Program on Extremism
THE GEORGE WASHINGTON UNIVERSITY

MODERATING EXTREMISM: THE STATE OF ONLINE TERRORIST CONTENT REMOVAL POLICY IN THE UNITED STATES

BENNETT CLIFFORD

December 2021

About the Program on Extremism

The Program on Extremism at George Washington University provides analysis on issues related to violent and non-violent extremism. The Program spearheads innovative and thoughtful academic inquiry, producing empirical work that strengthens extremism research as a distinct field of study. The Program aims to develop pragmatic policy solutions that resonate with policymakers, civic leaders, and the general public. The views expressed in this paper are solely those of the authors and not necessarily those of the Program on Extremism or George Washington University.

Executive Summary

Alongside a host of platform governance issues facing technology companies, the exploitation of social media platforms by terrorist and extremist groups is a major controversy in debates about how companies can combat harmful content online. In the United States and around the world, the shortcomings of social media providers in removing terrorist content have increased the frequency and intensity of calls by lawmakers and the public for governments to directly regulate social media companies' policies against terrorist and extremist content.

Advocates of direct governmental regulation present a straightforward narrative of companies failing to meet their responsibility to police terrorist content on their platforms, and governments intervening with strict parameters, hefty fines, and legal penalties to force them into compliance.[1] To push the U.S. government to act, advocates of government regulation cite examples of these measures adopted by governments around the world. Yet, oftentimes missing from these arguments are thorough evaluations of the state of terrorist and extremist content online, as well as historical assessments of the interplay between governments and social media providers on the question of how to manage online terrorist content.

By reviewing studies of how today's terrorist and extremist groups operate on social media in conjunction with an overview of U.S. government regulation of terrorist content online, this report finds that stricter U.S. regulation of social media providers may not be the most effective method of combating online terrorist and extremist content. Specifically:

- Direct governmental regulations that ignore other sources of regulation on content removal policies could disrupt growing intra-industry collaboration on countering terrorist content online.
- In many regards, the U.S. government defers to and depends on the private sector to conduct counterterrorism online. Many factors contribute to this arrangement, including limits on the government's authorities, expertise, staffpower, dexterity and political will to manage online terrorist content with the same efficacy as major social media companies.
- Attempts by other governments to strictly regulate social media companies' terrorist content removal policies hurt small companies, created double standards and redundancies, and raised concerns about censorship and free speech.
- Proposed regulations may only affect major U.S. social media providers; smaller and non-U.S. companies may be unable, unwilling, or not required to comply. Due to the proliferation of social media platforms exploited by terrorists and extremists, platforms that may be unaffected by U.S. government regulation currently host a large proportion of terrorist content online.
- In certain regards, major social media companies' content removal policies have more flexibility than the U.S. government to be able to account for new terrorist and extremist groups and actors and their respective tactics, techniques, and procedures online.

Introduction

The proliferation of online terrorist and violent extremist content, particularly on social media platforms, is one of the major policy issues facing U.S. counterterrorism authorities and digital communications technology providers. The advent of massive online social media services led to a range of terrorist and violent extremist groups exploiting these platforms for propaganda, recruitment, radicalization, and operational planning.[2] Initially, terrorist content was most plentiful on platforms operated by exponentially growing American companies, sparking society-wide debates about the role of these platforms, their approaches to harmful content, and industry regulation. From government officials to company shareholders, civil society organizations to media reporting, societal pressure to regulate digital communications service providers usually involves the question: "why is your company not doing more to stop terrorist content on your platform?"[3]

When the public perceives that major social media companies are failing to address terrorist content, many call for direct governmental regulation, or externally imposed laws that attempt to shape the behavior of the company in question. While government regulation can take an incentivizing form, pushes for regulation against major social media companies in the wake of violent extremist activity online almost always involves punitive action. For example, American lawmakers have threatened to fine companies, remove companies' immunity for hosting third-party content, charge companies with providing material support to terrorists, and threatened to break up companies.[4] Other debates on social media content moderation policies, particularly regarding hate speech, disinformation, and content harmful to children, have also influenced a more vocal call for the U.S. government to crack down on major social media companies.[5]

Calls for increased governmental regulation are understandably attractive in theory for lawmakers and the public, but if put into practice, the imposition of more stringent regulations on major service providers may not deliver the intended results. As this paper argues, direct U.S. government regulation of major social media companies' content removal efforts may not have a meaningful effect on either the amount of extremist content on those platforms or broader issues of online extremism and radicalization. As the landscape of extremist use of the internet has evolved in its architecture, major players, tools, and tactics, the public debate about content removal policy has largely remained stagnant, relying on the same tropes, axioms, and solutions that it did ten years ago. Some proposed regulations still fail to account for how terrorist and extremist content spreads online today and are therefore unlikely to be effective.

This paper details some limitations on the ability of the U.S. government to meaningfully regulate major social media companies' terrorist and extremist content removal policies. First, by defining the various forms of regulation that affect technology policy development, it argues that direct government regulation is not the only source of regulation influencing technology companies' terrorist content removal efforts. Then, evaluating the relationship between the U.S. government and major social media providers, it details how the U.S. government has effectively outsourced its online counterterrorism responsibilities to major social media companies. Not only would it take a herculean effort for the U.S. government to wrest these responsibilities back from the private sector, but doing so may inadvertently jeopardize efforts to confront the problem of terrorist content online. Finally, it charts the evolution of extremist use of social media, documenting the constellation of platforms and services popular with extremist groups today that would not be subject to U.S. government regulation against major social media providers. In sum, industry-led self-regulation by social media companies is imperfect, but is nonetheless more practical and promising than the U.S. government throwing its weight behind managing terrorist content removal or dictating standards to the private sector.

Sources of Regulation and Online Terrorist Content Removal Policy

Before delving into the effects and functions of regulation, it is important to outline the various modes under which technology—in this case social media—can be regulated. Oftentimes, observers tend to view the dynamic of regulation as a strictly push-pull affair between governments on one side and tech companies on the other. As the axiom entails, tech companies make decisions about how to run their platforms solely based on profitability, and it is solely the government's responsibility to constrain its behavior so that it benefits the public good. To some degree, this idea is correct insofar that it elucidates a few important ways that regulation can influence the behaviors of tech companies, including social media providers. But it is not a complete picture of how regulation operates.

In his 1999 book *Code and Other Laws of Cyberspace*, Lawrence Lessig proposes that four sources regulate the behavior of entities that use digital communications technologies—whether they are individual users, service providers, tech companies, or any user of the internet.[6] In this framework, the law, norms, the market, and architecture all play roles in regulating online behavior. Each source has a unique effect on regulation: "norms constrain [behavior] through a stigma that a community imposes; markets constrain through the price they exact; architectures constrain through the physical burdens they impose; and law constrains through the punishment it threatens."[7]

These four sources of regulation constitute the general modalities by which society governs online behavior, and users on the internet can observe each of the four sources at work in different ways. When law is the primary source of regulation, the regulator tends to be a government that utilizes legal penalties, grants, public/private partnerships, pressure, or the threat or promise thereof, to coerce or incentivize private entities to change their behavior.[8] When norms are influencing regulation, the regulators are the community of users; they self-regulate "through the threat of ex post sanctions imposed by a community."[9] When the market determines regulation, the primary incentive/disincentive structure is shaped through the lens of economic cost and benefit; users make decisions about how to use digital communications technologies based on potential profit and loss.[10] Finally, the architecture of cyberspace—in this case, the code, hardware and software shaping each platform—has the effect of constraining what is possible and what is not through the infrastructure of the technology.[11]

Applying Lessig's four sources of regulation to decisions by social media companies to adopt terrorist content removal policies problematizes the idea that the government—

through the law—is the only effective regulator of social media. As Lessig describes it, advocates of legal regulation who do not consider the effect of other regulations are engaging in "law-talk… speaking as if law must simply take the other three constraints as given and fashion itself to them."[12] Moreover, stricter government regulation does not always lead to enhanced efforts by social media companies to police terrorist content as expected by a law-centric approach to regulation. Attempting to constrain a company's decision-making merely through one form of regulation, without considering the effect of the other forms of regulation, can muddle or cancel out the effect of the imposed constraints.[13] Additionally, as the next section argues, non-legal forms of regulation may be equally important in shaping the decisions made by major social media companies to adopt certain policies against terrorist content.

