tech against terrorism

REPORT

PATTERNS OF ONLINE TERRORIST EXPLOITATION

TCAP INSIGHTS | APRIL 2023

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report evaluates the distribution of terrorist content across different platform types, sizes and locations between 25 November 2020 and 19 January 2023. The report draws on data taken from Tech Against Terrorism's Terrorist Content Analytics Platform (TCAP), the world's largest database of verified terrorist content utilised by tech platforms. This report provides an assessment of at-risk tech platforms and offers recommendations to improve the response to terrorist exploitation of the internet.

Our data analysis reveals that terrorists continue to exploit a wide range of smaller tech platforms for propaganda dissemination, with some platform types highlighted as especially vulnerable to exploitation. File-sharing and archiving platforms have become a particular area of concern: file-sharing platforms are most heavily exploited when assessed by the volume of content identified, although they prove effective in removing it, whereas archiving platforms, which host less content by volume, are least effective at removing terrorist content from their services. We provide specific guidance for these at-risk platforms. Policymakers too have a part to play in improving responses to terrorist exploitation of file-sharing and archiving services.

In the first study of its kind, we assessed terrorist exploitation by platform size, factoring in both a platform's user base and its capacity for content moderation as determined by its number of employees. This revealed a positive correlation between the size of content moderation teams and removal rates of terrorist content. We therefore call on policymakers to provide greater support for small tech platforms with fewer resources by supporting initiatives which are targeted to help smaller platforms, as well as by incorporating risk assessments for tech platforms when legislating for corporate regulatory liability for terrorist content.

Terrorist Content Analytics Platform

Launched in 2020, the Terrorism's Terrorist Content Analytics Platform (TCAP) is a secure online tool that detects and verifies terrorist content and then alerts technology companies to the presence of such material on their platforms. In the two and a half years since its launch, the TCAP has had an incredible impact on countering terrorist use of the internet, alerting over 100 different tech platforms to over 25,000 pieces of terrorist content, of which 94% is now offline.

KEY RISKS AND RECOMMENDATIONS

PLATFORM CATEGORY	RISKS	RECOMMENDATIONS
Small	• Small and medium-sized tech platforms were the most highly exploited tech platform sizes, based on the volume of content identified on their platforms. Smaller platforms also averaged a lower removal rate of alerted terrorist content than large and medium-sized platforms. • Earlier stage tech platforms — those with 0-50 employees — averaged a higher volume of terrorist content on their platforms than platforms with more employees. Earlier stage platforms also averaged a lower removal rate of alerted terrorist content.	• Publish information about the platform's resources and capacity, particularly regarding the capacity of moderation and Trust and Safety teams. Platforms could include this information in transparency reports, 'About' and 'Support' pages, or 'Help Centres'. • Encourage user reporting within the Terms of Service and/or Community Guidelines.
File-sharing	• Over half of the tech companies on which we identified terrorist content are file-sharing platforms (106 out of 187 platforms). The dispersal of terrorist content across file-sharing platforms makes targeted intervention to support tech platform moderation (e.g., through TCAP alerts) more difficult as it requires engagement and cooperation from a larger number of platforms. • File-sharing platforms were by far the most exploited platform type, with 28,724 URLs of terrorist content (72%) submitted to the TCAP.	• Prohibit terrorist content in line with international designation lists – such as the United Nations Security Council list – and/or in line with comparable lists from democratic countries such as the UK, Canada, and the US. • Consider implementing automated content moderation processes, with support from Tech Against Terrorism's Knowledge Sharing Platform. When doing so, human rights should be safeguarded, for example by means of hash-based detection methods.
Archiving	• Archiving sites were the second most heavily exploited platform type based on the volume of terrorist content identified (4795 submissions or 12% of total). • Only archiving sites were exploited to a significant extent by both Islamist and far-right terrorist actors. These sites perform a similar function for both ideologies, allowing their content to maintain a stable presence online. • Archiving sites had the lowest removal rate of alerted terrorist content at 57%, meaning 1687 URLs (out of 3936 URLs) remained online at the time of writing.	• Moderate terrorist content in line with national and international designation lists. • Use simple detection tools that flag key words and logos linked with terrorist entities for review. • Make archived terrorist content private and consider screening applicants to ensure the safeguarding of users, as well as screening the content.

PLATFORM CATEGORY	RISKS	RECOMMENDATIONS
Messaging Platforms	• The majority of far-right terrorist content was found on messaging platforms (52%). • Far-right terrorist actors have migrated towards more niche platforms in the past few years, partly due to increased moderation on larger platforms as well as the anonymity and audience reach provided by some of these alternative platforms. • In their role as beacons to far-right terrorist content, messaging platforms are uniquely placed to act quickly to remove this content before it can be widely disseminated. However, the average removal rate for messenger platforms was one of the lowest of all platform types, at 67%.	• Prohibit violent extremism in the Terms of Service and/or Community Guidelines. • Uphold human rights principles when engaging in content moderation practices, with particular concern for users' right to privacy online. • Maintain a list of the key words, phrases, images, etc. which are linked with far-right terrorist entities for content moderation purposes.

General Recommendations

- Make use of free tools to disrupt terrorist use of the internet, such as the Terrorist Content Analytics Platform.
- Provide an easily accessible contact point for users, law enforcement, governments, and initiatives such as the TCAP.
- Obstruct the dissemination of URL banks by utilising behaviour-based cues and moderation techniques already in place for spam material. Behaviour-based cues include abnormal posting volume, which can be picked up more easily by automated systems and require less human review than scanning outlinks for terrorist-related content.
- Use simple detection tools that flag key words and logos linked with terrorist entities for manual review.
- Mitigate against content moderation evasion tactics by introducing two-factor authentication and time limits to the validity of join links to private servers and channels.
- Consider involvement in cross-platform initiatives to counter terrorist use of the internet, such as the Christchurch Call to Action.
- Consider off-platform factors (persons on the platform engaging in severe abuse off service and/or on other platforms) when making moderation decisions related to terrorist content and behaviour.

Recommendations for Policymakers

- Consider smaller tech platforms when introducing online regulation and legislation. Policymakers should also commit to better supporting newer platforms whose user base rapidly grows beyond their content moderation means.
- Consider tech platforms' resources when categorising platforms by size.
- Incorporate risk assessments as a first step to online regulation, to aid platforms' understanding of the threat they are facing from Terrorist and Violent Extremist (TVE) actors.
- Implement targeted initiatives to support file-sharing, archiving, and messaging platforms.

INTRODUCTION

In November 2020, with support from Public Safety Canada,[1] Tech Against Terrorism launched the Terrorist Content Analytics Platform (TCAP),[2] a secure online tool that detects and verifies terrorist content and then alerts technology companies to the presence of such material on their platforms. In the two and a half years since, the TCAP has had an incredible impact on countering terrorist use of the internet, alerting over 100 different tech platforms to over 25,000 pieces of terrorist content, 94% of which is now offline.

Despite the TCAP's successes to date, there remains significant scope for improving the impact of the tool. This is the first research report harnessing the analytical power of the TCAP's unique dataset to identify at-risk tech platforms and provide them with data-driven recommendations to improve their response to terrorist exploitation of their services. This report seeks to understand patterns of online terrorist exploitation through three data sets:

1) By tech platform type

2) By tech platform size

3) By geographic hosting of terrorist content

We examine our findings firstly by providing an analysis of the data and highlighting at-risk platform sizes, types, and locations. The second part of the report delves into the implications of our findings, in particular for tech platforms and policymakers.

Finally, we outline our strategy for the future, with especial concern for how the report's findings can drive TCAP's development and policies, to better support subscribing platforms that are most at-risk.

1 Tech Against Terrorism (2019), Press release: Tech Against Terrorism awarded grant by the Government of Canada to build Terrorist Content Analytics Platform.

2 Terrorist Content Analytics Platform.

ANALYSIS OF DATA FROM THE TERRORIST CONTENT ANALYTICS PLATFORM

Understanding Terrorist Exploitation by Tech Platform Type

Terrorists exploit a wide online ecosystem of tech platforms to disseminate propaganda, spanning many different platform types which vary in purpose and functionality. Indeed, terrorist actors rely on a multiplatform approach to ensure both the rapid sharing of content and a resilient online presence.[3] To date, the Terrorist Content Analytics Platform (TCAP) has identified terrorist content on 13 different types of platforms. The table below (*Figure 1*) highlights these platform types and their respective core functionality. Where a platform has more than one functionality in practice, we examined the platform's own branding, as well as the main purpose for which it is exploited by terrorists.

Platform Type	Functionality Provided
File-sharing	Access to digital media such as photos, videos, and documents.
Archiving	Storage of information from defunct webpages or documents for anyone to view publicly.
Forum	Discussion site for conversations in the form of posted messages.
Video-hosting	Posting videos online.
Video-sharing	Uploading, conversion, storage, and later consumption of video content.
Link Shortener	Conversion of any URL into a shorter, more readable link.
Social Media	Creation and sharing of information through virtual communities and networks.
Messaging	Online chat in real time with individuals or larger groups and communities.
Photo-sharing	Uploading, conversion, storage, and later consumption of photo content on the internet.
Audio Streaming	Uploading, conversion, storage, and later consumption of audio content on the internet.
Paste Site	Uploading and sharing of text online, often used for sharing source code.
Search Engine	Performing web searches using key words or phrases.
Book Subscription	Subscription to officially published and user-published books and documents.

Figure 1: Types of platforms exploited by terrorists and alerted by the TCAP.

3 Ali Fisher, Nico Prucha, Emily Winterbotham (2019), "Mapping the Jihadist information ecosystem: Towards the next generation of disruption capability", *Global Research Network on Terrorism and Technology,* Paper No. 6

Analysis: Quantifying tech platforms by type

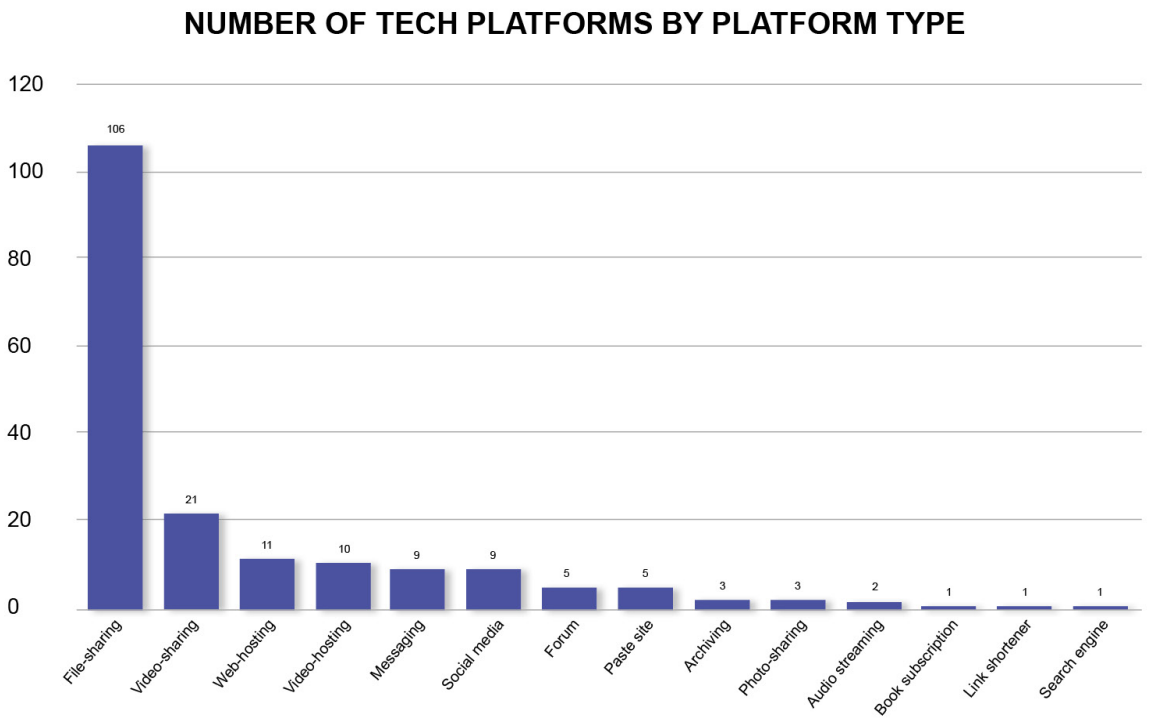

Figure 2: Number of tech platforms by platform type on which terrorist content has been identified.

