Digital Aftershocks:

Online Mobilization and Violence in the United States

MARIANA OLAIZOLA ROSENBLAT AND LUKE BARNES

NYU STERN
Center for Business and Human Rights

HFG HARRY FRANK GUGGENHEIM FOUNDATION

Carnegie CORPORATION OF NEW YORK

October 2025

Contents

Disclaimer: This report contains direct quotations and references to content, some of which may include offensive, hateful, or violent language. These materials are presented solely for research and analysis purposes and do not reflect the views of the authors or publishers. Reader discretion is advised.

We are grateful to Carnegie Corporation of New York and the Harry Frank Guggenheim Foundation for their generous philanthropic support for this report.

Authors

Mariana Olaizola Rosenblat is a policy advisor on technology and law at the NYU Stern Center for Business and Human Rights.

Luke Barnes is a senior research scientist at the NYU Stern Center for Business and Human Rights.

Nicole Gaouette is the Center's consulting editor and provided editorial assistance for this report.

Executive Summary

Political violence in the United States has increased in recent years and shows no signs of declining.[1] This trend was underscored in September 2025 by the assassination of conservative activist Charlie Kirk at Utah Valley University. In the two weeks before and after Kirk's killing, shooting incidents in Colorado, Minneapolis, and Dallas seized public attention.[2]

Amid growing concern about the relationship between online rhetoric and real-world violence, this report examines how violent extremist actors across the ideological spectrum use digital platforms to respond to, amplify, and exploit acts of political violence in the United States. Drawing on open-source intelligence (OSINT) gathered initially from March 24 to June 6, 2025, and then extended to include a period following Kirk's assassination, this analysis reveals sophisticated cross-platform strategies employed by far-right, far-left, violent Islamist, and nihilistic violent extremist (NVE) actors.

This report uses "violent extremist" to refer to individuals who support or commit ideologically motivated violence to further political goals, as well as those who commit violence driven by generalized hatred rather than a coherent ideology.

Key Findings

- Violent extremist groups systematically exploit trigger events—high-profile incidents of violence—to recruit supporters, justify their ideologies, and call for retaliatory action.
- These groups employ multi-platform strategies, using mainstream sites like X for visibility and recruitment while maintaining a presence on private or semi-private platforms for coordination and more extreme content.
- Far-right groups capitalized on cases like the Austin Metcalf stabbing and the Iryna Zarutska killing to advance narratives of White victimhood and justify threats against perceived enemies.
- Activities of both far-left and far-right networks revealed a troubling convergence around antisemitic targeting.
- Violent Islamic groups are more aggressively monitored than domestic groups espousing similar levels of violence.
- Violent Islamist groups, facing stricter moderation than domestic extremists, have migrated to decentralized platforms like Rocket.Chat while disseminating symbolic propaganda elsewhere.
- Nihilistic Violent Extremist (NVE) communities glorify violence across ideological lines for shock value and digital notoriety, making their threats harder to predict based on political triggers.

This report aims to bring clarity to a conversation clouded by vagueness and partisanship. It first maps the domestic threat landscape, offering timely examples of online violent discourse from across the ideological spectrum targeting US individuals or institutions, and sets out a clear definitional framework for types of speech that carry legal significance under US constitutional doctrine. It closes with practical recommendations for online service providers and policymakers.

Recommendations In Brief

For Online Platforms

1. **Adopt precise policies on threats and incitement and demonstrate willingness and capacity to enforce those policies.** Clearly define prohibited conduct involving threats of violence and incitement, and report publicly on their enforcement actions and outcomes.

2. **Implement user-friendly reporting tools compatible with encryption.** Any platform enabling user communication should allow users to flag illegal conduct and content they believe violates platform policies. Those reports should be examined swiftly and escalated as appropriate.

3. **Use metadata responsibly to disrupt networks.** When collecting metadata to detect abusive behavior, limit collection to what is necessary for specific safety purposes, be transparent about its use, and delete it after a set period.

4. **Cooperate with other services to monitor and combat dangerous cross-platform activity.** Participate actively in cross-industry initiatives to identify migration patterns and disrupt attempts by dangerous actors to exploit harder-to-monitor encrypted environments.

For US Legislators and Policymakers

5. **Recognize the limits of legal remedies.** Distinguish between harmful speech that is lawful and speech that is illegal under the First Amendment when setting out platform obligations.

6. **Clarify protocols for platform-law enforcement cooperation.** Establish clear standards for when and how platforms should share information related to threats or incitement with law enforcement.

7. **Revisit extremist and terrorist designation frameworks.** Re-examine the criteria used to designate terrorist organizations and apply them consistently across ideologies.

8. **Mandate transparency, design, and procedural standards without undermining encryption.** Require platforms to publish transparency reports that explain their abuse-detection and reporting goals, processes, and outcomes.

9. **Support research on effective counter-speech initiatives.** Explore partnerships with civil society to counteract violent narratives through counter-speech campaigns.

1. Introduction

On September 10, 2025, Charlie Kirk, the conservative influencer and founder of Turning Point USA, was assassinated at Utah Valley University. According to charging documents, the suspect, 22-year-old Tyler Robinson, was motivated by the belief that Kirk was "spreading hate."[3] Friends report that Robinson was highly active on Discord, a communication app, and other video game platforms, and etchings on the murder weapon's bullets—containing popular video game phrases—suggested he sought the fame of digital virality.[4]

> **"Recent attacks demonstrate a resurgent threat of political violence in the United States emerging from multiple ideological sources: the far-left, the far-right, and actors whose motivations defy neat ideological categories."**

Kirk's murder was part of a series of politically motivated acts of violence that struck the United States in the first nine months of 2025. On May 21, two young Israeli embassy staffers were gunned down outside the Capital Jewish Museum in Washington, D.C. The assailant posted a manifesto shortly afterward on X titled "*Escalate for Gaza, Bring the War Home*,"[5] echoing language from radical forums and online influencers who for months had encouraged political violence under the guise of resistance. In August, a Catholic school in Minneapolis was attacked by an assailant believed to be motivated by niche digital ideologies, resulting in two children's deaths.[6] In each case, content produced by or celebrating the attackers circulated widely online in the immediate aftermath.

These attacks demonstrate a resurgent threat of political violence in the United States emerging from multiple ideological sources: the far-left, the far-right, and actors whose motivations defy neat ideological categories, as potentially illustrated by the Charlie Kirk assassination. Yet as the threat landscape grows more volatile, responses remain fragmented. Public polling suggests declining support for efforts to remove "false or violent" content from online platforms.[7] Meanwhile, US government infrastructure meant to counter violent extremism is being dismantled or deprioritized.[8]

Calls for stronger platform action against violent online rhetoric have intensified, yet these demands often lack specificity about what content should be addressed, how platforms should respond, and what evidence demonstrates effectiveness. The debate has become highly polarized, with accusations frequently aimed at political opponents rather than grounded in systematic analysis.

The Purpose and Scope of This Report

This report aims to bring clarity to this fragmented and often politicized conversation by examining how violent extremist actors across the ideological spectrum use online platforms in relation to acts of violence in the United States. This report uses "violent extremist" to refer to individuals who support or commit ideologically-motivated violence to further political goals, as well as those who commit violence driven by generalized hatred rather than a coherent ideology.[9] Rather than claiming to prove that online discourse directly causes offline violence—a causal chain difficult to establish with the available methodology—this analysis documents how extremist networks exploit violence to advance their ideologies, recruit supporters, and create climates that may facilitate future radicalization and violence.

Drawing on open-source intelligence gathering, the report maps the network of accounts behind current violence-promoting online campaigns, as well as the digital infrastructure that enables them, which spans open platforms, encrypted messaging apps, and semi-encrypted or hybrid spaces. It examines their responses to trigger events, their cross-platform strategies for maximizing reach while evading moderation, and the rhetorical patterns they employ to mobilize sympathizers.

The analysis then provides practical recommendations for platforms and policymakers. These recommendations focus on disrupting propaganda networks and improving content moderation while respecting fundamental rights, particularly, freedom of expression. They acknowledge both the documented patterns revealed by this research and the limitations inherent in studying online extremism using open-source methodologies.

Unpacking "Violent Online Discourse"

Violent online discourse in this report refers to speech that either threatens violence or significantly increases the likelihood of violence. Most jurisdictions worldwide, including the US, prohibit direct threats and calls for violence. While criminalization thresholds differ across jurisdictions, the near-universal condemnation of purposeful threats and incitement justifies focused attention to these types of online content.

- **Threats** refer to statements or actions that convey an intention by the speaker to cause harm to a specific person, group, or institution, with the aim of terrorizing or intimidating them. Under US constitutional doctrine, a threat crosses the threshold of illegality if the speaker was reckless in expressing such an intention to cause harm.[1] Some US states also penalize "terroristic threats," which typically involve threats to commit a violent crime or cause serious harm with the intent to intimidate a civilian population, a government, or to cause major public disruption.[2]

  Threats can be explicit, such as when the speaker specifies both the intended action and the subject of the threat, or they can be indirect or implicit, taking the form of intimidation. Some instances of doxxing—the publication of the intended target's physical location—qualify as intimidation.[3]

- **Incitement** refers to speech that encourages others to commit violence or other unlawful acts. Incitement ranges from propagandistic speech justifying violence generally, to specific calls for imminent violent acts that are actually likely to transpire. Under US constitutional doctrine, only the latter are considered illegal.[4]

  This report adopts a broader analytical framework that aligns with international human rights standards and academic research on mass violence. International human rights law and academic frameworks take a more expansive view of incitement, noting that even speech that does not call for imminent lawlessness can still significantly raise the likelihood of mass violence.[5] The UN Rabat Plan of Action, for example, outlines a six-part threshold test for identifying incitement to hatred or violence, which includes analysis of the sociopolitical context, the speaker's status or position, and the form or style of the speech, in addition to the content itself.[6]

  Similarly, Susan Benesch's "dangerous speech" framework identifies markers of speech that significantly raise the likelihood of atrocities, drawing on extensive historical evidence of discourse preceding mass violence.[7] For example, statements dehumanizing a racial minority, when repeated over time in a context primed for racial conflict, might qualify as "dangerous speech," even without explicitly calling for imminent harm. Under this framework, some instances of "hate speech"—statements that denigrate or attack people based on shared characteristics such as race or religion[8]—could constitute incitement under Benesch's framework, depending on the factors that make them likely to catalyze violence.

The distinction between threats and incitement, while present in US constitutional jurisprudence, is often not clear-cut in practice. Many statements or pieces of content, including several examples referred to in this report, serve to terrorize specific individuals or groups and mobilize others toward violence—either against that specific target or others seen as similar, or both.[9] Moreover, categorizing violent online discourse under any framework requires contextual analysis. In the online realm, where speakers are often anonymous and audiences diffuse, contextual knowledge of threat actors, their rhetorical tendencies, and usage patterns, is key to discerning intent and estimating likelihood of harm. This is particularly challenging in closed platforms, where "in-group" language often leans on irony and memes, creating ambiguity about whether a post is an actionable command or an example of "edgelording"—deliberately using controversial, shocking, or taboo language to garner digital attention.[10]

---

[1] See *Virginia v. Black*, 538 U.S. 343 at 359 (True threats are "those statements where the speaker means to communicate a serious expression of an intent to commit an act of unlawful violence to a particular individual or group of individuals."). In *Counterman v. Colorado*, 600 U.S. 66 (2023), the Supreme Court held that, to convict a person of making true threats, a state must show that the speaker "'consciously disregard[ed] a substantial [and unjustifiable] risk that the conduct [would] cause harm to another.'" The Court explained the *mens rea* or mental state of recklessness would suffice for this showing.

[2] See e.g., Nebraska Revised Statute 28-311.01; Georgia Code § 16-11-37 (2024); Mississippi Code § 97-7-75 (2024).

[3] David L. Hudson, Is Doxxing Illegal? The Foundation for Individual Rights and Expression, February 28, 2024.

[4] Illegal incitement under US constitutional law, refers to speech that is intended to incite or produce *imminent* lawless action, and is *likely* to incite or produce such action. "Imminent" means that the illegal action would happen immediately or soon after the speech, and "likely" means that there is a real and substantial probability that the speech would result in the illegal action. See *Brandenburg v. Ohio*, 395 U.S. 444 (1969).

[5] Some US constitutional scholars, e.g., Alexander Tsesis, also argue that the imminence test is too narrow for the online context, where the audience is diffuse and the timing of encountering a piece of content which then inspired violence is uncertain. See Tsesis, "Inflammatory Speech: Offense Versus Incitement," supra, at 1170.