Which Sources of Regulation Matter and How?

Legal regulation can and does have a significant effect on major social media companies' decisions about content removal. Governments employ two forms of legal regulation: they can directly regulate behavior through legal incentives and disincentives, or they can indirectly attempt to regulate behavior through passing laws to influence the other sources of regulation.[14] To illustrate the difference between direct and indirect legal regulation, Lessig uses the example of regulating seat belt usage in vehicles. The government can pass laws that require seatbelt use, punishable by fine (direct legal regulation), it could start a public awareness campaign (indirect legal regulation through norms), fine automakers who don't install seatbelts (indirect legal regulation through the market), or mandate the installation of seatbelts in all vehicles (indirect legal regulation through architecture).[15]

Nevertheless, direct legal regulation has dominated the discussion in the U.S. about how the government responds to terrorist content online. Certain parts of the U.S. government, often in response to public pressure, frequently consider both direct and indirect regulations on social media companies due to the prevalence of terrorist content on their platforms. While the next section of this paper expands on this history, the early forms of regulatory behavior by the U.S. government in this arena during the mid-2010s typically were indirect. They came in the form of proposed public/private partnerships, awareness campaigns, memoranda of understanding, expertise transfer, and delineation of responsibilities between the U.S. government and major social media companies.[16] Over time, however, dissatisfaction with this state of affairs led to some to call for direct regulation, including removing legal immunity for social media companies for hosting terrorist content (pursuant to Section 230 of the Communications Decency Act), fining companies that fail to remove terrorist content within a particular timeframe, and even threatening to use anti-trust law to break up major social media providers.[17]

Yet as the debate over proposed legal regulations escalated in the U.S., other sources of regulation quietly grew in their ability to constrain and influence content removal policies. The first and most obvious non-legal source of regulation in this arena is the market. The major social media providers in the U.S. are all publicly traded companies with responsibilities to their shareholders, who are increasingly concerned about the use of platforms to promote terrorist and extremist activities.[18] In the wake of major terrorist attacks, when the perpetrators are found to have utilized certain social media applications, it is not uncommon for the ensuing public relations crisis to cause a downturn in stock prices for companies—particularly if they believe greater public scrutiny is forthcoming.[19] As a result, companies are also interested in demonstrating their ability to remove terrorist content at a greater rate than their competitors to show shareholders that they are taking the problem seriously.[20]

Because of the role of new technological developments in online content removal, architecture also plays a major role in regulating major social media companies' policies towards terrorist and extremist content. For instance, the mid-2010s heralded significant developments in algorithmic detection and machine learning, drastically increasing the propensity for a company to automatically detect terrorist content through the use of code, rather than sole reliance on human review.[21] Without these technological advances, it would have been incredibly difficult for companies to pursue expanded enforcement of their terms of service (ToS) against terrorist and extremist users of their platforms. That notwithstanding, limitations to new architecture also constrain content removal policy. Human review remains necessary to both determine the inputs for what constitutes a ToS violation (e.g., what qualifies as "terrorist" or "extremist" content), as well as make final decisions in borderline cases, in which computerized review is inconclusive.[22] However, the use of human content reviewers is also controversial: industry watchdogs report frequent and rapid turnover among social media companies' contracted employees tasked with reviewing terrorist and other harmful content, including high rates of post-traumatic stress disorder, depression, and suicide.[23] As a result, most companies' policies on terrorist content are constrained both by the technological limits on machine-learning algorithms as well as moral limits on the use of human reviews.

The last sources of regulation on technology companies in their efforts to remove terrorist and extremist concept are norms. Within international civil society, arguments for service providers' collective responsibility to remove terrorist and extremist content from the internet for moral and ethical reasons are increasingly popular. This norm has two implications for companies. First, individual companies are not only responsible for content on their own platform but on the internet as a whole. Therefore, as interconnected nodes in a network of service providers, companies (as stewards of the internet) should collaborate with one another to share best practices, coordinate ToS enforcement, and assist smaller companies with the know-how and resources necessary to address terrorist content online.[24]

In many ways, normative regulation is the most underestimated source of regulation on content removal policy today. Many observers, viewing tech companies as soulless behemoths driven solely by profit, tend to doubt the existence and/or impact of norms on their decision-making.[25] This conflates morals with social norms, which merely requires a group of actors to adopt a common set of understandings that govern their behavior regardless of their morality. More to the point, several landmark decisions by social media companies in the field of terrorist content removal during the past few years are impossible to understand without the emergence of the norm amongst major social media providers that they have a collective responsibility to remove terrorist content online. The concluding sub-section highlights the role of norms in these decisions and argues that blindly adopting increased legal regulation might have the unintended consequence of disrupting normative regulation.

Norms and Online Terrorist Content Removal Policy

As U.S. government officials consider methods of direct intervention to manage how social media companies use ToS enforcement against terrorist content, it must account for the growing role of intra-industry norms in company decision making. One important development of the past few years is a growing consensus among major American social media companies that they have a normative, collective responsibility to address terrorist content on their own platforms.[26] It would be easy to dismiss this norm as an effect of another source of regulation, such as the market or the law, if it were not for several decisions made collectively by major social media providers that support this norm without furthering the companies' economic interests or complying with legislation.

A growing consensus among major social media providers is that they have a collective responsibility, independent of legal, market, and architecture requirements, to remove terrorist content from their platforms and from the internet as a whole. This consensus is evident in efforts by the major providers towards intra-industry collaboration, coordination, and sharing of best practices amongst competitors large and small. They are also legally codified within the 2019 Christchurch Call to Eliminate Terrorist and Violent Extremist Content Online.[27] The Christchurch Call, launched by the government of New Zealand in the wake of a 2019 terrorist attack against a mosque in Christchurch, New Zealand in which the perpetrator live-streamed his attack online, includes signatories from government and the tech industry who pledge to:

> take transparent, specific measures seeking to prevent the upload of terrorist and violent extremist content and to prevent its dissemination on social media and similar content-sharing services, including its immediate and permanent removal, without prejudice to law enforcement and user appeals requirements, in a manner consistent with human rights and fundamental freedoms.[28]

Interestingly, when the Christchurch Call was launched in May 2019, the U.S. government was not among the initial signatories; the Biden Administration only recently signed the document in May 2021.[29] Instead, the first U.S. entities to sign the pledge were Facebook, Twitter, Google, Microsoft, and Amazon.[30] This timeline signifies the development of an industry-wide norm two years prior to a governmental norm, and is important because it establishes the fact that major social media providers made the decision to codify the norm even though there were no explicit requirements to do so issued by the U.S. government

Two aspects of the Christchurch Call are exemplars of the norms shaping companies' behavior in the field of terrorist content removal. First, normative guidelines tend to have the lofty objective of completely eliminating terrorist content from the internet as a whole, across websites, applications, and platforms. This goal is logistically impossible at the tactical level, but creates a linkage between terrorist content and other types of harmful content for which existing norms promote total elimination.[31] Analogies between terrorist content and child sexual abuse material (CSAM) are especially common in this regard; the latter type of harmful content is subject to a tech industry community-wide consensus that it should be prohibited and completely removed from the internet.[32] The argument from tech companies is less that the two types of content are analogous in their moral harm or in the demanded response, but instead that they are analogous in terms of the responsibilities that major companies have to remove them from the internet.