File-sharing platforms were by far the most common type of platform on which we found official terrorist content, more than five times the second most common (video-sharing, see *Figure 2* above). In fact, over half of the platforms on which we identified terrorist content were file-sharing in their functionality (106 out of 187 platforms). File-sharing sites are used by terrorist actors to host content such as text, images, and videos, which can then be accessed through aggregated outlinks on beacon platforms.[4]

Given terrorist content is spread across a wide number of different file-based platforms (106 platforms are counted in our dataset), the challenge is one of both volume and dispersal. This makes targeted intervention to support tech platform moderation (e.g., through TCAP alerts) more difficult as it requires engagement and cooperation from a larger number of platforms. Additionally, these platforms lack in-site search functions, meaning users can only access the terrorist content via a search engine or outlinks from elsewhere. This might mean that the general public are less likely to accidentally come across terrorist content, but it also makes platform moderation through user reporting less likely.

Terrorist exploitation of online services for disseminating videos was dispersed across a wide range of different platforms in our dataset (31 different video-sharing or video-hosting platforms). Video-sharing platforms are particularly attractive to terrorist actors due to search functions which allow wider audience reach and typically large file size limits.

Conversely, terrorist content was found on only a very small number of different archiving and pasting

4 Beacons act as centrally located lighthouses that signpost viewers to where content may be found, which is often done through outlinks posting to content stores. Terrorists often use these beacon platforms and have official channels on them which aggregate their central communications.

platforms (3 and 5 respectively). However, a large volume of terrorist content was found on these sites (12% and 5% of total TCAP submissions respectively) meaning terrorist content was heavily concentrated on a small number of platforms. This, in theory, suggests targeted intervention and successful engagement with these platforms would have a significant impact on reducing terrorist content online.

Archiving and pasting platforms are likely to be popular with terrorist actors due to their multifunctional nature. Archiving sites are used to aggregate outlinks to content stores as well as providing access to historic content stores following removal by moderators.[5] It is likely that terrorist actors are abusing the purpose of archiving sites as content preservers to host terrorist content on them for indefinite periods of time. Meanwhile, pasting sites are used to store content and aggregate information, such as lists of URLs which link to further content and are not immediately identifiable as necessarily terrorist in nature.

Analysis: Volume of terrorist content across platform types

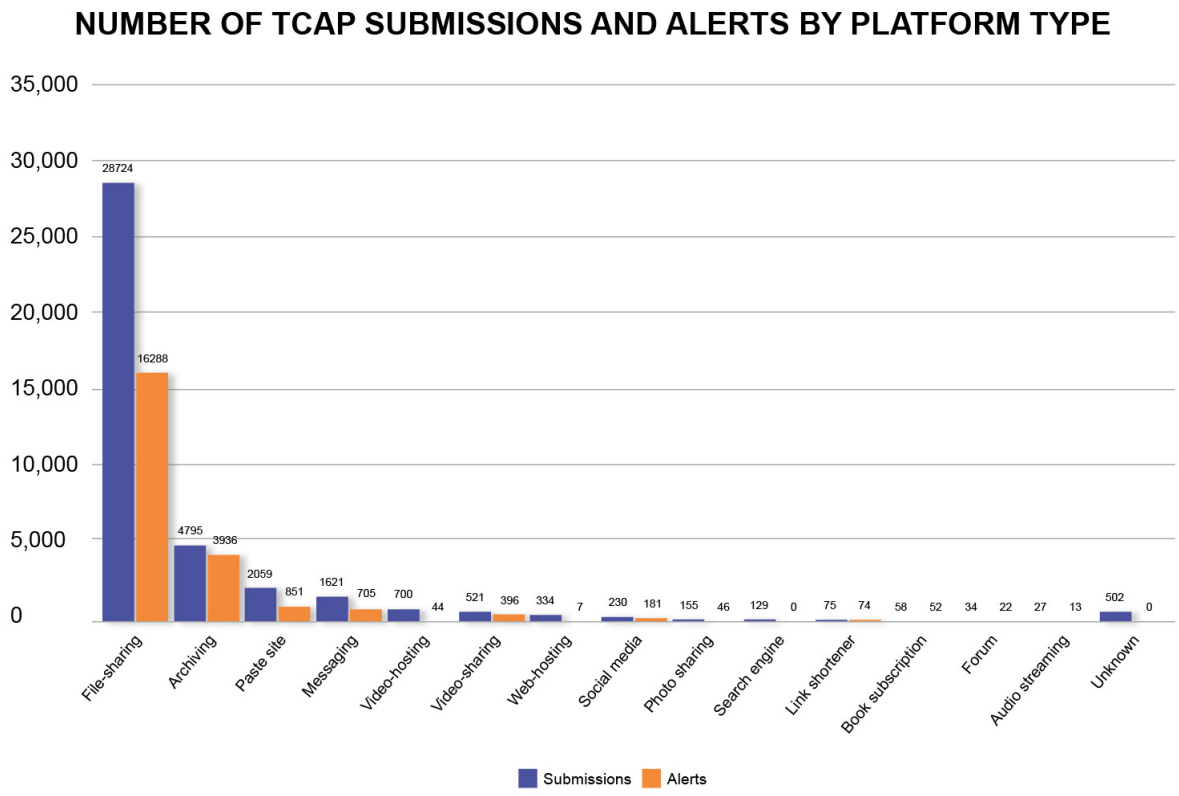

Figure 3: Number of TCAP submissions and alerts by platform type.

File-sharing platforms were by far the most exploited platform type in terms of volume of content, with 28,724 URLs submitted to the TCAP, representing 72% of all submissions. This was followed by archiving sites with 4795 submissions (12%) and paste sites with 2059 (5%). For file-sharing platforms, only 16,288 alerts had been sent out of 28,724 submissions (57%) meaning over 10,000 URLs containing terrorist content have not been notified to file-sharing platforms.

5 'Content stores' are where terrorist content is stored, including text and audio files, as well as images and videos. These are used as online libraries of content. Terrorists rely on content storage platforms and pasting sites, as well as archive services.

One explanation for this discrepancy is the removal of the content prior to submission to the TCAP, hence alerts not being sent. However, this gap is also due to the TCAP team being unable to get in contact with platforms due to a lack of accessible contact information, which in turn impedes subscription to TCAP alerts. Based on this data, Tech Against Terrorism is prioritising engaging with these platforms through a concerted outreach campaign.

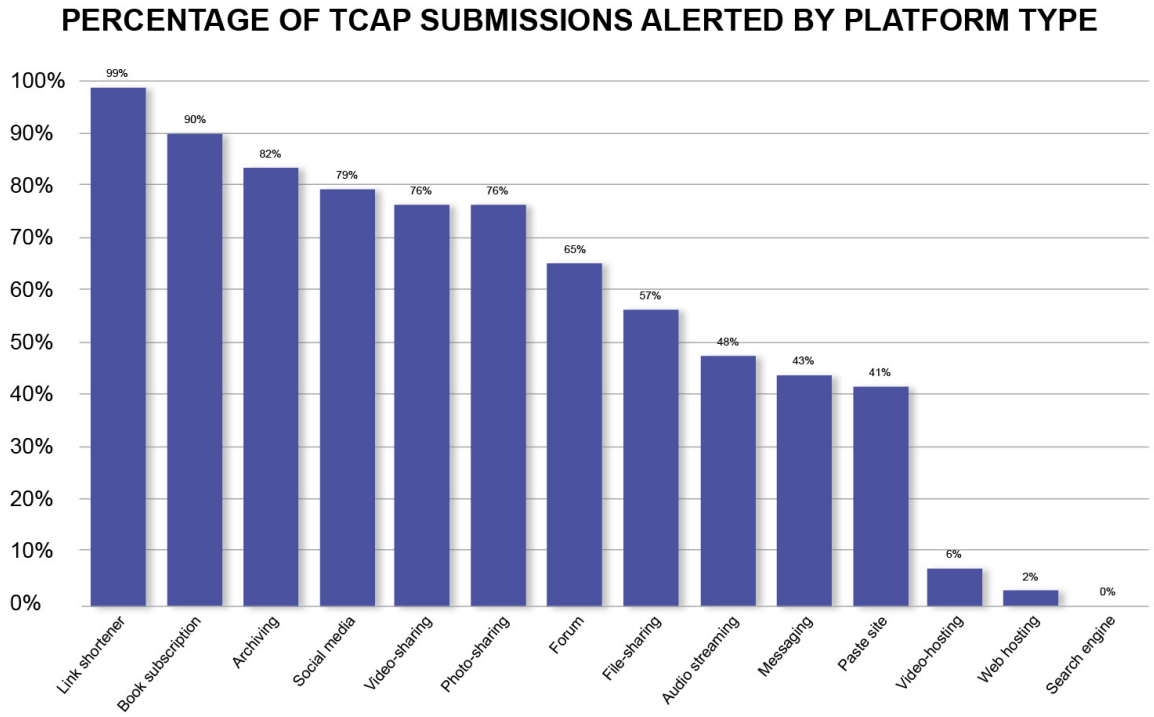

Figure 4: Percentage of TCAP alerts sent from total submissions by platform type.

Breaking down the data further, *Figure 4* shows the percentage of submissions to the TCAP that results in alerts. This illustrates the levels of successful engagement with tech platforms, split by platform type. It is important to note that engagement is determined by both outreach from Tech Against Terrorism, but also is dependent on the willingness of tech platforms to engage. The platform types with the lowest levels of engagement were search engines (0%), web-hosting (2%), and video-hosting platforms (6%). These low levels of alerts can be attributed to the complexities involved in alerting websites,[6] and a significant proportion of video-hosting sites being micro or small platforms making successful outreach more difficult. However, our open-source intelligence (OSINT) team has been developing and implementing a mitigation strategy to disrupt terrorist operated websites at the infrastructure level, more details of which can be found in a 2022 Strategy Paper.[7]

The alert percentage for the four most prolific platform types for terrorist content were file-sharing (57%), archiving (82%), paste sites (41%), and messaging platforms (43%). From this sample, there was a particularly low level of engagement with paste sites and messaging platforms. This is because much of this content is on platforms that are likely to be operated by terrorists or their supporters, and that we therefore

6 Tech Against Terrorism (2022), The Threat of Terrorist and Violent Extremist Operated Websites.

7 Tech Against Terrorism (2022), Strategy Paper: Responding to Terrorist Operated Websites; Tech Against Terrorism (2021), Major Al Qaeda in Arabian Peninsula (AQAP) website disrupted striking significant blow to its online operations.

do not engage with. This includes, for example, RocketChat servers linked to Islamic State and Al-Qaeda, respectively, and two paste sites we believe to be likely run by violent Islamist supporter networks. We have submitted more than 1500 URLs of terrorist content hosted on these three hostile platforms, but been able to send zero alerts.

Analysis: Ideological breakdown

There are clear differences in the ideological patterns of tech platform exploitation. This discrepancy is visible in terms of volume of content collected by the TCAP, as well as the types of platforms exploited. In terms of volume, we have identified much less far-right terrorist content (1264 submissions) than Islamist terrorist content (38,115 submissions) due to several factors including, with regard to far-right terrorist actors, online dissemination techniques that target a smaller number of platforms, a less frequent official propaganda output, difficulties of verification and attribution, and a lack of timely official designation of far-right groups that advocate for, or are engaged in, terrorism. We explain this discrepancy in more detail in a previous blogpost.[8]

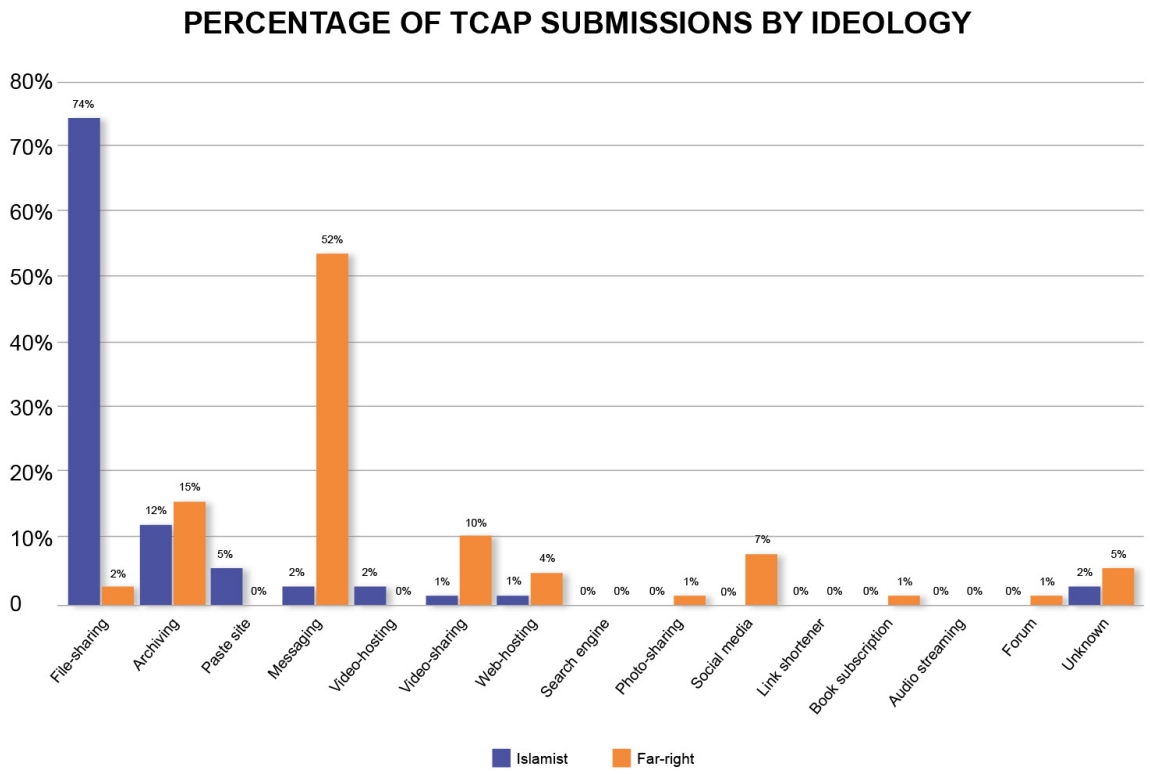

Figure 5: Percentage of TCAP submissions per tech platform type split by ideology.

