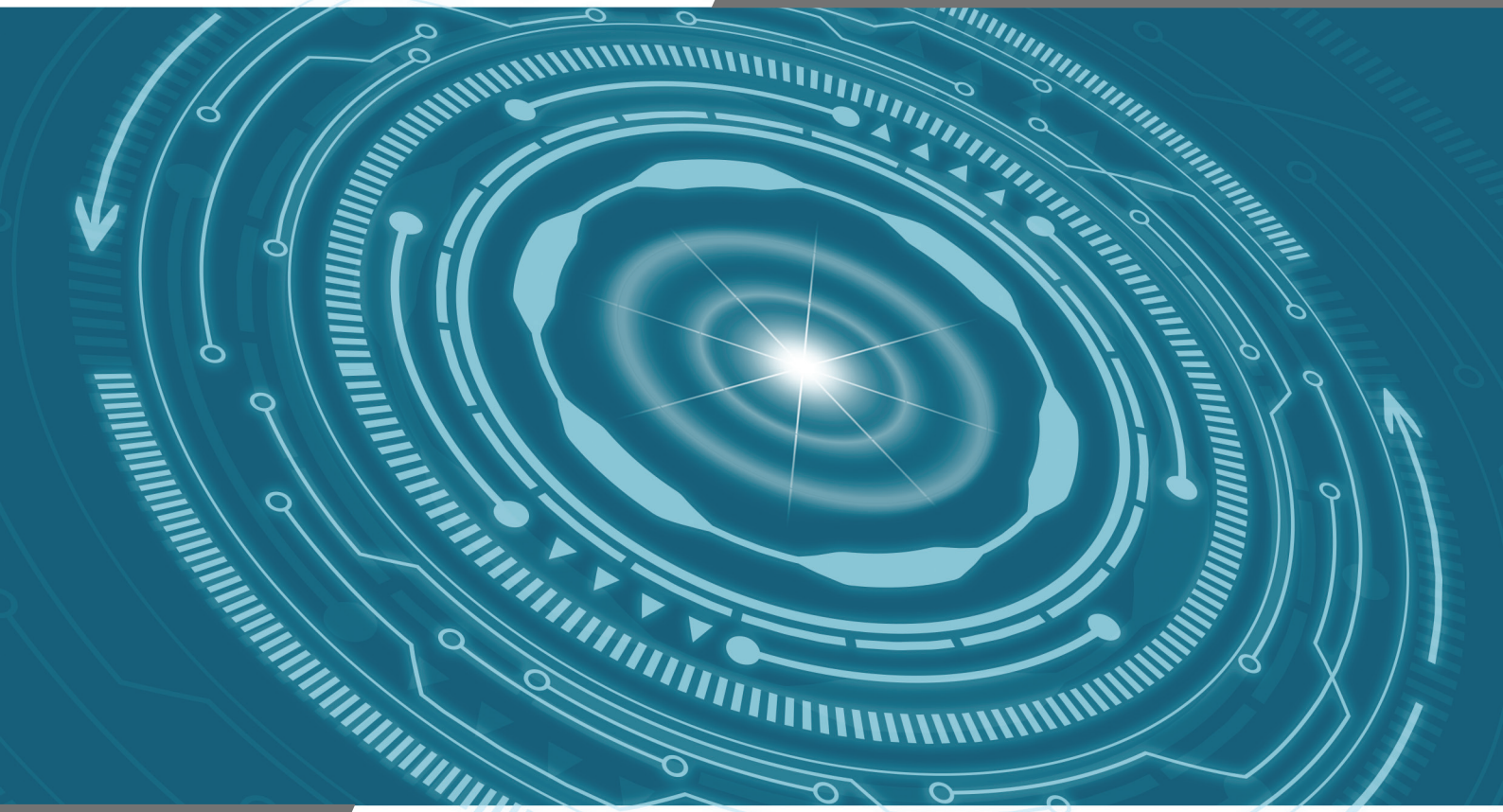
Centre for Economics and Foreign Policy Studies

edam

Cyber Governance and Digital Democracy 2010/01/EN



February 2020

EVOLUTION OF VIOLENT EXTREMIST AND TERRORIST THREATS ON SOCIAL WEB

Baris Kirdemir | EDAM & R. Bosch Cyber Policy Fellow

EVOLUTION OF VIOLENT EXTREMIST AND TERRORIST THREATS ON SOCIAL WEB

Baris Kirdemir | EDAM & R. Bosch Cyber Policy Fellow

INTRODUCTION

Social media represents an undeniable milestone in human civilization and progress. Yet, a constant battle takes place, either openly or behind the scenes, to save it from a committed crowd of malicious actors. Violent extremists and terrorist groups continue to develop new strategies to sustain their presence throughout online social networks. Twitter, Facebook, YouTube, and other social media platforms have employed various countermeasures in recent years. Automated and semi-automated systems utilize better algorithmic frameworks to effectively detect and eliminate terrorism, violent extremism, targeted hate speech, misogyny, racism, xenophobia, and a set of crimes including terrorism financing and arms smuggling. However, an evolving violent extremist ecosystem across the virtual infosphere still poses major national security threats.

This paper outlines current trends in the counterterrorism and counter extremism efforts online. The first section overviews the impact of actions taken by major social media companies, concentrating on content removal, suspension, and deplatforming. The second section explores the evolution of a broader extremist ecosystem online, with an emphasis on its adaptability and multi-layered structure. The third section then discusses the lone-actor terrorism and hate crimes, as well as the overarching connections between online and offline manifestations of violent extremist behavior. The final section surveys the transformation of ISIS's virtual presence, as it effectively illustrates the dynamic nature and resilience of the outlined security threats throughout the Internet.