Moreover, by signing the Christchurch Call, major social media providers recognize that terrorist content online is a tragedy of the commons. Larger platforms have the ability to hire review teams, deploy top-of-the-line algorithmic detectors, and procure terrorism-related expertise.[33] If implemented solely among individual companies, the individual responses may be sufficient to remove terrorist content from the larger platforms, but in so doing, may displace it onto smaller platforms. Smaller entities may lack the will, resources, or wherewithal to employ removal efforts, creating the commons problem.[34] The Christchurch Call also binds companies to "support smaller platforms as they build capacity to remove terrorist and violent extremist content, including through sharing technical solutions and relevant databases."[35] This norm-based regulation encourages companies to share best practices in large, international fora that bring together large and smaller companies alike, in the hope of evenly distributing capacity to remove terrorist content.[36]

If not targeted towards specific ends to avoid collateral damage, enhanced governmental regulations of social media providers could force companies to divert resources, workstreams and personnel away from collaborative endeavors and create an "every-platform-for-themselves" mentality to removing terrorist content online.[37] The goal of compliance—and avoiding legal penalties—might subsume norms that flourished in an environment without direct legal regulations and leave smaller platforms to fend for themselves, which is inimical to the spirit and nature of the Christchurch Call. Therefore, legal regulation—particularly direct legal regulation—could disrupt emerging norms by limiting the factors necessary for their creation and development.

The U.S. Government's Outsourcing of Online Terrorist Content Removal Policy

While private companies have been tangibly involved in U.S. counterterrorism efforts for decades, the development and rapid growth of major social media platforms has substantially boosted the role of the private sector in what was previously considered a core governmental responsibility.[38] Today, both the American public and the U.S. government consider major social media companies not as auxiliary actors in online counterterrorism, but as the primary entities responsible for countering terrorist content online. This shift occurred due to a litany of factors. Social media companies are viewed as more adept and more knowledgeable than the government in managing content on their own platforms, certain parts of the government are perceived as unable to adapt to new technologies, and the problem of terrorist content online became a transnational problem, not subject to the jurisdiction of any single government or regulatory entity.[39] Recognizing its own problems, the U.S. government has deferred responsibility to regulate terrorist content online to major social media companies.

To reverse this outsourcing of responsibilities would require a sea change in how the U.S. government operates: namely, massive efforts to dictate the terms of regulation to social media companies, investments in new agencies, bureaucracies, and departments tasked with regulating content, and a host of legal, ethical, and Constitutional challenges to combat.[40] The U.S. government abhors massive overnight change, but oftentimes faces public pressure to take a more active role in regulating social media. Thus, senior national security officials and Members of Congress often opt for the middle ground, issuing largely empty threats to punish major social media companies for hosting terrorist content on their platforms. The likelihood of the U.S. government making good on these threats is minimal, which is intentional. The constant threat of massive regulatory action functions as a sword of Damocles, provoking companies into increased steps to remove content.[41] In turn, the companies make minor overtures to the U.S. government, modifying their policies on terrorist content to sufficiently appease lawmakers and delay regulation.

From one vantage point, this repeated cycle of government threats to regulate and concessions by social media companies seems tantamount to kicking the can down the road. But in many ways, this process is responsible for many of the major efforts that social media companies launched in order to remove terrorist content from their platforms in recent history and is therefore beneficial. Considering the alternatives to the cycle, it is unreasonable to expect that if the U.S. government had either adopted a *laissez-faire* approach or followed through on their threats to regulate, that it would have resulted in sufficient incentives for social media companies to develop effective content removal policies. To discuss this dynamic, the paper evaluates the effects of an informal series of engagements with social media companies that the U.S. government referred to as the "Madison Valleywood Project" and its role in subsequent efforts by those companies to create architectures for responding to online terrorist content.

Increasing Government and Industry Engagement

In the U.S. around 2015, political, societal, and market pressures converged on major social media companies to take greater action against terrorist content on their platforms.[42] U.S. law enforcement and intelligence agencies had understandably large stakes in encouraging social media companies to ramp up enforcement of their ToS against terrorist content. Attacks by supporters of the Islamic State of Iraq and Syria (ISIS) in the West, particularly in Paris in November 2015 and in San Bernardino, California in December 2015, ramped up the stakes for U.S. counterterrorism agencies, who viewed the perpetrators' use of digital communications technology as essential elements to their plots.[43] In early 2016, representatives from the Department of Justice's National Security Division, the National Security Council, and the National Counterterrorism Center set up a series of meetings with Silicon Valley executives to discuss the role of tech companies in countering violent extremist groups online.[44] During these meetings, referred to as the "Madison Valleywood Project," the U.S. government senior officials encouraged a range of companies, including Facebook, Google, Twitter, Instagram, Snapchat, Tumblr, and Microsoft, to assist the U.S. government in its counterterrorism mission by helping to counter extremist exploitation of their services.[45] In negotiations, the U.S. government promised its partners in Silicon Valley the opportunity to access and receive briefings from counterterrorism experts within the U.S. government and other resources in exchange for their cooperation.[46]

Yet, at the same time that segments of the U.S. government sought outreach to social media companies and incentivized them to remove terrorist content more efficiently, other parts of the U.S. government engaged in public feuds with tech companies about responding to government requests for information In criminal investigations against ISIS supporters and other homegrown violent extremists in the United States, the FBI frequently subpoenaed records from major social media companies, threatening similar suits in court for non-compliance.[47] The seemingly contradictory approach can be partially explained by the vociferous debate within government on whether content removal was net beneficial to the government's counterterrorism strategy. Policymakers tended to view content removal as necessary to reduce the number of individuals who could potentially be exposed to terrorist content online and radicalize to violence as a result. Meanwhile, their counterparts in intelligence collection and operations saw the existence of terrorist content online as a "bug-light."[48] Online terrorist propaganda was not capable of large-scale radicalization for terrorist groups, according to the practitioners' arguments, but instead helped law enforcement and intelligence agencies track who was producing and accessing content. If a mass-scale removal of terrorist content took place, this side of the debate maintained that U.S. counterterrorism agencies would lose their window of access to critical evidence necessary to interdict terrorist supporters.[49]

Against this backdrop of public and governmental pressure, the major American social media companies each began stepping up their efforts to reduce the amount of terrorist content on their platforms. The so-called "Big Four," referring to Facebook, Twitter, Google, and Microsoft, each engaged in individual efforts to systematically remove accounts promoting terrorist content from their platforms and take down content, with an emphasis on more strictly defining their respective ToS and increasing resources available to in-house enforcement teams.[50] More importantly, the Big Four began processes for intra-industry collaboration on content removal. This began with the establishment of Facebook, Twitter, Google, and Microsoft's internal hash sharing database, a jointly managed databank of unique image and video hashes identified as terrorist content by the platforms in question.[51] From there, each hash representing terrorist propaganda—such as the unique code for an ISIS video—would be installed into automatic algorithmic detectors that could identify and immediately delete any content matching that hash across the four platforms. This dramatically increased the scope and alacrity of terrorist content removal efforts.[52]

The hash-sharing database and the establishment of cooperative efforts between the Big Four led to the creation of the Global Internet Forum to Counter Terrorism (GIFCT) in 2017. The GIFCT was spearheaded by "a group of companies, dedicated to disrupting terrorist abuse of members' digital platforms."[53] To this end, the GIFCT's primary workstreams are the management of the internal hash-sharing database and the Content Incident Protocol, a rapid-response mechanism for GIFCT's member companies to collaboratively react in real time after major terrorist attacks, when social media is often flooded with propaganda, misinformation, and other harmful content.[54]

At its foundation, the GIFCT was an industry-led forum, where the four founding entities rotated leadership positions; in 2019, it was reorganized as an independent entity with its own executive director, operating board, and independent advisory committee.[55] During that time period, the GIFCT also expanded to include 19 new member and partner companies in addition to its original founding stakeholders.[56] Organizations that seek to join the GIFCT agree to the organization's Membership Pillars. They must have internal content standards (ToS or privacy policies) that explicitly prohibit content that promotes terrorism and violent extremism, create or receive reports that outline violations of ToS related to terrorist content, employ technical solutions to respond to terrorist exploitation, and most importantly, pledge their commitment to transparency about content removal decisions, "respecting human rights, particularly free expression and privacy, when implementing content removal policies," and support civil society organizations that are engaged in efforts to counter violent extremism.[57]