[6] Following Article 20 of the ICCPR, the Rabat Plan of Action focuses not only on identifying and preventing incitement to violence but also to discrimination and hostility. See United Nations Office of the High Commissioner for Human Rights, Rabat Plan of Action on the Prohibition of Advocacy of National, Racial or Religious Hatred That Constitutes Incitement to Discrimination, Hostility or Violence, October 2012.

[7] Susan Benesch's "dangerous speech" framework is one of the most influential typologies for distinguishing forms of harmful speech that may lead to violence. Susan Benesch, *Dangerous Speech: A Proposal to Prevent Group Violence*, World Policy Institute, 2012.

[8] Susan Benesch, Proposals for Improved Regulation of Harmful Online Content, Dangerous Speech Project, June 19, 2020.

[9] See Tsesis, "Inflammatory Speech: Offense Versus Incitement," supra, at 1173-74.

[10] Cambridge Dictionary, Edgelord, Cambridge University Press, accessed October 2, 2025.

The relationship between online rhetoric and violence is complex and bidirectional. Violent incidents trigger online activity; online networks exploit that violence for recruitment; the resulting propaganda may influence future actors; and the cycle continues. This report documents primarily the middle portion of that cycle: how extremist networks respond to and amplify violence after it occurs.

This focus reflects both the methodology's capabilities and its constraints. By monitoring semi-public channels on platforms like Telegram and X, analysts could observe propaganda, recruitment tactics, and explicit threats. What they could not observe—due to ethical constraints on infiltrating closed groups—was operational planning that may occur in truly private spaces. This gap is significant and shapes the report's conclusions.

In an era of deepening polarization, one imperative should transcend partisan divides: the urgent need to understand and address the rising tide of political violence in American society. To the extent that online activity plays a role in such violence—whether by causing it, celebrating it, or exploiting it for radicalization—relevant stakeholders should seek to understand that dynamic and explore practical responses. This report aims to provide impetus for that urgent effort.

"In an era of deepening polarization, one imperative should transcend partisan divides: the urgent need to understand and address the rising tide of political violence in American society."

2. Contemporary Threat Landscape: A Snapshot

During the reporting period (March 24 to June 6, 2025, expanded to include September 10 to 24), open-source intelligence analysts from Tech Against Terrorism tracked threats and calls for violence against US persons and institutions disseminated by actors across the ideological spectrum.

> **Extremist groups strategically select platforms based on their tactical needs, using mainstream sites for recruitment and visibility while maintaining a presence on encrypted or decentralized platforms for coordination.**

The monitoring initially focused on encrypted or semi-encrypted messaging platforms—Telegram, WhatsApp, Viber, and Signal—based on their presumed utility for operational planning. However, fully encrypted platforms like Signal and WhatsApp yielded minimal data due to access limitations inherent in the methodology. Instead, the most substantial findings emerged from semi-public platforms such as Telegram, where extremist actors balance reach with reduced moderation, and from decentralized platforms such as Rocket.Chat, where Islamic groups have established persistent infrastructure.

This pattern itself represents a significant finding: extremist groups strategically select platforms based on their tactical needs, using mainstream sites for recruitment and visibility while maintaining a presence on encrypted or decentralized platforms for coordination. The following platform ecosystem reflects where observable extremist activity concentrated during the monitoring period.

Monitoring across these platforms revealed distinct patterns of activity by ideological category. Far-right actors operated relatively openly on Telegram and X, capitalizing on trigger events to advance narratives of White victimhood. Far-left networks, dominated by pro-Palestine activism during the monitoring period, similarly used relatively public channels for propaganda and doxxing. Violent Islamist groups, facing stricter moderation, relied more heavily on decentralized platforms and out-linking strategies. NVE communities reportedly gravitate toward semi-private platforms like Discord, though access limitations prevented comprehensive documentation of this activity during the monitoring period. The following sections detail these patterns.

Far-right

During the monitoring period, analysts documented a steady escalation of online intimidation and threats originating from far-right actors—individuals and groups promoting violence against perceived enemies such as immigrants, women, LGBTQ+ people, and other minorities. The far-right ecosystem is ideologically diffuse but includes three prominent subgroups: White supremacists advancing theories of racial superiority, neo-Nazis advocating explicitly fascist and antisemitic ideology, and accelerationists seeking to hasten the collapse of liberal democracy through violent disruption.

Key Platforms for Violent Mobilization during the Monitoring Period

Messaging applications

Element: An open-source, end-to-end encrypted messaging app commonly exploited by violent Islamist entities. Built as a decentralized structure, it allows users to host their own servers and maintain control over their data.

GemSpace: A messaging and community platform designed for private, secure group communication, often exploited by violent Islamist actors.

SimpleX: A messaging platform that operates without user identifiers (no phone numbers, emails, or usernames are required). Communication is facilitated via invitation links or QR codes, and the platform uses end-to-end encryption by default.

Discord: A social networking platform, initially designed for people playing online video games and composed primarily of private (although not end-to-end encrypted) chatrooms, called "servers." Users join these private servers via invite links and are often submitted to vetting processes by the server's owner.

Alternative (alt-tech) platforms

Rocket.Chat: An open-source, decentralized communication platform in which users can host and manage their own "instances," or servers. Its support for encrypted messaging, file sharing, and private group formation, coupled with limited external monitoring, have made it useful for operational coordination and ideological reinforcement among violent Islamist groups, such as the Islamic State, which has established a hub on an instance called "TechHaven."

ChirpWire: An encrypted social media platform known for minimal content moderation and user anonymity. It does not require email or phone verification, allowing users to register with only a username and password.

4chan: An anonymous imageboard forum known for its minimal moderation and time-limited content.[1] While not inherently extremist, its /pol/ (politically incorrect) board has been a consistent launch site for narratives and memes that amplify and celebrate far-right beliefs, propagating this ideology to a wide audience.[2]

Video-sharing platforms

Odysee: a video-hosting platform that promotes itself as a "free-speech" alternative to YouTube.

TikTok: a Chinese-based short-form video platform with a vast global user base and highly effective algorithmic content delivery. While primarily entertainment-focused, it has been exploited by extremist and conspiracy groups to push subtle messaging, reach younger audiences, and capitalize on viral trends.

YouTube: the largest video-sharing platform globally. While it has strong moderation policies, violent actors use coded language or migrate audiences from YouTube to less regulated platforms when content is removed.

Social media platforms

X: a platform known for real-time news sharing and public discourse, with strong virality-promoting features. A decrease in attention to content moderation has led to a more permissive environment for dissemination of violent narratives disguised as political commentary.

Telegram: a hybrid messaging and social media platform with minimal moderation and limited end-to-end encrypted features—i.e., users must explicitly choose to create a one-on-one "secret chat" in order for content to be end-to-end encrypted.

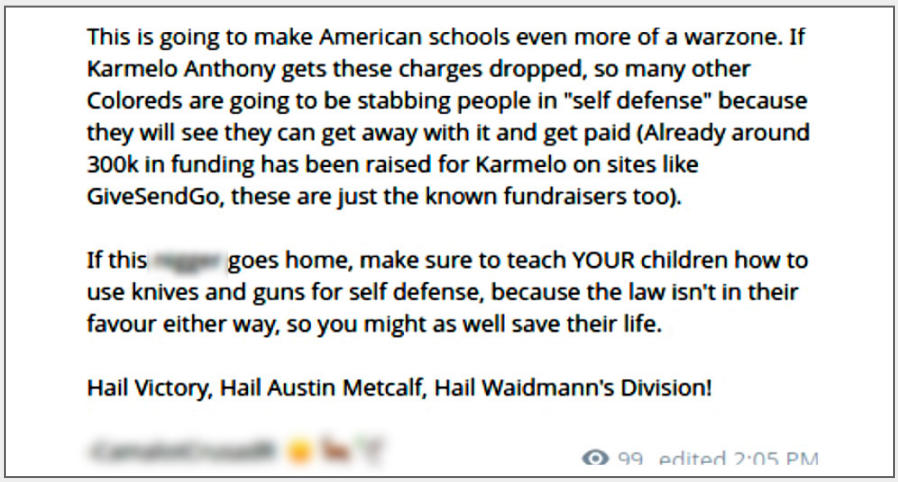
Instagram: a visual-first social media platform primarily used for sharing images and videos. Despite increased moderation, coded imagery and symbolic messaging suggesting violence continue to circulate, often linking out to less-regulated platforms for deeper engagement.

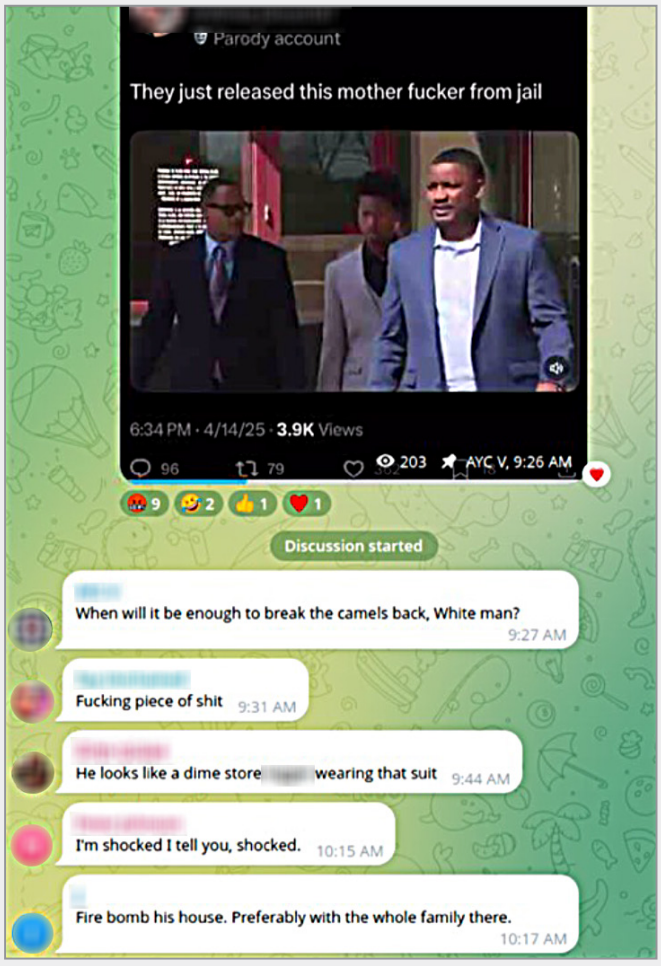
[1] 4chan only has space for a set number of threads on each board, so as newer threads are created older ones are pushed off the board and are eventually deleted, or in 4chan slang "404'd" – which refers to the HTTP 404 error ("page not found").

[2] Gabriel Emile Hine et al., Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web, Cornell University, October 11, 2016.

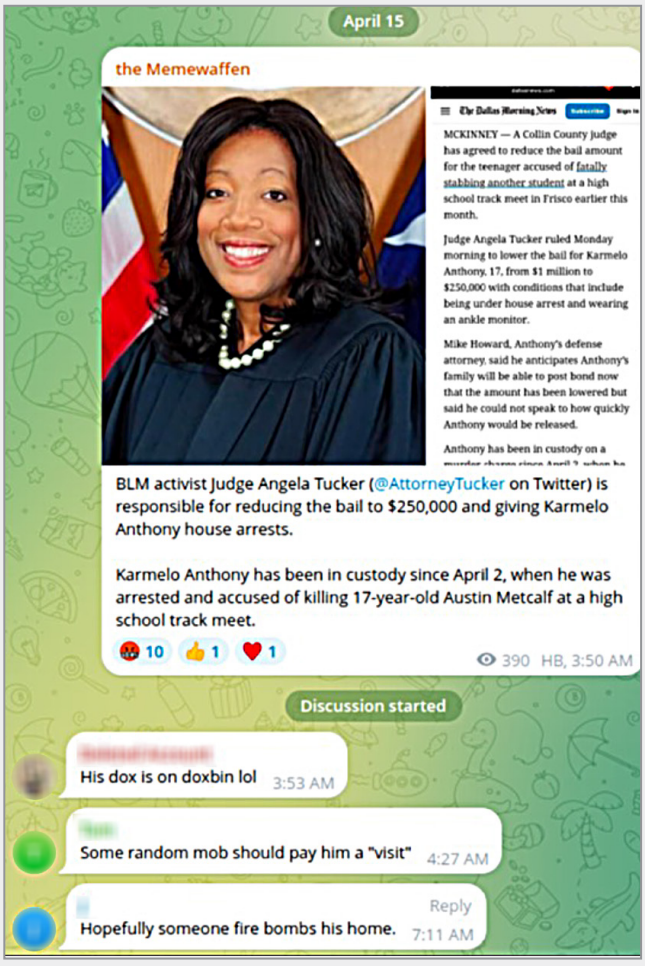
Sample of Far-right Posts During the Monitoring Period

Screenshot of a post on the White nationalist "Waidmann's Division Hub" Telegram channel using the Metcalf case to justify calling on White parents to teach their children how to use knives and guns to defend themselves against Black Americans. Posted on April 14, 2025.

