it is by design. In fact, some outside companies even advertise that they provide "discrete services" to Silicon Valley.[91] Thankfully, there is at least some public documentation of some companies and institutes that have been known to collaborate with technology firms. In 2017, Monika Bickert, Meta's Head of Global Policy Management, and Brian Fishman published a blog post titled "Hard Questions: Are We Winning the War on Terrorism Online?"[92] They noted how over two years ago, they "started meeting with more than a dozen other technology companies to discuss the best ways to counter terrorists' attempts to use our services."[93] They go on to say that among the different tactics Meta has employed to tackle the problem, one has been to "tap expertise from inside the company and from outside, partnering with those who can help address extremism across the internet."[94] Their internal experts—who come from intelligence communities, academia, and law enforcement[95]—help Meta "build stronger relationships with experts outside the company who can help us [Meta] more quickly spot changes in how terror groups are attempting to use the internet."[96] These outside experts, Meta touts, have:

> [E]xpertise in global terrorism or cyber intelligence to help us in our efforts. These partners – which include Flashpoint, the Middle East Media Research Institute (MEMRI), the SITE Intelligence Group, and the University of Alabama at Birmingham's Computer Forensics Research Lab – flag Pages, profiles and groups on Meta potentially associated with terrorist groups for us to review.[97]

It is critical to point out that firms such as Flashpoint offer services falling under risk intelligence or risk analysis, an offshoot of security-related work. Though savvy firms have been quick to make their work technologically relevant for Big (and small) Tech, risk intelligence is nothing new. Edward Herman traces the long history between corporations, intelligence agencies, and private security firms, arguing that this relationship first began out of a concern to crack down on labor union mobilization around the turn of the 20th century.[98] These third-party outside experts in the Meta article are a mix of established organizations, such as MEMRI, as well as newer companies that have found ways to monetize services, like risk intelligence, to social media platforms. The paper will examine several outside experts below.

---

91. *See*, *e.g.*, *Actor Risk Intelligence*, CRISP THINKING, https://perma.cc/PY8L-V8QY.
92. Monika Bickert & Brian Fishman, *Hard Questions: Are We Winning the War on Terrorism Online?*, META NEWSROOM (Nov. 28, 2017), https://perma.cc/A7QQ-KW36.
93. *Id.*
94. *Id.*
95. *Id.*
96. *Id.*
97. *Id.*
98. *See* Herman & O'Sullivan, *supra* note 31, at 119.

1. Flashpoint

Flashpoint, headquartered in New York City, is the "globally trusted leader in risk intelligence for organizations that demand the fastest, most comprehensive coverage of threatening activity on the internet."[99] Its "team of experts"[100]—led by the organization's board of directors who are "an experienced team of experts"[101]—has "tradecraft skills honed during years of operating in the most austere online environments, training in elite government and corporate environments, and building and leading intelligence programs across all sectors."[102] Flashpoint's Senior Vice President of Intelligence Tom Hofmann "has been at the forefront of cyber intelligence operations in the commercial, government, and military sectors"[103] and other members of the rest of the executive leadership also boast intelligence experience.[104] The firm appears to straddle multiple sectors as clients—public sector, health, retail, finance, and, most relevant here, technology.[105]

The services it offers for the technology industry mimics a common strategy of political risk analysis firms. Many security companies that provide risk assessment and risk intelligence services offer "regularly updated data bases and make their findings known to private subscribers and government agencies."[106] In fact, Flashpoint seems to embrace this tactic for the technology sector when it advertises—after warning technology firms (in case they forgot) that extremist actors seek to use their platforms for malicious purposes[107]—its "Intelligence Platform," which is its "archive of . . . Intelligence reports and technical data."[108] What exactly is in this database, and how much a company pays to access it, is not available.[109]

While database services may not be an innovation, what is novel in Flashpoint's suite of services is a product it develops and sells called Flashpoint Alerting.[110] This service alerts Flashpoint's customers when "relevant information is uncovered in threat-actor discussions," though the firm breaks down alerting options into four categories.[111] The first and fourth alerting services highlighted already seem to potentially overlap with "Intelligence Desk" roles: Automated Alerting "matches conversations from illicit online communities with

99. Flashpoint, LINKEDIN, https://perma.cc/9DW7-7KDY.
100. *Our People*, FLASHPOINT, https://perma.cc/5CK3-A6TE.
101. *Our Board,* FLASHPOINT, https://perma.cc/9LN5-3UPL.
102. *Our People*, *supra* note 100.
103. *Tom Hofmann*, FLASHPOINT, https://perma.cc/JD6B-S8FC.
104. *See, e.g.*, *Jake Wells*, FLASHPOINT, https://perma.cc/H8VV-TPJ8.
105. *Industries*, FLASHPOINT, https://perma.cc/4DCS-HVQM.
106. Herman & O'Sullivan, *supra* note 31, at 123.
107. *See Technology Providers*, FLASHPOINT, https://perma.cc/B3DH-SY26.
108. *Id.*
109. Edward Herman provides some examples of database subscriptions in other risk analysis companies back in the late 1980s. In the risk analysis companies he studied, some charged clients for a weekly subscription to the firm's database, others charged an annual pass. *See* Herman & O'Sullivan, *supra* note 31, at 123-24.
110. *Flashpoint Alerting*, FLASHPOINT, https://perma.cc/KLG2-Y6JM.
111. *Id.*

a client's areas of concerns, and automatically provides these matches directly" while Industry Alerting "provides customers tactical information derived from conversations from illicit online communities to users."[112] The firm also provides native language translations of these exchanges on discussion fora.[113]

2. Crisp Thinking

While Flashpoint's work focuses on some traditional tactics embraced by risk assessment companies in the form of database access, there is another company that works in the social media space that deserves a closer examination. Crisp Thinking, a company based in Leeds, England,[114] offers a more "back-to-basics" approach for risk intelligence: monitoring services on the internet.[115] In particular, Crisp's staff appears to specialize in surveillance in the Deep Web, promising that its "global team of experts"[116] will "go where no one else goes"[117] as they "analyze billions of instigator and influencer signals from the **open, deep, dark web** and **closed messaging apps**."[118] Companies like Meta and YouTube—with its Intelligence Desk that works with third-party consultants—stand to benefit immensely from Crisp's entire suite of "Platform Trust and Safety"[119] services that monitor the web and report its intelligence to Silicon Valley.

Crisp Thinking has an entire practice aimed at companies' content policy teams. From the beginning, Crisp soaks its advertisement in fear-mongering language, telling viewers that nefarious actors "including violent or hateful extremists" are "creating millions of pieces of harmful content"[120] on the internet. The defense and intelligence work moves not only from the "open, deep, dark web" but to the platforms' content themselves. Listed under the defense solutions, Crisp offers "Platform Defense" and "Live Streaming Defense" to their services. By monitoring, say, Meta Live, a livestream on YouTube, or Twitter feeds, Crisp promises to report new trends, concerns, and potential violations lurking in the recesses of these platforms.

---

112. *Id.*

113. *See id.*

114. *About Us*, CRISP THINKING, https://perma.cc/G4Z3-4ATD.

115. As discussed in Part II.B(1), *infra*, Flashpoint also engages in monitoring work, as reflected in its "Flashpoint Alerting" product. However, it is only one service out of an entire, varied docket of services.