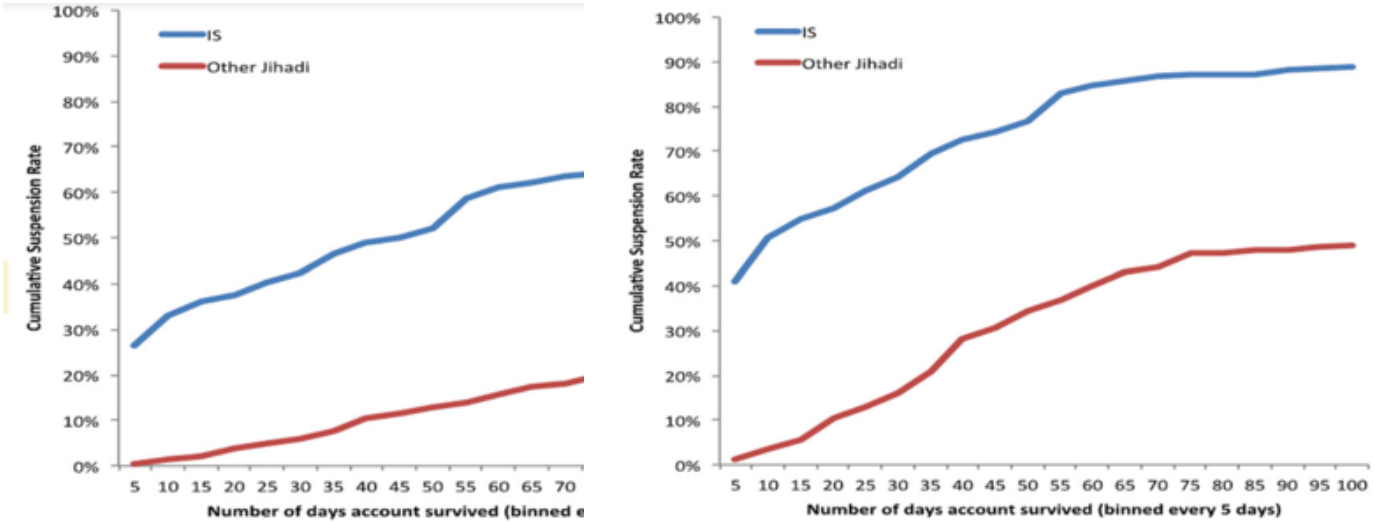
Robert Bosch Stiftung The research has been made possible by funding obtained from Robert Bosch Stiftung.

Content Removal: Is Social Web Safer Now?

Internet Organized Crime Threat Assessment (IOCTA) report by EUROPOL's European Cybercrime Centre (EC3) mentions extremist influence as a major security threat on the Internet.[1] Terrorist groups exploit new technologies and weaknesses in online service providers' (OSPs) systems. According to the report, extremists' early adoption of new tools and adaptability complicate the law enforcement and policy implementation by avoiding timely countermeasures. Smaller companies with fewer resources are especially vulnerable against swarms of extremists and organized criminals.

Telegram has been a platform of choice for many extremist groups. The broader impact of its recent promise to remove terror-related content remains to be seen.[2] Following the large-scale crackdown and suspension by mainstream platforms such as Facebook, Twitter, and YouTube, terror outlets and extremist groups diversify their operations the tools they use, as in the case of "decentralized" platforms. Thus, a cross-platform approach and collaboration of multiple stakeholders is a requirement for understanding the scope of online extremism and radicalization.

In recent years, major social media platforms have increased the extent of account and content removal to mitigate terrorist propaganda. In particular, the numbers and durability of pro-ISIS content on Twitter and YouTube decreased dramatically. Yet, the results of the ongoing content removal policies are nuanced and complicated. Firstly, studies suggest that countermeasures heavily concentrated on the pro-ISIS activity, while some other terror outlets continued to enjoy more freedom. Secondly, terrorist communication is not limited to major platforms such as Twitter. Often, such groups employ coordinated strategies and methods in a larger digital ecosystem.[3]



Conway et al. demonstrated the differences between the removal rates of terrorism-related accounts associated with ISIS and other jihadi groups on Twitter. Accounts other than ISIS survive significantly longer. Source: Maura Conway et al. Disrupting Daesh: Measuring Takedown of Online Terrorist Material and Its Impacts, Studies in Conflict & Terrorism, 42(1-2), 141-160, 2019.

1 IOCTA Internet Organized Crime Threat Assessment, EUROPOL European Cybercrime Centre, 2019.

2 Ibid.

3 Maura Conway et al. Disrupting Daesh: Measuring Takedown of Online Terrorist Material and Its Impacts, Studies in Conflict & Terrorism, 42(1-2), 141-160, 2019.

Social media companies' removal of terrorism-related content has its own limitations and side effects. The effectiveness seems to be mixed and is yet to be fully understood. Apart from the violent and pro-terror content that is easy to flag, it is often difficult to define a clear border line between legitimate content and others that violate the rules. Removal mostly relies on the automated or semi-automated analysis of the content itself, or the behavior of the accounts that spread the harmful material. Content-based decisions mostly rely on the analysis of several objects in a post, such as *"linguistic characteristics, word use, images, and URLs."*[4] Other signals include user complaints, ties to other suspicious accounts, bot-like behavior and so on. Companies use combinations of human content moderators and automated detection systems for removal. Smaller startups may have fewer capabilities and more challenges than major social media platforms.

Social media platforms' initiatives to counter terrorist, violent extremist, or hateful content and associated accounts vary in terms of their focus, methods they employ, and effectiveness. Often, these initiatives are subject to scrutiny. For example, YouTube's (Google/Jigsaw) "*Redirect Method Pilot Program*"[5], announced as a countermeasure against groups such as ISIS. aims to redirect users to "counter-narrative videos" when they see or search for relevant harmful content. The Counter Extremism Project (CEP), a non-profit organization specializing on counter-extremism and online radicalization, conducted a small study on the impact of Google's program and concluded that the results are nuanced. The researchers examined videos that were associated with ISIS, the Nusra Front, the Taliban, Hezbollah, and others, and concluded that the number of counter-narrative videos was significantly less than extremist videos in their sample.[6]

Although there are a number of published studies on radicalization and extremism on social media, they are mostly limited in terms of the sample size or the algorithmic factors they consider. Online radicalization and counter-terrorism are difficult subjects to study comprehensively. The real world impacts normally emerge in time and due to many intertwined factors. For example, the algorithmic effects of the YouTube recommendation and personalization systems as a gateway to radicalization is one of the hot-topics in the research community. However, the phenomenon is yet to be understood in detail.