Threatened legal penalties, fines, or even modifications of liability protection laws (like Section 230 of the Communications Decency Act) to make social media companies accountable for terrorist content never came to fruition in the U.S.[58] But, independent of the threat of regulation, social media companies took steps in furtherance of a dual moral responsibility to confront terrorism on their platforms and assist the government in counterterrorism prosecutions. More importantly, the government's "sticks" came with "carrots" for social media companies that engaged with the U.S. government in its counterterrorism priorities, such as access to expertise and public/private partnerships.[59] In short, rather than taking the route of direct regulation, the government "[outsourced] the responsibility to prevent and confront terrorists' use of the internet to private companies."[60]

Why Outsourcing May Be Preferable to Direct Regulation

As the previous section demonstrates, sometimes the threat of regulation is sufficient to hold major social media companies to account, and some delegation of online counterterrorism responsibilities can benefit both the U.S. government and social media companies. It is reasonable, especially after instances of significant failure by social media companies to enforce their ToS against terrorist actors on their platforms, for the U.S. government to threaten crackdowns. However, an actual crackdown—in the form of fines, legal penalties, anti-trust actions, or removal of third-party hosting immunity—may hamper any leverage that the U.S. government has over social media companies. While maximizing disincentives for non-cooperation, it would also destroy any incentive that a social media company would have to cooperate with the U.S. government. At best, increased regulation or antitrust actions would not affect a company's interest to improve their policies against terrorist content; at worst, it would force the U.S. government, who have not maintained primary responsibility in the policy arena for decades, to take the driver's seat without the skills, expertise, and architecture necessary to do so.

What if the U.S. government had followed through on their threats to punish social media companies for hosting terrorist content in the mid-2010s? Some answers to this question can be found in the European regulatory environment, where some countries responded to the developments above by tightening the screws on social media companies.[61] Germany's Network Enforcement Act (NetzDG), passed in the Bundestag in the summer of 2017, places a 24-hour time window for social media providers to delete terrorist and other illegal content after it is posted, before it levies up to five-million Euro fines against the provider.[62] The French parliament passed a similar law in May 2020, but with a time limit of only one hour.[63] The United Kingdom authorized criminal penalties on the demand-side of terrorist content; the UK Counter-Terrorism and Border Security Act 2019 made accessing or viewing terrorist content online a chargeable offense in some cases.[64]

Due to U.S. tech companies' multinational operations, they are often forced to comply with much stricter regulation about terrorist content in other jurisdictions, alongside other laws regulating the behavior of social media companies.[65] Independent reviewers found that for major social media companies, the policy tools utilized in strict regulations (e.g., time limits for removing content, fines, required installation of algorithmic detectors, and criminal penalties) were often ineffective in improving content removal policies and also led to several negative externalities.[66] First, while major social media companies were easily able to afford compliance with European regulations without changing their behavior, other social media platforms were not.[67] As the next section shows, smaller social media providers today are equally if not more important in confronting terrorist content online, and strict regulation "[risks] penalizing small platforms with heavy fines and leaving them behind, instead of offering them the support needed to counter the threat."[68]

As a result, direct governmental regulation for terrorist content could bifurcate platforms into "haves" and "have-nots"; or those with the resources to comply with increased regulations and those without.[69] Potential new social media platforms, facing significant financial burdens from compliance, may either fail to successfully enter the market or degrade other aspects of their operations due to resource strains.[70] This creates two potential limits on the ability of social media providers as a whole to effectively counter terrorist content. First, it limits the number of social media providers to the major players only, allowing their ToS enforcement against terrorist content (for better or for worse) to play an even more outsized role in determining content removal policies as a whole.[71] Second, it limits innovation and creativity in developing content removal and moderation policies, as fewer sources of new ideas and experiences have inputs into determining industry-wide norms and policies against terrorist content.[72]

Second, in order to clarify their regulatory policies to companies, governments that adopted strict regulation were forced to define the terms of regulation and/or develop the infrastructure necessary for social media companies to comply.[73] For instance, to institute a legal mandate that social media companies are required to remove "terrorist content" from their platforms, it becomes the government's responsibility to define what "terrorist content" is. In certain cases, this created massive legislative debates.[74] The pre-existing terrorism designation processes in many countries were insufficient to cover all forms of "terrorist content" extant on the internet, and partisan politics seeped into the discussion of which content should be banned. In some cases, this led to unclear definitions of terrorist content that left companies unsure about the terms of compliance.[75]

In other cases, governments that required social media companies to undertake a particular content detection process—such as requiring them to use algorithms to identify terrorist content—had to develop the algorithms themselves.[76] Governments developing definitions of terrorist content and/or algorithms to detect and remove it are in effect replicating what most major social media companies have done in-house already, and there is no guarantee that governmental definitions or algorithms will be more precise or effective than major social media companies'.[77] The United Kingdom Home Office's artificial intelligence-based detector of terrorist content, for instance, was rejected by dozens of companies for redundancy with their own algorithmic detectors, and could only reliably detect official propaganda videos produced by one terrorist group.[78]

Lastly, the development of harsh government regulation against social media providers' terrorist content removal policies encourages modeling by other countries. With regard to removing terrorist content, democratic states may have legitimate interests in balancing the right to free speech and expression online with the need to avoid harms to public safety. Unfortunately, other states have modeled anti-terrorist content regulations to crack down on their political opponents online.[79] Danielle Keats Citron argues that the harsh European Union regulation of social media providers as an example of direct regulation poses the potential for "censorship creep…whereby a wide array of protected speech…may end up being removed from online platforms on a global scale."[80] If the U.S. chooses to add its own direct regulations to the global list of standards for social media companies on terrorist content, it may not only add to the confusion that the current patchwork system entails, but also encourage less-democratically inclined actors to use content removal regulation as a cudgel for policing thought.

Evolutions in Terrorist Exploitation of Social Media

On top of the challenges for governments and social media companies associated with developing policies, the information environment is rapidly evolving due to changes in how terrorists and extremists utilize the internet. In a 2019 article, the then-director of Facebook's counterterrorism and dangerous organizations team Brian Fishman observed that "generally speaking, terrorists use the internet in much the same way as other people."[81] While this seems self-evident, incorporating this observation into public policy and the discourse surrounding it has proved immensely difficult. The idea that terrorists and violent extremists use social media in the same ways (if not for the same purposes) as an average person has several important implications. First, it would mean that terrorists utilize different social media platforms for different ends, applying a multi-vector strategy of disseminating content by selecting whichever platforms work best in getting the message out to their audiences.[82] Secondly, no two terrorist groups—or even actors within a terrorist group—are likely to use the same social media platforms in the same way.[83] Finally, and most importantly, terrorists are just as willing as the average person to experiment with new social media platforms and stop using others if they cease to fulfill their purposes.[84]

Despite these implications, supported by a mountain of evidence on online violent extremist behavior in the past decade, the assumption that the bulk of terrorist content is concentrated on a few major platforms continues to drive lawmakers and the public towards regulation. However, this is no longer the case. Efforts by major social media companies to detect and remove terrorist content caused a dispersal of violent extremist material onto a constellation of different platforms, to the extent that the bulk of terrorist content online today is not on the major social media providers but on smaller platforms.[85] For instance, a 2019 longitudinal study by the United Nations Counter-Terrorism Executive Directorate-backed initiative Tech Against Terrorism analyzed over 45,000 URLs posted by ISIS supporters on over 330 different platforms between 2014 and 2019. They found that a majority of the URLs were spread out across 322 out of the 330 platforms, and a majority of the top 50 sites used were "small or micro-platforms."[86] A Program on Extremism study during the same timeframe analyzed a corpus of over 46,000 URLs posted by ISIS supporters, finding that they directed to over 730 unique base domains.[87]

Simply put, increased regulations and antit-trust actions against major social media companies would likely have no effect on a large proportion of the terrorist content that exists on the internet today. The new staging ground for extremist propaganda is not

Facebook, Twitter, YouTube, or any major American social media company; it is comprised of a collection of smaller social media providers, many of which are incorporated outside of the jurisdiction of U.S. regulators.[88] As the next section details, extremist groups of multiple persuasions have already adapted their strategies of social media engagement to weather increased steps by major companies to crack down on terrorist content. During the past decade, as most of the major social media providers took increased steps to remove this content from their platforms, terrorist groups online rapidly and fluidly adapted new platforms and tactics to avoid removal.[89] Therefore, even if a stricter legal regime was able to marginally limit the amount of terrorist content on major American social media platforms, it would still be unlikely to make a significant dent in the vast, cross-platform ecosystem of online terrorist content.