We also identified Islamist terrorist content across a wider range of platform types than far-right terrorist content, with no far-right terrorist content found on search engines, audio streaming platforms, or link shortening sites. This is likely due to Islamist terrorist actors being forced to experiment with a more diverse range of platforms to circumvent improved and more proactive content moderation efforts. In terms of overlap, only archiving sites were exploited to a significant extent by both Islamist and far-right terrorists. Archiving performs a similar function for both ideologies, allowing their content to maintain a stable presence online, as well as providing an opportunity for intervention through engagement with these platforms.

8 Terrorist Content Analytics Platform (2022), Comparative Analysis of the TCAP Transparency Report Statistics on Content Collection and Removal Rates.

Islamist terrorist exploitation

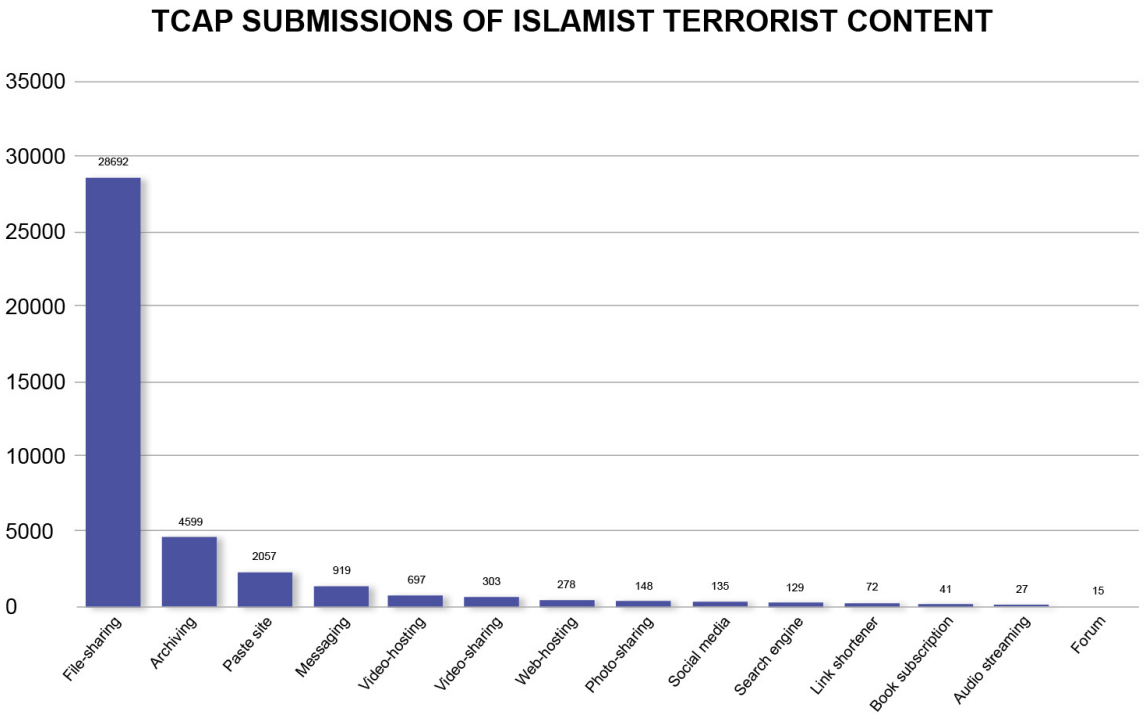

Figure 6: Islamist terrorist TCAP submissions by tech platform type.

In response to broadly improved tech platform moderation, our data indicates that Islamist terrorists have become more creative in how they disseminate propaganda. The propaganda output from the two most prominent groups, al-Qaeda and IS, is mostly structured around centralised channels or accounts (beacons) which release content and supporter networks which re-upload and re-share it. This 'swarmcast' model is used to maintain the online presence of violent Islamists accounts and channels across platforms.[9]

The vast majority of Islamist terrorist content submitted to the TCAP was found on file-sharing platforms (74%), followed by archiving (12%), paste sites (5%) and messaging platforms (2%). This is because Islamist terrorist groups within the TCAP's scope often disseminate each piece of propaganda content (e.g., a video) with large lists of URLs that link to multiple file-sharing, as well as archiving and paste sites. These platforms all act as content stores,[10] explaining why these platform types host the highest volume of content.

Messaging platforms only represented 2% of Islamist terrorist submissions because they most frequently act as beacon channels where official content is often signposted to and disseminated from via outlinks, rather than hosted. When we do identify public channels on messaging applications that host official Islamist terrorist content, we alert the whole channel rather than individual posts, resulting in fewer TCAP

9 See: Ali Fisher (2015), "Swarmcast: How Jihadist Networks Maintain a Persistent Online Presence", *Perspectives on Terrorism*, Vol. 9, Issue 3; Ali Fisher, Nico Prucha, Emily Winterbotham, (2019), "Mapping the Jihadist information ecosystem: Towards the next generation of disruption capability", Global Research Network on Terrorism and Technology, Paper No. 6.

10 'Content stores' are where terrorist content is stored, including text and audio files, as well as images and videos. These are used as online libraries of content. Terrorists rely on content storage platforms and pasting sites, as well as archive services.

alerts. Paste sites (5% of submissions) also function as aggregators, where long lists of URLs are hosted in a centralised location directing to where the propaganda content is hosted. Some of these sites are likely operated by terrorists to avoid content moderation. Archive sites (12%) also function as circumventors,[11] allowing Islamist terrorist supporter networks to back up propaganda content from elsewhere, making it more resistant to content moderation efforts. Archive sites are consistently exploited by a wide variety of Islamist terrorist groups, and this content tends to stay up for a long time.

Far-right terrorist exploitation

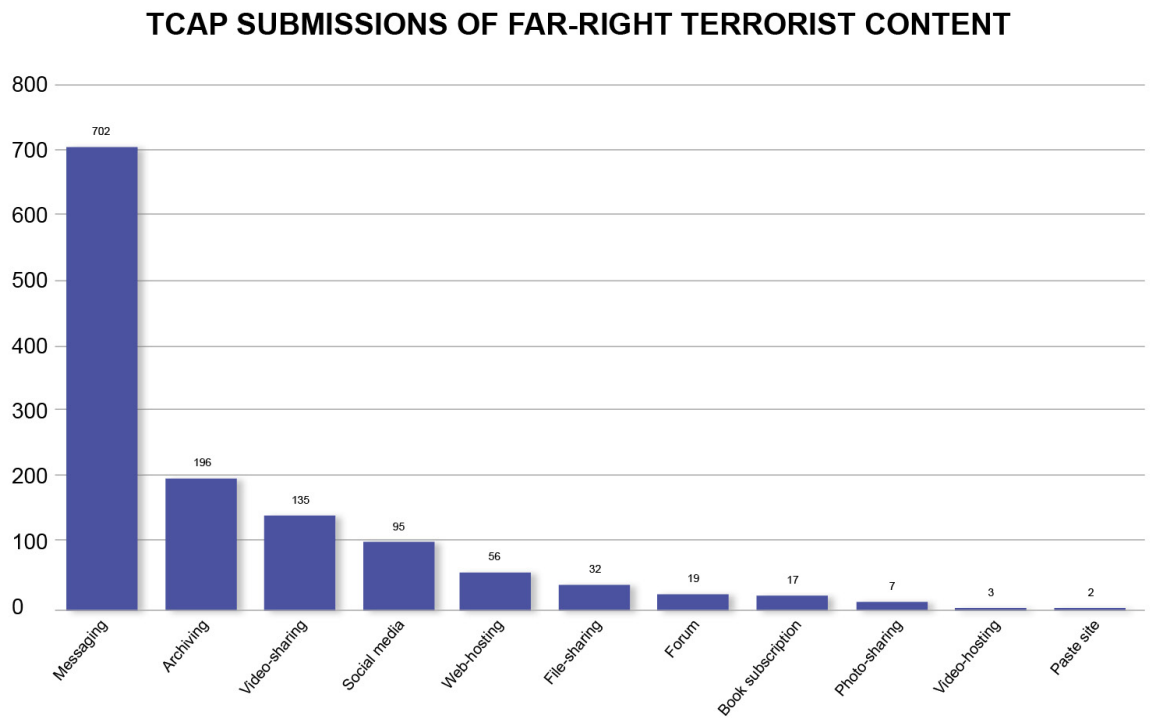

Figure 7: Far-right terrorist TCAP submissions by tech platform type.

Islamist terrorists' pattern of propaganda dissemination is very different from far-right terrorist networks, who often post propaganda and other material in-app, without sharing it as an external URL. This means, for example, that the TCAP might collect 100 URLs for one video produced by Islamic State and promoted by its supporter networks, but only one or two URLs for a single video produced by a far-right terrorist group such as National Socialist Order.

The majority of far-right terrorist content was found on messaging platforms (52%), followed by archiving sites (15%), video-sharing platforms (10%), and social media (7%). Far-right terrorist networks have migrated towards more niche platforms in the past few years, partly due to increased moderation on larger platforms, as well as the anonymity and audience reach provided by some of these alternative platforms. In particular, some of the 'free speech' video-sharing platforms have attracted large numbers of far-right extremists due to their resistance to strict content moderation.

11 Circumventors are online services and platforms used to circumvent content moderation and deplatforming measures, and include VPNs, which can enable nefarious actors to access content that has been blocked in specific countries. Another example of circumventors is the use of decentralised web technologies, which avoid website takedowns.

Whereas Islamist terrorist propaganda is disseminated in a fairly standardised and sophisticated way from centralised sources for each group, there is a less clear pattern of distribution for far-right terrorist content covered by the TCAP. Named, cohesive terrorist organisations do not form centralised nodes in the online far-right terrorist ecosystem in the same way that IS and al-Qaeda do.

Official propaganda produced by a terrorist organisation (for example, an Atomwaffen Division (AWD) video) tends to be released on an official channel on a messaging app, a video-sharing platform or website, then reposted by violent extremist supporter networks across multiple channels and on different platforms (especially video-sharing and social media platforms). Once posted, this official content typically stays online longer without the need for sophisticated content moderation avoidance techniques (such as outlinking).[12] We assess this lower takedown rate of far-right terrorist content may be down to being relatively unidentifiable to moderators, legal and jurisdictional confusion, and the hesitancy of platforms to undertake removal due to libertarian policies on freedom of speech.[13] When it is removed, the content is reposted periodically elsewhere by supporters (often on the same platforms) or backed up on archiving sites or in messaging chats.

Crisis material

The dissemination strategies discussed above differ from the distribution patterns of content produced by terrorist attack perpetrators, such as livestreams and manifestos. We have identified the use of outlinks from larger beacon platforms to smaller file-sharing sites to disseminate the attacker's manifesto. As outlined in a recent blog, this reflects a new typology of behaviour in a subset of far-right violent extremist online networks, namely lone-actor attackers who seek to exploit file-sharing platforms to host their crisis material.[14]

This method of dissemination is in line with the adversarial shift of Islamist terrorist organisations' online strategies, who have long targeted smaller file-sharing platforms for propaganda hosting. This strategy disperses content across a wide range of unique platforms to complicate content moderation efforts. Crisis content relating to historical attacks is reposted periodically by supporters across messaging and video-sharing platforms and is often backed up on archiving sites giving it a stable presence online.