Many Actors, Many Platforms: Why Extremist Infosphere is so Resilient?

Terrorist and violent extremist groups' online activities take place in an ecosystem of different types of platforms that facilitate direct or indirect communication, coordination, propaganda, planning, direction, financing, recruitment, purchases of weaponry and other types of inventory, and almost all other types of information utilization. Platforms in the information ecosystem range from open and centralized intermediaries such as Twitter, Facebook, or YouTube to encrypted channels and groups on Telegram, "decentralized" platforms, and to darknet[7] (dark web).

In particular, the use of **encrypted, decentralized**, or **darknet-based** platforms facilitate secrecy and further concealment of activities. This trend also complicates the efforts to monitor online activities that are associated with violent groups and cause significant security threats. Investigative reports show that facing the crackdown by major platforms, the tendency among extremist groups to use such alternative frameworks increase. Therefore, countering online terrorist and extremist activities is more complicated than the enforcement of terms of services by social media companies.

Within the given context, coordination is key. Mass

4 Isabelle van der Vegt et al. Shedding Light on Terrorist and Extremist Content Removal, RUSI. 2019.

5 Andy Greenberg, Google's Clever Plan to Stop Aspiring ISIS Recruits, WIRED, 2016, https://www.wired.com/2016/09/googles-clever-plan-stop-aspiring-isis-recruits/, Accessed on: January 10, 2020.

6 OK Google, Show Me Extremism: Analysis of YouTube's Extremist Video Takedown Policy and Counter-Narrative Program, Counter Extremism Project (CEP), 2018.

7 Nikita Malik, Terror in the Dark: How Terrorists Use Encryption, The Darknet, and Cryptocurrencies, Centre for the Response to Radicalization and Terrorism at The Henry Jackson Society, 2018.

propaganda mostly takes place in larger platforms that facilitate the amplification of a narrative to largest possible groups in the receiving end. However, other platforms, either encrypted, decentralized, or fully on the dark web, attract committed individuals, and are used to launch or coordinate actions in the first category of online social networks. Vice versa, posts in popular platforms sometimes encourage people to visit or move to closed channels to ensure continuity of communication. Most importantly, this information ecosystem is highly dynamic. Terror outlets and extremist groups try to remain adaptive while facing bans, massive content removals, crackdowns, and monitoring.

Darknet, also named as dark web, is the most closed and difficult to reach layer in the Web, coming after surface web and deep web categories. Although it hosts lower amounts of data and information flow, activities in the darknet are significant for several reasons. On the positive side, darknet have been used by journalists and human rights activists to ensure privacy and secrecy to overcome security threats or surveillance by totalitarian states that continuously abuse their citizens. On the other hand, the darknet facilitates many criminal and terrorist activities, as well as the coordination of violent terrorist groups. Although security agencies continue to monitor and act upon the information they gather on the darknet, anecdotal evidence suggests a greater requirement to focus on such activities.[8]

A recent report by the Centre for the Response to Radicalisation and Terrorism (CRT) at The Henry Jackson Society outlines several ways terrorist groups use or may use the darknet. In particular, the report emphasizes attack planning, coordination with open platforms, recruitment, direct interaction, indoctrination, propaganda layering, and financial coordination using other technologies such as cryptocurrencies as common practices.[9] Besides, the darknet further facilitates personal radicalization pathways that play a role in lone-actor terror attacks.

Moreover, terror groups and extremists particularly try to improve the utilization of "decentralized social networks". Usually, decentralized platforms are not governed by any central hosting company. Thus they further enable ways to avoid service terms, policies, and deplatforming. They are often "open-source" and can be installed and run on private servers. In addition, they can be scaled across multiple servers and connect to other hubs in a larger ecosystem. Some platforms move beyond server-based mechanisms and run on blockchain-enabled or P2P (peer-to-peer) frameworks to achieve even greater levels of decentralization. Blockchain and P2P allow a distribution of information flow across *"a global network of computers"*.[10] Therefore, decentralization increases terror groups' online resilience. However, the use of such platforms is not as large-scale as Twitter, YouTube, or Facebook, as they do not provide the same reach the popular platforms offer.

As briefly described above, online violent extremism, terrorism, and radicalization take place in a highly interactive digital information environment. Interaction and dynamism are the key features of such a vertically multi-layered and horizontally compartmentalized system. Yet, how **intergroup dynamics** function across the extremist networks has not been fully discovered as of today. There are various levels of interaction, coordination, rivalry, alignment, and influencing among similar-minded groups. On the other hand, the interaction between seemingly opposing groups and how they influence each others' behavior is less known but equally important. **Moreover, such opposite violent ideologies may even facilitate each others' resilience and survival.**

A few research projects have recently examined similar intergroup dynamics. To illustrate, in 2018, a group of researchers analyzed *"the interactional dynamics between anti-Muslim extremists and radical Islamists in Germany and beyond,"* offering *"direct evidence showing that Islamist and far-right movements converge at different levels and mutually amplify one another."* The amplification effect manifests itself through indoctrination, propaganda, and commitment inside each camp and that is observable in their activities on social media. Furthermore, the study outlined several resembling behavioral patterns and narrative characteristics within both groups: *"the demonization of enemies," "the victimization of one's own group",* and *"conspiracy."* [11]

8 Ibid.

9 Ibid.

10 Ben Pierce Peter King, Extremists Experiment with Decentralised Social Networks, Jane's Intelligence Review, 2019.

11 Maik Fielitz et al. Loving Hate: Anti-Muslim Extremism, Radical Islamism and The Spiral of Polarization, Institut Für Demokratie und Zivilgesellschaft (IDZ), 2018.