116. *About Us*, *supra* note 114.

117. *Id.*

118. *Id.* (emphasis added).

119. *Platform Trust and Safety Defense and Intelligence*, CRISP THINKING, https://perma.cc/H4P5-TFPK [hereinafter *Platform Trust*]. Many companies' content policy and content moderation efforts are carried out by Trust & Safety teams inside companies. While the vast majority of content moderation work is done by third-party contractors, Trust & Safety employees in the technology companies are full-time staff that write the policies and enforcement guidelines, work with the contracting companies, and also collaborate with Legal, Public Policy, and Engineering teams. *See* SARAH ROBERTS, BEHIND THE VEIL: CONTENT MODERATION IN THE SHADOWS OF SOCIAL MEDIA 41-43 (2019) (detailing the different types of content moderation employment structures).

120. *Platform Trust*, *supra* note 119.

In fact, Crisp's business model seems to revolve largely on supplementing the work that is supposed to be left to the companies: reviewing content for violations of their Terms of Services (ToS) or community guidelines. Crisp now sells "proprietary filtering software that purports to effectively filter hate speech and terrorist content"[121] although Meta, Twitter, and YouTube already all boast of their impressive algorithmic interference to detect and deter the dissemination of terrorist material on their platforms. The concerns of this type of business model (as well as all of these expert services) will be discussed in more detail in Part III.

3. SITE Intelligence Group

Rita Katz, SITE (Search for International Terrorist Entities) Intelligence Group's founder and executive director, is a self-professed addict of following terrorist chatter online. In a profile in *The New Yorker,* she is described as being so relentless in her pursuit that she even leaves a computer open during dinner parties she throws for guests.[122] Katz, originally from Iraq,[123] grew up in Israel and served in the Israel Defense Forces before later moving to America and starting SITE's work. Her efforts have been praised by law enforcement agencies in the United States, as Benjamin Wallace-Wells writes that Katz "keeps copies of letters from officials whose investigations into terrorism she has assisted,"[124] with Katz telling Wallace-Wells that the letters are useful "when she meets with skepticism or lack of interest; **they are her establishment bona fides**."[125]

The fact that law enforcement and intelligence officials have expressed their gratitude may be the only real qualification Katz, a terrorism expert, has. (Though as discussed in Part I, this form of praise from individuals in the terrorism industry is an effective way to signal "expertise" in this space.) SITE Intelligence Group has been active for over twenty years and adds its value to the space by tracking chatter online, translating propaganda materials, and offering the organization's own analyses. All of their output—from translations to reports—are only available for subscribers, whether individuals, governments, law enforcement, or private sector.[126] While SITE first began with Al-Qaeda's online activity, it has expanded not only to other Islamist terrorist organizations like the Islamic State, but also now to track white supremacy movements and other far-right groups.[127]

Though a subscription portal has been SITE's longest standing service, it also has a set of services aimed at tech companies. For example, SITE's SourceFeed enables:

121. Hannah Bloch-Wehba, *Automation in Moderation*, 53 CORNELL INT'L L. J. 41, 86 (2020).

122. *See* Benjamin Wallace-Wells, *Private Jihad*, NEW YORKER (May 22, 2006), https://perma.cc/389U-QE3Q.

123. *Id.*

124. *Id.*

125. *Id.* (emphasis added).

126. *Our Services*, SITE INTELLIGENCE GROUP, https://perma.cc/5K8T-XEM8.

127. *Id.*

> ICT [Information and Communications Technology] companies to easily and rapidly locate designated terrorist and violent extremist material on platforms. SourceFeed contains the largest commercially available data set of confirmed terrorist and violent extremist material online multimedia content, spanning statements, video/audio, and online chatter from extremist entities around the globe.[128]

Flashpoint and SITE both offer many of the same services, such as a database service and online chatter monitoring. Crisp Thinking differs remarkably by focusing heavily on the social media platforms directly, monitoring not only the broader online ecosystem for "risks" but also scouring the tweets, posts, and videos on popular social networks directly.

Despite Rita Katz's public appearances, leadership over SITE's work, and deep connections with law enforcement in the US and around the world, she and SITE Intelligence still garner significant criticism. The former manager of the Central Intelligence Agency's Osama bin Laden unit described SITE as saying, "[m]uch as Al Jazeera underplays terrorist threats, the SITE Institute at times overhypes them."[129] Katz's embellishment has many other critics as well, with some saying her enthusiasm can "make her too eager to find plots where they don't exist"[130] citing one example where "she publicized a manual for using botulinum in terror attacks, for example, which experts later concluded was not linked to any serious threat."[131]

4. Middle East Media Research Institute (MEMRI)

Yigal Carmon, a former member of Israeli military intelligence, founded the Middle East Media Research Institute (MEMRI) in Washington, DC in 1997. MEMRI "provides its readers worldwide with comprehensive access to the primary source content from the Arab and Muslim world" tracking media outlets in Arabic, Turkish, Pashto, Urdu, Dari, and Farsi.[132] Their monitoring and analyses are produced in English, French, Polish, Japanese, Spanish, and Hebrew.[133] Much of MEMRI's framing is in an "us" versus "them" mentality, saying their work is to bridge the "West and the Middle East and South Asia."[134]

Unlike SITE Intelligence Group, Flashpoint, and Crisp Thinking, however, MEMRI does not seem to charge for its analyses, translations, or services. Instead, it relies heavily on donations and allows parties to subscribe to free email updates of its translations. Of the many members of its Board of Directors and Board of Advisors, it only takes a cursory glance to see the number who come from Israeli and American government, military, and intelligence agencies.[135]

---

128. *Id.*
129. Wallace-Wells, *supra* note 122.
130. *Id.*
131. *Id.*
132. *Our Languages*, MEMRI, https://perma.cc/XA92-6Z75.
133. *Id.*
134. *About*, MEMRI, https://perma.cc/6KF4-XY6B.
135. *See id.*

Although MEMRI asserts it is an independent organization, many have voiced discomfort with the organization's translations and how it frames political issues. Many critics believe MEMRI focuses on finding content that is the most "dangerous-sounding."[136] A leader of the Council on American-Islamic Relations (CAIR) blasted MEMRI, saying their intent "is to find the worst possible quotes from the Muslim world and disseminate them as widely as possible."[137] Even setting aside the editorial choices that may generate these criticisms, other critics point to MEMRI's role in perpetuating terrorist tropes onto Arab communities through their translation apparatus.[138]

Though this section details the constellations of inside and outside experts in the tech sector and its terrorist content moderation efforts, a discussion on the overall risks must follow. Part III will expand on the threats.