12 Between December 2020 and November 2021, the average removal rate by tech companies following alerts of far-right terrorist content was 50% compared to 94% for Islamist terrorist content. (Source: TCAP Transparency Report, 2021)

13 Terrorist Content Analytics Platform (2022), Comparative Analysis of the TCAP Transparency Report Statistics on Content Collection and Removal Rates.

14 Terrorist Content Analytics Platform (2023), Far-Right Lone-Actor Terrorist Attacks and Violent Extremist Use of File-Sharing Platforms.

Analysis: Content removal rates by platform type

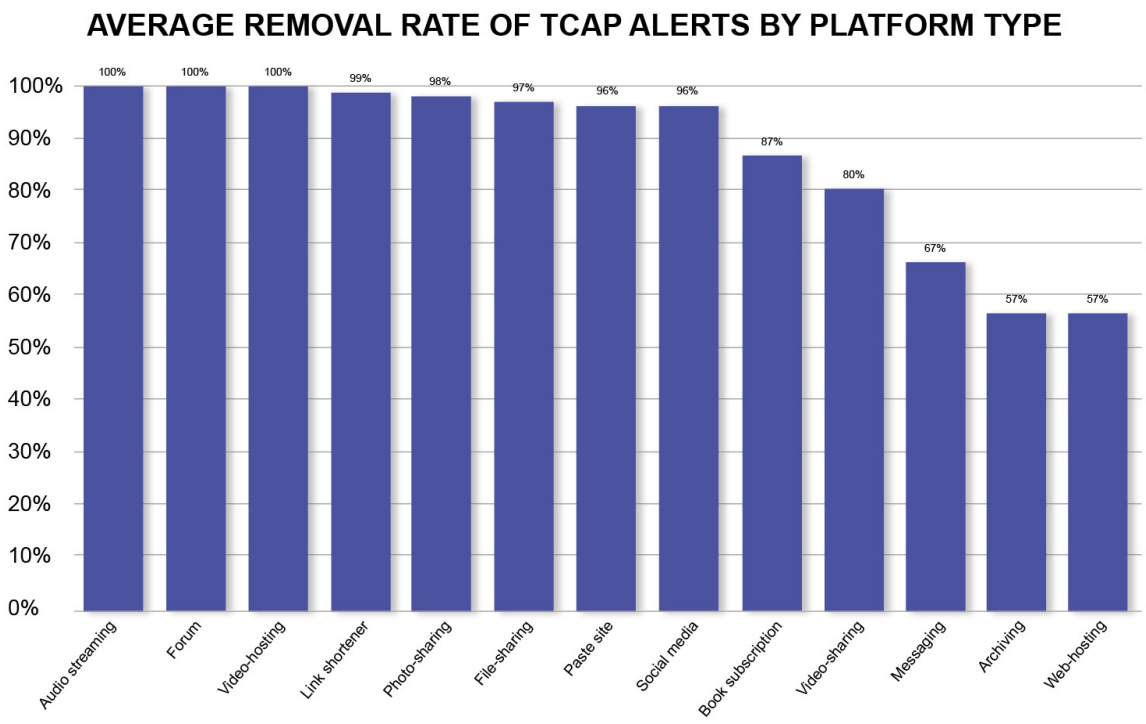

Figure 8: Average percentage of TCAP alerts offline by tech platform type.

The outstanding finding from analysing removal rates of TCAP-alerted terrorist content is that in total 94% of alerts resulted in removal. There was a consistently high removal rate across most tech platform types, with 8 out of 13 having removed over 90% of alerted terrorist content and no platform type removing less than 50%. However, the data also suggests certain at-risk platform types are not effectively tackling terrorist content through removal. Of particular concern in this regard are video-sharing platforms (80% removal rate), messaging applications (67%), and archiving services (57%). The low removal rate of terrorist content on web hosting, while concerning, is based on the exploitation of only one website with a total of 7 alerts.

File-sharing platforms, which, as discussed, were by far the most heavily exploited platform type as measured by TCAP alerts, also removed terrorist content at a very high rate (97% of alerts offline). Archiving sites, on the other hand, removed 57% of terrorist content alerted to them by the TCAP, meaning 1687 URLs (out of 3936 URLs) remained online at the time of writing. This is perhaps unsurprising given that archiving sites are intended to archive and therefore preserve online content. However, it is likely that both far-right and Islamist terrorist actors are abusing this purpose to host terrorist content on archiving sites for indefinite periods of time as they know it will not be removed.

The lower removal rates of terrorist content on video-sharing (80% of alerts offline) and messaging platforms (67% of alerts offline) are also notable. These platform types are both heavily exploited by far-right terrorist actors, which likely explains the lower removal rates.[15] Much of the far-right terrorist content alerted through the TCAP is hosted on alt-tech video-sharing platforms.

15 This is supported by previous TCAP data, in which 69.39% of far-right terrorist content alerted to video-sharing platforms was removed (compared to 95.52% of Islamist). (Source: TCAP Blog: Comparative Analysis of the TCAP Transparency Report Statistics on Content Collection and Removal Rates, 2022)

These platforms have a higher threshold for content removal due to libertarian policies on freedom of speech - in line with US laws - and therefore show lower takedown rates of TCAP URLs.

Furthermore, due to the historical prioritisation of Islamist terrorism in counterterrorism policies the far-right has often been overlooked, which has resulted in confusion for tech platforms over the legality of such content. This has also meant that content produced by Islamist groups in scope tends to be more easily recognisable to platform moderators, whereas often far-right content may contain symbols of designated groups which are not well known.

UNDERSTANDING TERRORIST EXPLOITATION BY TECH PLATFORM SIZE

This section analyses TCAP data by the size of tech platforms, with reference both to a platform's average user base, as well as to its resources where this information is available. When utilising the internet, terrorists exploit platforms of all sizes for specific reasons. The size of a tech platform, and particularly the availability of resources both financial and human determines its ability to moderate terrorist content on its services. It is, therefore, useful to understand how platforms of different sizes are both exploited by terrorists, and how they respond to terrorist exploitation.

We would note that TCAP collection methodology focuses on collecting terrorist URLs from smaller-sized platforms, which may, in part, account for the disparity in TCAP submissions and alerts between small and large platforms. The TCAP also focuses on the core networks of terrorist actors online, rather than supporters who are not directly linked to these networks, and does so with the intention of disrupting terrorist content at its source.

Analysis: Terrorist exploitation by tech platform size

Micro	< 100,000 average users per month
Small	> 100,000 average users per month
Medium	> 10 million average users per month
Large	> 1 billion average users per month

Figure 9: Classification of platform size by average user base.

Terrorists exploit tech platforms of all sizes, from micro platforms with less than 100,000 average users, to large platforms with over one billion average users. To date, the TCAP has identified terrorist content on 27 micro platforms, 51 small platforms, 33 medium platforms, and 16 large platforms, with one platform size unknown (see *Figure 10*).

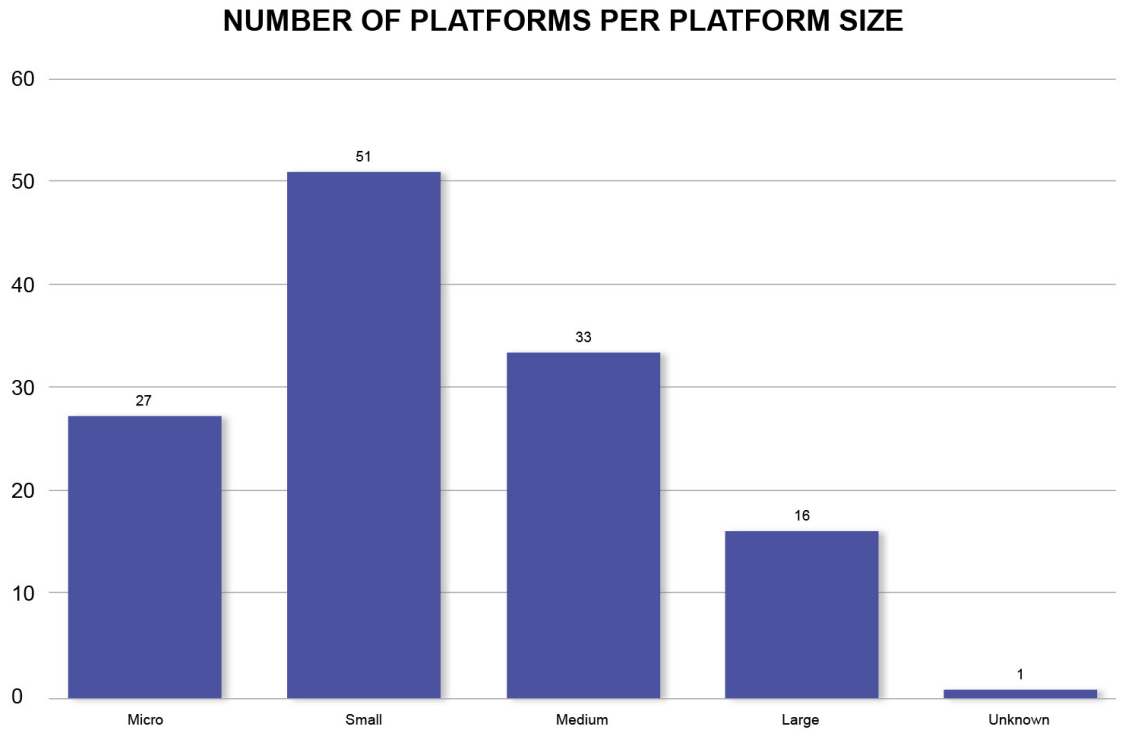

Figure 10: The number of tech platforms at varying sizes with terrorist content submitted to the TCAP.

Of all tech platform sizes, the TCAP has collected content most frequently from small platforms (51 platforms in total), indicating that terrorist content may be more dispersed across different small platforms rather than accumulating on fewer, bigger platforms. This corroborates Tech Against Terrorism's assessment that terrorist content is often disseminated from a small number of beacons (typically large platforms) to a much wider ecosystem of smaller platforms.[16]

Whilst terrorist actors utilise platforms of all sizes, exploitation of online platforms for dissemination of propaganda content is highly concentrated amongst medium and small-sized platforms (see *Figure 11*). Medium-sized platforms had the most submissions (17,521 or 48% of all TCAP submissions) and alerts (8591 or 57% of overall TCAP alerts). This is followed by small platforms, with the second highest number of submissions (14,079 or 38% of overall TCAP submissions) and alerts (8591 or 38% of overall TCAP alerts).

Tech Against Terrorism assesses that terrorist actors consider four factors when deciding on platforms to exploit: security, stability, audience reach, and usability.[17] Terrorist exploitation of platforms of different sizes can partially be attributed to terrorists' evaluation of these different factors, and how different platform sizes offer different benefits and serve different purposes.

The large disparity in TCAP submissions and alerts between small-medium and large platforms may be due, in part, to the consideration of security and stability. Generally, there has been an adversarial shift in terrorist exploitation of the internet towards smaller platforms, partly due to greater attention, resources and

16 Beacons act as centrally located lighthouses that signpost viewers to where content may be found, which is often done through outlinks posting to content stores. Terrorists and violent extremists often use these beacon platforms and have official channels on them that signify their central communications.

17 This framework is based on the key conclusions of the European Counter Terrorism Centre's (ECTC) Advisory Network Conference at Europol.