Findings of the study mentioned above are striking. The authors further demonstrate how seemingly opposing extremist movements learn from each other and their communication strategies, how they *"adopt"* each other's *"strategic references"*, and how a violent or hate-related real-world event amplifies radicalization and a continual spiral of radical messaging. For example, lone-actor terror attacks by far-right extremists motivate further hateful posts in jihadi communication channels and encourage actions to take revenge. These and a handful other factors lead to a *"mutual dependency"* and a *"symbiotic"* relationship between two camps.[12] Thus, endeavors to understand and counter online radicalization, violent extremism, and terrorism need to adopt a broad approach not only for the digital infrastructure but also for overarching and complicated intergroup influence dynamics.

A visualization of the connections between different tools used by terrorist propaganda outlets. Outlinks are often used to move the followers to different platforms and to avoid detection systems. ***Source:*** *RUSI*[13]

This paper does not adopt the view that the internet is the sole *"prerequisite"* for the contemporary non-state armed groups' terror campaigns and extremist groups to achieve a large base of committed followers, or it is the only *"pathway"* for radicalization. There are many other *"offline"* factors that interact in complex ways. Yet, as briefly introduced above, online social networks and digital infrastructure play major amplifier roles for such groups' communication efforts.[14]

Extremist networks tend to achieve adaptability and high survival rates as well as an undeniable ability to learn from others. A widely studied topic proving this phenomena is the international recruitment achievements of the terrorist outlet ISIS[15], which was able to strengthen its ranks and files with many recruits from the Western hemisphere. Beyond this overwhelming international focus, groups other than ISIS also possess similar characteristics and achieve varying levels of success in digital communication.

12 Ibid.

13 Ali Fisher, Nico Prucha and Emily Winterbotham. Mapping the Jihadist Information Ecosystem, RUSI, 2019.

14 Stefan Goertz and Alexander E. Streitparth. New Technology in the Hands of of the New Terrorism, in Stefan Goertz, Alexander E. Streitparth, The New Terrorism: Actors, Strategies and Tactics, 85-115, Springer, 2019.

15 Ibid.

Lone-Actors and Online Hate Speech: Is There a Connection?

Lone-actor terrorism is still a major security threat. Observations suggest that a significant number of lone actor terrorists interact with extremist information on the Internet before they carry out the attacks. Moreover, lone actor terrorism threat is also intertwined with other security challenges, such as the return of so-called "foreign figthers" from conflict zones.[16] Online messaging across extremist channels encourages the followers to carry out independent attacks throughout the world. ISIS, Al Qaeda, and other groups openly call for continuous terror attacks in Western countries. On the other end of the violent extremist spectrum, far-right extremists openly applaud mass atrocities on the platforms they use.

In sum, online social networks and other platforms on the Internet relate to lone-actor terrorist attacks in multiple ways. First, they multiply the dissemination of extremist information and terrorism propaganda. As outlined in other sections, this infosphere is adaptive to countermeasures and difficult to eliminate as a whole. Second, potential attackers are able to reach information about how to plan and prepare for the attacks, and to acquire weapons and other inventory. When they also possess a prior technical capability to hide their identities and location, they can also avoid some of the automated systems used by counter-terrorism agencies that constantly monitor violent threats. Third, psychologically, individuals develop strong ties to the wider network and their political cause, crossing a key milestone in their radicalization process. The sense of belonging develops in time while they consume the extremist information and interact with other like-minded people. Finally, if they also have a prior or current experience in interacting with their violent extremist network offline, such as in the forms of physical training or indoctrination, they potentially become even more capable in terms of the harm they can cause.[17]

According to the figures extracted by RUSI from a lone-actor terrorism database that covers the incidents in European countries between 2000 and 2014, the attackers' engagement with and use of mainstream social media platforms have increased in time. Such a finding is not surprising as the internet and online social networks had gradually become prominent globally within the same timeframe. On the other hand, details of the internet use by the lone-actor terrorists offer additional insights, especially for further research that can inform policy. Accordingly, *"two-thirds of the perpetrators (67 percent) had never been active in an extremist group."*[18] Most of the relationship on mainstream online social networks was *"one-way"*, as the lone-actor terrorists remained as consumers and amplifiers of information while had very limited direct interaction with other individuals. They used the internet for tactical research in 33 percent of the cases in the database, including *"downloading manuals, watching training videos, or undertaking basic reconnaissance".*[19]

Overall, lone-actor terrorists *"rely on"* the internet more than the individuals who operate in the organized non-virtual groups. Also, mental disorders, prior criminal records, behavioral similarities with mass shooters are among the general patterns.[20] Another study that focused on the UK-based lone-actor terrorism cases between 1995 and 2015 found similar results. Almost half of the cases included previous criminal records and approximately one-third of the individuals had *"a history of mental illness or personality disorder."* Other major themes included religiosity, ideological factors, and social isolation. **Over 87% consumed online extremist content, and almost 60% made virtual connections.** These numbers exceed real-world connections between the lone-actors and wider violent extremist networks outside platforms on the Internet.[21]

Two additional prominent trends of online terrorist behavior are related to the dissemination of the video footage and imagery that glorify armed attacks to the global audience. Terror groups have been using edited and curated footage

16 Raffaello Pantucci, Clare Ellis and Lorien Chaplais. Lone-Actor Terrorism: Literature Review, RUSI, 2015.

17 Ibid.

18 Clare Ellis et al. Lone Actor Terrorism: Analysis Paper, RUSI, 2016.

19 Ibid.

20 Paul Gill et al. What do Closed Source Data Tell Us About Lone Actor Terrorist Behavior? A Research Note, Terrorism and Political Violence, 2019.