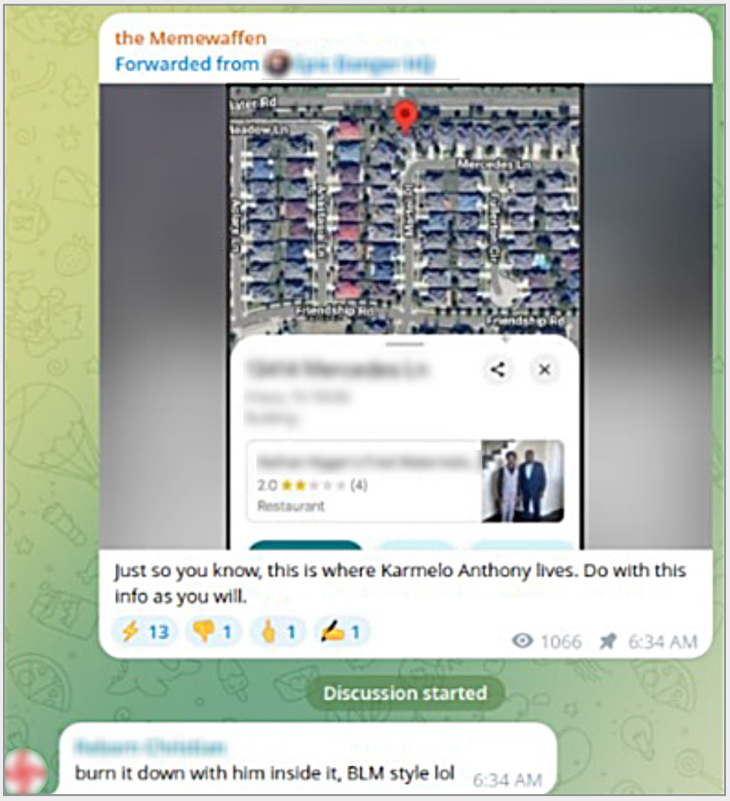


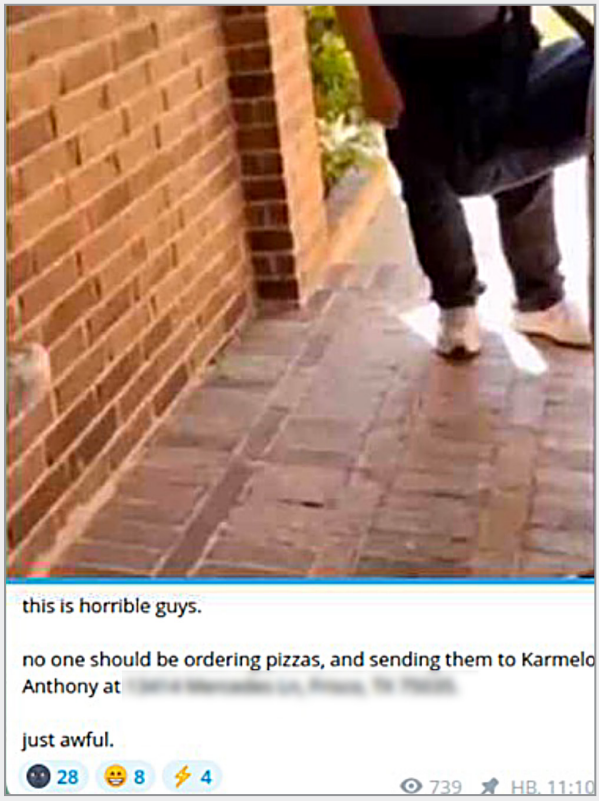


Screenshot of a post on "The Memewaffen" Telegram channel depicting Anthony being released from custody with threatening comments such as "Firebomb his house. Preferably with the whole family there." Posted on April 15, 2025.

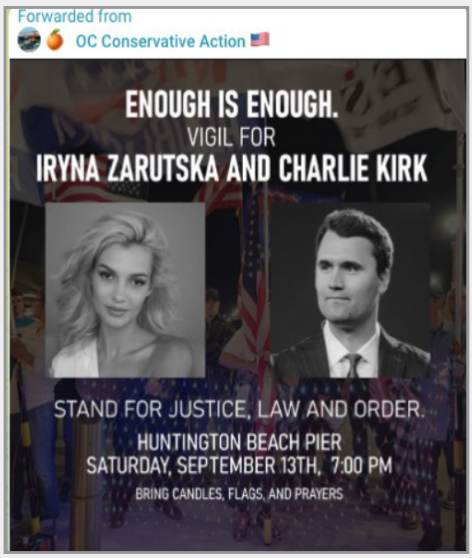


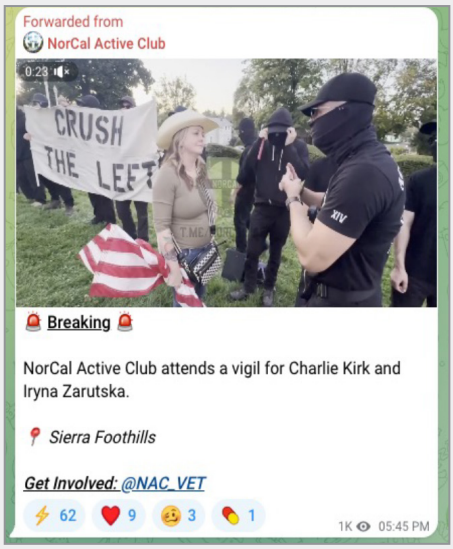
Screenshot of a post on Telegram depicting the presiding judge alongside threatening commentary about Anthony, illustrating how far-right actors expanded their targeting beyond the perpetrator to include officials involved in legal proceedings. Posted on April 15, 2025.





This series of posts from April 18-23, 2025, shows sustained doxxing activity across platforms, with Anthony's address appearing on both X and Telegram. Analysts documented the posts being used for recruitment messaging (urging supporters to join "Active Clubs"); ideological reinforcement (framing it as proof of racial conflict); and as a touchstone for discussing subsequent incidents.





Left to right: A September 11 post from a far-right Telegram account calling for a vigil in Huntington Beach, CA, for the murder of Charlie Kirk and Iryna Zarutska; a picture of the vigil on Saturday 13, including members of the extremist group Patriot Front (masked, bottom left); a September 14 vigil by the far-right NorCal Active club.

> **Monitoring across platforms revealed distinct patterns of activity by ideological category.**

Analysts observed several hundred posts containing threats or calls for violence from far-right accounts, primarily concentrated on Telegram and X. These posts exhibited common tactical patterns: exploitation of trigger events involving interracial violence; doxxing of perpetrators, their families, and officials involved in legal proceedings; justification of threats as "self-defense" against systemic abandonment; and cross-platform amplification to maximize reach.

The fatal stabbing of Austin Metcalf, a 17-year-old White student, by Karmelo Anthony, a 17-year-old Black student, during an April 2, 2025 track meet in Frisco, Texas, demonstrates the far-right exploitation cycle in detail.

Within hours of the incident, far-right Telegram channels began framing the stabbing as evidence of systematic anti-White violence rather than an isolated altercation. As analysts monitored activity over subsequent weeks, each legal development triggered fresh content: Anthony's arrest, his bail hearing, pretrial proceedings, and the eventual first-degree murder charge all became focal points for renewed threats and propaganda.[10]

Two additional incidents in September 2025 followed similar exploitation patterns. On September 5, local media in Charlotte, NC, released video footage of the August 22 stabbing of Iryana Zarutska, a 23-year-old Ukrainian woman, on a light rail train.[11,12] The alleged perpetrator, Decarlos Brown, was African American and well-known to law enforcement. When the footage went viral the following week, far-right accounts used it to advance narratives similar to those developed around the Metcalf case.

Then, on September 10, Charlie Kirk was assassinated. Again, footage of the murder spread rapidly across social media networks. Far-right accounts framed both the Zarutska and Kirk killings as evidence of escalating violence against White conservatives, despite significant differences between the cases. Following these murders, far-right groups organized a rally in Huntington Beach, CA, explicitly positioning Kirk and Zarutska as "martyrs," and displaying banners urging people to "Crush the Left."

Far-left

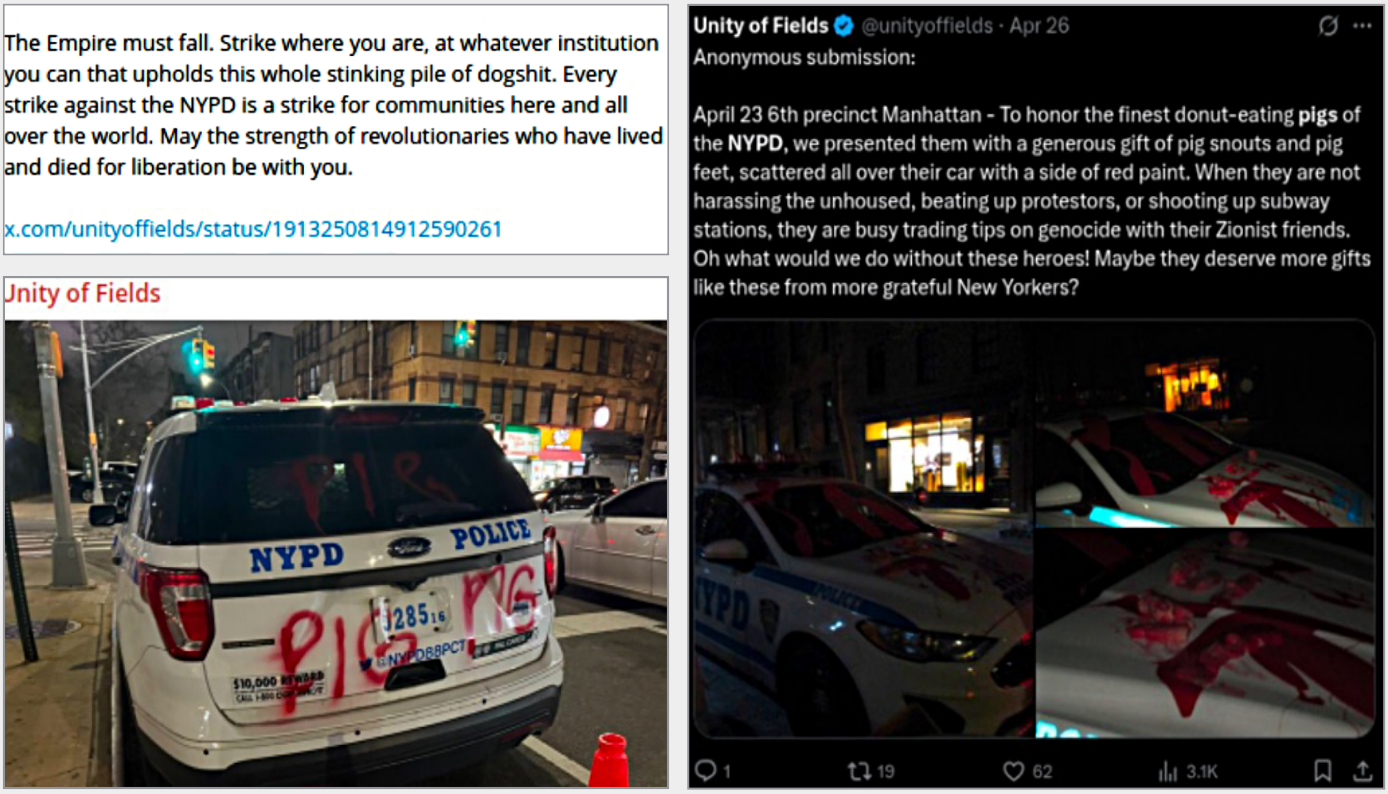
Far-left groups in the US advocate for transformative or revolutionary changes to political, economic, and social systems, going beyond mainstream progressive or liberal positions to call for alternatives to capitalism (e.g., socialism, anarchism, or communism), deep structural reforms to government, or abolition of institutions they view as oppressive (policing, prisons, immigration enforcement). These groups range from small anarchist collectives to Marxist-Leninist parties, socialist organizations, armed leftist groups, and antifascist ("antifa") networks[13]—the latter recently designated as a terrorist organization by the Trump administration.[14]

During the monitoring period, pro-Palestine activism dominated far-left online activity, reflecting the timing of the study during heightened conflict in Gaza. Analysts observed hundreds of posts from far-left accounts containing threats, doxxing, and calls for violence against US individuals and institutions perceived as supporting Israel. Frequent targets included law enforcement agencies, universities with financial ties to Israel, and corporations accused of facilitating Israeli military operations in Gaza. Notably, many of these far-left groups operated relatively openly.