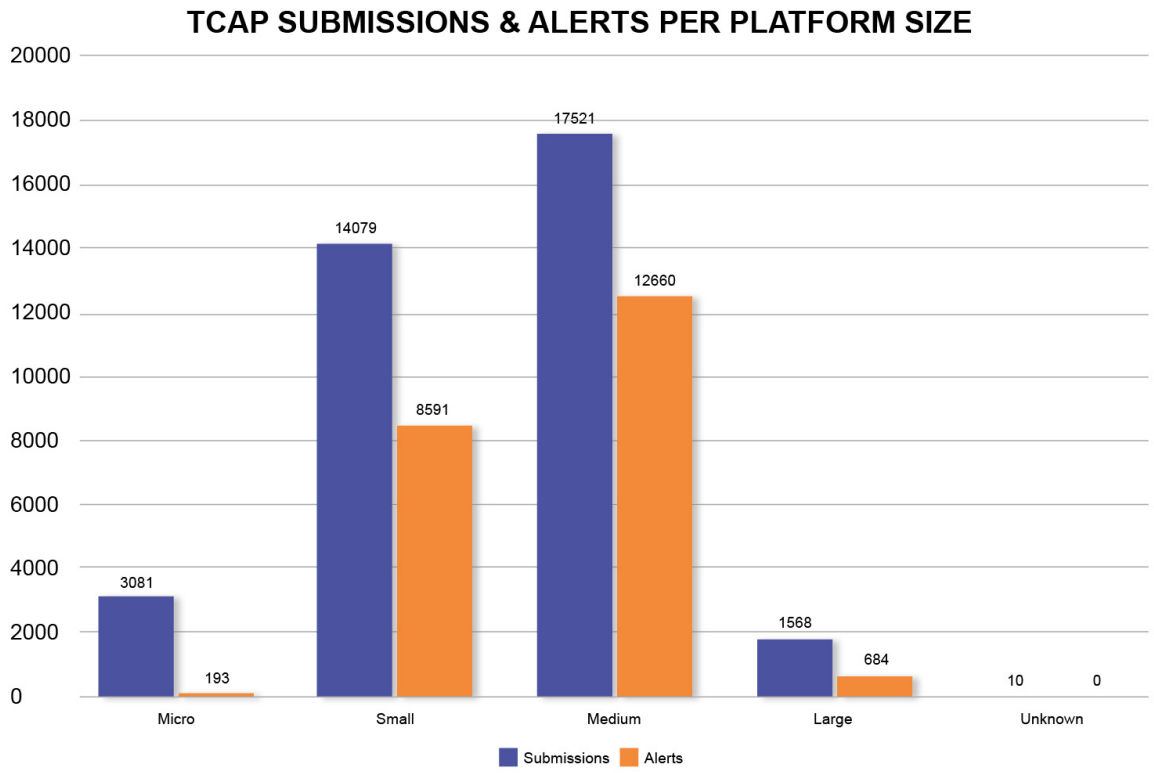

Figure 11: The number of TCAP submissions and alerts to varying platform sizes.

capability directed towards counterterrorism policy and content moderation amongst larger platforms.[18] In response, terrorist actors have adopted a fairly sophisticated and well-established propaganda dissemination pattern, particularly among Islamist terrorist organisations such as Islamic State (IS) and Al-Qaeda.

Terrorist actors often utilise smaller platforms to store propaganda, rather than large platforms, as they typically offer greater stability.[19] Smaller platforms have fewer resources and a lesser capacity (or sometimes willingness) to moderate terrorist content, meaning such content can stay online for longer than if it was posted to a larger platform. After uploading their material to smaller platforms, terrorist actors will then post links to the content in aggregated form on large platforms which provide a greater audience reach (such as social media or messaging platforms).

The content identification process of the TCAP focuses on proactively tracing terrorist entities to their beacon platform: the platform or channel where they signpost to all official content.[20] As terrorist actors are less likely to directly upload content on larger platforms, the TCAP focuses on identifying the beacon which signposts to where the content is actually hosted (such as file-sharing, paste sites, etc.). The TCAP then

18 To read more on this shift, see Arthur Bradley and Deeba Shadnia (2022), Examining Online Migration to Terrorist and Violent Extremist-Owned Domains, Program on Extremism, The George Washington University and Tech Against Terrorism.

19 It is important to note that the analysis here focuses on designated, and often official terrorist content and propaganda. Terrorist supporter content is currently outside of the TCAP's scope, and therefore outside of the scope of this report.

20 Terrorist Content Analytics Platform, How It Works.

alerts these smaller content-hosting platforms to remove content at its source. Once content is removed, outlinks to source material on larger beacon platforms are nullified, as they direct users to content which is no longer online. The intended focus on beacon platforms, and disrupting terrorist content at its source, also results in a higher number of submissions and alerts to these smaller platforms which host the content.

Terrorist actors can also use micro platforms as content stores, as they offer similar security and stability as small platforms. However, the reason for fewer TCAP submissions and alerts could be attributed to their lesser usability compared to small or medium-sized platforms with greater capacity and resources. When choosing between two platforms with similar feature sets, terrorists are likely to choose the platform that is more user-friendly. Their choice of platform may also be influenced by those platforms listed on file mirroring services, which enable users to upload material to multiple platforms simultaneously.

Analysis: How tech platform resources impact terrorist exploitation and content removal

Very Early	0 – 10 employees
Early	11– 49 employees
Mid	50 – 249 employees
Enterprise	250+ employees

Figure 12: Classification of platform stage, based on the eSafety Commissioner's categorisation of tech company size.[21]

Generally, there is a lack of publicly available or easily accessible information about platforms' resources, particularly the size and capabilities of Trust and Safety teams.

Less than half of our data sample (43.75%) had publicly available information about the number of employees working at a given platform (see *Figure 13*).

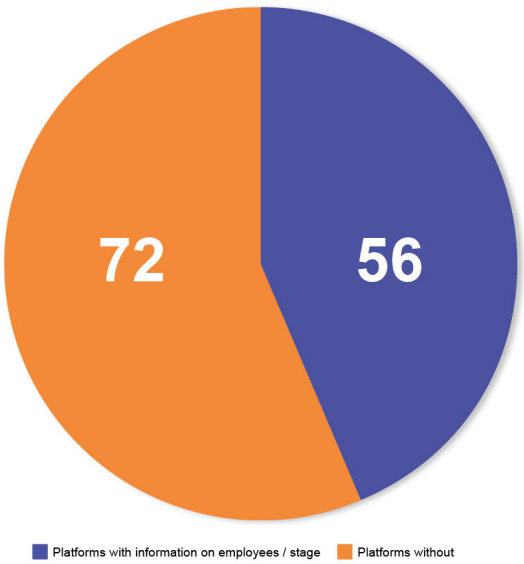

Figure 13: Number of platforms with publicly available information on number of employees or stage of the platform.

21 eSafety Commissioner (2021), Development of industry codes under the Online Safety Act.

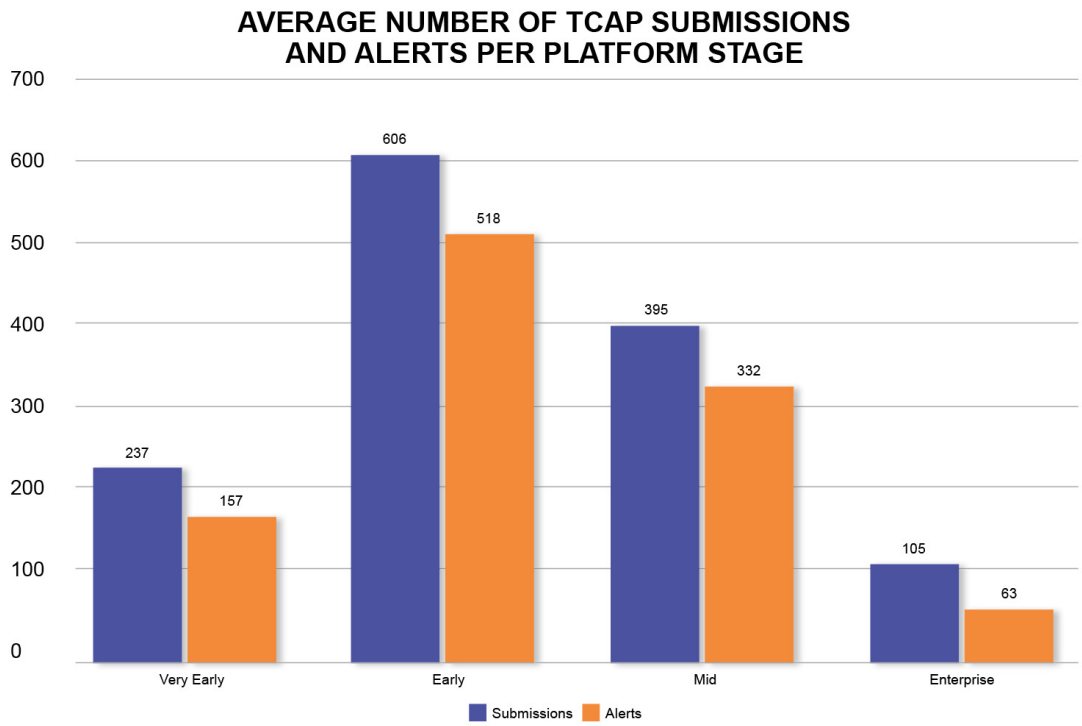

Figure 14: Average number of TCAP submissions and alerts by platform stage.

To explore terrorist exploitation of tech platform size – with regard to a platform's employees, resources, and capacity – we created a smaller, secondary dataset from the 43.75% of platforms which did have information on their resources (see *Figure 14*).[22]

Early-stage platforms (those with fewer employees) average the most submissions per platform (606 or 61.76% of TCAP submissions) and alerts per platform (156 or 65.73% of TCAP alerts) (see *Figure 14*). This is followed by mid-stage platforms, which average the second-highest number of submissions per platform (394 or 12.40%) and alerts (332 or 17.74%). Very early-stage platforms, with fewer than 10 employees, average more than twice the submissions and alerts for 'enterprise'-stage platforms with over 250 employees.

Our data shows that platforms with fewer employees are, therefore, on average most heavily exploited by terrorist actors; they also have the fewest resources to cope with this exploitation. As large platforms scale up efforts to automatically identify and remove content, terrorists have increasingly shifted to, and depend upon, smaller platforms, with lower capacity for moderation. This highlights the need for continued tailored and practical support for smaller platforms, such as through actionable alerts via the TCAP, as well as through Tech Against Terrorism's Mentorship Programme which aims to provide bespoke support to smaller tech companies in developing their counterterrorism policies and response.

22 In this report, we used the labels for a platform's 'stage' ('very early', 'early', 'mid' and 'enterprise') to classify the number of employees at a given platform, thus giving an indication of its resources and capacity to moderate content. We used these labels to differentiate from those used to classify a platform's size as per its user base. A platform's 'stage', therefore, is not in reference to how many years a platform has been established.

Analysis: Removal rate of terrorist content by tech platform size

Generally, platforms of all sizes are receptive to TCAP alerts, with all platform sizes removing, on average, more than 85% of URLs alerted to them (see *Figure 15*).

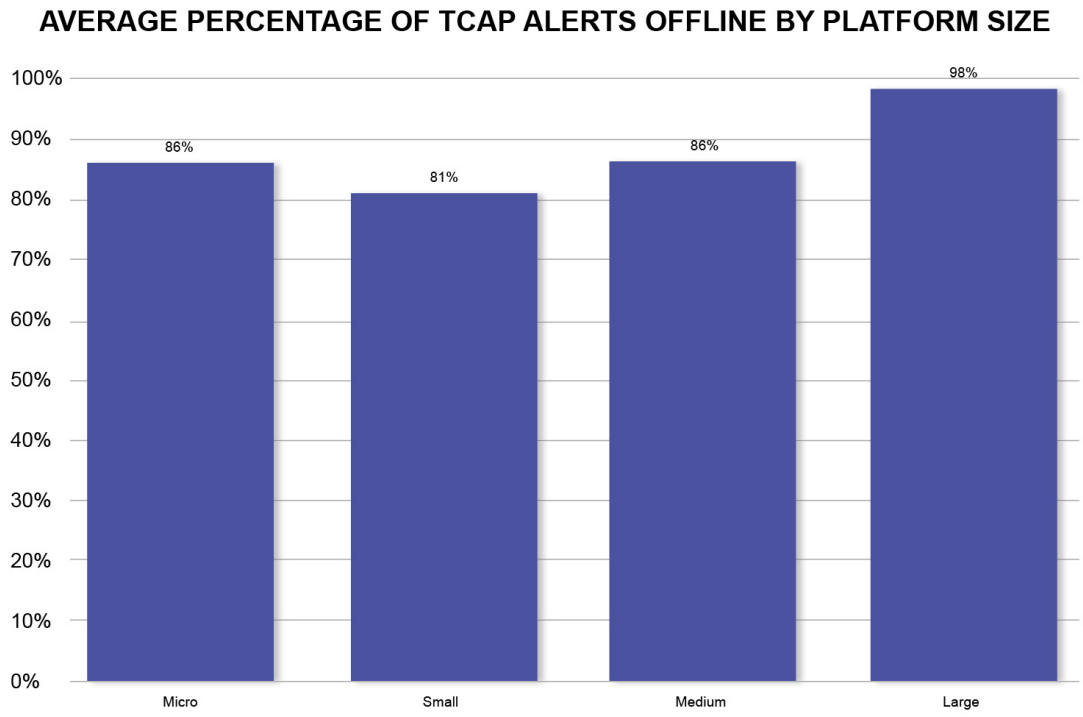

Figure 15: Average percentage of alerts removed per platform size.