21 Ibid.

of attacks for some time. For example, ISIS used high-quality videos extensively, especially during the peak of its presence in Iraq and Syria. Live streaming, however, is a newer phenomenon that triggers major shockwaves after major atrocities throughout the world. Most importantly, a little is known about whether such horrifying levels of violent coverage may cause further propagation of violence.

Live streams, relying either on bystanders' uploads or the use of mounted cameras and social media tools by the attackers themselves, have proven to be particularly difficult to eliminate in real time. The video of the Christchurch attack in New Zealand, for example, remained accessible across the Web long after the attack itself ended.

Most of the detection systems that target terrorism and violent extremism rely on prior data that sometimes do not provide useful input to identify live streams. A further complication is potential *"false positives"*. Detection systems that rely on machine learning falsely flag a small number of legitimate items as harmful. Given the enormous amounts of live streaming that takes place at any moment, most of the "detected" content would be irrelevant in a counterterrorism context.[22]

Online hate speech is another relevant category of mounting challenges and it is connected to hate crimes, terrorism, extremism, and violence. Studies show a reciprocal connection between offline events and online hatred, that is, hate speech at both individual and group levels is *"a process"* that takes place in a loop of online and offline crimes.[23] Therefore, online hate speech is tied to the general sociopolitical problems that need to be tackled.

Both online and offline, hate crime rates are often correlated with important events such as elections and terrorist attacks. On the other hand, studies suggest that *"online hate speech targeting race and religion and offline racially and religiously aggravated crimes"* are associated with each other even without such *"triggers."*[24] Even more troublesome, lone-actor attackers often shift their behavior from online hate speech to offline violence, as previously documented after the far-right terrorists attacks in the US, UK, Norway, and New Zealand. In other cases, longstanding hatred across online social networks can quickly turn into communal violence based on false information. Repeated lynching incidents in India are the prominent graphic examples in this category. Thus, it will be crucial in the near future to better understand the online-offline dynamics of hate and violence.

Apart from violence and crimes that appear to be correlated with online hatred, the direct effects of hate speech on victims are significant even without the offline connection. As previously documented, *"fear, anger, sadness, depression, and a newfound prejudice against the attacker's group, as well as physical effect including behavioral changes and isolation"* are among the common direct effects of online hate speech. Offline crimes *"intensify the effects"*[25] of severe hatred that occurs on virtual spaces.

Similar to extremism and malicious information operations, the digital ecosystem of hate speech is also resilient against countermeasures. According to a widely circulated paper by a group of scientists on global hate *"network of networks", "the current hate network rapidly rewires and self-repairs at the micro level when attacked"*. Moreover, platform-centric measures such as curbing hate speech content solely on Facebook may make the problem even worse. Accordingly, online hate clusters self-organize and evolve in time.[26]

Hate speech networks closely interact with other topic groups such as sports fans and also communities that speak different languages, enabling themselves to attract more members from other parts of the online infosphere. Also, hate-centric networks are strongly tied within, especially when a group identity is also formed. Followers of violent extremist groups such as ISIS or KKK develop strong communication bonds within their online networks. This similarity, despite the ideological differences, indicates the presence of a general behavioral pattern among the extremist groups.

Among global trends that are intertwined with online

22 Maura Conway and Joseph Dillon. Case Study Future Trends: Live-Streaming Terrorist Attacks?, VOX Pol, 2016.

23 Matthew L. Williams et al. Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggrevated Crime, Oxford University Press, Centre for Crime and Justice Studies (ISTD), 2019.

24 Ibid.

25 Matthew Williams. Hatred Behind the Screens: A Report on the Rise of Online Hate Speech, Mishcon Academy, 2019.

26 N. F. Johnson et al. Hidden Resilience and Adaptive Dynamics of the Global Online Hate Ecology, Nature, 573, 261-265, 2019.

extremism are ongoing hostilities and violence in different parts of the world, polarization across democratic societies, rise of populism, xenophobia, extremist political parties, and the prominence of aggressive, emotive political discourse. One can also follow the traces of economic inequality and evolving geopolitical tensions. Disinformation campaigns remain prominent on online social networks, threatening modern societies by altering the availability and factuality of information as well as the nature of conflict in general. In the digital information space, effects of algorithms that lead users to more and more emotive and extreme content are yet to be fully understood. All in all, the rise and resilience of extremism as well as the coordinated activities of terror groups across online social networks will continue to pose security challenges in the foreseeable future, in connection to the other troublesome trends.

Evolution of ISIS's Virtual War

In the last six years, a significant amount of research on online extremism has focused on ISIS and some other similar groups. The correlation between the conflict trajectory in the battlefield and the reach, productivity, and quality of online propaganda has become apparent as the wars in Syria and Iraq shifted to new phases. When the ISIS military and terror campaign reached its peak, its online propaganda and recruitment also were at unprecedented levels. The success and effectiveness of ISIS media operations attracted greater public attention within the same timeframe. However, the media operations and information maneuvers lost their previous momentum along with the military defeat in the battlefield. In the meantime, social media companies intensified their efforts to prevent the terror outlet's presence on their platforms.

That being said, ISIS has been attempting to adapt to new circumstances and extensive actions taken by social media companies by diversifying the tools and platforms it uses and transforming its coordination across the information ecosystem, as briefly outlined in the previous sections. Besides, while the countermeasures and content removal policies overwhelmingly focused on ISIS, other violent extremist groups and terrorist organizations continue to pose significant security threats. Exploratory studies suggest that the content such groups push through online social networks remain accessible for longer periods than pro-ISIS content and accounts.