The content patterns showed tactical sophistication. Posts combined explicit calls for action ("act against the NYPD") with doxxing information that enabled targeting. Vandalism and property destruction were celebrated and documented. Employee information from defense contractors was systematically collected and disseminated.

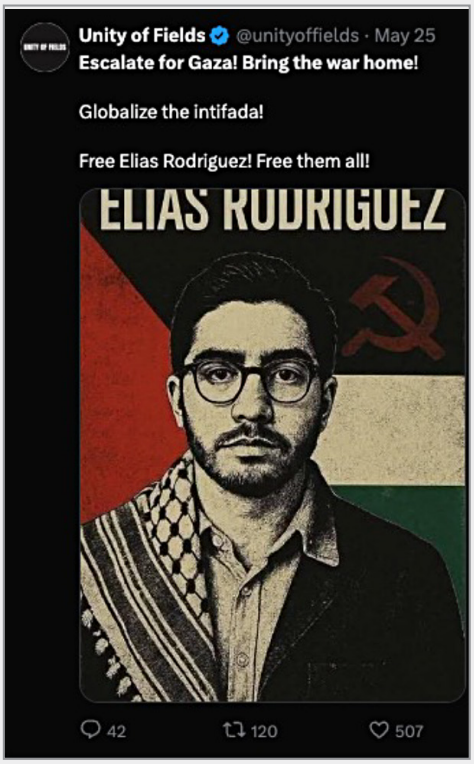
Sample of Far-left Posts During the Monitoring Period

Screenshot of a post on Telegram documenting vandalism at Northwestern University including the phrases, “Death to Israel” and “Intifada Now.” Posted on April 15, 2025.

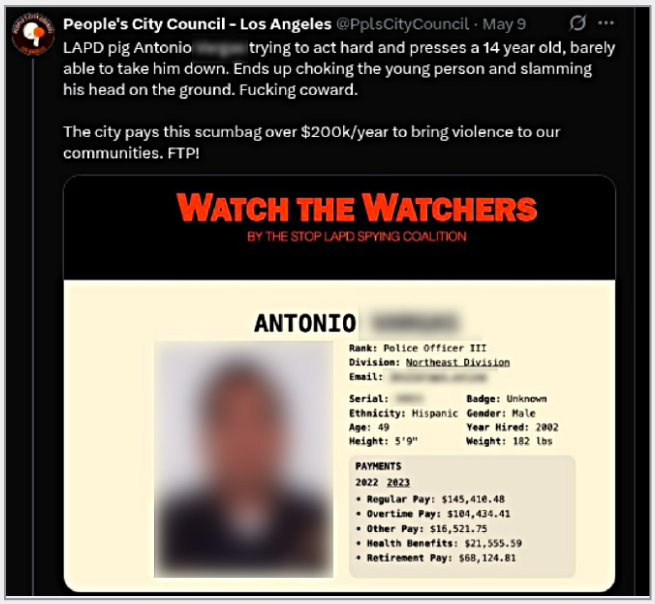


Screenshots of posts on the anti-Zionist group Unity of Fields’ Telegram channel demonstrate the sustained targeting of the New York Police Department (NYPD). The posts use dehumanizing language (“pigs”) while calling for supporters to act against police. Posted on April 19, 22 and 26, 2025, respectively.

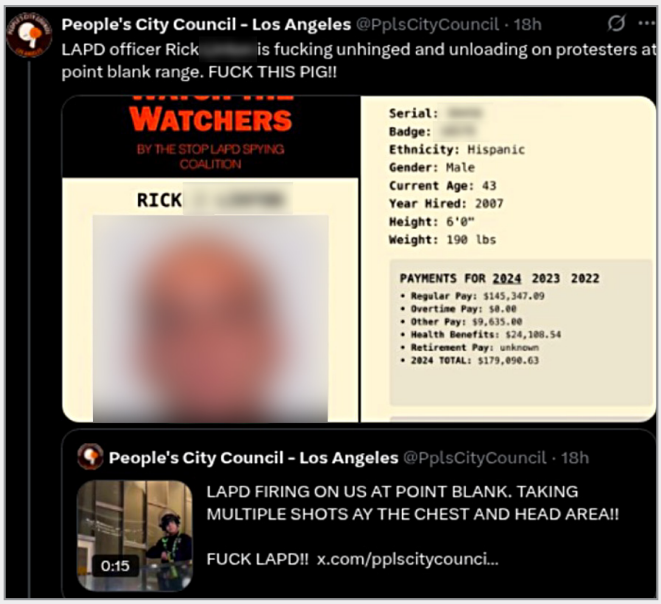
Screenshot of a post on Unity of Fields' Telegram channel linking to Pastebin.com containing personal information of US-based employees of Israeli company Elbit Systems. Posted on April 22, 2025, the post received 7,347 views and 79 reactions within hours, demonstrating both the reach of doxxing content and the active engagement it can generate.



A May 25 screenshot of Unity of Fields' X account, celebrating Elias Rodriguez's attack on the Israeli embassy staffers.



Screenshot of X post by People's City Council – Los Angeles, depicting the doxxing of a Los Angeles Police Department Officer. Posted on May 9, 2025. The post received 15.8k views, 731 likes, 13 comments, and 295 reshares as of May 12, 2025.



Screenshot of a second doxxing post by People's City Council account targeting a different LAPD police officer. Posted on June 10, 2025, this received significantly higher engagement: 1.5M views, 24k likes, 453 comments, 6.1k reshares, and 1.3k saves as of June 11.

Law enforcement officers became particular targets, with detailed personal information shared alongside dehumanizing rhetoric ("pigs") and implicit or explicit threats.

Engagement metrics revealed significant reach. Posts doxxing police officers garnered hundreds of thousands to over a million views, with thousands of likes and shares, indicating both large sympathetic audiences and effective algorithmic amplification. The viral spread of this content created cascading effects, with information initially posted on Telegram channels migrating to X where it reached broader audiences.

Nihilistic Violent Extremists (NVEs)

Nihilist Violent Extremists glorify, advocate for, or actively commit violence driven by a generalized hatred of society rather than coherent political ideology. The term entered federal law enforcement usage only in 2025, although extremism researchers had begun tracking these communities earlier.[15] NVEs present a particularly challenging category for threat assessment because their motivations are diffuse, their online activity often blends irony with genuine intent, and their glorification of violence transcends ideological boundaries.[16]

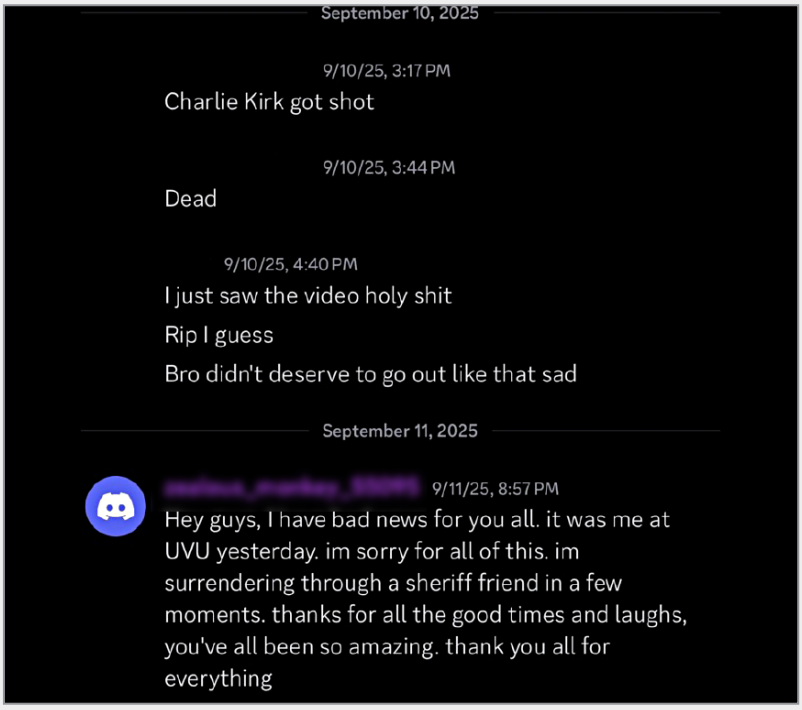
Unlike far-right or far-left extremists who justify violence through political frameworks, NVEs are motivated by nihilism, misanthropy, and a desire for notoriety. They often emulate or glorify previous extremists, regardless of those attackers' political leanings, treating mass violence as performance art designed to achieve "internet clout."[17]

This cross-ideological borrowing is distinctive: NVE communities celebrate far-right mass shooters such as Anders Breivik, the Norwegian behind the 2011 mass shooting in Utoya, or Brenton Tarrant, the Australian who live-streamed his 2019 attack on a New Zealand mosque. The glorification of these attacks is not about a political manifesto, but about the scale and spectacle of the violence.

This copycat, performative tendency was evidenced in the August 27 attack on the Annunciation Catholic School in Minneapolis, MN. Prior to the shooting, the suspect uploaded a manifesto to his YouTube page, as well as video depicting weapons decorated with a mix of ideological symbols. This aesthetic choice deliberately echoes the attack of the Australian far-right extremist Tarrant—in fact, one of the phrases inscribed on the weapon was "Brenton 4 ever."[18]

NVEs' drive for performative infamy makes both public and private social platforms crucial hubs for propaganda and networking. The YouTube channel of the Annunciation School attacker was quickly disseminated across social networks. A TikTok account belonging to Desmond Holly, the shooter at Evergreen High School in Colorado, posted content that venerated previous mass shooters, again including Tarrant. As was the case in the Annunciation shooting, content from Holly's social media accounts circulated after he was confirmed as the attacker, increasing the likelihood of copycat attacks.[19]

Additionally, in the wake of the Charlie Kirk assassination, police revealed some characteristics of Tyler Robinson's attack that seem to align with NVE patterns—his reported Discord activity, desire for virality evidenced by inscriptions on the ammunition, and potential lack of connection to organized political networks—although the extent of his online footprint remains under investigation. If Robinson was motivated less by ideology than by a nihilistic desire for notoriety, it underscores the broader analytical challenge NVEs pose: their attacks may appear political while being fundamentally performative.



Screenshot of leaked Discord server from Tyler Robinson confessing to the Charlie Kirk assassination. Posted on September 11. Source: Ken Klippenstein, September 16, 2025.

The Effect of Terrorist Designations on Violent Islamist Online Activity

Violent Islamist online activity during the monitoring period was noticeably more furtive than that of domestic extremist actors, operating in more restricted spaces and employing greater operational security. This pattern largely reflects stricter moderation and legal enforcement against Islamist extremism—particularly entities designated as Foreign Terrorist Organizations by the US State Department—compared to domestic ideological movements.[1]

In response, violent Islamist actors have gravitated toward decentralized platforms that support encrypted messaging. Rocket.Chat emerged as particularly significant during the monitoring period. This open-source communication platform allows users to create and manage their own "instances" or chat servers, with very limited external control or monitoring. Even so, material tended to focus on generic anti-US propaganda rather than explicit operational content. When more operational or explicitly threatening content appeared, it was typically shared through "out-links" (URLs directing users to another site) to encrypted or semi-encrypted spaces such as SimpleX and Gem Space.

The examples below are consistent with other patterns in violent Islamist online propaganda, including the frequent use of symbolic messaging featuring imagery and text designed to inspire violence rather than explicit calls for violence. While not directly linked to specific incidents, this type of content has been present in previous cases of radicalization, including the Pulse Nightclub mass shooting in Orlando, FL, in 2016.[2]

The violent Islamist online presence demonstrates that terrorist designations combined with platform enforcement can force groups into smaller online spaces, reducing their recruitment reach and propaganda dissemination. That framework currently applies to foreign organizations, creating an enforcement asymmetry with domestic groups engaging in similar behavior.