In addition, despite the early media and governmental focus on Salafi-jihadists, a range of extremist groups exploit social media platforms for their own ends. To combat these actors, major social media companies have in some cases deviated from government guidance and devised their own designation procedures, in an attempt to define more holistically what constitutes terrorist content or actors online.[90] These efforts are not without their flaws, but it is highly unlikely that increased U.S. government intervention to dictate standards for social media companies' designation processes would improve them. Today's U.S. government lacks a designation process for U.S.-based terrorist groups, as well as the dexterity to quickly adjust standards to account for new groups and actors and the ability to consistently apply them.

New Platforms

The current landscape of online violent extremist content was shaped in large part by decisions made by major social media companies to alter their content removal policies. Many of these shifts, such as the creation of the GIFCT and altering companies' parameters for defining, detecting, and removing terrorist content on their platforms, were detailed in preceding sections. Different extremist groups reacted to this heightened enforcement in a number of ways, but centered their strategies for survival online around adaptation and migration.[91] Audrey Alexander writes that supporters of extremist groups "demonstrate tremendous agility across multiple platforms" in reacting to major social media companies' increased enforcement of ToS, noting that "some accounts rallied in the face of shutdowns [while] others expressed interest in migrating to online environments that were more hospitable or optimal for extremist users."[92]

Despite the investments of major platforms in ToS enforcement, extremist groups struggle to maintain footholds on major social media platforms because they are the only avenue to ensure access to global audiences.[93] Nevertheless, it is an uphill battle for most extremist groups, with the pace of takedowns of content overwhelming the pace of uploads on some platforms.[94] Yet, extremist groups have generally been successful in migrating to other social media platforms which they can exploit as alternatives when the major services are inaccessible.[95] Using alternative platforms can be disadvantageous for terrorist and extremist groups: they rarely attain the same audience engagement as they would on a major platform, and due to their obscurity and relative lack of resources, they are subject to service disruptions. [96]

One clear advantage of many of these platforms, however, is that they are relatively more hospitable environments for extremist content. This is a function of several traits that are common amongst companies that manage smaller platforms. The first and most common trait is that many smaller providers lack the personnel, resources, and expertise necessary to institute a broad-based terrorist content removal paradigm on their platforms.[97] A commonly cited example of this type of company is JustPaste.it, a file-sharing site operated by a Polish social media startup.[98] Due to the platform's simple design and accessibility features, including operating with right-to-left alphabets like Arabic, ISIS supporters exploited the platform to host multimedia propaganda releases.[99] In the early days of ISIS social media campaigns on JustPaste.it, the company had one staff member and a minimal budget; it simply could not keep up with the influx of violent extremist content posted to the platform.[100]

Second, many of the platforms that violent extremists prefer today provide some technological affordances to users to protect them from takedowns. After the Big Four launched their campaign against terrorist content on their platforms, many extremist groups migrated to using text-based instant messaging applications that provide encryption.[101] Chief among these platforms is Telegram, an online instant messenger that combines a unique suite of features (including direct messaging, group messaging, file-sharing, and encrypted communications).[102] While Telegram has engaged in efforts to remove terrorist content from its platform, it remains immensely popular among extremists, including supporters of the global jihadist movement, the extreme right-wing, and conspiracy groups.[103] Because of the structure of Telegram's service and its provision of enhanced security and privacy protections, much of the extremist content that is present on the platform is outside the reach of ToS enforcement.[104] More recently, extremist groups have experimented with a host of platforms offering decentralized servers and/or data storage, which would theoretically make content hosted on the platforms immune from any effort by the service provider to remove it.[105]

Finally, there are platforms that have little to no interest in complying with U.S. government regulations, or alternatively will only work with governments behind closed doors. Their reasons for non-compliance are multiple. Some platforms are immune from binding American government regulations because they are not based in the U.S.[106] This applies to the two platforms mentioned by name above, although as EU-based entities, they are subject to European laws governing terrorist content removal policies.[107] Others, regardless of their jurisdiction, may have an ideological inclination against content removal altogether. Some U.S. social media firms cite their perspectives on First Amendment protections or government censorship as their justification for avoiding content removal policies.[108] Others, however, are motivated by malignant ends. During the past several years, a variety of extremist groups have experimented with creating their own social media platforms to avoid content removal altogether.[109]

Today, at least one of the categories above applies to many of the platforms on which terrorist content is concentrated. Even if the U.S. were to adopt more stringent regulations against major American social media companies, many other social media providers either could not, would not, or are not required to comply. The dynamic of extremists preferring smaller platforms could metastasize as a result of stricter regulations. Harsher crackdowns by major social media providers may encourage terrorist and extremist groups to move even more of their online activities to platforms that are not covered under the new regulations. In effect, this would achieve the goal only of further dispersing extremists throughout the internet, not reducing their ability to operate online.[110]

New Actors

In addition to the diverse platforms that terrorist and extremist groups exploit, there is also an increasing diversity of terrorist and extremist groups with a substantial social media presence. This is also not a new phenomenon. A plethora of extremist groups of various ideological persuasions were initial adopters of social media, as they saw their potential for recruitment and propaganda operations.[111] Nonetheless, until recently, the efforts of social media companies to remove terrorist content were primarily concentrated on a handful of groups.[112] Major social media companies' efforts were predominantly focused on groups designated as foreign terrorist organizations by the U.S. Department of State, especially Salafi-jihadist groups, oftentimes at the expense of other groups that did not fit either label.[113] This led to criticisms of major social media providers that they were implicitly permitting other actors, in particular right-wing extremists, to operate with impunity online.[114]

This is a legitimate criticism of social media companies, unless it is paired with the suggestion that increased governmental regulation would help remedy this disparity. Indeed, the record of major social media providers in enforcing ToS against American domestic violent extremist groups has been mixed at best.[115] Unfortunately, despite the numerous updates to privacy policies and ToS, enhanced terrorist content designation processes, and the increases in resources for policy enforcement and intra-industry collaboration, domestic terrorists in the U.S. frequently exploit major social media platforms to recruit and disseminate propaganda. However, these dynamics would be unlikely to change if the primary responsibility for policing the platforms fell on the U.S. federal government, whose own strategy against domestic terrorism exhibits many of the same gaps as social media providers.[116]

The difference between the government managing the standards for content removal versus the private sector is threefold. First, most major social media companies have a complicated, multi-tier process for defining and designating "terrorist content," regardless of its ideology of origin. For instance, Facebook's Dangerous Individuals and Organizations policy classes individuals and organizations into three separate tiers, including state-designated foreign terrorist organizations in the first tier and a range of other actors, including domestic violent extremists, criminal groups, militia groups, and other violent actors into several additional tiers.[117] Each tier corresponds to a certain action for algorithms and human reviewers when content associated with the group appears online.[118] The U.S. government, in contrast, lacks a domestic terrorism designation process.[119] In fact, most of the major social media companies developed their own process specifically because the U.S. government had no guidance or policy on how to address these actors and groups.[120] If the U.S. wanted to more closely regulate how major social media companies formulate their designation processes, it would either have to modify its own process or attempt to hold companies to a different standard than it holds itself to. Either of these options could run the risk of the standards being unclear to companies, less effective in removing terrorist content than existing efforts, or would simply reinvent the wheel.[121]