III. THE HARMS AND CONCERNS OF EXPERT INFLUENCE ON SOCIAL MEDIA CONTENT MODERATION

Until now, this paper has discussed the critiques surrounding the generation, propagation, and legitimation of terrorism expertise and examined the "inside" experts and "outside" experts that influence Silicon Valley. However, considering the dominant nature of social media in contemporary communication and discourse, there are harms and concerns unique to content moderation that arise from expert involvement—including the privatization of international counterterrorism work, erasure of human rights content, codification of expert engagement in content moderation regulations, and expert involvement in shaping the companies' own terrorism policies. This section does not mean to suggest that the larger problem of orthodoxy highlighted in Parts I and II is not a problem in Silicon Valley's growing dependency on terrorism experts. Adhering to one particular narrative threatens to continue to marginalize other valid viewpoints, critiques of terrorism studies, and alternative, unorthodox proposals for how best to engage with terrorism. In fact, many industry-specific harms in this part are the direct result of orthodox narratives created by these experts. The individuals and companies with the "ear" of these intermediary platforms produce and reproduce dominant narratives about terrorism and what is needed to "defeat" this foe. These experts:

> [H]ave revolving-door relationships with governments and government intelligence agencies, and many are connected with private security firms. They therefore reflect official views and a state agenda, and they rarely depart from

136. Robert Worth, *Mideast Analysis, Fast and Furious*, N.Y. TIMES (June 18, 2006), https://perma.cc/WL2E-AX6F.

137. Brian Whitaker, *Selective Memri*, THE GUARDIAN (Aug. 12, 2002, 6:29 AM), https://perma.cc/6TYS-BRFG.

138. *See*, *e.g.*, Mona Baker, *Narratives of Terrorism and Security: 'Accurate' Translations, Suspicious Frames*, 3 CRITICAL STUD. ON TERRORISM 347, 353-60 (2010).

> the assumptions of the Western model of terrorism [. . .] **[E]xperts have a material interest in "threat inflation."**[139]

If the threat is diminished or reframed, then experts face the risk that their databases, subscription services, products, and perspectives are no longer needed. Additionally, if companies focus on groups or individuals that the expert community does not think qualifies as terrorist entities, then experts can lose out on business deals with social media platforms. Instead, experts engage in a process of maintaining a narrative of orthodoxy of who is a terrorist and what the "threat landscape" looks like on the horizon, which in turn conditions companies to interpret and re-interpret the very policies that they write, rely on additional abuse-fighting tools, and double-down on their partnerships with these experts.

By generating expertise and propagating a particular narrative, experts can move seamlessly with their peers from government bodies and research centers to the private sector—whether inside companies directly or to a third party that consults, advises, and sells products and information to Silicon Valley. Potential solutions to the perniciousness of expert involvement in the technology sector will be examined in more detail in Part IV, though the following sub-sections identify the risks and implications of the dominance of orthodoxy and the use of inside and outside experts in terrorist content moderation. This section identifies and discusses the primary problems with expert involvement in content moderation: the privatization of public counterterrorism functions, expert shaping of the policies that guide platform enforcement and the conditioning of companies to re-interpret their own terms of service, pressure on companies to rely on questionable filtering and database technology, the regulation of company collaboration with experts as part of governments' growing oversight efforts, and, finally, operating with little transparency.

A. The Privatization of Public Law Functions in the Counterterrorism Space

One of the most alarming risks of social media companies' involvement in the counterterrorism space—and by proxy the experts that influence and lobby these platforms—is the increasingly public roles they assume as privatized public actors in the counterterrorism landscape. Traditionally counterterrorism work was the responsibility of governments and intergovernmental organizations, with an abundance of legal frameworks and law enforcement mechanisms that they relied on to think about and counter terrorism. Of course, while private industry involvement in the counterterrorism landscape may not be new, what is new is the transformation and elevation of social media platforms from news, communication, and entertainment sources to harsh instruments in combating terrorism. The key drivers behind social media's uncomfortable role as a tool to fight terrorism stem from two main constraints in the public law sector: the lack of a consistent definition for terrorism at the national and international levels and

139. Herman & O'Sullivan, *supra* note 31, at 190 (emphasis added).

constitutional constraints that limit the range of options available to a governing body to respond to a purported threat. These limitations ultimately create conditions that generate a symbiotic relationship between governments and companies; social media platforms are appealing vectors because they can carry out public law functions in combatting terrorism without being bound by constitutional norms, and in carrying out the will of the state they can avoid the threat of excessive regulation.

Central to any effort to respond to terrorism are definitional questions: how do institutions define who is a terrorist and what constitutes terrorist conduct? Compounding these already difficult questions is a national-international axis, with countries fluctuating from country-specific needs and problems with terrorism on one end to international collaboration on the other.[140] Despite the global interconnectedness of counterterrorism intelligence and operations, terrorism definitions vary by country. Consider, for example, three Western nations' attempts to elucidate terrorism and terrorist activity in their counterterrorism legislation and statutes. France's counterterrorism legislation defines terrorism as "an individual or collective enterprise intending to gravely trouble public order by means of intimidation or terror."[141] The United States defines "international terrorism" as violent acts that are "dangerous to human life" and "appear to be intended to intimidate or coerce a civilian population; to influence the policy of a government by intimidation or coercion; or to affect the conduct of a government by mass destruction, assassination, or kidnapping."[142] In a third example, the United Kingdom's (UK) Terrorism Act of 2001 defines terrorism as "the use or threat of action" that is "designed to influence the government or an international governmental organization or to intimidate the public or section of the public and the use or threat is made for the purpose of advancing a political, religious, racial or ideological cause."[143] The range of perspectives—from a public safety focus in France to an international scope in America and an ideology-centered understanding in the UK—belies not only the difficulty of identifying terrorism but also the national priorities when it comes to framing the problem.

At the international level, the United Nations (UN)—as an intergovernmental organization—also plays a strong role in considering terrorism, with 19 counterterrorism legal instruments that are designed to prevent terrorist acts from the hijacking of civilian aircraft to hostage situations and terrorist financing.[144] UN Security Council Resolution 1373 was a substantial step forward regarding

140. A useful illustration: Country A is focused on combatting Actor X's presence in its borders. It may not be as concerned by Country B's battle with Actor Y unless X and Y are connected in a particular manner or Y threatens, or may threaten, Country A's security. If so, then Country A's response may shift from one end of this axis (national) to the other end (international collaboration).

141. Calliope Makedon Sudborough, *The War Against Fundamental Rights: French Counterterrorism Policy and the Need to Integrate International Security and Human Rights Agreements*, 30 SUFFOLK TRANSNAT'L L. REV. 459, 464 (2012).

142. 18 U.S.C. § 2331.

143. Terrorism Act 2000 § 1(1) (U.K.).

144. *International Legal Instruments*, UN SEC. COUNCIL, https://perma.cc/42R8-7Q9B.

international efforts to combat terrorism, making it binding on all member states.[145] Instead of wading into the politics of defining terrorism,[146] the Security Council's goal was to "raise the average level of government performance against terrorism across the globe"[147] and to establish "uniform obligations for all 191 member states to the UN, thus going beyond the existing international counterterrorism conventions and protocols binding only those that have become parties to them."[148] Because the UN does not define terrorism, and instead delegates this task to member states' governments, there is a spectrum of definitional possibilities across countries. Interestingly, as well, is that the resolution creates the Security Council's Counterterrorism Committee (CTC) which oversees member states' efforts and "consults with independent 'Expert Advisers' in advising member states on such issues as legislative drafting, various areas of law, and law enforcement."[149] While the CTC may rely on outside experts to help their work in supervising member states' adoption of the resolution, one can see the concerning effects of expert involvement in counterterrorism legislation. The influence of expertise in shaping social media companies' terrorism content policies, a fascinating analogous situation, is explored in Part III B.