Screenshot of a graphic from the Al-Battar Foundation posted on TechHaven, a server within Rocket.Chat, stating in Arabic and Turkish: "It is a Religious War! Not Individual Actions!" The messaging frames violence as collective religious duty while avoiding explicit calls for specific attacks—a common approach designed to inspire while maintaining plausible deniability. Posted on April 8, 2025.

Screenshot of a graphic uploaded by Halummu Media to TechHaven stating: "By Allah's permission, will be your end, destruction, and elimination." Posted on April 22, 2025.

[1] These include ISIS, al-Qaeda, and their affiliates. U.S. Department of State, *Foreign Terrorist Organizations*, Bureau of Counterterrorism, accessed October 2, 2025.

[2] Rebecca Shabad, FBI Director Comey 'Highly Confident' Orlando Shooter Radicalized Online, CBS News, June 13, 2016.

Key Accounts and Networks[1]

Far-right	Far-left
US-based Active Clubs A network of channels disseminating online intimidation and threats, often in coordination. Framing themselves as innocuous fitness groups, Active Clubs use both mainstream and niche decentralized networks to recruit, radicalize and prepare members for racist violence.[2] The channels engage in significant cross-platform activity.	**Unity of Fields** The US-based far-left "direct action network," operates primarily through its X account, where it has over 11,000 followers, and Telegram channel, where it has over 10,000 subscribers. The group frequently engages in online intimidation and issues violent threats, paired with calls for direct action, primarily targeting individuals and institutions accused of supporting Israel. Content typically receives high to very high levels of engagement.
The Memewaffen With nearly 3,000 subscribers, the channel regularly disseminates and promotes far-right content related to the US and receives high levels of engagement.	**Bronx Anti-War** Operating on Telegram (691 subscribers) and X (3,813 followers), this group engages in online intimidation and threatening behavior, including the doxxing of law enforcement officers, primarily motivated by pro-Palestine narratives.
US Department of Women Haters This channel has over 4,000 subscribers and is likely operated by users based in the US. It promotes far-right and misogynist ideologies, frequently resharing material from other like-minded channels, including The Memewaffen.	**People's City Council – Los Angeles** This group operates through an X account with almost 80,000 followers. It frequently engages in online intimidation and threats justified as in defense of Palestine, as well as rhetoric expressing broader anti-government and anti-state sentiments. Members have engaged in doxxing police officers and other government officials.
/pol/ 4chan With over 12,000 subscribers, this channel is closely linked with the "imageboard" (or image-centric) forum of the same name. Content tends to feature anti-immigration, anti-left, and anti-semitic narratives.	

[1] A more complete list of key accounts and channels labeled as far-right and far-left, respectively, along with their current online status and user base, is contained in the Appendix.

[2] Art Jipson, *Active Clubs Are White Supremacy's New, Dangerous Frontie*r, TalkingPointsMemo, August 31, 2025.

3. Intersections Between Online Discourse And Offline Incidents

The relationship between online discourse and offline violence is complex, bidirectional, and difficult to establish with certainty. While it is theoretically clear that exposure to violent content can influence behavior, demonstrating that specific online activity caused specific violent acts requires evidence typically unavailable to external researchers: perpetrators' browsing histories, private communications, psychological assessments, and detailed investigative findings.

> "Observed patterns reveal the mechanisms through which online spaces may contribute to climates conducive to violence, even when direct causal chains cannot be established."

This project, constrained by ethical limitations on infiltrating closed groups and reliant on observation of semi-public channels, cannot definitively prove causation. What it can document is the observable patterns of how networks respond to and exploit violence, how they attempt to mobilize audiences, and what rhetorical and tactical approaches they employ. These patterns reveal the mechanisms through which online spaces may contribute to climates conducive to violence, even when direct causal chains cannot be established.

The monitoring period revealed three primary patterns in the relationship between online activity and real-world incidents: exploitation of trigger events for violent propaganda; cross-ideological convergence in the targeting of certain groups; and spirals of retaliation and online harassment, leading to deepening hostility.

Trigger events

The dominant pattern observed was extremist networks seizing upon violent incidents to advance existing narratives, recruit supporters, and justify further violence. This pattern appeared consistently across ideological categories:

- Far-right networks exploited the Metcalf stabbing, Zarutska killing, and Kirk assassination to advance White victimhood narratives.
- Far-left networks celebrated the Capital Jewish Museum shooting to glorify anti-Israel violence.
- NVE communities disseminated content from the Annunciation and Evergreen attackers to glorify violence generally.

In these cases, violence occurred first, then online networks rapidly mobilized to extract propaganda value. The online activity did not cause the initial incident but

worked to ensure maximum impact, normalize the violence, inspire potential future actors, and recruit sympathizers.

The fatal stabbing of Austin Metcalf by Karmelo Anthony on April 2, 2025, demonstrates the full exploitation cycle. Far-right networks immediately reframed an isolated incident as evidence of systemic anti-White violence. They sustained this narrative through each legal development (arrest, bail hearing, court appearances) and kept it active for weeks as a touchstone for recruitment ("join your local Active Club") and to justify threats against Anthony, his family, and court officials.

The week of September 8-10, 2025, provided multiple trigger events in rapid succession. The September 5 release of footage showing Iryna Zarutska's August 22 stabbing, followed by Charlie Kirk's September 10 assassination, created what analysts described as a "compounding effect." Far-right channels linked the incidents, despite their differences, framing both as evidence of increasing violence targeting White, conservative Americans.

This compounding pattern appeared to intensify mobilization. Following the murders, far-right groups (notably those active in Huntington Beach, CA) attempted to use both Kirk and Zarutska as martyrs requiring retaliation, with banners calling on people to "Crush the Left." A separate video posted on September 11 called for a similar "White Lives Matter" protest in Miami, FL.[20]

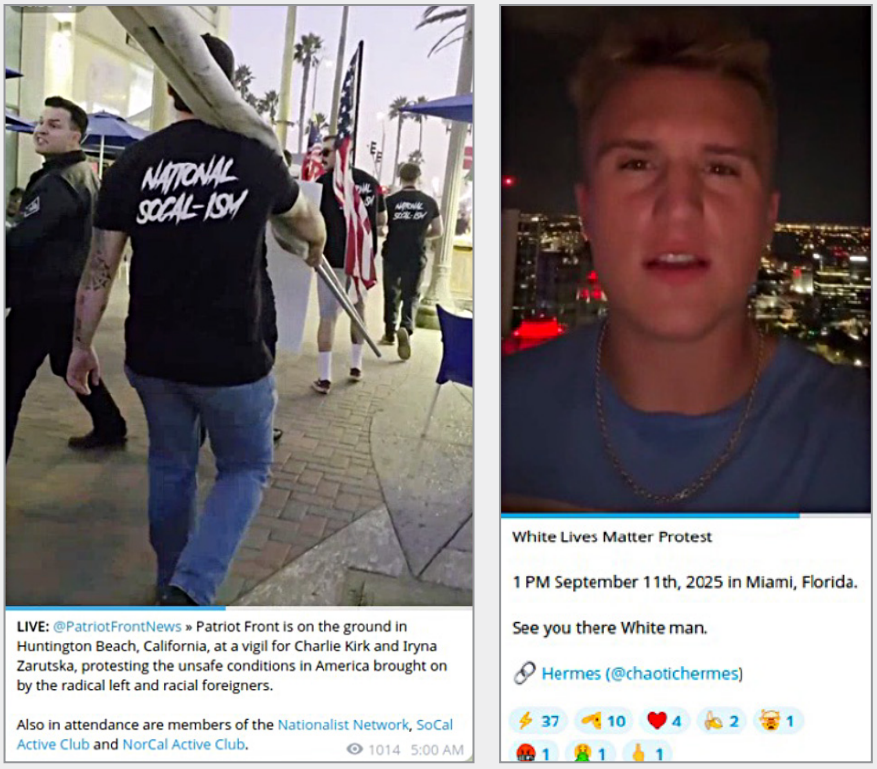
The back-and-forth dynamic between online rhetoric and real-world altercations makes it difficult to attribute offline violence solely to online rhetoric. Rather, the two realms are mutually reinforcing. For example, the "Protect White Americans" rally in Texas on April 19, 2025, was organized online and drew participants through digital recruitment. Although it was not a call for targeted violence, altercations at the event then triggered renewed online activity calling for escalation, demonstrating the bi-directional dynamic.

For NVEs, the triggering events are less explicitly political and consequently more difficult for researchers and law-enforcement to detect and disrupt. Their generalized obsession with violence, however, makes it possible for unrelated political violence to be seen as an opportunity by an NVE to increase their personal infamy and generate further chaos.

Cross-ideological convergence

The Capital Jewish Museum shooting of the Jewish couple on May 21, 2025, triggered activity from both far-left and far-right networks, revealing a troubling convergence around antisemitic targeting. Following the murders, there was a marked spike in explicit calls for violence against Jewish communities. On far-left channels, analysts recorded high engagement with content praising the attacker, Elias Rodriguez, hailing him as a political prisoner and ideological martyr, with posts encouraging others to adopt similarly violent tactics.

The attack also drew toxic, albeit less explicit, responses from far-right extremists. Some celebrated the murders and echoed antisemitic tropes, demonstrating that violent ideologies, despite their differing justifications, often converge in their outcomes and targets.



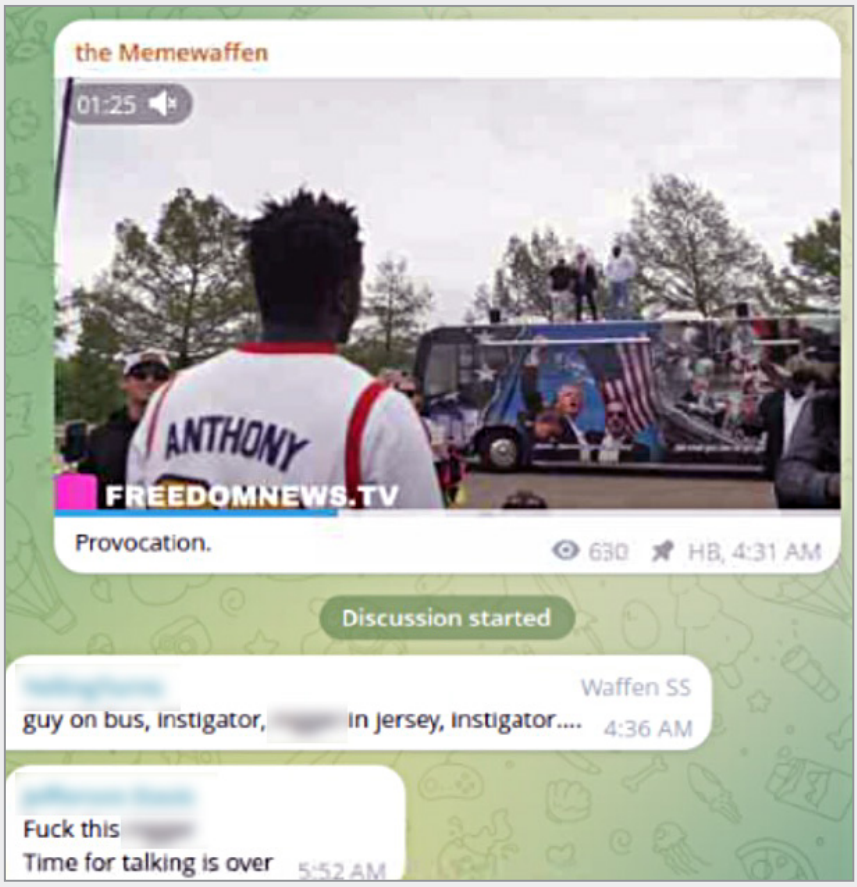
Left, a September 14 Telegram post showing members of the far-right group Patriot Front at the Huntington Beach "Crush the Left" protest. Right, a September 11 Telegram post purporting to organize a White Lives Matter Protest in Miami, FL on the same day.

On Telegram, a post made to The Memewaffen channel lamented that ANTIFA had been more successful than "American Nazis" in damaging "the image of Jews." It received 1,985 views, 80 reactions, and 30 comments on the same day it was posted.

Spirals of retaliation

During the monitoring period, tensions between far-left and far-right groups in the US remained a prominent feature of the online and offline threat landscape. These groups frequently engaged in hostile exchanges, coordinated harassment, and direct confrontations, contributing to a cycle of deepening hostility. The retaliatory exchanges, which were primarily observed on X and Telegram, often took the form of doxxing, harassment campaigns, and online "raids," which consist of infiltrating online groups and spamming them with offensive slurs and other pejoratives.