Telegram has been one of the prominent intermediaries for ISIS to sustain its online propaganda, coordination, and recruitment efforts following the military degradation in the battlefield and the crackdown by other service providers. The platform has quickly become a communication hub of choice for terror outlets. Statistics indicate periodic increases in the numbers of activity on Telegram within the last two years.[27] In the meantime, during the intense fighting that led to its military defeat, the themes of ISIS's online messaging and narratives shifted from the utopia of the so-called caliphate, victimhood, and brutality against the enemy to war fighting and simply proving its continued military relevance. Similarly, ISIS's propaganda characteristics further changed during the same time period, failing to create and disseminate *"non-Arabic language magazines since its loss of Raqqa."*[28] The terror group lost its production facilities and personnel along with the territorial hold. This trend was also observable in the numbers of high-end, well curated videos that were once central to its online propaganda and psychological warfare.[29]

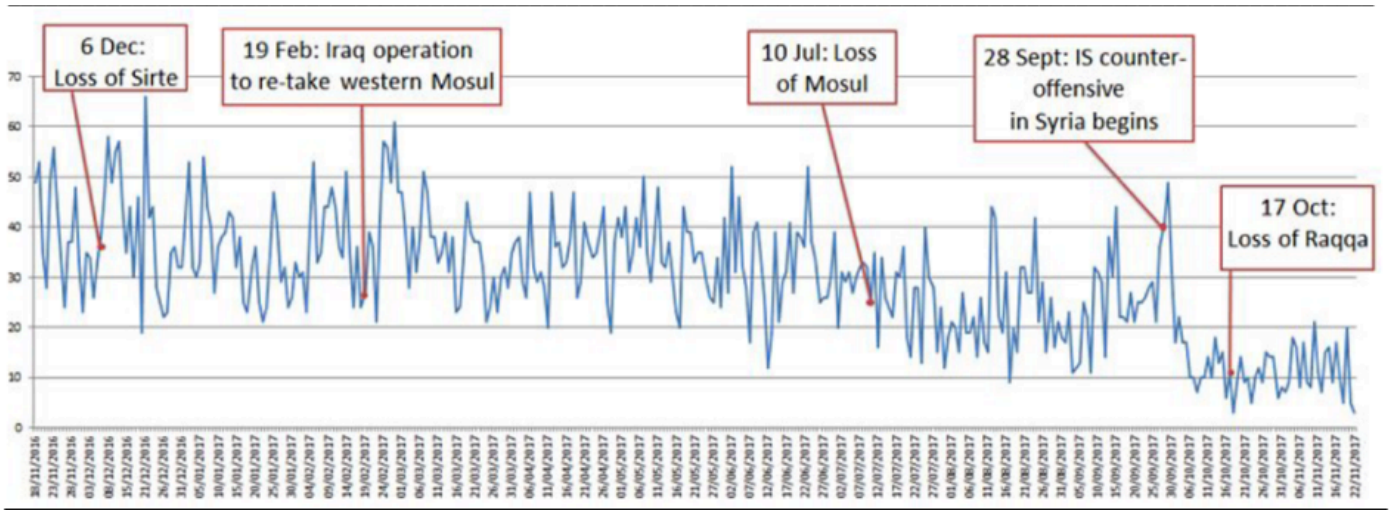
To attract followers and readership to Telegram-hosted channels and groups, violent extremist groups post links on mainstream social media, including Twitter, YouTube, and Facebook. Analyses of ISIS's Telegram activity demonstrated that the group posts outinks to a diverse list of open sites including YouTube, Google Drive, JustPaste. It, Google Photos, Sendvid, Archive.org, Archive.is, and Medium.[30] It shows the fact that even ISIS, facing limitations resulting from the large-scale countermeasures, can still use such open tools to run its communication operations.

27 Maura Conway, Violent Extremism and Terorism Online in 2018: The Year in Review, VOX Pol, 2019.

28 Ibid.

29 Ibid.

30 Maura Conway and Michael Courtney, Violent Extremism and Terrorism Online in 2017: The Year in Review, VOX Pol, 2018.



The daily media output of ISIS changed during the intense battles in Syria and Iraq. Source: VOX Pol and BBC Monitoring [31]

According to a study by the Institute for Strategic Dialogue (ISD), ISIS and its supporters were able to run a wide-scale propaganda campaign right after the elimination of its leader Abu Bakr al Baghdadi in Syria by the US Special Forces. The network of pro-ISIS accounts, called the "Baghdadi Net' by the author, used various tactics such as hijacking trending topics and mentioning popular persona on Twitter. A significant amount of accounts were either automated or semi-automated (bots and cyborgs) according to ISD, with many new accounts *"regenerating"* on a daily basis. The extent of this effort to stay resilient was so high that *"new accounts were being launched every five minutes."* Most significantly, the study suggests that Twitter, as one of the most popular and capable companies, *"is still struggling with how to deal with terrorist accounts on its platform,"* especially in non-English languages.[32] To note, Facebook, YouTube, and Twitter have been curbing such inauthentic and inorganic behavior for several years. The findings of the study mentioned here are disturbing as they indicate a significant capability gap in detection systems that are being employed.

Some evidence suggests intergroup rivalries between jihadist groups to dominate the social media space as an important catalyst that enabled ISIS's success on the Internet. Effective use of online social networks by terror groups such as ISIS and Jabhat al-Nusra goes back, at least, to the beginning of the previous decade. Most often, the characteristics and nature of the communication dynamics on social media and the new era of hyperconnectivity are mentioned as the primary factors behind the strategic communication capabilities of violent extremist groups. According to this tech-centric view, easy, fast, and accessible nature of the virtual hyperconnectivity leads to such side effects.

The complete picture is probably more complicated than that. Social, political, and conflict-related dynamics in the offline environment often interact with the online developments in different ways. For example, according to recent evidence, intergroup rivalries between terror organizations may play major roles in the effectiveness and success of their online communication efforts. Evidence shows that in 2013-2014 period ISIS and Jabhat al-Nusra, engaged in a fierce rivalry within the battlefield and broader jihadist ecosystem, focused on their information campaign on Twitter partly because they did not want to leave the entire platform and the propaganda opportunities to the competitor.[33] In sum, both terror outlets pursued the ownership of the jihadi agenda across online social networks.

The number of studies on pro-ISIS information campaigns increased significantly in recent years. However, there are still open questions about how much, how wide-scale, how far, and how effective the pro-ISIS communication was able

31 Ibid.

32 Moustafa Ayad. 'The Baghdadi Net': How A Network of ISIL-Supporting Accounts Spread Across Twitter, Institute for Strategic Dialogue (ISD), 2019.