Public international law's mandate regarding terrorism is to consider it along two avenues: (1) bolstering national capacity to respond to terrorism through legislation law enforcement and (2) identifying particular acts of terrorism and their corresponding motives. While public international law does occasionally wade into the tense, controversial work of proscribing terrorist actors, the UN focuses only on certain individuals or groups like ISIS or Al-Qaeda that garner relatively universal consensus. International law and domestic law (e.g., constitutional restraints on speech, law enforcement) are limited in their response to terrorism—to say nothing of the bureaucratic inefficiencies of national and international bodies—hence, the nimble social media companies and their expert army have become effectively privatized responders in a public law realm.

Despite social media's massive influence in the world, courts are still reluctant to classify them as public entities.[150] In an American context, for example, the

145. S.C. Res. 1373 (Sept. 28, 2001).

146. *See* Eric Rosand, *Security Council Resolution 1373, The Counter-Terrorism Committee, and the Fight Against Terrorism*, 97 AMER. J. INT'L. L. 333, 334 (2003) (explaining the "principal reason Resolution 1373 did not attempt to define terrorism was to avoid the divisive debate in the Security Council . . . [t]he sponsors . . . wanted a resolution that would pass quickly.").

147. *Id.* at 334.

148. *Id.*

149. Sudborough, *supra* note 143, at 469.

150. In Germany, for example, the Federal Constitutional Court (BVerfG) in the case *Der III. Weg* stopped short of saying Facebook had the same fundamental rights obligations as the state. BVerfG, 1 BvQ 42/19, May 22, 2019, https://perma.cc/UD47-RFU7 (holding in specific circumstances under the "doctrine of indirect third-party effect of fundamental rights (*mittelbare Drittwirkung*)" that Facebook had to reinstate a political party's content). *See also* Matthias Kettemann & Anna Sophie Tiedeke, *Back Up: Can Users Sue Platforms to Reinstate Deleted Content?*, 9 INTERNET POL'Y REV. 8-9 (2020). In Italy, a neo-fascist group sued Meta after it deleted its content under Facebook's violence and hate speech policies. The Tribunale di Roma (the ordinary court of first instance of Rome) in *CasaPound v.*

Ninth Circuit held that "[d]espite YouTube's ubiquity and its role as a public-facing platform, it remains a private forum, not a public forum subject to judicial scrutiny under the First Amendment."[151] Furthermore, "[t]o characterize YouTube as a public forum would be a paradigm shift."[152] This insistence of classifying companies as private entities allows states to think about private companies, even the influential social media platforms, as fundamentally different from a government, thus keeping "private law" as the appropriate realm in which these institutions should operate, despite how integral or fundamental a platform is to make a public space for the global community to utilize. As I have written elsewhere with my co-author Rabea Eghbariah:

> Online content intermediaries and governments, therefore, function in different legal realms and are governed by different legal norms. A direct consequence of this distinction is the bifurcation between user and citizen. While the former is largely governed by private contractual norms—like a platform's Terms of Service—the latter is governed by public law norms.[153]

Ultimately because companies can enjoy more freedom to do as they please without the constraints of public law norms, treating all of us not as citizens with constitutional rights but rather as users bound by terms of service is disconcerting. Moreover, with this division, private companies can—either of their own accord or through public pressure—go further than a government or intergovernmental organization in responding to terrorism, despite the fact that many have said content moderation should not curtail freedom of expression.[154] For example, some companies' community guidelines specify what flags, symbols, or logos actually belong to a terrorist organization. Platforms can also adopt a flair of strict liability and ban a user for possessing *any* content featuring a terrorist organization, an idea that would generate immense scrutiny if a government tried to mimic a similar stance. For example, the UK recently proposed criminalizing possession of extremist material—a remarkable move since currently only distribution is a crime under its terrorism law *if* "material is useful in commission of a terrorist act."[155] The proposal has galvanized the public and civil society groups, with Liberty, a prominent civil liberties organization, going so far as to say that:

Facebook similarly held that Facebook had to reinstate CasaPound's account on the basis of constitutional right to political participation but did not go as far as saying that Facebook is a public actor. Trib. di Roma, 29 aprile 2020 (It.). *See also* Kettemann & Riedke, *supra*, at 9.

151. Prager Univ. v. Google LLC, 951 F.3d 991, 995 (9th Cir. 2020).

152. *Id*. at 998.

153. Rabea Eghbariah & Amre Metwally, *Informal Governance: Internet Referral Units and the Rise of State Interpretations of Terms of Service*, 23 YALE J.L. & TECH. 542, 551 (2021).

154. *See, e.g*., David Kaye (Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression), *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, U.N. Doc. A/HRC/38/35 (Apr. 6, 2018).

155. Haroon Siddique & Jamie Grierson, *Home Office Proposes Offence of Possessing Terrorist Propaganda*, GUARDIAN (Jan. 14, 2020 11:53 AM), https://perma.cc/222B-ZT3C.

> The UK already has oppressive counter-terror laws which put our freedom to think, debate and learn in jeopardy. Making the law even more heavy-handed would undermine our freedom of thought and our right to free expression, without making us any safer. It is critical that the government focuses on creating counter-terror strategies that keep us safe and free, protecting the rights terrorists seek to destroy.[156]

While public law is focused on scoping terrorism to be an ideologically motivated violent act, these private "laws" that shape online public squares for billions of users are far more aggressive when it comes to thinking about terrorism—a sort of "muscular" public law function. For example, YouTube's content policies (modified after I left) now go so far as forbidding supportive intent, writing that videos featuring insignia belonging to a terrorist group that is intended to praise the organization are not allowed on the platform.[157] These aggressive counterterrorism stances are developed through the content policies that companies write, a process that is shaped heavily by terrorism experts, which is detailed below.

B. Shaping Platforms' Content Policies and Enforcement

These experts matter a great deal as well when one considers that these companies try to define what is and is not terrorism. All three large platforms—Meta, YouTube, and Twitter—have content policies aimed at terrorist and extremist actors, though they label the policy differently. No company's extremism policy is titled counterterrorism outright. Twitter's content policy is called Violent Organizations policy which defines extremist groups using three criteria:

- identify through their stated purpose, publications, or actions as an extremist group;
- have engaged in, or currently engage in, violence and/or the promotion of violence as a means to further their cause; and
- target civilians in their acts and/or promotion of violence.[158]

Meta's Dangerous Individuals and Organizations policy states:

> In an effort to prevent and disrupt real-world harm, we do not allow any organizations or individuals that proclaim a violent mission or are engaged in violence to have a presence on Meta. This includes organizations or individuals involved in the following:

156. *Id.*

157. *Violent Criminal Organizations Policy*, YOUTUBE HELP CENTER, https://perma.cc/MN6Y-K378.

158. *Violent Organizations Policy*, TWITTER HELP CENTER (Oct. 2020), https://perma.cc/TRM3-LDFX.