While these online conflicts did not directly cause major violent incidents during the monitoring period, they contribute to a climate of escalating hostility where violence can become increasingly normalized as acceptable political expression.

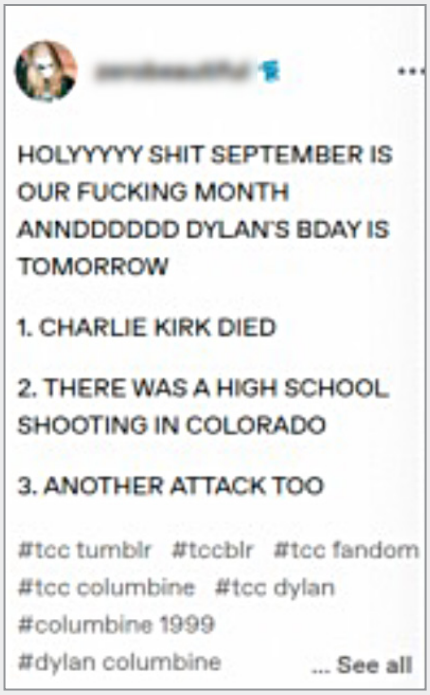


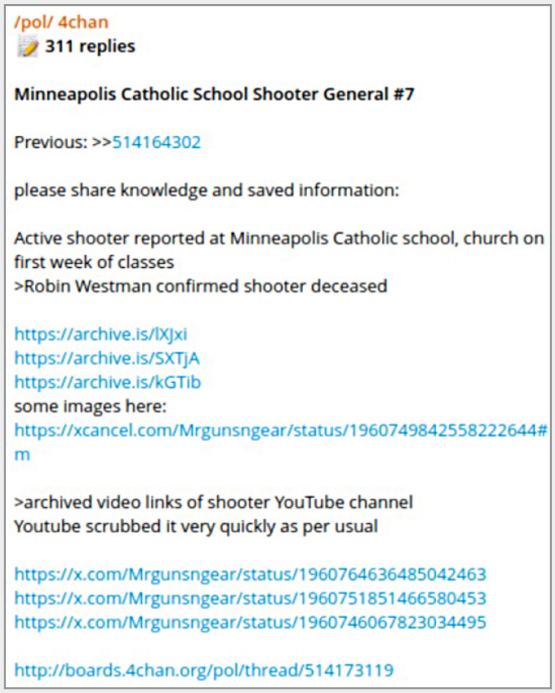
Screenshot of a post made to The Memewaffen's Telegram channel depicting a video of a verbal altercation that occurred at the "Protect White Americans" rally alongside a veiled threat stating, "Time for talking is over." Posted on April 20, 2025, it received 630 views and 9 comments as of April 22, 2025.

Rhetorical indicators of potential escalation

The development of reliable and nuanced rhetorical indicators of escalation is critical for identifying threats before they manifest into acts of violence. The following indicators are drawn from Tech Against Terrorism's broader analytical framework based on patterns observed across multiple research projects, not exclusively from this 13-week monitoring period. They represent risk markers that analysts use to assess potential escalation, although they have not been validated as predictive tools. The more indicators present and the more synchronized or intensified they become, the greater the likelihood that an individual or network is advancing towards actionable violence.

- **Increased specificity:** A shift from ambiguous or ideologically broad threats to concrete, specific targeting. This may include the naming of individual people, government officials, institutions, or critical infrastructure, often alongside details that suggest capability, access, or information gathering, such as the identification of work schedules, blueprints, or security gaps. In some cases, this specificity is embedded within layered or coded language, making it essential for analysts to contextualize posts carefully. Specific threats can escalate rapidly when accompanied by real-world preparatory activities, such as weapons purchases, conducting online or physical surveillance, and sharing attack strategies.
- **Indications of operational planning:** Clear evidence of operational preparation, including mention of weapons, tactics, timelines, or training regimens, is a critical indicator of imminent threat. Planning may appear in forum discussions, digital notebooks, private channels, or screenshots





Far left, a Tumblr (captured September 12) post from an NVE-aligned community, celebrating the murder of Kirk and the same-day mass shooting in Evergreen, CO.

Left, an August 28 Telegram post archiving and amplifying content made by the Annunciation School shooter. These examples show how acts of NVE violence are used by the community to further galvanize each other.

of schedules and equipment. Even when discussed abstractly, conversations such as, "Which rifle is better for close quarters," or "best routes in and out of [location]" may indicate a user is beyond ideological endorsement and is entering the logistical phase of mobilization. Sharing information about GPS coordinates, mapping escape routes, or asking about past attack methods are signs of escalated risk, especially when coinciding with evidence of group coordination or platform migration.

- **Shifts in tone or urgency:** A noticeable shift toward apocalyptic, doom-laden, or desperate language can indicate that users believe a "point of no return" has been reached. Expressions such as "this is the last straw," or "no more waiting," and posts that frame violent action as the only remaining moral or practical option, are high-risk indicators. When combined with other indicators—such as explicit operational planning or discussion of martyrdom and legacy—this tonal escalation can serve as a warning sign of imminent mobilization.

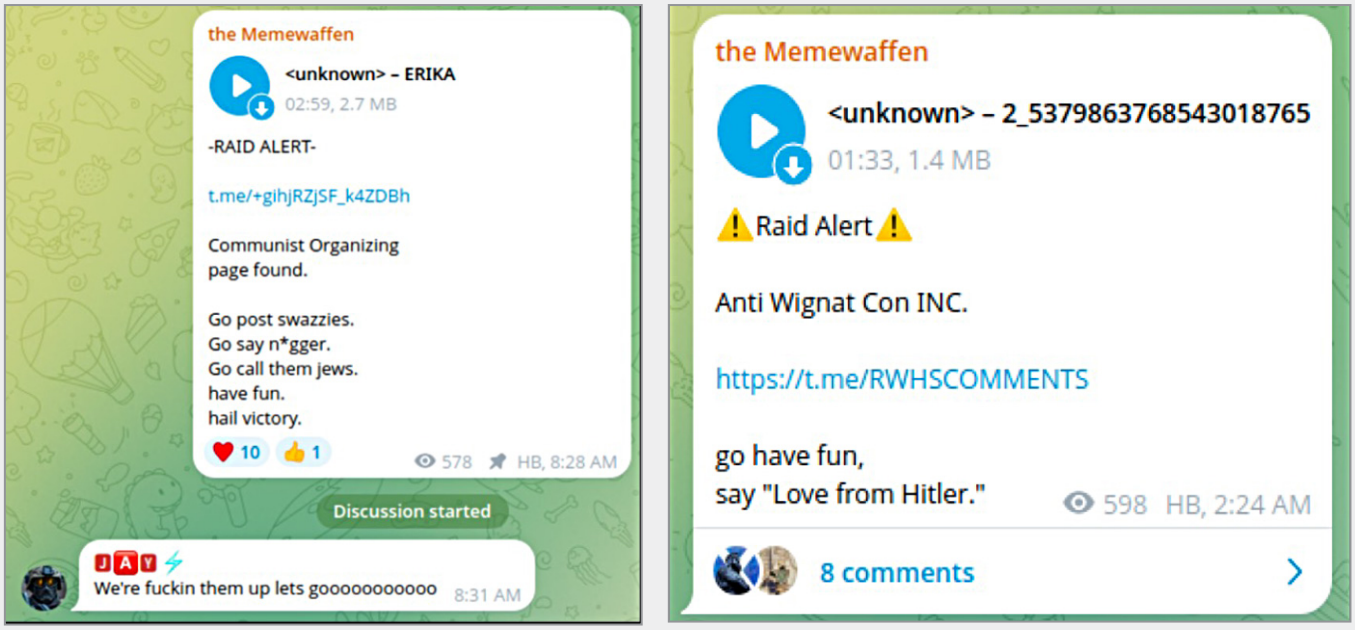
- **Platform migration:** Noticeable movement to more encrypted or private channels with limited public access could indicate intent to organize and operationalize. These movements are especially indicative when conducted shortly after a major event, such as a protest, law enforcement operation, or controversial political development.

- **Cross-ideological messaging:** Tactical alignment by actors across the ideological spectrum towards a collective target or goal increases the risk of larger, more volatile mobilizations and the possibility of complex, multi-actor incidents. This convergence can manifest in shared language, coordinated actions, or mutual amplification of propaganda. For example, during the monitoring period, far-left and far-right accounts usually converged around antisemitic narratives, notably following the murder of the young couple outside the Capital Jewish Museum and an incendiary device attack on a pro-Israel demonstration in Boulder, Colorado that killed one woman and injured others.

- **Energy bursts:** A sudden spike in digital activity, sometimes referred to as a "pre-attack energy burst," can serve as an early warning sign of violence. Spikes may consist of rapid posting, sudden account creation across platforms, or an increased rate of private messaging.

These indicators are observational patterns, not validated predictive tools. Many instances of this rhetoric do not result in violence. Conversely, some attacks occur without observable warning signs in public channels.

Posts from Unity of Fields and Bronx Anti-War X accounts celebrating the killing of a Jewish couple and encouraging similar action. Posted on May 23, 2025.



Screenshots of posts made to The Memewaffen Telegram channel calling on users to "raid" far-left Telegram channels. Posted on April 18 (left) and April 26, 2025 (below).

4. Strategic Use of the Online Platform Ecosystem

> As long as platforms exist with complementary features—some offering reach, others security, others monetization—actors will be able to reconstitute their presence and preserve the connective tissue of their networks.

Online threat analysts have noted the strategic and simultaneous use of various types of online platforms by violent extremist actors.[21]

Each platform comes embedded with a different set of features, ranging from algorithmic feeds that favor wide audience reach and rapid dissemination of propaganda to encrypted messaging that enables secure and private channels for operational planning. The simultaneous use of multiple platforms not only provides these actors with a diverse toolkit but also ensures their continued online presence in the face of possible expulsion, or "de-platforming."[22] Understanding these strategic choices is essential for effective intervention.

This multi-platform strategy was observed during the monitoring period as well, with a key practice involving "out-linking—the use of posts that include a URL directing users to another platform or website. The following snapshots on page 22, from a randomly selected week in May illustrate observed out-linking patterns by ideological category. Posts with out-links were often direct mirrors of one another, serving as backups for content, or adapted based on the target audience and content moderation policies of each platform.

The left-hand column shows where posts first appeared. The columns to the right show the number of times the original posts were shared via out-links on the other listed platforms.

These patterns reveal how extremist actors use different platforms for distinct tactical purposes, with clear platform-specific roles emerging. For domestic extremists (far-right and far-left), Telegram, which is perceived by extremist groups as a relatively safe place for hardened rhetoric and movement coordination, serves as the primary content hub. The social media site X functions as the main public-facing amplifier due to its larger user base and viral features. Together, these two platforms enable resilience, audience segmentation, and enhanced virality. Violent Islamist groups show more aggressive out-linking to encrypted platforms (Gem Space, SimpleX), reflecting their need for operational security given stricter enforcement pressure.

Snapshot of far-right out-linking behavior

Origin platform	X	TikTok	YouTube	4chan	Instagram
The Memewaffen channel (Telegram)	11	3	4	0	1
US Department of Women Haters channel (Telegram)	15	4	3	2	2
/pol/ 4chan channel (Telegram)	29	3	17	42	0

Snapshot of far-left out-linking behavior

Origin platform	X	TikTok	YouTube	WhatsApp	Instagram
Unity of Fields channel (Telegram)	11	1	4	0	2
Bronx Anti-War channel (Telegram)	9	1	3	0	3
Resistance News Network channel (Telegram)	7	2	3	1	1

Snapshot of violent Islamist out-linking behavior

Origin platform	X	Telegram	Gem Space	Chirpwire	SimpleX
TechHaven	2	32	3	1	2
GeoNews	3	6	1	2	1
Element	1	8	1	4	2

Scholarly research supports these observed patterns. Tamar Mitts, a Columbia University professor, argues that actors who face heightened moderation "gravitate to platforms of a particular kind: those that have less restrictive moderation policies, but that also reach a sufficiently large audience."[23] Both Telegram and X fit these criteria during the monitoring period. At the same time, Mitts argues, extremists tend to maintain a presence in smaller, less moderated platforms where they focus their recruitment and mobilization.

Beyond moderation level and audience size, several platform characteristics emerged as significant "push" and "pull" factors during the monitoring period.