33 Gunnar J. Weimann. Competition and Innovation in a Hostile Environment: How Jabhat Al-Nusra and Islamic State Moved to Twitter in 2013-2014, Studies in Conflict & Terrorism, 42(1-2), 25-42, 2019.

to influence others on a major platform such as Twitter. This limitation emanates from the difficulties in acquiring the data that cover the entire pro-ISIS activity. One of the studies that was able to analyze such a large-scale dataset (platform-specific) reached important conclusions about pro-ISIS Twitter activity and its overall impact that took place in 2015. First of all, the researchers documented high-level activity of pro-ISIS accounts, which pushed significant amounts of Tweets before being suspended. On the other hand, analyzing more than 340 million Tweets and 173 thousand accounts, the study suggested that the accounts used by ISIS had limited influence in overall Twitter space, despite the high levels of activity. The evidence demonstrates that, in the mentioned dataset, most of the engagement to ISIS-led Twitter activity originated from other pro-ISIS accounts, which were also suspended eventually by the platform.[34]

Furthermore, in terms of the pro-ISIS reach and activity on Twitter, the crackdown and aggressive countermeasures worked by limiting the groups' online influence.[35] However, as mentioned in other sections, platform-specific studies are able to explore only a small fraction of an adaptive system that develops varying communication strategies in an ecology of many platforms and tools. Moreover, as a general complication of social media analytics, how we define and measure the impact of such malicious activities may define the conclusions. With the given analytical tools and techniques, it might be the right approach to measure the influence through measurements of retweets, likes, and shares, but it still remains short of showing the overall visibility and cognitive impact of extremist communication among different groups of people.

Pro-ISIS accounts also have been active on Facebook. Some experts suggested that suspension rates of pro-ISIS accounts on Facebook are lower than other platforms. Moreover, the content and account removal process may be slower, which in turn can increase the durability of terror-related posts. In sum, what is known is that operators and supporters of ISIS originate from many different countries, their posts are in many different languages, and they utilize fake accounts at great scales. Overall, ISIS enjoyed a *"global support network"* on Facebook[36], with distinct local communities that connect to each other via influential "propagandist" accounts.

Pro-ISIS accounts demonstrate coordinated propaganda activities on Facebook. The groups and accounts are usually removed by the platform. However, they are able to regenerate and continue to coordinate with a global network of supporters. Source: Gregory Waters and Robert Postings, Spiders of the Caliphate: Mapping the Islamic State's Global Suppoort Network on Facebook, Counter Terrorusm Project (CEP), 2018.

34 Majid Alfifi et al. A Large-Scale Study of ISIS Social Media Strategy: Community Size, Collective Influence, and Behavioral Impact, Proceedings of the Thirteenth International Conference on Web and Social Media, Vol 13, 2019.

35 Ibid.

36 Gregory Waters and Robert Postings, Spiders of the Caliphate: Mapping the Islamic State's Global Suppoort Network on Facebook, Counter Terrorusm Project (CEP), 2018.

As mentioned in earlier sections, ISIS content has been visible on YouTube. An illustrative analysis by the Counter Extremism Project (CEP) in 2018 found that *"hundreds of ISIS videos are uploaded to YouTube every month, which in turn attract thousands of views."* Videos that were collected and analyzed in the mentioned study were specifically terror-related and created by either the operators or highly committed supporters of the group. Out of 1,348 videos CEP collected, only 24% stayed online over two hours. Nevertheless, they created more than 160,000 views in three months. Moreover, *"60 percent of accounts remained live after uploaded videos had been removed for content violations."*[37]

Correlation between the extremist activity on YouTube and radicalization as well as real world terror attacks is a long-lasting hot topic among the relevant policy and research communities. A significant number of people who charged with terror-related crimes in Western countries reported watching terrorist propaganda content on YouTube and some other platforms prior to committing criminal actions.

Major social media companies' approaches to counter extremism and terrorist propaganda on their platforms have evolved in response to public scrutiny and violent attacks. Also, there is some collaboration among major companies to counter terrorist propaganda. Google (YouTube), Facebook, and Microsoft announced a decision to use a shared database to detect and remove terror-related content. On the other hand, there are many remaining unknowns regarding how such systems function and whether they focus on specific groups more than others. Moreover, as mentioned before, what types of online messaging leads to violence, how radicalization "pathways" work, and where the borderline between harmful and legitimate activities is located remain as difficult many-piece puzzles.

The anecdotal and platform-specific evidence we refer to in previous paragraphs does not even remotely capture the scope of the extremist and terrorist infosphere. The digital ecosystem in which creators, disseminators, and receivers of extremist content interact has been growing, covering a large spectrum of online social networks, blogs, file sharing systems, and many other types of platforms on surface, decentralized, deep, or dark web layers. Jihadist groups utilize a constantly evolving ecosystem to more effectively conduct communication and influence their target audiences. As mentioned in earlier sections, Telegram has become an important hub for such groups, not only to communicate with their followers and spread direct and straightforward propaganda, but also to coordinate information campaigns and maneuvers on other platforms, including Twitter, YouTube, Facebook on the mainstream side, and decentralized or dark web platforms on the other.