- Terrorist activity
- Organized hate
- Mass murder (including attempts) or multiple murder
- Human trafficking
- Organized violence or criminal activity

> We also remove content that expresses support or praise for groups, leaders, or individuals involved in these activities.[159]

YouTube's Violent Criminal Organizations Policy states the following descriptions as grounds for removal:

- Content produced by violent criminal or terrorist organizations
- Content praising or memorializing prominent terrorist or criminal figures in order to encourage others to carry out acts of violence
- Content praising or justifying violent acts carried out by violent criminal or terrorist organizations
- Content aimed at recruiting new members to violent criminal or terrorist organizations
- Content depicting hostages or posted with the intent to solicit, threaten, or intimidate on behalf of a violent criminal or terrorist organization
- Content that depicts the insignia, logos, or symbols of violent criminal or terrorist organizations in order to praise or promote them[160]

We know that outside experts can, and do, influence tech companies as they write content policies and make enforcement decisions. Meta's content policy team holds meetings twice a month to discuss and debate potential changes to their content policies. As Meta notes, to prepare for these meetings, employees "reach out to internal and external experts"[161] and meeting notes are kept. Twenty-eight meeting documents are publicly available beginning from November 13, 2018.[162]

An examination of each of the publicly available files reveals that a total of six of the Meta policy staff meetings discussed terrorism, counterterrorism, or terrorism and violations of International Humanitarian Law. In a November 13, 2018 meeting, Meta's Dangerous Criminal Organizations staff (led by Brian Fishman) discussed the

159. *Dangerous Individuals and Organizations Policy*, META TRANSPARENCY CENTER, https://perma.cc/JF46-SWPM.

160. *Violent Criminal Organizations Policy*, YOUTUBE HELP CENTER, https://perma.cc/SE63-XNPU.

161. *Writing Meta's Rulebook*, *supra* note 26.

162. *Product Policy Forum Minutes*, META NEWSROOM (Nov. 15, 2018), https://perma.cc/XHT9-QS24.

need to revisit the designations Meta uses to add an organization to its own internal list of terrorist actors.[163] In this particular document, they cite the key questions they want to use to develop better this policy, including how to create a more consistent structure for designating organizations with input from outside organizations.[164] They cite the need to consult with "law enforcement experts" in this process.[165]

This question of designation signals continued with a follow-up meeting on January 15, 2019.[166] They worked with nine experts—six based in the West, with three in North America, two in Europe, and one in Israel[167]—and ultimately recommended to designate a person or organization as terrorist if and only if the entity was charged or convicted on terrorism grounds.[168] Unfortunately Meta does not reveal the names of the experts; instead, it only provides the geographic location (e.g., Asia-Pacific, or APAC) and general institutional affiliation (e.g., NGO, international organization, academic, journalist).[169]

On December 11, 2018, Brian Fishman's staff discussed with the broader Meta policy team the potential options for the company's new definition of terrorism.[170] They "[c]onsulted with 16 external stakeholders"[171] to ultimately produce four different options for terrorism. One option ("Option 1") defines a non-state actor as one who "engages in or advocates and lends substantial support to purposive and planned acts of violence"[172] which Meta notes would "be considered over-broad as it may remove 'freedom fighters' and other orgs that are parties to an armed conflict."[173] In another expert-informed definition, ("Option 2"), Meta proposes defining a non-state actor as anyone who uses violence to coerce a government or organization or deliver a message and causes harm or death or property destruction.[174] This alarming definition, Meta notes, would not "distinguish between different types of groups and organizations among them 'freedom fighters,' insurgents, militias and separatist actors based on the tactics and the target of their violence. As such we would be labeling non-state actors engaged in violence as terrorists."[175] Even if only an option, the range of possibilities reflects not only how difficult it is to define terrorism, but the ways in which expert influence could potentially steer Meta to one option over others.

163. *Content Standards Forum – November 13, 2018*, META NEWSROOM, https://perma.cc/E9HV-TEQC.
164. *Id.*
165. *Id.*
166. *Content Standards Forum – January 15, 2019*, META NEWSROOM, https://perma.cc/32AF-6NLJ.
167. *Id.*
168. *Id.*
169. *Id.*
170. *Content Standards Forum – December 11, 2018*, META NEWSROOM, https://perma.cc/WHS2-8ZUX.
171. *Id.*
172. *Id.*
173. *Id.*
174. *Id.*
175. *Id.*

In the remaining documents that center on terrorism, it was revealed that in one meeting Meta employees consulted with 22 experts as they considered how best to incorporate International Humanitarian Law (IHL) principles into their content policies.[176] Further, on March 26, 2019, a document traced Meta's struggle on whether to classify white nationalism and separatism as hate speech or under their Dangerous Criminal Organizations policy. They reached the conclusion that their "discussion with 20 experts across the globe and our own research shows that white nationalism & white separatism is tied to organized hate groups."[177] The last reference to terrorism or extremism in these public documents was on April 9, 2019 to discuss how Meta should delist individuals or organizations from its own Dangerous Criminal Organizations policy if they are no longer designated as terrorist entities by the US or UN.[178]

C. Conditioning Companies to Re-Interpret their Own Terms of Service

As discussed above, inside and outside experts have been directly involved in defining impermissible speech, a task that normally rests squarely with the companies. I previously wrote about how state involvement in content moderation creates conditions to exert government interpretations of a company's terms of service.[179] Expert involvement, however, invites yet another external body into what should be a company's responsibility to interpret its policies. The influence of "outside experts" signals to companies that their *interpretation* of their own terms of service needs to be adjusted to the experts' satisfaction. Although companies ostensibly have the final decision on a piece of potentially violative content, expert input could be provided that is an alternative reading of companies' often vague community guidelines around terrorist or extremist content. This "definitional ambiguity"[180] is no trivial matter. There are already significant differences in how each country defines and labels terrorism in its own legislative apparatus; the companies themselves also think about their respective terrorism policies differently from one another, offering broad and poorly scoped terms in their guidelines.[181] David Kaye, the former UN Special Rapporteur on Freedom of Expression, has discussed how these varying definitions can cause people to "take advantage of open-ended platform standards to insist upon takedowns."[182]

Even if there is close alignment among every expert and company on what is "terrorist" material, "shades of alternative interpretations can be harmful."[183] For example, as one critique about terrorism experts at the UK's Counterterrorism

176. *Content Standards Forum – February 12, 2019*, META NEWSROOM, https://perma.cc/EAW4-7FU4.

177. *Id.*

178. *Content Standards Forum – March 26, 2019*, META NEWSROOM, https://perma.cc/SFK4-ZP9Y.

179. *See generally* Eghbariah & Metwally, *supra* note 153.

180. Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035, 1052 (2018).

181. *See supra* Part III.A for a more detailed discussion.