Even if the U.S. government were to establish a more cohesive list of terrorist organizations and content for the private sector to use, legislation and complicated bureaucratic machinations would be necessary to add or subtract groups or individuals from the list.[122] Ecosystems of terrorist supporters online are constantly in flux: new actors and groups come to the fore or dissolve into oblivion on a regular basis.[123] In this realm, there are notable benefits to having the private sector to bear the primary responsibility for addressing new online actors. Companies are more closely able to monitor developments on their platforms in real time, corroborate the developments with real-world trends in violence and extremism, and quickly make decisions to designate a new group or actor.[124] In a world of increased regulations, if the U.S. government's designation process for "terrorist content" mirrored its process for designating terrorist organizations, the legislative process would be too slow and arduous to effectively respond to online terrorist content. [125]

Furthermore, the framing of this problem set used by the private sector (e.g., centered around "dangerous organizations" or "harmful content") is ideally suited for situations where terrorist content has similarities or overlaps with other types of actors and content that are subject to ToS enforcement. For instance, in many cases there can be significant overlaps between terrorist propaganda and coordinated inauthentic behavior, or crossovers between terrorist groups and conspiracy theory networks.[126] Within the bureaucratic structure of most major social media companies, when content or an actor straddles the line between two types of harmful content or two types of dangerous organization, one specific team is responsible and able to apply standards from both types onto the content in question and decide on the course of action.[127] Within the U.S. government, the responsibilities for different types of actors or different problem sets might be spread out across several agencies, meaning that decisions about borderline cases could be different depending on which U.S. government agency is responsible for determining standards for content removal.[128] In effective schemes for managing harmful content, consistency and agility are key. A government-managed process, due to the range of agencies and bureaucracies responsible for counterterrorism, may heighten the risk of inconsistency and lead to increased criticism of the effectiveness of terrorist content removal.[129]

None of these criticisms should preclude the U.S. government from being more involved in scrutinizing social media companies' standards or processes for terrorist content designations. If the government does want an increased role, however, they must first remedy the disparities present in their own counterterrorism infrastructure. Doing otherwise is placing the cart before the horse. Without a clear set of standards, methods, and bureaucracies within the U.S. government for managing a range of terrorist actors of different ideological persuasions—particularly U.S.-based violent extremists—there are no consistent or effective standards that can be used effectively to hold social media companies to account.

Conclusion

The debate about U.S. government regulation and anti-trust action of social media platforms is among the principal issues facing policymakers, practitioners, and scholars involved with technology policy. In recent years, a number of flashpoints have substantially increased the intensity and recurrence of calls for the U.S. government to take a broader role in managing how social media platforms run by private companies should govern content. Terrorist and extremist content is just one area of consideration within this ongoing debate, which is extremely unlikely to abate anytime soon. The issues of hate speech, mis- and disinformation, political polarization, targeted advertising, child sexual abuse material (CSAM), other content harmful to children, inauthentic behavior, and criminal content on social media are just a few of the many flashpoints for calls from the public for increased regulation.[130] Relatedly, there are a range of other means for policymakers, companies, and the American public to hold social media companies to account through indirect legal means or regulations on the market, architecture, and norms.[131]

As regulation pertains to terrorist content, however, there is insufficient evidence to suggest that stricter direct legal constraints issued by the U.S. government towards social media providers would have significantly improve the collective response to online terrorist and extremist content.[132] This paper found that enhanced direct legal regulation could threaten the collaborative efforts between major social media companies, driven by normative self-regulation, that have yielded the most fruit in improving the capacity of platforms to address terrorist content. Meanwhile, the development of government-private sector relations over the past decade have left the U.S. government in an extremely limited position to take responsibility for the management of terrorist content removal policies on social media. While the policies of social media companies in this endeavor are imperfect, they are nonetheless preferable to a situation where the U.S. government dictates standards to the private sector. The American federal government lacks the technical expertise, bureaucratic architecture, know-how, dexterity, authority, and ability to conduct terrorist content removal policy on social media with the same effectiveness as the private sector.

Advocates of increased American governmental regulation as the means to combat terrorist and extremist exploitation of the internet must recognize that the U.S. government in this space is effectively starting from "square zero." For better or for worse, the U.S. government has almost entirely outsourced its responsibility for content removal to the private sector, and there are few incentives for the government in taking back control. Until it develops similar capacities as the private sector to conduct this work, it could not, would not, and should not attempt to retake the driver's seat for determining social media companies' guidelines, policies, and ToS enforcement mechanisms against terrorist content.

References

[1] "Examining Social Media Companies' Efforts to Counter On-Line Terror Content and Misinformation." 2019. Hearing Before the Committee on Homeland Security, U.S. House of Representatives. June 26, 2019. https://www.govinfo.gov/content/pkg/CHRG-116hhrg38783/html/CHRG-116hhrg38783.htm
[2] Watts, Clint. 2018. "The Rise and Fall of the Virtual Caliphate." In Messing with the Enemy: Surviving in a Social Media World of Hackers, Terrorists, Russians, and Fake News, First edition. New York, NY: Harper Collins.
[3] Pandith, Farah. 2019. "Sheikh Google." In How We Win: How Cutting-Edge Entrepreneurs, Political Visionaries, Enlightened Business Leaders, and Social Media Mavens Can Defeat the Extremist Threat. New York: HarperCollins.
[4] "Examining Social Media Companies' Efforts to Counter On-Line Terror Content and Misinformation." 2019. Hearing Before the Committee on Homeland Security, U.S. House of Representatives. June 26, 2019. https://www.govinfo.gov/content/pkg/CHRG-116hhrg38783/html/CHRG-116hhrg38783.htm
[5] Berntsson, Jacob, and Maygane Janin. 2021. "Online Regulation of Terrorist and Harmful Content." Lawfare. October 14, 2021. https://www.lawfareblog.com/online-regulation-terrorist-and-harmful-content.
[6] Lessig, Lawrence. 1999. "What Things Regulate." In Code: And Other Laws Of Cyberspace. New York: Basic Books, 86-88
[7] *Ibid*., 88
[8] *Ibid*., 89
[9] *Ibid*., 89
[10] *Ibid*., 89
[11] *Ibid*., 89
[12] *Ibid*., 91
[13] *Ibid*., 93-95
[14] *Ibid*., 91-93
[15] *Ibid*., 93-94
[16] Hughes, Seamus. 2018. "Whose Responsibility Is It to Confront Terrorism Online?" Lawfare. April 27, 2018. https://www.lawfareblog.com/whose-responsibility-it-confront-terrorism-online.
[17] Tsesis, Alexander. 2017. "Social Media Accountability for Terrorist Propaganda Symposium: Terrorist Incitement on the Internet." Fordham Law Review 86 (2): 605–32.
[18] Softness, Nicole. 2017. "Terrorist Communications: Are Facebook, Twitter, and Google Responsible for the Islamic State's Actions?" SIPA Journal of International Affairs 70 (1). https://jia.sipa.columbia.edu/terrorist-communications.
[19] Bogage, Jacob. 2016. "Family of ISIS Paris Attack Victim Sues Google, Facebook and Twitter." Washington Post, June 16, 2016. https://www.washingtonpost.com/news/the-switch/wp/2016/06/16/family-of-isis-paris-attack-victim-sues-google-facebook-and-twitter/.
[20] Lomas, Natasha. 2018. "Twitter Claims More Progress on Squeezing Terrorist Content." TechCrunch, April 5, 2018. https://social.techcrunch.com/2018/04/05/twitter-transparency-report-12/.
[21] Macdonald, Stuart, Sara Giro Correia, and Amy-Louise Watkin. 2019. "Regulating Terrorist Content on Social Media: Automation and the Rule of Law." International Journal of Law in Context 15 (2): 183–97. https://doi.org/10.1017/S1744552319000119. 185-188
[22] *Ibid*., 189-194
[23] Garcia, Sandra E. 2018. "Ex-Content Moderator Sues Facebook, Saying Violent Images Caused Her PTSD." The New York Times, September 25, 2018, sec. Technology. https://www.nytimes.com/2018/09/25/technology/facebook-moderator-job-ptsd-lawsuit.html.
[24] Alexander, Audrey, and Bill Braniff. 2018. "Marginalizing Violent Extremism Online." Lawfare. January 21, 2018. https://www.lawfareblog.com/marginalizing-violent-extremism-online.
[25] Grygiel, Jennifer, and Nina Brown. 2019. "Are Social Media Companies Motivated to Be Good Corporate Citizens? Examination of the Connection between Corporate Social Responsibility and Social Media Safety." Telecommunications Policy 43 (5): 445–60. https://doi.org/10.1016/j.telpol.2018.12.003.
[26] *Ibid*.
[27] "The Call." 2019. Christchurch Call to Eliminate Terrorist and Violent Extremist Content Online. 2019. https://www.christchurchcall.com/christchurch-call.pdf.
[28] *Ibid*.
[29] "Statement by Press Secretary Jen Psaki on the Occasion of the United States Joining the Christchurch Call to Action to Eliminate Terrorist and Violent Extremist Content Online." 2021. The White House. May

7, 2021. https://www.whitehouse.gov/briefing-room/statements-releases/2021/05/07/statement-by-press-secretary-jen-psaki-on-the-occasion-of-the-united-states-joining-the-christchurch-call-to-action-to-eliminate-terrorist-and-violent-extremist-content-online/.