Micro-sized platforms have, on average, removed 86% of the URLs alerted by the TCAP, the same removal rate as medium-sized platforms. However, it is important to qualify this statistic by noting the smaller size of the data set: only 95 alerts were sent to 6 different micro-sized platforms. The removal rate of TCAP-alerted terrorist content by the other-sized platforms (small, medium, and large) increases with platform size. Platforms with the largest user base were most effective at removing terrorist content (98% removal rate) followed by medium-sized (86%), followed by small-sized (81%).

Based on Tech Against Terrorism's experience of mentoring tech platforms, smaller platforms are often very receptive to mentorship on how to improve their counterterrorism policies and any opportunity to learn how to minimise the terrorist threat online.[23] Smaller tech platforms often want to remove terrorist content, but lack the expertise, resources, or capacity to do so. The TCAP bridges this gap by finding, verifying and alerting terrorist content to smaller tech platforms, and providing them a direct avenue to triage and assess it against their policies or the law. Our data suggests that where micro platforms are supported through TCAP alerts, this can result in a relatively high removal rate of terrorist content (86%). However, given small-sized platforms have the lowest average removal rate, further engagement and support should continue to be targeted towards smaller platforms.

23 The Tech Against Terrorism Mentorship Programme is a capacity-building endeavour, aimed at supporting tech platforms in improving their counterterrorism, counter violent extremism, and content moderation policies whilst respecting human rights. As of March 2023, Tech Against Terrorism has mentored 50 tech platforms on a one-to-one advisory basis.

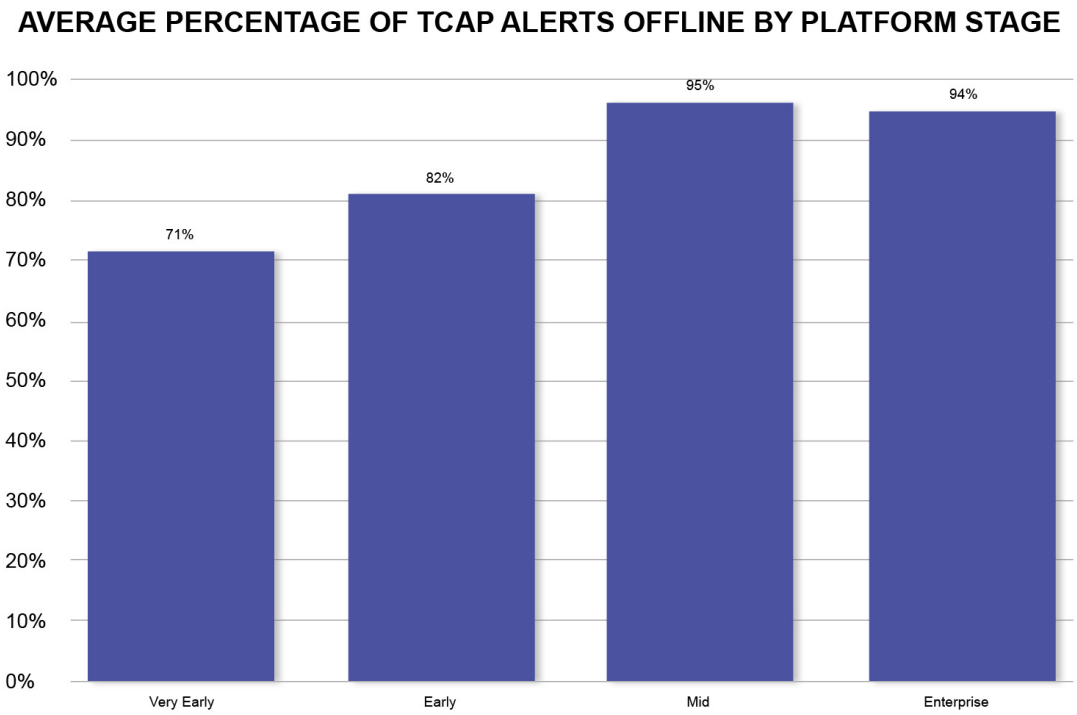

Figure 16: Average percentage of alerts offline per platform stage.

Further breaking down the average percentage of TCAP alerts removed by platform stage, very early-stage platforms had on average the lowest removal rate of URLs alerted to them via the TCAP (71%), followed by early-stage platforms (82%) (see *Figure 16*). The lower average removal rates of terrorist content for platforms with fewer employees is unsurprising given the direct correlation with fewer resources and capacity for human moderation. The data suggests that very early-stage platforms (those with fewer than 10 employees) have the weakest capacity to moderate terrorist content through content removal. Content moderation tools and initiatives aimed at supporting tech platforms in removing terrorist content should, therefore, keep smaller platforms in mind, with particular focus on those platforms with the fewest resources available to moderate such content.

Both mid and enterprise stage platforms have a high average percentage of TCAP-alerted content offline (95% and 94% respectively). This is unsurprising given that larger platforms have a higher in-house capacity to moderate terrorist content due to larger Trust and Safety teams with more resources and expertise.

UNDERSTANDING TERRORIST EXPLOITATION BY GEOGRAPHIC REGION

This section analyses TCAP data in relation to the geographic location of tech platforms, including the quantity of platforms per country and region, the volume of terrorist content per region, and the removal rate of terrorist content. The geographic location in which a tech company is legally based determines the jurisdiction it operates in, and therefore the laws it must follow. The global growth in online regulation which governs tech platforms' responsibilities in tackling harmful content (including terrorist content), makes regional differences in terrorist exploitation and associated tech platform liability useful to understand.

Analysis: Quantifying tech platforms by geographic distribution

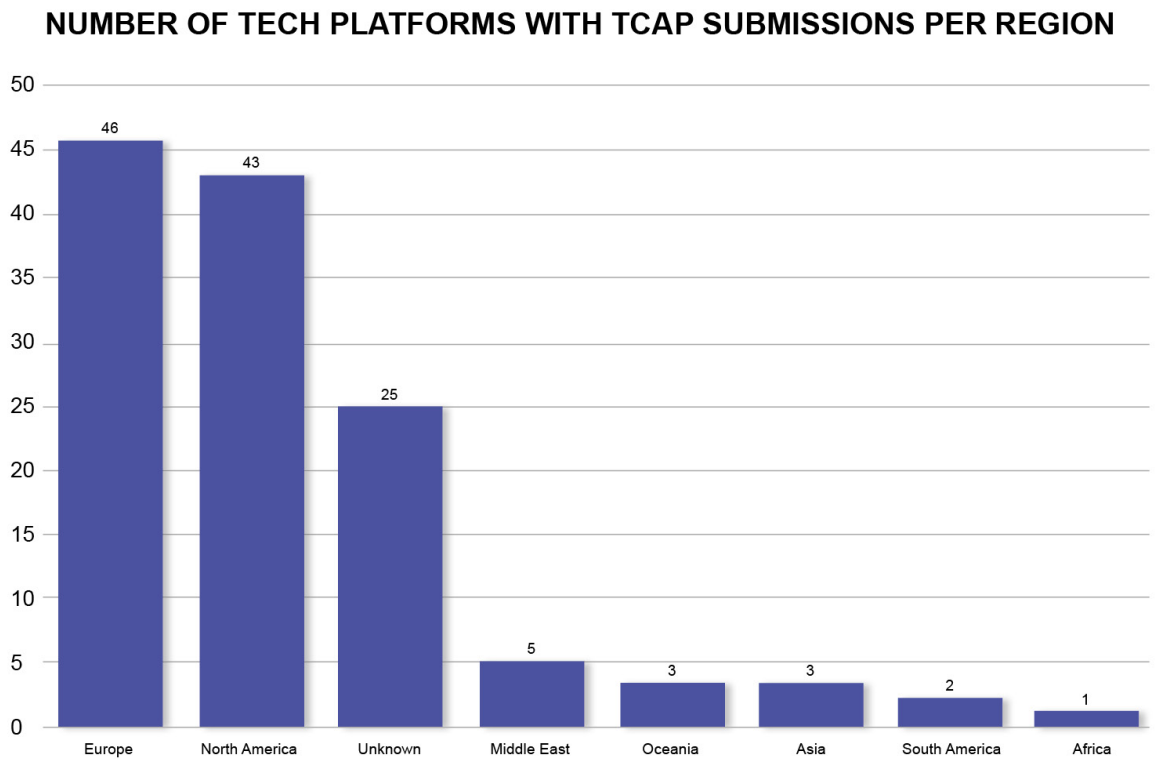

Figure 17: Number of platforms with TCAP submissions by geographic region.

Though we found that terrorists exploit tech companies located across the globe, the overwhelming majority of platforms on which we found terrorist content are located in Europe or North America, with a similar number in each region (46 and 43 respectively). The heavy concentration of platforms in these two regions reflects a broader trend of western domination of the global tech sector. We found terrorist content within the scope of TCAP to be based most commonly on platforms in the United States (43 platforms), which is perhaps unsurprising given that Silicon Valley remains one of the largest technology hubs in the world.[24] However, this pattern may also reflect a collection bias given the TCAP team primarily engages with tech platforms based in the west. The global diversity of the tech companies we work with is likely to increase as we seek to expand our outreach and as we include within the TCAP more localised terrorist groups within the TCAP who may exploit local platforms.

24 Jennifer Martin (2022), Is Silicon Valley a Technology Hub?, University of Silicon Valley.

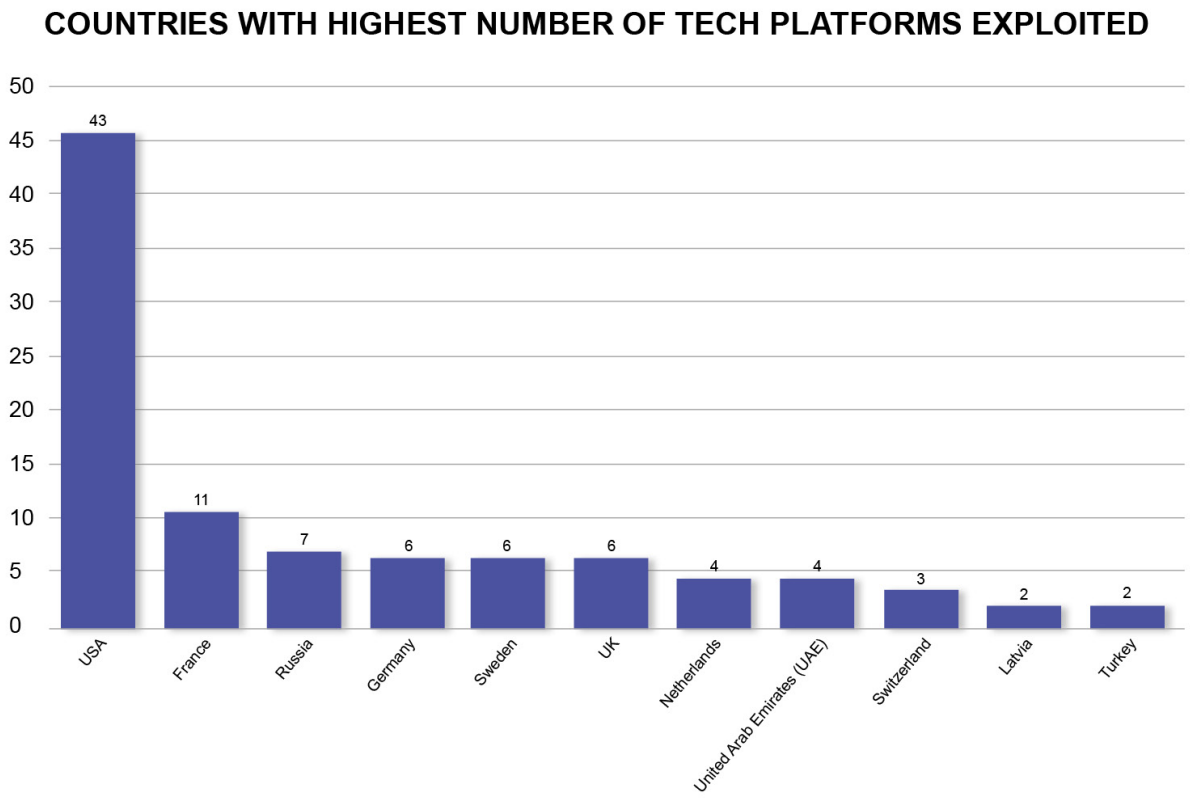

Figure 18: Countries with highest number of platforms with TCAP submissions.