182. DAVID KAYE, SPEECH POLICE: THE GLOBAL STRUGGLE TO GOVERN THE INTERNET 81 (2019).

183. Eghbariah & Metwally, *supra* note 153, at 601.

Internet Referral Unit (CTIRU) put it the "true relationship between CTIRU content removals and matters of national security and crime preventions is likely to be subtle, rather than direct and instrumental."[184] Considering that companies review millions of pieces of content, "each video, tweet, or post invites multiple interpretations, effectively an open invitation" for companies to adopt experts' interpretations of the companies' terms of service.[185]

Outside experts wield the media to broadcast both their dissatisfaction and satisfaction of companies' own interpretations of content policies. Rita Katz herself praised YouTube's work on combatting terrorist content online, writing:

> YouTube's strategy appears to be working. My previous research, along with an exhaustive new SITE Intelligence Group report analyzing nearly 30,000 verified ISIS and al-Qaeda online artifacts published between April and August 2018, show a steep decline—and, in at least one case, complete halt of YouTube use. ISIS and al-Qaeda's shifts away from YouTube are promising developments in the online fight against terrorist propaganda. [. . .] Online vigilance remains as critical as ever in combatting violent extremism. Stifling terrorist propagandists and recruiters demands a far more collaborative, coordinated approach between governments, tech companies, and third-party entities.[186]

In another example, the Counter Extremism Project (CEP)[187]—an international policy arm created by former US Ambassador to the UN Mark Wallace, former Senator Joseph Lieberman, and former senior Homeland Security advisor Frances Townsend—raised repeated, scathing criticism of YouTube's failure to remove Anwar al-Awlaki sermons.[188] CEP published a fear-mongering report on Anwar al-Awlaki content on YouTube in August 2017, sparking intense media scrutiny and waves of criticism that YouTube was not doing enough to combat the proliferation of extremist material at a time when the company was already facing an advertiser exodus and mounting governmental pressure over the topic.[189] Three months later, YouTube made the decision to declare that any and all Awlaki content must be removed, a decision that CEP CEO called a "watershed moment" that reflected YouTube's genuine effort to "clean" the platform.[190]

---

184. Jim Killock, *Informal Internet Censorship: The UK's Counter Terrorism Internet Referral Unit (CTIRU)*, VOXPOL (July 31, 2019), https://perma.cc/PF9U-EDPW.

185. Eghbariah & Metwally, *supra* note 153, at 601.

186. Rita Katz, *To Curb Terrorist Propaganda Online, Look to YouTube. No, Really.*, WIRED (Oct. 20, 2018, 8:00 AM), https://perma.cc/7LZX-6ZE9.

187. *Senior Leadership*, COUNTER EXTREMISM PROJECT, https://perma.cc/9DVL-9DFN.

188. Anwar al-Awlaki was a key recruiter for Al-Qaeda and was also the first American citizen to be killed by an American drone strike as part of his involvement with the terrorist organization. *See* Scott Shane, *Internet Firms Urged to Limit Work of Anwar al-Awlaki*, N.Y. TIMES (Dec. 18, 2015), https://perma.cc/9D9G-K2VZ.

189. *See generally* COUNTER EXTREMISM PROJECT, ANWAR AL-AWLAKI ONLINE (2017), https://perma.cc/89M5-9TAP.

190. Scott Shane, *In "Watershed Moment," YouTube Blocks Extremist Cleric's Message*, N.Y. TIMES (Nov. 12, 2017), https://perma.cc/7UXY-N2MP.

Experts are not only engaging in an exercise that challenges specific posts, tweets, or videos, they are also involved in a more subtle effort of getting others to see the issue the way they do. The study of the relationship between sources and journalists provides a helpful conceptualization. In violent crises, there is a struggle "between the sources which seek to gain access to the public and to shape the news, on one hand, and the journalists who try to obtain certain information, on the other."[191] When journalists are not able to witness a violent conflict themselves, they are often at the mercy of sources that seek to frame particular narratives on how a story unfolds.[192] The sources' control over information is, in other words, power.

The leverage experts wield in reframing how companies look at their own policies and enforcement seems to be working. With calls from Katz and other experts to the terrorism dangers lurking in other platforms,[193] praising the large content intermediaries produces a shift, involving the companies' "successful" interpretation and enforcement of their terms of service. Ultimately, it suggests that experts are not only involved inside and around Silicon Valley but also that their praise and assessments are evidence that they are satisfied with how particular platforms interpret their policies.

D. The Problem with Databases and Other Outside Experts' Products

SITE Intelligence Group and Flashpoint offer database services and Crisp Thinking sells proactive filtering software for companies to root out hate speech and extremist content, even though all major social media companies already deploy extensive filtering through their algorithmic software.[194]

The most notable database of terrorist content belongs to the Global Internet Forum to Combat Terrorism (GIFCT) hash-sharing database. Hashes, or "digital fingerprints," belong to unique posts, videos, or images on a site that can then be used to search for identical copies on other platforms. YouTube, Meta, Microsoft, and Twitter created a database of hashes as part of their creation of the GIFCT. This repository is composed of content removed by social media companies for violating a company's specific policies on terrorism.[195] Any GIFCT-participating company can then check its corpus of content against the database and contribute to it as well. The ultimate goal of this database is to prevent the emergence of content removed by one platform to appear on other social media sites. The popularity of the GIFCT and its database has exploded in recent years, with the Christchurch Call embracing the

191. Yonatan Gonen, *Journalists-Sources Relationship in Violent Conflicts Coverage: Shifting Dynamics*, 12 SOC. COMPASS 1, 2 (2018). For the purposes of "sources" and "journalists" in this analogy, the "sources" are experts and "journalists" are platforms that present the events of the world online for users.

192. *See id.* at 8.

193. *See* Katz, *supra* note 186.

194. *See infra* Part II B(1) for Flashpoint, Part II B(2) for Crisp Thinking, and Part II B(3) for SITE Intelligence Group.

195. *To Stop Terror Content Online, Tech Companies Need to Work Together*, GOOGLE PUBLIC POLICY BLOG (Dec. 20, 2018), https://www.blog.google/outreach-initiatives/public-policy/stop-terror-content-online-tech-companies-need-work-together/.

organization[196] as well as the upcoming European Union Regulation on Preventing the Dissemination of Terrorist Content Online including automated tools to detect violent extremist material quickly.[197] Despite governments' adoration for the hash-sharing terrorism database, civil society organizations and academics alike have voiced concerns around the lack of transparency,[198] anti-Muslim and anti-Arab bias,[199] collaboration with law enforcement,[200] and potential for false positives that silence speech.[201] They have also expressed fear that these databases and filtering technologies can further harm human rights communities and efforts to document war crime evidence using social media.[202]

With databases like SITE Intelligence and Flashpoint, however, companies could be directed to content that, if reviewed incorrectly, could be removed under an intermediary's terrorist content policies and added to the GIFCT database. (Any content removed for terrorism on the large platforms is automatically added to the GIFCT database.) It even appears that databases may work together: the GIFCT announced a 12-month pilot program to expand sharing URLs of known terrorist material to platforms.[203] The GIFCT noted that it partnered with SITE Intelligence to collect 24,000 URLs, with the "majority of new URLs shared amongst GIFCT member companies" coming from SITE Intelligence.[204] Other companies can then use this industry-wide database to remove material on their respective platforms. If mistakes are added to a database, they then amplify and ripple across platforms. Even if something is not a "mistake" but rather an aggressive interpretation (which may be the case with SITE's content in its own database), then companies could be over-enforcing. As the Center for Democracy and Technology notes, "[t]here is evidence that processes intended to remove terrorist content have the counter-productive effect of removing anti-terrorism counterspeech, satire, journalistic material,

196. PRIYAL PANDEY, ONE YEAR SINCE THE CHRISTCHURCH CALL TO ACTION: A REVIEW 3 (2020), https://perma.cc/N8WN-9WFG.

197. *See* Courtney Radsch, *GIFCT: Possibly the Most Important Acronym You've Never Heard Of*, JUST SECURITY (Sept. 30, 2020), https://perma.cc/94F5-XQRM.