[30] Christchurch Call, "The Call."

[31]Keats Citron, Danielle. 2018. "Extremist Speech, Compelled Conformity, and Censorship Creep." Notre Dame Law Review 93 (3): 1035–79." 1055-1057

[32] *Ibid.*, 1047-1049

[33] Bernstsson and Janin, "Online Regulation of Terrorist and Harmful Content."

[34] *Ibid.*

[35] Christchurch Call, "The Call."

[36] *Ibid.*

[37] Bernstsson and Janin, "Online Regulation of Terrorist and Harmful Content."

[38] *Ibid.*

[39] Borelli, Marguerite. 2021. "Social Media Corporations as Actors of Counter-Terrorism." New Media & Society, August, 14614448211035120. https://doi.org/10.1177/14614448211035121.

[40] Samples, John. 2019. "Why the Government Should Not Regulate Content Moderation of Social Media." 865. Washington: CATO Institute. https://www.cato.org/policy-analysis/why-government-should-not-regulate-content-moderation-social-media.

[41] Hughes, "Whose Responsibility Is It To Confront Terrorism Online?"

[42] Macdonald, Stuart. 2018. "How Tech Companies Are Trying to Disrupt Terrorist Social Media Activity." Scientific American. June 26, 2018. https://www.scientificamerican.com/article/how-tech-companies-are-trying-to-disrupt-terrorist-social-media-activity/.

[43] Martelle, Michael, and Audrey Alexander. 2020. "Operation Glowing Symphony: The Missing Piece in the U.S. Online Counter-ISIS Campaign." In Cyber War & Cyber Peace in the Middle East: Digital Conflict in the Cradle of Civilization, edited by Michael Sexton and Eliza Campbell. Washington: Middle East Institute. https://www.mei.edu/publications/book/cyber-war-cyber-peace.

[44] Carlin, John P. 2016. "Remarks by Assistant Attorney General John Carlin: Opening of Madison Valleywood Project." February 24. https://epic.org/foia/MadisonValleywood_2.pdf.

[45] *Ibid.*

[46] *Ibid.*

[47] "Digital Counterterrorism: Fighting Jihadists Online." 2018. Washington, DC: Bipartisan Policy Center. https://bipartisanpolicy.org/wp-content/uploads/2019/03/BPC-National-Security-Digital-Counterterrorism.pdf.

[48] *Ibid.*

[49] Alexander, Audrey. 2017. "Digital Decay: Tracing Change Over Time Among English-Language Islamic State Sympathizers on Twitter." Program on Extremism. https://extremism.gwu.edu/sites/extremism.gwu.edu/files/DigitalDecayFinal_0.pdf: 47-48

[50] Bipartisan Policy Center, "Digital Counterterrorism."

[51] *Ibid.*

[52] *Ibid.*

[53] "About." n.d. Global Internet Forum to Counter Terrorism (GIFCT). Accessed December 20, 2020. https://gifct.org/about/.

[54] "Joint Tech Innovation." n.d. Global Internet Forum to Counter Terrorism (GIFCT). Accessed December 20, 2020. https://gifct.org/joint-tech-innovation/.

[55] "Transparency." n.d. Global Internet Forum to Counter Terrorism (GIFCT). Accessed December 20, 2020. https://gifct.org/transparency/.

[56] "Membership." n.d. Global Internet Forum to Counter Terrorism (GIFCT). Accessed December 20, 2020. https://gifct.org/membership/.

[57] *Ibid.*

[58] Hughes, "Whose Responsibility is it to Confront Terrorism Online?"

[59] *Ibid.*

[60] *Ibid.*

[61] Hughes, "Whose Responsibility Is It To Confront Terrorism Online?"

[62] BBC News. 2018. "Germany Starts Enforcing Hate Speech Law," January 1, 2018, sec. Technology. https://www.bbc.com/news/technology-42510868.

[63] BBC News. 2020. "France Gives Online Firms One Hour to Pull 'terrorist' Content," May 14, 2020, sec. Technology. https://www.bbc.com/news/technology-52664609.

[64] "Counter-Terrorism and Border Security Act 2019." 2019. Queen's Printer of Acts of Parliament. 2019. https://www.legislation.gov.uk/ukpga/2019/3/pdfs/ukpga_20190003_en.pdf.

[65] Hughes, “Whose Responsibility Is It To Confront Terrorism Online?”
[66]
[67] “The Online Regulation Series: The Handbook.” 2021. Tech Against Terrorism. https://www.techagainstterrorism.org/research/?__cf_chl_tk=3V71XEKRhE6pmuERqtXSaZCrLfPr.VxnP0tqDwpsVQg-1636477839-0-gaNycGzNCKU.
[68] *Ibid.*
[69] *Ibid.*
[70] *Ibid.*
[71] *Ibid.*
[72] *Ibid.*
[73] *Ibid.*
[74] *Ibid.*
[75] *Ibid.*
[76] See for instance, UK Home Office, “New Technology Revealed”
[77] Tech Against Terrorism, “The Online Regulation Series”
[78] Temperton, James. 2018. “Isis Could Easily Dodge the UK’s AI-Powered Propaganda Blockade.” Wired UK, February 13, 2018. https://www.wired.co.uk/article/isis-propaganda-home-office-algorithm-asi.
[79] Tech Against Terrorism, “The Online Regulation Series”
[80] Keats Citron, Danielle. 2018. “Extremist Speech, Compelled Conformity, and Censorship Creep.” Notre Dame Law Review 93 (3): 1035–79.
[81] Fishman, Brian. 2019. “Crossroads: Counter-Terrorism and the Internet.” Texas National Security Review 2 (2). https://tnsr.org/2019/02/crossroads-counter-terrorism-and-the-internet/.
[82] *Ibid.*
[83] *Ibid.*
[84] *Ibid.*
[85] Alexander, “Digital Decay”; Alexander and Braniff, “Marginalizing Violent Extremism Online,” Berntsson and Janin, “Online Regulation of Terrorist and Harmful Content”; Conway, Maura, Moign Khawaja, Suraj Lakhani, Jeremy Reffin, Andrew Robertson, and David Weir. 2019. “Disrupting Daesh: Measuring Takedown of Online Terrorist Material and Its Impacts.” Studies in Conflict & Terrorism 42 (1–2): 141–60. https://doi.org/10.1080/1057610X.2018.1513984.
[86] “Analysis: ISIS Use of Smaller Platforms and the DWeb to Share Terrorist Content.” 2019. Tech Against Terrorism. April 29, 2019. https://www.techagainstterrorism.org/2019/04/29/analysis-isis-use-of-smaller-platforms-and-the-dweb-to-share-terrorist-content-april-2019/, https://www.techagainstterrorism.org/2019/04/29/analysis-isis-use-of-smaller-platforms-and-the-dweb-to-share-terrorist-content-april-2019/.
[87] Clifford, Bennett, and Helen Powell. 2019. “Encrypted Extremism: Inside the English-Speaking Islamic State Ecosystem on Telegram.” Washington: George Washington University Program on Extremism. https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/EncryptedExtremism.pdf.
[88] Tech Against Terrorism, “Analysis: ISIS Use of Smaller Platforms;” Conway, Maura, Ryan Scrivens, and Logan Macnair. 2019. “Right-Wing Extremists’ Persistent Online Presence: History and Contemporary Trends.” The Hague, Netherlands: International Centre for Counter-terrorism. https://icct.nl/wp-content/uploads/2019/11/Right-Wing-Extremists-Persistent-Online-Presence.pdf; Hadley, Adam. 2021. “Terrorists Will Move to Where They Can’t Be Moderated.” Wired UK, May 31, 2021. https://www.wired.co.uk/article/terrorists-dweb.
[89] Alexander, “Digital Decay”
[90] Meserole, Chris, and Daniel Byman. 2019. “Terrorist Definitions and Designations Lists: What Technology Companies Need to Know.” Global Research Network on Terrorism and Technology Paper 7. London: RUSI. https://www.brookings.edu/wp-content/uploads/2019/07/GRNTT-Paper-No.-7.pdf.
[91] Alexander, “Digital Decay”; Alexander and Braniff, “Marginalizing Violent Extremists Online”; Clifford, Bennett. 2020. “Migration Moments: Extremist Adoption of Text-Based Instant Messaging Applications.” Global Network on Extremism and Technology. November 2020. https://gnet-research.org/wp-content/uploads/2020/11/GNET-Report-Migration-Moments-Extremist-Adoption-of-Text%E2%80%91Based-Instant-Messaging-Applications_V2.pdf
[92] Alexander, “Digital Decay.”
[93] Clifford, Bennett, and Helen Powell. 2019. “De-Platforming and the Online Extremist’s Dilemma.” Lawfare. June 6, 2019. https://www.lawfareblog.com/de-platforming-and-online-extremists-dilemma.
[94] Conway et. al., “Disrupting Daesh”
[95] *Ibid.*
[96] Clifford and Powell, “De-Platforming and the Online Extremist’s Dilemma”