A recent comprehensive analysis by RUSI attempts to capture a broader picture of the jihadist information ecosystem mentioned above. Accordingly, accounts on Facebook, Telegram, and Twitter mostly use outlinks to encourage their followers to move to other platforms. That being said, the amount of posts on these three platforms still constitute the largest *"sources of the traffic."*[38] Therefore, by adopting this tactic of using the major platforms to signpost other mediums where the actual content is located, extremist groups avoid detection systems and rapid removal of the posts they promote. In the meantime, such malicious groups constantly regenerate the content and accounts when they are suspended or removed. Furthermore, use of non-English languages, mostly Arabic for jihadist groups, in the form of simple text, PDF, or Microsoft Word documents further complicates the tasks of detection systems.[39]

Telegram, as an online social network and messaging app that prioritizes encryption and privacy, has attracted a significant amount of users worldwide. The features of Telegram has eventually made it the central communication hub for many terrorist organizations and extremist groups. The platform recently announced a large-scale crackdown against terrorist content and accounts to reverse the trend and avoid an increasingly negative image. The effects of Telegram's platform-wide crackdown are still developing, with some clues already being revealed. For example, researchers specializing on online extremism and ISIS propaganda point out that the group is experimenting with alternative platforms to operate on, while openly promoting the new mediums to their committed sympathizers,

37 The EGLYPH Web Crawler: ISIS Content on YouTube, Counter Extremism Project, 2018.

38 Ali Fisher, Nico Prucha and Emily Winterbotham. Mapping the Jihadist Information Ecosystem, RUSI, 2019.

39 Ibid.

propagandists, and eventually broader target audiences. Alternative smaller platforms such as TamTam and Hoop messenger are only two of the new tools in the evolving information ecosystem.[40]

Secondly, from a broader perspective, experts also emphasize potential unintended consequences of deplatforming and aggressive enforcement of policies that aim to mitigate such malicious activities. On the cognitive and psychological side, by following the trends and also directions from such groups' operators, the sense of commitment and belonging may strengthen among the supporters. People who constantly move between different platforms in a jihadist information ecosystem may become even more committed to "the cause" that drives their online and offline behavior,[41] promoting themselves from simple listeners to participants. Therefore, although deplatforming and enforcement of the terms and policies disrupt the malicious networks online, they also lead to further complications in counterterrorism and counter-extremism efforts.

ISIS has been disseminating propaganda on a decentralized platform called RocketChat since 2018 as a *"potential back-up"* to Telegram.[42] As mentioned above, ISIS and other terror groups try decentralized platforms as they offer more secrecy and resilience against scrutiny or large-scale deplatforming. When Telegram intensified the removal of terror-related accounts and channels, ISIS's Nashir propaganda outlet encouraged its supporters to join the channels on the new platform.[43]

RocketChat and similar open-source communication networks allow users to host the entire network on their own servers, with the option to stay connected to other nodes in the ecosystem. Separate hosting prevents the control and takedown of the network by a central host that would enforce previously accepted terms and policies. Thus, decentralization creates even more challenges to counterterrorism efforts online, compared to Twitter, Facebook, YouTube, and Telegram, which, despite limitations, are technically able to take action against terrorist propaganda that takes place on their platforms. ISIS and other terror outlets are expected to use similar tools in the near future.[44]

Policy Implications

The ongoing cooperation between various major social media companies, government agencies, the research community, and non-governmental organizations at national and international scales partially limited the visibility of violent extremist content and terrorist propaganda on mainstream platforms. However, the ecosystem of the mentioned online threats is highly dynamic and it continues to evolve.

In particular, such groups experiment with, and eventually employ, an increasing number of tools and platforms. Violent extremist groups have proven their resilience and a consistent capability to adapt to changing circumstances while facing the most comprehensive countermeasures to date. To ensure the safety of social media and digital information environment, policy priorities should better address the dynamic cross-platform nature of the threats this report briefly summarized.

Due to the combination of parameters outlined above, there is no silver bullet solution to end terrorist and violent extremist activities all over the Internet, especially in the short term. In particular, blocking, suspending, removing, criminalizing, or hacking the platforms and sources that disseminate harmful content are extremely unlikely to end with a decisive victory in favor of counterterrorism agencies. To complement such measures and the potential new steps we have forecasted in the previous paragraph, the stakeholders should improve their efforts to implement a comprehensive and sensible strategic communications strategy.

40 Amarnath Amarasingam, Telegram Deplatforming ISIS Has Given Them Something to Fight For, Vox POL (Web), 2020, https://www.voxpol.eu/telegram-deplatforming-isis-has-given-them-something-to-fight-for/, Accessed on: January 10, 2020.

41 Ibid.

42 Peter King. RocketChat Platform Offers Potential Telegram Back-Up for Islamic State, Jane's Intelligence Review, 2019.

43 Ibid.

44 Ibid.

The proposed strategic communications strategy should address both online and offline factors that relate to the set of security threats outlined in this report. Online and offline worlds are densely connected. Terror attacks, hate crimes, lone-actor terrorism, intense and high-stakes political campaigns, and other significant security-related political events correlate with the social behavior online. Furthermore, violent extremist and terrorist groups utilize existing sociopolitical vulnerabilities. Thus, the formation and implementation of strategic communications should cover a wide area ranging from interagency and intersectoral collaboration to understanding the vulnerable groups, evolving challenges, threat monitoring, communicating, and, finally, to fighting the malicious groups online.

Similar to other relevant practices in counterterrorism and counter-extremism, online efforts should "always" prioritize the uninterrupted continuation of democratic processes, sense of security, and accessibility of reliable information. Authoritarian governments tend to implement drastic measures such as cutting off the entire Internet when they face imminent security threats or confront mass protests. In addition, many governments throughout the world create their own disinformation ecosystem to use "noise" as a countermeasure against internal or external "hostile actors" and strengthen their grip of the domestic information environment.

Both directions eventually lead to severe problems. As the evidence this report outlined suggests, seemingly opposing extremist ideologies may mutually reinforce each other, creating a feedback loop that further strengthens the cycle of radicalization. Accessible and truthful information is at the core of any sociopolitical system that relies on democratic fundamentals and a social contract. Using a different kind of extremism and facilitating further polarization of society remains the most counterproductive policy option with potential destructive outcomes.

Cyber Governance and Digital Democracy 2020/01/EN

February 2020

EVOLUTION OF VIOLENT EXTREMIST AND TERRORIST THREATS ON SOCIAL WEB

Baris Kirdemir | EDAM & R. Bosch Cyber Policy Fellow