198. Svea Windwehr & Jillian C. York, *One Database to Rule Them All: The Invisible Content Cartel that Undermines the Freedom of Expression Online*, ELECT.FRONTIER FOUND.: DEEPLINKS BLOG (Aug. 27, 2020), https://perma.cc/CY2Z-3EKP.

199. Ángel Díaz, *Global Internet Forum to Counter Terrorism's 'Transparency Report' Raises More Questions Than Answers*, JUST SECURITY (Sept. 25, 2019), https://perma.cc/DC5B-9VHB.

200. Emma Llansó, *Human Rights NGOs in Coalition Letter to GIFCT*, CTR. FOR DEMOCRACY & TECH. (July 30, 2020), https://cdt.org/insights/human-rights-ngos-in-coalition-letter-to-gifct/.

201. *Id.*

202. *See, e.g.*, HUMAN RIGHTS WATCH, "VIDEO UNAVAILABLE": SOCIAL MEDIA PLATFORMS REMOVE EVIDENCE OF WAR CRIMES 9-10 (2020), https://perma.cc/9DPL-7RCH; *Social-media Platforms are Destroying Evidence of War Crimes*, ECONOMIST (Sept. 21, 2020), https://perma.cc/FVK4-AXB4; Dia Kayyali & Raja Althaibani, *Vital Human Rights Evidence in Syria is Disappearing from YouTube*, WITNESS (Aug. 2017), https://perma.cc/99GF-S3YY; Kate O'Flaherty, *YouTube Keeps Deleting Evidence of Syrian Chemical Weapons Attacks*, WIRED (June 26, 2018, 7:00 AM), https://perma.cc/63XZ-EFYY.

203. GIFCT TRANSPARENCY REPORT JULY 2020 5 (2020), https://perma.cc/3N8Q-BCVX [hereinafter GIFCT 2020].

204. *Id.*

and other content that would, under most democratic legal frameworks, be considered legitimate speech."[205] If a database, whether produced by GIFCT or by an outside expert, cannot or fails to differentiate between political or satirical commentary and actual terrorist content, then there are significant implications to free speech.

In addition to the concerns about databases, the use of automated software for content moderation raises questions about surveillance, control, and censorship. Deploying extra filtering technology—particularly when we have no understanding of the data, quality, or accuracy of the software—can have disastrous consequences for speech. Platforms "are willing to tolerate higher error costs for speech that is identified as a priority for removal"[206] but the over-enforcement of content to "play it safe" is a choice, not a necessity.[207]

E. Cementing Expert Influence and Faulty Detection Technology in Social Media Content Regulation

Governments have grown to praise industry collaboration with expert bodies and inter-industry efforts to combat terrorism through automated tools such as the GIFCT database. Their expectations that these become norms for the social media industry are reflected in regulatory initiatives around the world.

As part of the UK's efforts to establish a duty of care on companies, Ofcom—the regulatory body that will oversee the platforms' compliance—is expected to also issue "codes of practice which outline the systems and processes that companies need to adopt to fulfil their duty of care."[208] The British government published a non-binding code in December 2020 for terrorist content that outlines the "good practices" companies will be expected to implement.[209] Among the various obligations outlined, an entire section is devoted to collaboration, chiefly the expectation that companies work with governments, the technology industry, civil society, and academia.[210] This government crystallizes this expectation through a series of examples, including "use existing cross-industry capabilities such as hash lists."[211]

The Interim Code also specifies that companies should take steps to "commit to collaborative working with industry and with governments, academia and civil society" and also "engage with relevant industry bodies, which enable the sharing of knowledge and expertise to improve companies' capability

205. *See* Llansó, *supra* note 200.

206. Bloch-Wehba, *supra* note 121, at 44.

207. *See id*. at 45.

208. SECRETARY OF STATE FOR DIGITAL, CULTURE, MEDIA & SPORT & SECRETARY OF STATE FOR THE HOME DEPARTMENT, ONLINE HARMS WHITE PAPER: FULL GOVERNMENT RESPONSE TO THE CONSULTATION, 2020, CP 354, at 11 (UK) [hereinafter ONLINE HARMS WHITE PAPER].

209. DEPARTMENT FOR DIGITAL, CULTURE, MEDIA & SPORT, INTERIM CODE OF PRACTICE ON TERRORIST CONTENT AND ACTIVITY ONLINE (Dec. 15, 2020), https://perma.cc/WNX3-85TG [hereinafter INTERIM CODE].

210. *Id.*

211. *Id.*

to respond to terrorist content and activity online."[212] Despite the concerns of automated tools in content moderation, governments are increasingly mandating the adoption of this technology. More concerning is the loosely defined, yet legally enforceable, standard that companies work with experts. Because the UK's regulatory approach ties a company's duty of care to codes of practice with which companies must comply, one can quickly imagine scenarios where companies are found liable not because they worked with an expert but because they collaborated with ones that the UK government found unsatisfactory. In other words, codifying expert collaboration creates situations where invisible colleges can sway regulatory efforts and push expert consultation onto the same individuals that maintain a stronghold on terrorism scholarship and frame the narrative that many rely on in the military, think tanks, and the private sector.

F. Transparency

It is critical to note that so much of how these experts work is still not known. Even with the Meta staff meeting notes, for example, we are lucky to have only a sense of geographic range, at least, and general work industry, at most. Even with the four companies examined in Part II B, three of them are only known because of a Meta blog post from four years ago. Instead, Crisp Thinking's promise to offer "discreet" services to platforms suggests that Silicon Valley and outside experts alike prefer to keep working arrangements covered.

With this lack of transparency, it is harder to know whether companies are asking for assistance, and from whom. When experts "are veiled in deep secrecy" then "civil society has no effective means to determine the contents of their expert information which impacts the . . . policies in the 'the War on Terror.'"[213]

IV. CHECKS AND BALANCES: HOW DO WE CONTROL THIS PROBLEM?

This paper, has illuminated the problem with expertise—first in the broader context of terrorism studies then in an examination of the inside and outside experts that seek to influence and manipulate how companies create, interpret, and enforce their terrorism policies. Much of this work has been rooted in critique, and while it is needed, it is as important to devote some space to "solutions."

When it comes to the terrorism, and counterterrorism, space, what is abundantly clear is that we have a crisis of expertise. As these individuals and organizations resort to "world-making"[214] by telling companies who is a terrorist and who must be removed from the internet, it may be easy to think away the problem by saying that only "terrorists" are silenced in this process. The truth that is harder

212. *Id.*

213. Reetta Toivanen, *Counterterrorism and Expert Regimes: Some Human Rights Concerns*, 3 CRITICAL STUD. ON TERRORISM 277, 278 (2010).

214. DAVID KENNEDY, A WORLD OF STRUGGLE: HOW POWER, LAW, AND EXPERTISE SHAPE GLOBAL POLITICAL ECONOMY 39-50 (2016).

to acknowledge is that we all are: when decisions, particularly ones as politically and socially consequential as who can participate in political and public discourse, are made *for* us instead of *by* us, we suffer a collective loss. And while this paper is concerned by the ways in which expert influence shapes political extremism and content moderation's response to it, the likelihood that we are facing (or will face) an expert tsunami in other thorny content moderation speech problems is high.