[97] Tech Against Terrorism, "The Online Regulation Series"
[98] Fishwick, Carmen. 2014. "How a Polish Student's Website Became an Isis Propaganda Tool." The Guardian, August 15, 2014, sec. World news. https://www.theguardian.com/world/2014/aug/15/-sp-polish-man-website-isis-propaganda-tool.
[99] *Ibid.*
[100] *Ibid.*
[101] Clifford and Powell, "Encrpyted Extremism"
[102] *Ibid.*
[103] *Ibid.*; Gordon, Steven, and Sarah Ashraf. 2021. "Are Telegram and Signal Havens for Right-Wing Extremists?" Foreign Policy. March 13, 2021. https://foreignpolicy.com/2021/03/13/telegram-signal-apps-right-wing-extremism-islamic-state-terrorism-violence-europol-encrypted/.
[104] Volpicelli, Gian M. 2021. "Telegram Is Becoming a Cesspool of Anti-Semitic Content." Wired, October 13, 2021. https://www.wired.co.uk/article/telegram-antisemitism-hopenothate.
[105] King, Peter. 2019. "Islamic State Group's Experiments with the Decentralized Web." Europol. https://www.europol.europa.eu/publications-documents/islamic-state-group%E2%80%99s-experiments-decentralised-web; Bodo, Lorand. 2018. "Decentralised Terrorism: The Next Big Step for the so-Called Islamic State (IS)?" VOX - Pol. December 12, 2018. https://www.voxpol.eu/decentralised-terrorism-the-next-big-step-for-the-so-called-islamic-state-is/.
[106] Macdonald et al., "Regulating Terrorist Content on Social Media"
[107] Amarasingam, Amarnath. 2020. "A View from the CT Foxhole: An Interview with an Official at Europol's EU Internet Referral Unit." CTC Sentinel 13 (2). https://ctc.usma.edu/view-ct-foxhole-interview-official-europols-eu-internet-referral-unit/.
[108] See for instance, Torba, Andrew. 2019. "Gab's Policies, Positions, and Procedures for Unlawful Content And Activity On Our Social Network." Gab News (blog). August 23, 2019. https://news.gab.com/2019/08/23/gabs-policies-positions-and-procedures-for-unlawful-content-and-activity-on-our-social-network/.
[109] Veilleux-Lepage, Yannick. 2016. "Paradigmatic Shifts in Jihadism in Cyberspace: The Emerging Role of Unaffiliated Sympathizers in Islamic State's Social Media Strategy." Journal of Terrorism Research 7 (February).
[110] Alexander and Braniff, "Marginalizing Violent Extremists Online."
[111] Conway, Maura, Ryan Scrivens, and Logan Macnair. 2019. "Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends." The Hague, Netherlands: International Centre for Counter-terrorism. https://icct.nl/wp-content/uploads/2019/11/Right-Wing-Extremists-Persistent-Online-Presence.pdf; Rudner, Martin. 2017. "'Electronic Jihad': The Internet as Al Qaeda's Catalyst for Global Terror." Studies in Conflict & Terrorism 40 (1): 10–23. https://doi.org/10.1080/1057610X.2016.1157403.
[112] Hughes, "Whose Responsibility Is It to Confront Terrorism Online?"
[113] Díaz, Ángel, and Laura Hecht-Felella. 2021. "Double Standards in Social Media Content Moderation." New York: Brennan Center for Justice. https://www.brennancenter.org/sites/default/files/2021-08/Double_Standards_Content_Moderation.pdf; Meier, Anna. 2019. "Why Do Facebook and Twitter's Anti-Extremist Guidelines Allow Right-Wingers More Freedom than Islamists?" Washington Post, August 1, 2019. https://www.washingtonpost.com/politics/2019/08/01/why-do-facebook-twitters-anti-extremist-guidelines-allow-right-wingers-more-freedom-than-islamists/.
[114] *Ibid.* all.
[115] *Ibid.* all.
[116] Alexander, Audrey, and Kristina Hummel. 2021. "A View from the CT Foxhole: Mary McCord, Executive Director, Institute for Constitutional Advocacy and Protection, Georgetown University Law Center." CTC Sentinel 14 (3). https://ctc.usma.edu/a-view-from-the-ct-foxhole-mary-mccord-executive-director-institute-for-constitutional-advocacy-and-protection-georgetown-university-law-center/.
[117] "Dangerous Individuals and Organizations." n.d. Facebook Community Standards. Accessed December 20, 2020. https://www.facebook.com/communitystandards/recentupdates/dangerous_individuals_organizations.
[118] *Ibid.*
[119] Lewis, Jon, Seamus Hughes, Ryan Greer, and Oren Segal. 2020. "White Supremacist Terror: Modernizing Our Approach to Today's Threat." Washington: Program on Extremism-Anti Defamation League Joint Report. https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/White%20Supremacist%20Terror%20final.pdf.
[120] Meserole and Byman, "Terrorist Definitions and Designations Lists"
[121] *Ibid.*

[122] Blazakis, Jason. 2018. “American Terrorists: Why Current Laws Are Inadequate for Violent Extremists at Home.” Lawfare. December 2, 2018. https://www.lawfareblog.com/american-terrorists-why-current-laws-are-inadequate-violent-extremists-home.
[123] Alexander and Braniff, “Marginalizing Violent Extremists Online.”
[124] Meserole and Byman, “Terrorist Definitions and Designations Lists”
[125] Samples, “Why the Government Should Not Regulate Content Moderation of Social Media.”
[126] Fishman, “Crossroads”
[127] *Ibid.*
[128] *Ibid.*
[129] *Ibid.*
[130] Douek, Evelyn. 2020. “The Rise of Content Cartels.” Columbia University: Knight First Amendment Institute. https://knightcolumbia.org/content/the-rise-of-content-cartels.
[131] Ibid.
[132] Keller, Daphne. 2019. “Three Constitutional Thickets: Why Regulating Online Violent Extremism is Hard.” Program on Extremism Legal Perspectives on Tech Paper, September 2019. https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/Three%20Constitutional%20Thickets.pdf