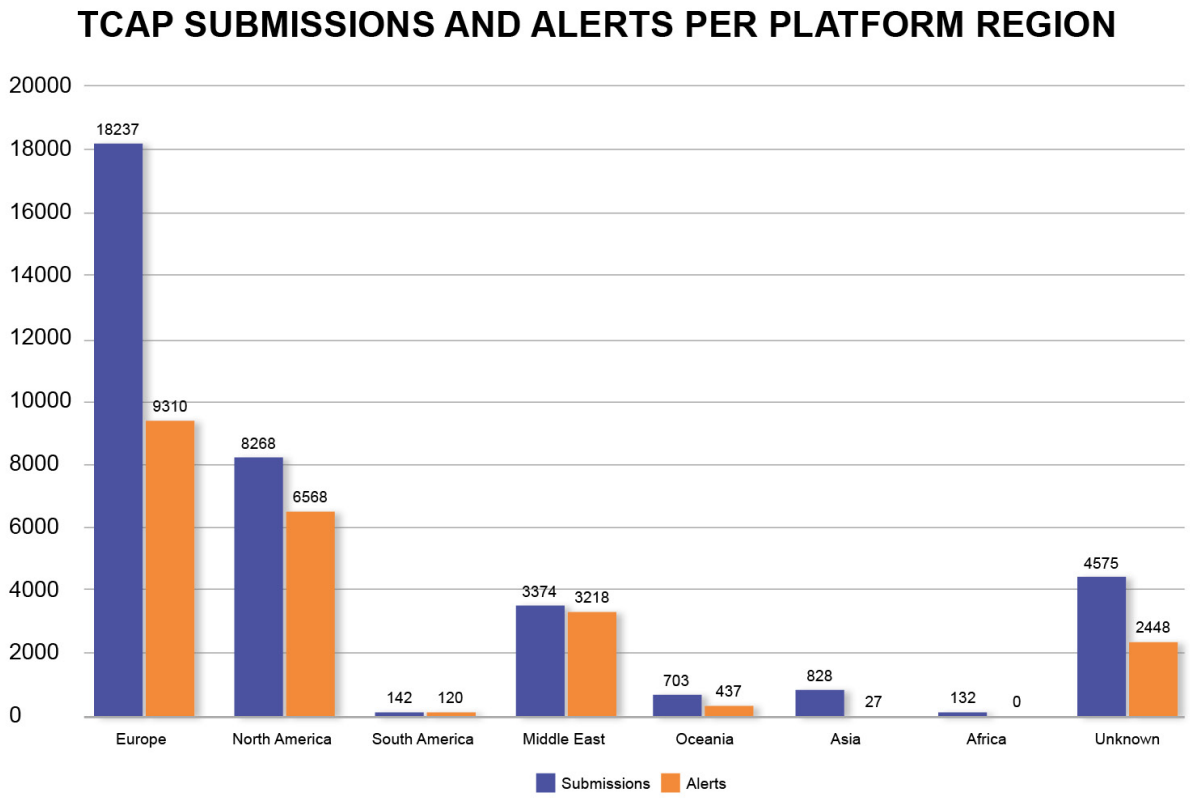

Figure 19: TCAP Submissions and alerts by platform region.

A significant percentage of platforms in our sample (20%) do not have publicly available information on where the platform is based. Even if platforms do not publicly disclose where they are geographically based, they are still required to register to a location and are liable to that country's online regulatory laws. The platforms without publicly available information tend to be small or micro platforms which therefore slightly skews the data, as well as making it more difficult to reach out to these companies to offer support.

Analysis: Volume of terrorist content across geographic regions

We found the highest volume of terrorist content on Europe-based platforms (18,237 submissions: see Figure 19). This is more than double that on North America-based platforms (8,268 submissions). We also sent a higher proportion of alerts (in relation to submissions) to North America-based platforms (79%) than Europe-based platforms (51%), meaning our engagement with North America-based companies is much higher than for Europe-based platforms. This exposes a clear gap in which Europe-based platforms are being heavily exploited by terrorist actors for propaganda dissemination, and could benefit from receiving more targeted and increased support from the TCAP. As part of the Tech Against Terrorism Europe (TATE) project, Tech Against Terrorism will be providing tailored support to European based platforms, including through TCAP alerts, to support them tackle terrorist content that is illegal under the European Union's terrorist content online (TCO) regulation.

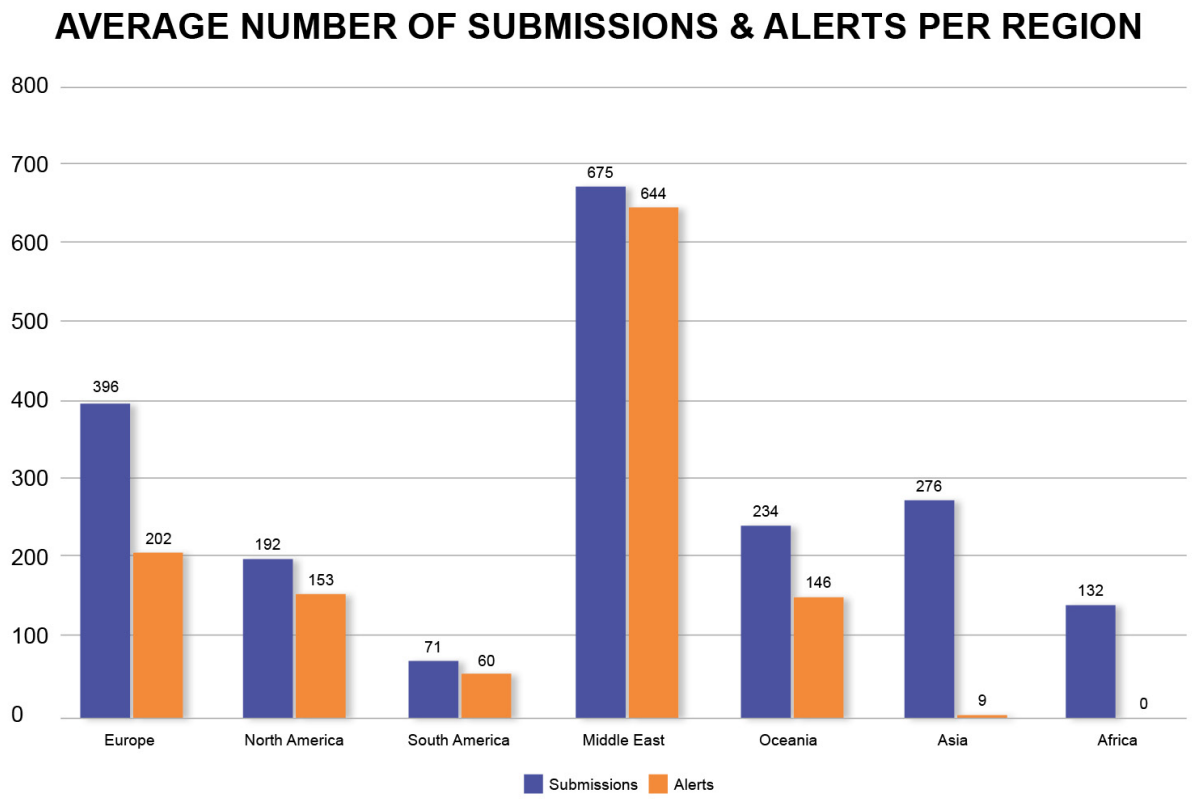

Figure 20: Average number of TCAP submissions and alerts per platform by geographic region.

On average, platforms based in the Middle East were exploited most heavily, with the highest number of TCAP submissions and alerts per platform (see *Figure 20* above). Whilst this was a small sample size comprising only 5 platforms based in the region, it demonstrates a heavy concentration of terrorist content on just a few platforms.

Analysis: Tech platform content removal rates by geographic region

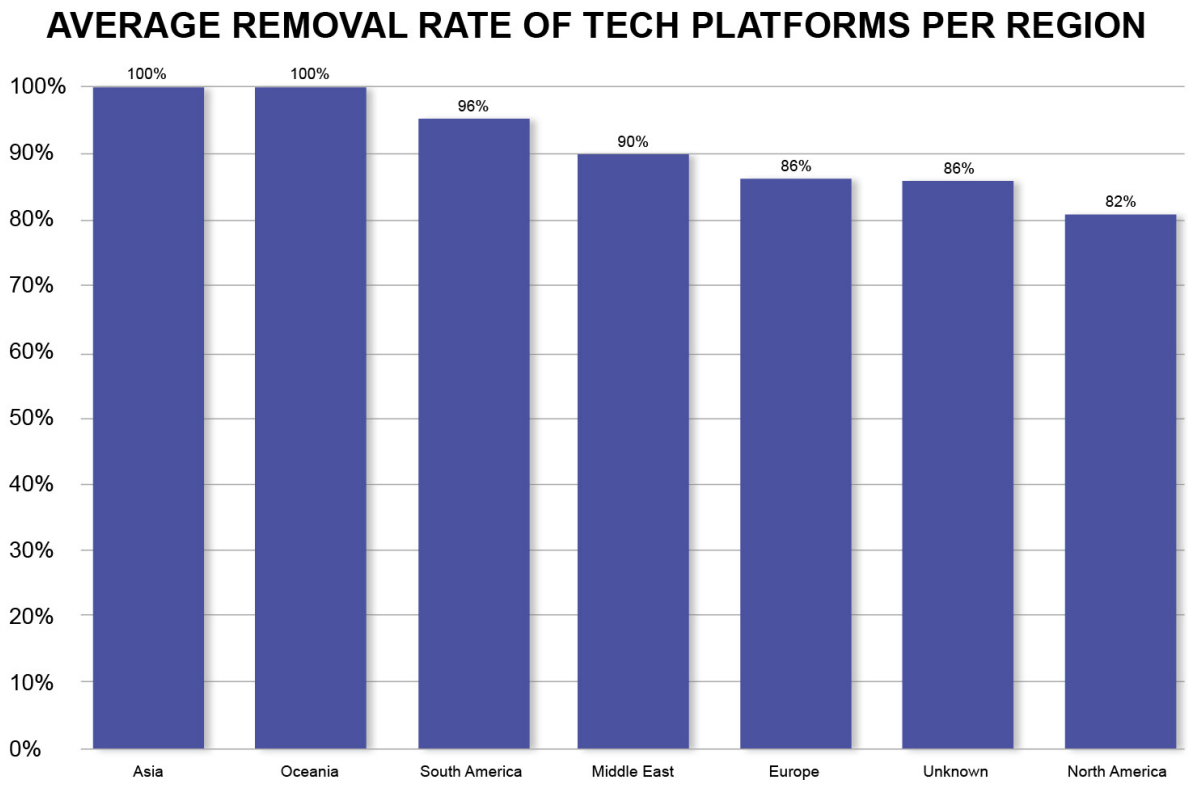

Figure 21: Average content removal rate of tech platforms by geographic region.

There is a consistently high tech platform removal rate across all the regions in which terrorist content is hosted. It is worth noting that tech platforms based in the three regions with the highest average removal rates (Asia, Oceania and South America) received a negligible number of alerts (584 between them) in comparison to Europe and North America (15,878 alerts between them).

When comparing Europe- and North America-based platforms, to which the majority of TCAP alerts were sent, North America-based platforms averaged a lower removal rate (82% compared to 86% of alerted URLs offline). This may be attributable to more libertarian attitudes towards freedom of speech in the United States, which have limited the burden on tech platforms to remove harmful content through online regulation. Of note is that tech companies whose location was not publicly available did not, on average, remove terrorist content at a lower rate than where the location was available.

POLICY IMPLICATIONS AND RECOMMENDATIONS

Recommendations for At-Risk Tech Platforms

PLATFORM CATEGORY	RISKS	RECOMMENDATIONS
Small	• Small and medium-sized tech platforms were the most highly exploited tech platform sizes, based on the volume of content identified on their platforms. Smaller platforms also averaged a lower removal rate of alerted terrorist content than large and medium-sized platforms. • Earlier stage tech platforms - those with 0-50 employees - averaged a higher volume of terrorist content on their platforms than platforms with more employees. Earlier stage platforms also averaged a lower removal rate of alerted terrorist content.	• Publish information about the platform's resources and capacity, particularly regarding the capacity of moderation and Trust and Safety teams. This will help to mitigate expectations of what realistic content moderation looks like across the tech ecosystem, paving the way for more proportionate regulation of tech platforms. Platforms could include this information in transparency reports, 'About' and 'Support' pages, or 'Help Centres'. • Encourage user reporting within the Terms of Service and/or Community Guidelines to counterbalance smaller platforms' more limited capacity to proactively moderate content. • Allow users to report content under a category for terrorism and violent extremism to enable moderators to prioritise dangerous content.