- **Discoverability features**
  Algorithmic recommendation systems, trending-topic feeds, and search indexing draw extremist actors toward platforms where content can reach beyond immediate followers. X's recommendation algorithm and "For You" page exemplify high-discoverability features. Conversely, the absence of such features on platforms like Discord, which mostly consists of invite-only servers, can limit reach, but offers other advantages such as exclusivity and obscurity.

- **Privacy-enhancing features**
  End-to-end encryption, self-destructing messages, and anonymous account creation act as strong pull factors for actors prioritizing operational security. Telegram's perceived insulation from external scrutiny has made it a favored coordination hub for a variety of violent actors. Signal's comprehensive encryption, and SimpleX's identity-free structure each appeal to different security priorities.

- **Decentralization**
  Platforms built on decentralized infrastructure (e.g., Rocket.Chat, Element) reduce reliance on central intermediaries who could remove content or ban users. These qualities not only influence where extremist actors congregate—attracting groups that have faced more stringent moderation, such as Violent Islamists—but also how they sequence their platform use: leveraging discoverable,

centralized spaces for outreach and recruitment, then shifting to encrypted or decentralized environments for planning and propaganda archiving.

- **Monetization opportunities** Economic incentives act as significant pull factors for some actors. Platforms that offer monetization—whether through advertising revenue shares, paid subscriptions, or live-streaming donations—can reward high-engagement content, including inflammatory or controversial material. During the monitoring period, multiple X accounts linked to far-right and far-left networks monetized high-engagement posts about the Austin Metcalf stabbing and the Capital Jewish Museum shooting. The combination of monetization with algorithms that boost engagement-driven content creates feedback loops where provocative material becomes both more visible and more profitable.

By weaving together these various platform characteristics—discoverability, privacy, decentralization, moderation thresholds, audience reach, and monetization opportunities—extremist actors create cross-platform networks that are both adaptive and durable. This adaptive capacity to calibrate content and migrate audiences undermines single-platform enforcement efforts. As long as platforms exist with complementary features—some offering reach, others security, others monetization—actors will be able to reconstitute their presence and preserve the connective tissue of their networks. The problem becomes systemic rather than platform specific. Any effective policy response will need to address these interlocking dynamics.

The Relationship between Telegram and X

A notable pattern among both far-right and far-left accounts is the frequent use of out-linking between Telegram and X, which provide complementary advantages.

Telegram provides
• Semi-private channels for committed followers
• Minimal content moderation
• Channel structure enabling one-to-many communication
• Relative stability against de-platforming
• Community consolidation

X provides
• Massive potential reach (hundreds of millions of users)
• Algorithmic amplification of engaging content
• Real-time virality during breaking events
• Legitimacy through mainstream platform presence
• Recruitment pool of passive sympathizers

This out-linking strategy serves at least three strategic purposes. It allows groups to establish a following on two influential platforms, creates resilience against possible de-platforming, and allows them to segment their audience and tailor their messaging.

Telegram posts, usually more explicit or incendiary in nature, tend to include out-links to X, where more casual followers or unwitting users encounter more sanitized versions. If X responds by deleting a post or suspending an account, those actors can usually maintain their Telegram following and attempt to rebuild their presence on X at a later stage.

5. Conclusion and Recommendations

The 13-week monitoring period documented extensive exploitation of violent incidents by extremist networks across the ideological spectrum.

Far-right actors used the Metcalf stabbing, Zarutska killing, and Kirk assassination to advance White victimhood narratives. Far-left networks celebrated the Capital Jewish Museum shooting and engaged in systematic doxxing of law enforcement. Violent Islamist groups maintained symbolic propaganda presence despite enforcement pressure. Content from NVE-motivated attackers circulated after their attacks, glorifying violence for shock value and notoriety.

These networks employed sophisticated cross-platform strategies: Telegram for consolidating committed followers, radicalizing, and coordinating behind digital closed doors; X for high-visibility provocation and audience expansion; and decentralized servers for resilience when moderation closes other avenues. The choice of platform in each instance reflected tactical calculation—balancing reach against moderation risk, visibility against operational security.

Confronting the documented patterns should not be partisan. The deliberate online targeting of individuals and institutions with threats and incitement undermines public safety and democratic discourse regardless of political affiliation. While free speech protections remain paramount, they do not preclude measured, rights-respecting interventions to disrupt networks that systematically exploit violence and threaten specific targets.

Recommendations

For Platforms and Online Service Providers

1 Adopt precise policies on threats and incitement and demonstrate willingness and capacity to enforce them consistently.

Platforms and service providers should move beyond vaguely defined policies against "harmful" behavior or "extremism" to clearly define prohibited conduct involving threats of violence and incitement. They should communicate these policies clearly to users, publicly commit to acting on them, and take visible steps to deter misuse, including by publishing frequent updates on how these platforms are staying ahead of bad actors' evolving tactics.

Violent actors strategically choose platforms based on their perceptions of platforms' willingness and ability to enforce policies.[24] A strong publicly stated stance against threats and incitement by platform leadership, combined with reliable enforcement, can deter exploitation. Conversely, a platform's persistent failure to communicate robust policies and demonstrate effective content moderation typically turns it into a de facto safe haven for actors seeking to evade scrutiny.[25]

2 Implement user-friendly reporting tools compatible with encryption.

At a minimum, platforms and service providers should offer in-app reporting tools that allow users to flag content that they believe violates platform policies. These tools can be designed to be fully compatible with end-to-end encryption using tools such as message franking—a cryptographic technique that enables user reporting without compromising message security or privacy.[26]

Effective reporting systems require (1) intuitive interfaces allowing users to easily locate and use reporting functions without technical expertise; (2) prompt review and timely responses, with clear communication about outcomes, (3) safeguards against the exploitation of the reporting system itself, for example, to silence legitimate speech; and (4) transparency around how reports are evaluated and the possible responses.[27]

3 Use metadata responsibly to disrupt networks.

Platforms, including encrypted messaging services, may collect and analyze metadata—information about a message, file, or user, rather than the message content itself.[28] Some platforms justify this collection by noting that it helps them detect abusive behavior proactively. For example, WhatsApp says it uses metadata to combat spam-like activity, such as bulk and automated messaging.[29] However, metadata collection carries privacy risks. Responsible use requires transparency about how much metadata is collected and how it is used, a commitment to only use metadata for specific safety purposes, collecting the minimum metadata necessary for these purposes, and eliminating the metadata after a certain period of time.

4 Cooperate with other services to monitor and combat dangerous cross-platform activity.

Given extremist actors' strategic use of multiple platforms, efforts to curb violent discourse are far more effective when companies act in concert. To close loopholes and maintain a level playing field across the industry, platforms should actively participate in knowledge-sharing

initiatives such as the Global Internet Forum to Counter Terrorism (GIFCT)[30] and the Christchurch Call.[31] In particular, platforms should share information about emerging tactics, coded language, migration patterns, and coordinated recruitment campaigns. Where possible, without compromising encryption, platforms should try to block violent extremist actors' entry into harder-to-monitor encrypted spaces.[32] They can do so by tracking out-links and invite links posted in public channels and limiting the dissemination of these links when there are signals of potential violent mobilization.

At the same time, cross-platform cooperation must be carefully managed to avoid "censorship creep," the gradual expansion of prohibited content categories, often driven by regulatory or political pressure.[33] Platforms should ground their policies on threats and incitement in international human rights standards, while also implementing geofencing, or location-based, protocols to respect jurisdictional differences in legal thresholds.

For US Legislators and Policymakers

5 Mandate transparency, design, and procedural standards without undermining encryption.

Lawmakers should require platforms, including those with end-to-end encrypted features and those which offer semi-private spaces (such as Discord servers), to publish transparency reports that explain, among other things, their abuse-detection, reporting processes, and government data access requests. They should also require platforms to implement user-friendly reporting tools and actually review and address user reports in a timely fashion.[34]

Regulations should be nuanced and tailored to platform features. Lawmakers should require that platforms deploy proactive detection and reporting of illegal content, but only with respect to features that are not end-to-end encrypted.[35] Critically, regulations should not create backdoors for law enforcement access to encrypted messages, which would fundamentally compromise security for all users, including journalists, activists, dissidents, and vulnerable populations.

6 Revisit extremist and terrorist designation frameworks to ensure enforcement consistency.

The monitoring period revealed that violent Islamist groups are, in many cases, more cautious about making explicit calls for violence than domestic far-right and far-left groups, even on semi-closed platforms like Telegram. This restraint likely reflects the moderation crackdowns and coordinated de-platforming efforts prompted by terrorist designation laws, which have historically prioritized foreign Islamist groups.[36] To close this enforcement gap, lawmakers should revisit the criteria used for inclusion on terrorist designation lists and apply them consistently across ideologies, ensuring that far-right and far-left violent extremist groups are subject to the same standards.

7 Recognize limits of legal remedies and respect constitutional boundaries.

A significant share of speech that many would find harmful or dangerous remains lawful under First Amendment doctrine. Policymakers should clearly distinguish between protected but harmful speech (hate speech, offensive content, and general calls for violence without imminence, which are protected under First Amendment doctrine), and illegal speech (true threats and incitement to imminent lawless action likely to occur, which are not protected).

Recommendations

Platforms should be required to remove illegal speech, but must not be compelled to police vaguely defined "harmful" categories, as some foreign regimes mandate.[37] Nor should platforms be prohibited from voluntarily setting and enforcing higher standards for acceptable content and conduct, as recent laws in Texas[38] and Florida[39] have attempted.

The September 2025 response to Charlie Kirk's assassination illustrates that public officials can have trouble differentiating types of speech. Attorney General Pam Bondi initially wrote on X that "Hate speech that crosses the line into threats of violence is NOT protected by the First Amendment,"[40] appearing to conflate hate speech (generally protected) with illegal threats (not protected). Although she walked back these comments, the episode reveals policymaker confusion about constitutional boundaries.

8 Clarify protocols for platform-law enforcement cooperation.

Policymakers should establish clear standards for when and how platforms should share information with law enforcement related to threats or incitement. By the same token, they should expand legal training of law enforcement to ensure officers can distinguish between protected speech and illegal threats or incitement under US law. Cooperation between these two entities should be consistent with constitutional due process guarantees, including requirements for proper judicial authorization, transparency in the handling of user data, and safeguards against overreach or politically motivated investigations.

9 Support research on effective counter-speech initiatives.

Given constitutional limits on restricting speech, policymakers should explore non-coercive alternatives, including counter-speech campaigns, partnering with civil society to counteract violent narratives.[41] Early studies looking at the effects of counter-speech reveal varying results, with some studies showing promising results depending on the method adopted.[42] Government agencies should support further research into the factors that make counter-speech effective and seek to learn from case studies from countries like Brazil, Colombia, and Sweden.[43]

Additionally, policymakers should encourage platforms to promote healthy interactions and content consumption through their design features and content amplification. YouTube's Creators for Change initiative[44] and Google/Moonshot's Redirect method[45]—which lead users towards positive content and away from problematic narratives respectively—are worthwhile models to build upon.

Methodology

This research employed open-source intelligence (OSINT) monitoring of online platforms to document discourse related to violence in the United States. The methodology's strengths came with significant limitations that fundamentally shaped both the research process and findings. As a result, the study documents propaganda, recruitment, and public-facing threats but could not observe operational planning or private radicalization occurring in truly closed spaces.

Research Design

The NYU Stern's Center for Business and Human Rights retained Tech Against Terrorism, an organization specializing in open-source intelligence (OSINT), to monitor online discourse for evidence of intimidation and violent threats against US persons or institutions. The monitoring initially covered eleven weeks from March 24 to June 6, 2025, then expanded to include two additional weeks between September 10-24, following Charlie Kirk's assassination, as well as the attacks in Minneapolis, MN and Evergreen, CO.

The project initially focused on encrypted or semi-encrypted messaging platforms—WhatsApp, Telegram, Signal, and Viber—chosen for their presumed utility for violent actors engaging in attack planning. This focus reflected the hypothesis that studying closed platforms would reveal coordination and planning invisible on public social media. Instead, fully encrypted platforms yielded minimal relevant data, while semi-public platforms provided extensive observable activity. This outcome required shifting the research focus from operational planning to propaganda and recruitment tactics.

Ethical Constraints and Access Limitations

Tech Against Terrorism operates under a policy of non-engagement: analysts did not interact with users or present false identities in order to gain access to exclusive groups or channels. Specifically, access was only possible in rare cases where public invite links were shared openly, such as in Telegram channels. Where these links led to groups with light or generic vetting procedures, such as yes/no gatekeeping questions, Tech Against Terrorism considered access on a case-by-case basis, ensuring that it could be obtained without misrepresentation or interaction. In cases where active vetting was required—such as providing ideological justification, personal data, skills credentials, or engaging directly with administrators—analysts did not pursue access. This strict non-engagement protocol, while a cornerstone of Tech Against Terrorism's OSINT methodology, likely led to blind spots in the research findings and analysis. The research therefore captured what extremist actors were willing to share semi-publicly, not what they discussed privately. The implications are substantial.