What I illustrate in this paper, however, is that we have collectively arrived here precisely because we venerate expertise without thinking for a moment whether we even like the world they are making for us. In light of this, it may be daunting to wonder how we begin to fix this pernicious problem. One of the most apparent root causes behind the influence experts wield in the terrorism community stems in part from failings at the international level. Some proponents of the current international counterterrorism landscape may point to the fact that the UN has a definition of terrorism—one that is intentionally narrow in scope and framed more on actions and less on actors. Others, however, see this issue differently and may criticize the organization's failure to define terrorism, instead delegating it to member states to resolve for themselves and creating immense variations across member states.

Companies, under immense pressure from governments to crack down on terrorist content, move from underenforcing if they rely only on UN Security Council ideas of terrorism to perhaps over-enforcing by shifting to their own restrictive definitions to make all governments happier. Even if there is a definition at the international level, though, the room for expert leverage remains because the problem of inconsistency *across companies* remains since no platform shares identical definitions or content policies about terrorism either. Though it is normatively preferable to stop these nebulous webs of terrorism expertise entirely, it is both incredibly naïve and wildly unrealistic to say that the use of these experts should stop altogether. Companies are now at a point—whether because of their own vague guidelines, inconsistent definitions across governments, or the absence of a clearer framework at the international level—where they feel they need this type of guidance, though whether they genuinely need it is a topic for another time. Further confounding the problem is that companies also have their own incentives to be tough on terror. Whether through supporting counterspeech efforts[215] to advertising nearly 100% removal rate for Islamic State material,[216] this online war on terror actually works for the companies too. It is good material for public relations teams, eases advertisers wary of their own brands, and soothes jittery governments that appear ready to legislate social media into an internet no one will like.

215. *Counterspeech*, META, https://perma.cc/5YP9-NR5M.

216. SPANDANA SINGH, TAKING DOWN TERRORISM: STRATEGIES FOR EVALUATING THE MODERATION AND REMOVAL OF EXTREMIST CONTENT AND ACCOUNTS 13 (2018), https://perma.cc/5AZG-KTCC.

Instead, the types of solutions that could help with this expertise problem seek to primarily bring much more transparency and diversity of thought into the interplay between content intermediaries and the terrorism industry. First, transparency: one question that has surfaced repeatedly in the research about outside experts is how Silicon Valley can be more transparent about when exactly they will reach out to experts. Yes, Meta's blog post from 2017 lists several expert organizations they worked with, but the problem is this level of detail is inconsistent across platforms.[217] Additionally, Silicon Valley could stand to benefit from a Request for Proposals (RFP) process that would publicly announce when a company wants outside input, and for what purpose. This is not a radical solution; in fact, companies, governments, and non-governmental organizations alike all use RFPs for bidding, consultancies, and everything in between. An RFP process could allow civil society groups, for example, to encourage experts with a different, or fresh, viewpoint to apply for the opportunity to have the "ear" of one of these large platforms. Instead, deals are most likely done through this invisible college of experts, friends tapping friends (or friends of friends) for help.

Another critical way to boost transparency is for companies to be more forthcoming about who precisely provides expert input. One former colleague of the author, who only agreed to speak on the condition of anonymity, said that "the moment you know your name will be public, people don't want to cooperate. It becomes a security risk, even a reputational risk." Perhaps the security risk is valid: if an outside expert has genuine concern for their safety to be named publicly, then it could be counterproductive to ask a company to "out" them. However, anonymity to protect one's reputation could actually encourage less thorough work. After all, imagine the types of recommendations one could champion if his or her name is never publicly known?

In addition to transparency, there is a desperate need to disrupt the orthodox, hegemonic narratives around terrorism to which many experts prescribe. The idea here is not to purge a company's payroll of employees or parties just because they adhere to a conventional point of view. Instead, companies must be more intentional about working with alternative and critical experts, as well as individuals who may study terrorism but do so from other disciplines that are often overlooked in this space—in particular anthropology, sociology, and peace studies. One idea could be to do what disciplines and professions do: create a set of learning objectives and a credentialing body that assesses whether someone can call themselves a terrorism expert.

One other way companies could promote a diversity of thought is through the research support Meta, Twitter, and YouTube already provide. In one blog post, Monika Bickert wrote that "[c]onducting and funding research to study counterterrorism and terrorism is a critical part of our [Meta's] work."[218] The GIFCT initiative has a research arm—the Global Network on Extremism and Technology

217. In the 20+ policy meeting notes available, not one actually revealed the experts that weigh in on, and help shape, Meta's enforcement strategy. *See infra* Part III.B.

218. Monika Bickert & Erin Saltman, *An Update on Our Efforts to Combat Terrorism Online*, META NEWSROOM (Dec. 20, 2019), https://perma.cc/6HKJ-BW7P.

(GNET).[219] GNET's first batch of reports was delivered by the Royal United Services Institute,[220] and GNET's research more broadly is carried out by the International Centre for the Study of Radicalisation (ICSR) in the Department of War Studies at King's College London.[221]

Unfortunately, this research initiative is already replicating the problems with orthodoxy that have been identified in this paper. KCL "illustrates more than any other higher education institution in the UK the blurring of the line between government, the military and academia."[222] The ICSR works very closely with Boaz Ganor and the Interdisciplinary Center Herzliya in Israel.[223] Ganor served the Israeli Prime Minister as the country's counterterrorism coordinator.[224] This "nexus of institutions shows more than connections between academia and the military; it shows that the organizational lines between academia and the military/government have been at minimum blurred, perhaps even erased together."[225] Though the large content intermediaries have a chance to fund research from an array of terrorism experts, the fact they only resort to institutions that perpetuate many of the problems, risks, and concerns illustrated throughout this paper is disquieting.

CONCLUSION

In law school, so much of my time has been spent parsing out the difference between "law" and "policy" arguments. I once joked to a friend that I instinctively knew the difference because policy arguments were the more consequential ones. The truth is that we all engage in this expertise business. We sell our names, our degrees, our institutions, and our voices—selling "expertise"—in an effort to shape policy the way we think it should be.

Often, these expertise transactions shape legislative policy that structures how wealth is distributed, what schools teach, and whether certain rights are civil rights. Sometimes it is deadly and in a way that we are unable to see on a daily basis like America's recently-ended "forever war" in Afghanistan. Other times, and increasingly lucratively so, this expertise takes root in this new type of policy that plays out behind what we watch on YouTube or read on Twitter and Meta: platform community guidelines.

By expanding the institutions, individuals, and voices that examine terrorism and its many layers, technology companies can help push the discussion of terrorist use of social media into new spaces of knowledge, or at the very least into subjugated ones. There is, and there will always be, room for orthodoxy. The voices, ideas, and discourses on the margins should be let in too.

219. *Research*, GLOB. INTERNET F. TO COUNTER TERRORISM, https://perma.cc/2TX3-96NN.
220. *Reports*, GLOB. NETWORK ON EXTREMISM & TECH., https://perma.cc/YGE6-TXQL.
221. *Partners*, GLOB. NETWORK ON EXTREMISM & TECH., https://perma.cc/5PKV-FP6U.
222. Miller & Mills, *supra* note 41, at 213.
223. *Id.* at 214.
224. *Id.*
225. *Id.*

***