Platform-Specific Findings

Telegram proved the most productive platform, yielding the majority of documented content. X also provided significant data due to its public nature and the bidirectional content flow between X and Telegram. The discovery of relevant content was typically initiated via keyword searches, hashtag monitoring, existing network mapping, and out-links from other platforms.

Rocket.Chat instances (particularly TechHaven used by violent Islamist groups) were accessible for monitoring, yielding approximately 2,000-3,000 posts.

WhatsApp yielded minimal relevant data: analysts accessed fewer than 10 group chats, none containing in-scope content. This likely reflects both the platform's fully encrypted nature and the researchers' inability to infiltrate operational groups.

Signal and Viber yielded zero relevant data. Analysts observed no promotion or mention of these platforms during the monitoring period, suggesting either: these platforms weren't being used by monitored networks; usage was occurring but entirely invisible due to encryption and access constraints; or networks using these platforms successfully avoided any public mention that would enable discovery.

Finally, Discord monitoring was limited despite the platform's known use by NVE communities. Discord's invite-only server structure and vetting processes prevented access to operational spaces without deceptive engagement.

Over the course of the thirteen-week project, analysts viewed approximately:

Telegram	10,000-20,500 posts
X	15,000-26,000 posts
Tech Haven (part of Rocket.Chat)	2,050 posts
Geo News	1,050 posts
WhatsApp	10 group chats
Viber	0 posts
Signal	0 posts
Discord	0 posts

Platform Engagement
The researchers shared preliminary findings with Telegram and X (formerly Twitter), the two platforms most frequently used for extremist content during the monitoring period. Neither platform provided comment or response to the research findings.

Appendix

TAT-NYU: Channels and Accounts Monitored During the NYU Project
Date: September 10, 2025

The table on page 30 is a compilation of key channels and accounts identified by TAT analysts as being operated or utilised by far-right and far-left entities or otherwise affiliated. These were monitored throughout the duration of the NYU project.

This list is not exhaustive and does not represent all accounts and channels that were monitored during the project. Many accounts and channels were only observed a limited number of times, often in response to major discourse events such as the Washington shooting or ICE protests.

Due to the scale of TAT's online mapping efforts, it is not feasible to provide a complete list. Instead, the accounts and channels included here represent those that were consistently monitored and served as primary sources for content collection throughout the project. Some of the listed channels and accounts have since been removed or suspended.

Appendix (cont.)

Far Left Channels and Accounts

Platform	Channel/Account Name	User Base as of Monitoring Date
Telegram	Unity of Fields	9,137 subscribers
Telegram	Bronx Anti-War	744 subscribers
Telegram	Voices Ignited	16,567 subscribers
Telegram	WOL Protest	8,121 subscribers
Telegram	Queer, Leftist, Anti Fascist Unity	No longer active
Telegram	Resistance News Network	160,408 subscribers
X	Unity of Fields	14,800 followers
X	Brox Anti-War	4,803 followers
X	People's City Council – Los Angeles	92,100 followers
X	Film the Police LA	86,900 followers
X	People's City Propaganda	10,100 followers

Far Right Channels and Accounts

Platform	Channel/Account Name	User Base as of Monitoring Date
Telegram	The Memewaffen	No longer active
Telegram	US Department of Women Haters	4,564 subscribers
Telegram	ZoomerWaffen	22,756 subscribers
Telegram	Totally Awesome Friends	1,442 subscribers
Telegram	Media of Inoculation	478 subscribers
Telegram	National Christian Resistance	226 subscribers
Telegram	Esoteric Thoughts	636 subscribers
Telegram	/pol/ 4chan	11,871 subscribers
Telegram	4chan -/POL/HIS/INT/	11,857 subscribers
Telegram	White Lives Matter Official	22,626 subscribers
Telegram	Sons of Virginia Active Club	723 subscribers
Telegram	Republic of Texas Proud Boys – Official Channel	895 subscribers
Telegram	Proud Boys	9,836 subscribers
Telegram	Proud Boys (Unrestricted)	688 subscribers
Telegram	Sunflower Society	8,326 subscribers
Telegram	Smoke Pit V	2,610 subscribers
Telegram	Murder the Media	15,139 subscribers
Telegram	The Western Chauvinist (Main)	12,418 subscribers
Telegram	Alt Skull's Charnel House	35,669 subscribers
Telegram	Nationalist Squads	3,391 subscribers
Telegram	Crew 94	No longer active
Telegram	Media of Flames	No longer active
Telegram	111v2	No longer active
Telegram	TRaKtifa	No longer active
Telegram	Great Lakes Active Clun	1,850 subscribers
Telegram	Lone Star Active Club	1,164 subscribers
Telegram	Pennsylvania AC	1,620 subscribers
Telegram	AC x Official	4,827 subscribers
X	National Christian Resistance (NCR)	131 followers
X	Mr Zoomer	311 followers
X	The Fourth Reich	95 followers
X	Zoomer Acceleratorr	1,037 followers
X	White Lives Matter	5,603 followers
X	Southern vindaler brigade	No longer active
X	Murder The Media	792 followers
X	NatSoc Cowboy	2,145 followers
X	Martyrdom Division	50 followers
4chan /	pol/ -Politically Incorrect	n/a

Endnotes

1 Government Accountability Office, Domestic Terrorism: Additional Actions Needed to Implement an Effective National Strategy, April 2025.

2 Zoe Sottile, Tyler Robinson – suspect in Charlie Kirk shooting, CNN, September 29, 2025. Jamie Stengle, A Mexican man is the second victim to die after shooting at Dallas ICE facility, Associated Press, September 30, 2025; Rachel Estabrook, Second Evergreen school shooting victim is 14-year-old boy, family says in first public statement, Colorado Public Radio, September 26, 2025. Ashley Grams, Annunciation community marks 1 month since mass shooting, CBS, 2025.

3 James Oliphant, What to know about the case against Tyler Robinson, accused of killing Charlie Kirk, Reuters, September 16, 2025.

4 Nick Robins-Early, Memes and nihilistic in-jokes: the online world of Charlie Kirk's alleged killer, The Guardian, September 17, 2025.

5 Julie Bosman, Israel Embassy shooting suspect arrested in U.S., The New York Times, May 22, 2025.

6 Tech Against Terrorism Monitoring Alert, August 28 (on file with authors).

7 Christpher St. Aubin & Michael Lipka, Support dips for U.S. Government, tech companies restricting false or violent online content, Pew Research Center, April 14, 2025.

8 Tess Owen, Trump administration is minimizing white supremacist threat, officials warn, The Guardian, May 24, 2025. Hannah Allam, "The Intern in Charge": Meet the 22 year-old Trump's team picked to lead terrorism prevention, Pro Publica, June 4, 2025.

9 This definition closely tracks the definition set forth by the US Department of Homeland Security in its August 2011 report ("violent extremists [are] individuals who support or commit ideologically-motivated violence to further political goals"). We expanded the definition to encompass extremist violence that is not ideologically motivated—including violence perpetrated by "nihilistic violent extremists" (NVEs). DHS, Empowering Local Partners to Prevent Violent Extremism in the United States, August 2011.

10 S.E. Jenkins et al., Trial date set for Karmelo Anthony, teen charged in fatal stabbing of Austin Metcalf at Frisco track meet, CBS, July 13, 2025.

11 Hank Lee et al., CATS releases surveillance video of deadly light rail stabbing, WCNC Charlotte, September 5, 2025.

12 The crime occurred on August 22, 2025, but CCTV footage was only released on September 5.

13 Jacob Guhl, Left Wing Extremism, Institute for Strategic Dialogue, January 8, 2025.

14 The White House, Designating Antifa as a Domestic Terrorist Organization, Presidential Action, September 22, 2025.

15 Jacob Ware, Nihilistic Violent Extremism: A Valuable Stride Forward in American Counterterrorism, Just Security, May 21, 2025; Marc-André Argentino, Barrett Gay, and Matt Bastin, Nihilism and Terror: How M.K.Y. Is Redefining Terrorism, Recruitment, and Mass Violence, Combating Terrorism Center at West Point, September 2024.

16 Marc-André Argentino, Beyond the Headlines: Arrest Data and Drivers of Nihilistic Violent Extremism in the COM Network, From the Depths, September 18, 2025.

17 Ibid.

18 TAT Monitoring Alert, Terrorist and Violent Extremist Content, August 28 (on file with authors).

19 TAT Monitoring Alert, Terrorist and Violent Extremist Content, September 11 (on file with authors).

20 @F0XXDENCHAT, Telegram post, September 11, 2025.

21 Tamar Mitts, Safe Havens for Hate: The Challenge of Moderating Online Extremism, Princeton University Press, 2025.

22 Tech Against Terrorism, Terrorist Use of E2EE: State of Play, Misconceptions, and Mitigation Strategies, September 2021.

23 Mitts, supra.

24 Tech Against Terrorism, Terrorist Use of E2EE: State of Play, Misconceptions, and Mitigation Strategies, supra.

25 Mitts, supra; Will Bedingfield, *Deplatforming works, but it's not enough to fix Facebook and Twitter*, Wired, January 15, 2021

26 Dhanaraj Thakur et al., Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems, Center for Democracy & Technology, August 2021.

27 Kat Lo et al., Shouting into the Void: Why Reporting Abuse to Social Media Platforms Is So Hard and How to Fix It, PEN America, June 29, 2023.

28 Riana Pfefferkorn, Content-Oblivious Trust and Safety Techniques: Results from a Survey of Online Service Providers, Journal of Online Trust and Safety, 1(2).

29 WhatsApp Privacy Policy, accessed October 10, 2025. While WhatsApp publicly justifies its metadata collection by citing the need to combat spam and terrorism, its privacy policy also permits sharing metadata with its parent company, Meta, for "showing relevant offers and ads across the Meta Company Products."

30 GIFCT, Global Internet Forum to Counter Terrorism (GIFCT), accessed October 2025.

31 Christchurch Call, Christchurch Call to Action, Christchurch Call, accessed October 2025.

32 Tech Against Terrorism, Terrorist Use of E2EE: State of Play, Misconceptions, and Mitigation Strategies, supra.

33 Danielle Citron, Extremist Speech, Compelled Conformity, and Censorship Creep, 93 Notre Dame L. Rev. 1035 (2018).

34 For e.g., see requirements for procedurally adequate moderation under the EU's Digital Services Act (Articles 16, 17, 20 and 22), the UK's Online Safety Act (Sections 20, 21, 31 and 32), and Australia's Online Safety Act RES Standard (s 16).

35 Many regulations propose adding backdoors for law enforcement to access encrypted messages, so this is an important clause to keep.

36 Mitts, supra.

37 Mariana Olaizola Rosenblat, Ayushi Agrawal & Isaac Yap, Online Safety Regulations Around the World: The State of Play and the Way Forward, NYU Stern Center for Business and Human Rights, April 2025.

38 HB00020F Bill Text, Texas State Legislature, accessed October 2025.

39 SB 7072 Bill Text, Florida Senate, 2021.

40 @AGPamBondi, post on X, September 16, 2025.

41 A Toolkit on Using Counterspeech to Tackle Online Hate Speech, Future Free Speech, accessed October 2025.

42 See, e.g., Joshua Garland et al., Impact and dynamics of hate and counter speech online, SpringerOpen, January 24, 2022.

43 Robert Faris et al., Understanding Harmful Speech Online, Berkman Klein Center for Internet & Society at Harvard University, December 2016. Catherine Buerger, Collective Counterspeech in Conflict and Fragile Settings: Risks and Possibilities, Dangerous Speech Project, April 6, 2022. Catherine Buerger, #iamhere: Collective Counterspeech and the Quest to Improve Online Discourse, Social Media + Society, December 2021.

44 Issie Lapowsky, YouTube's Push to Counter Toxic Videos with "Good" Creators, Wired, October 31, 2018.

45 Ryan Greer & Vidhya Ramalingham, The Search for Extremism: Deploying the Redirect Method, The Washington Institute for Near East Policy, February 27, 2020.

NYU Stern Center for Business and Human Rights
Leonard N. Stern School of Business

bhr@stern.nyu.edu
bhr.stern.nyu.edu

Center for Business and Human Rights