PLATFORM CATEGORY	RISKS	RECOMMENDATIONS
File-sharing	• Over half of the tech companies on which we identified terrorist content are file-sharing platforms (106 out of 187 platforms). The dispersal of terrorist content across many file-sharing platforms makes targeted intervention to support tech platform moderation more difficult as it requires engagement and cooperation from a larger number of platforms. • File-sharing platforms were by far the most exploited platform type, with 28,724 URLs of terrorist content (72%) submitted to the TCAP.	• Prohibit terrorist content in line with international designation lists – such as the United Nations Security Council list – and/or in line with comparable lists from democratic countries such as the UK, Canada, and the US. • Consider implementing automated content moderation processes[25] – such as searches of key words, phrases, images, logos, symbols, and so on, associated with known terrorist entities across files uploaded and shared on their platform – with support from Tech Against Terrorism's Knowledge Sharing Platform.[26] When doing so, human rights should be safeguarded, for example by means of hash-based detection methods.
Archiving	• Archiving sites were the second most heavily exploited platform type based on the volume of terrorist content identified (4795 submissions or 12% of total). • Only archiving sites were exploited to a significant extent by both Islamist and far-right terrorist actors. These sites perform a similar function for both ideologies, allowing their content to maintain a stable presence online. • Archiving sites had the lowest removal rate of alerted terrorist content at 57%, meaning 1687 URLs (out of 3936 URLs) remained online at the time of writing.	• Whilst archiving sites have an important mission statement to archive online content, as with all platforms, they have a responsibility to moderate illegal terrorist content in line with national and international designation lists. • Use simple detection tools that flag key words and logos linked with terrorist entities for review. A compendium of relevant words and symbols can be found on Tech Against Terrorism's Knowledge Sharing Platform, where this data can be downloaded and directly input into content moderation workflows.[27]

25 Given that users are only able to access terrorist content on file-sharing platforms via direct outlinks, most who interact with such links are likely to be empathetic to the contents' values, and are therefore less likely to report it.

26 Tech Against Terrorism, Knowledge Sharing Platform.

27 Based on monitoring of archiving sites by Tech Against Terrorism's Open-Source Intelligence (OSINT) team, files are often named based on the content's title and pro-IS accounts repeatedly use similar account names.

PLATFORM CATEGORY	RISKS	RECOMMENDATIONS
Messaging Platforms	• The majority of far-right terrorist content was found on messaging platforms (52%). Far-right terrorist actors have migrated towards more niche platforms in the past few years, partly due to increased moderation on larger platforms as well as the anonymity and audience reach provided by some of these alternative platforms. • In their role as beacons to far-right terrorist content, messaging platforms are uniquely placed to act quickly to remove this content before it can be widely disseminated. However, the average removal rate for messenger platforms was one of the lowest of all platform types, at 67%.	• Given that far-right terrorist actors and organisations are less comprehensively covered by terrorist designation lists, messaging platforms should prohibit violent extremism in the Terms of Service and/or Community Guidelines. • Uphold human rights principles when engaging in content moderation practices, with particular concern for users' right to privacy online, given the private nature of content on messaging platforms. • Maintain a list of the key words, phrases, images, etc. which are linked with far-right terrorist entities for content moderation purposes, with support from Tech Against Terrorism's Knowledge Sharing Platform.[28]

General Recommendations for Tech Platforms

RECOMMENDATION	EXPLANATION
Make use of free tools to disrupt terrorist use of the internet.	• Platforms should make use of the TCAP which alerts terrorist content to tech companies when found on their platform. It is particularly useful for smaller platforms, assisting them by identifying and flagging terrorist content on their services and helps them to make better content moderation decisions and to streamline content moderation processes. • If platforms are unable to develop in-house automated moderation capacity themselves, they could also engage with free third-party content moderation resources. One such option would be the new tool Tech Against Terrorism is creating with Google Jigsaw to help smaller platforms identify and take down terrorist content.[29]

28 Tech Against Terrorism, Knowledge Sharing Platform.

29 Tech Against Terrorism (2023), Tech Against Terrorism to Build Content Moderation Tool with Google Jigsaw.

RECOMMENDATION	EXPLANATION
Provide an easily accessible contact point for users, law enforcement, governments, and initiatives such as the Terrorist Content Analytics Platform.	• Platforms across the tech ecosystem should provide a contact point to ensure that dangerous content, as well as crises and threat to life incidents, can be actioned as swiftly as possible. Such a contact point could be as simple as a contact email address included within the Terms of Service or at the bottom of a platform's homepage, which should be reachable 24/7. This would also help to ensure that internet referral units, such as Europol's IRU, are able to action terrorist content and alert the relevant competent authorities.[30] • Platforms could also provide a feedback form that allows for the categorisation of enquiries, including a category for TVE-related enquiries which would allow platforms to prioritise and quickly address serious and dangerous content on their sites.
Obstruct the dissemination of banks of URLs by utilising behaviour-based cues and moderation techniques already in place for spam material.	• As discussed in a recent Resolve Network research report,[31] pro-IS users are likely to operate on the assumption that the outlinks they share to TVE content will be deactivated. Therefore, the core of their strategy lies in volume and speed, relying on the use of automation for rapid generation and dissemination of banks of URLs (such as through Telegram bots). • Platforms could, therefore, make use of behaviour-based cues such as abnormal posting volume, which can be picked up more easily by automated systems. Platforms could also make use of moderation mechanisms already in place for spam behaviour to disrupt terrorist use of their services.
Use simple detection tools that flag key words and logos linked with terrorist entities for review.	• Another mitigation strategy to disrupt terrorist use of outlinking is automated detection of keywords and logos present in terrorist propaganda. The authors of the previously cited Revolve Network research report recommend automated detection tools for keywords in the name of the uploaded file (such as the weekly Islamic State newsletter, al-naba) or by identifying relevant logos on the publication.[32] A compendium of relevant keywords and symbols can be accessed in Tech Against Terrorism's Knowledge Sharing Platform.

30 Europol (2022), EU Internet Referral Unit – EU IRU.

31 Stuart Macdonald, Connor Rees, and Joost S. (2022), Remove, Impede, Disrupt, Redirect: Understanding & Combating Pro-Islamic State Use of File-Sharing Platforms, Resolve Network.

32 Ibid.

RECOMMENDATION	EXPLANATION
Mitigate against content moderation evasion tactics.	• In order to circumvent improved content moderation efforts across various platforms, Islamist terrorist entities have experimented with and employed a range of circumvention or evasion tactics to maintain their presence online. Such tactics often include URL shortening, deliberate misspelling or lexical variation, and account or content mirroring. • Platforms can mitigate the effect of such evasion tactics in various ways, including by means of two-factor authentication and introducing time limits to the validity of join links to private servers and channels. To read more, see our blog on the topic on our Knowledge Sharing Platform.[33]
Consider involvement in cross-platform initiatives to counter terrorist use of the internet.	• As terrorist entities are likely to exploit a multitude of different platform types, platforms may find cross-platform co-operation useful in tackling terrorist exploitation of their services. For example, initiatives such as the Christchurch Call assemble an active community of civil society and government organisations together with tech platforms collaborating to eliminate terrorist content online.
Consider off-platform elements when making content moderation decisions related to terrorist content and behaviour.	• Platforms could incorporate off-platform monitoring when considering removing content or users from their services, with particular attention given to persons on the platform engaging in severe abuse off-service and/or on other platforms. If a platform chooses to do so, it should explain in public policy what weight is given to off-platform elements in moderation decisions. To see an example of off-platform monitoring, read Twitch's Off Service Conduct policy.[34]

33 Tech Against Terrorism: Knowledge Sharing Platform, Content Moderation Circumvention Tactics

34 Twitch, Off Service Conduct.

Recommendations for Policymakers

RECOMMENDATION	EXPLANATION
Consider smaller tech platforms when introducing online regulation and legislation.	• Regulation and debate on content moderation are often narrowly focused on big tech rather than smaller platforms.[35] Regulation and debate are therefore disproportionate to the threat from terrorist actors, who, as evidenced by our data, most heavily exploit smaller and medium sized platforms. Smaller platforms therefore require increased attention and support in tackling this threat and mitigating risk factors. • Policymakers should also commit to better supporting newer platforms whose user base rapidly grows beyond their content moderation means.[36]
Consider tech platforms' resources when categorising platforms by size.	• Both in legislative regulation and tech policy regarding content moderation, the categorisation of tech platforms by size must include a consideration of not only a platform's average user-base, but also platform resources (including financial, human, and technical). Considering a platform's resources, particularly the size of their content moderation or Trust and Safety team, provides crucial insights into how platforms can realistically moderate content, thus paving the way for more realistic regulation.
Incorporate risk assessments as a first step to online regulation.	• Larger platforms can hire counterterrorism experts, analysts, or entire teams who dedicate their time to counterterrorism policy and moderation. Such platforms are also able to spend more on automated content moderation systems to alleviate strain on human resources. Smaller platforms, however, often without counterterrorism experts, do not necessarily have the same understanding of terrorist exploitation of online platforms, or their services specifically, and certainly do not have the same ability or financial resources to dedicate to moderating terrorist content. • Policymakers should consider including risk assessments for platforms as a first step to online regulation, to aid platforms' understanding of the threat they are facing from TVE actors. It is only once platforms understand the threat that they can begin to counter it. Governments and policymakers could also encourage partnerships and knowledge-sharing endeavours for smaller platforms to facilitate further understanding.

35 Tarleton Gillespie (2020), Expanding the debate about content moderation: scholarly research agendas for the coming policy debates, Internet Policy Review.

36 Ysabel Gerrard (2020), Expanding the debate about content moderation: scholarly research agendas for the coming policy debates, Internet Policy Review.

RECOMMENDATION	EXPLANATION
Implement targeted initiatives to support file-sharing, archiving, and messaging platforms.	• File-sharing, archiving, and messaging platforms are three of the most heavily exploited of all platform types. When considering introducing regulation or legislation for terrorist use of the internet, policymakers and governments should introduce initiatives to support such platforms, such as knowledge-sharing endeavours and civil society support.

Tech Against Terrorism's Next Steps

This report has evaluated the distribution of terrorist content across different platform types, sizes, and locations using data taken from the TCAP – the world's largest database of verified terrorist content – between 25 November 2020 and 19 January 2023. Based on our findings, we have provided guidance for at-risk platforms, as well as tech platforms more generally, to disrupt this terrorist exploitation of their services. We have also called on policymakers to provide greater to support to these exploited tech platforms. Below, we outline the steps Tech Against Terrorism is taking to improve our tailored support to tech platforms.

We are pleased that the Government of Canada has awarded Tech Against Terrorism up to $1.9 million of funding over three years for Phase 2 of the TCAP.[37] This funding will cover archiving and classifying terrorist content to provide academics with a centralised knowledge base to inform counter-terrorism research, as well as providing hashes for smaller tech platforms to use as references. This funding will help us strengthen our capacity to support at-risk tech platforms through outreach and engagement. We are also planning to improve our monitoring and evaluation frameworks to better measure the impact of the TCAP.

Engagement

- The data reveals that terrorist actors exploit a massive number of different platforms, a significant proportion of which are not currently being supported by TCAP alerts (around 60 platforms not supported). Tech Against Terrorism's priority is to reach out to these platforms to register them to the TCAP, as well as provide an introduction to the platform. Since data collection for this report was finalised, we have added over 70 platforms to the TCAP's subscriber base and increased the number of platforms alerted to 100.
- The data exposes a clear gap in which Europe-based platforms are being heavily exploited by terrorist actors for propaganda dissemination, but a significant proportion of these platforms lack targeted support through TCAP alerts due to a lack of awareness or engagement with TCAP. Tech Against Terrorism Europe (TATE) is a project that commenced in January 2023 that seeks to address this problem by providing tailored support to European based platforms, including through TCAP alerts, to assist them in tackling terrorist content that is illegal under the European Union's terrorist content online (TCO) regulation.

37 Public Safety Canada (2022), Government of Canada announces up to $1.9 million in funding to combat online terrorist and violent extremist content, Government of Canada.

- The Tech Against Terrorism Policy and Response (PAR) team will focus outreach efforts on file-sharing, archiving and messaging platforms as a priority for mentorship, as these were identified as the most at-risk platform types. If you are a tech company that needs support, please get in touch with us at contact@techagainstterrorism.org, where we will be able to support you through our Mentorship Programme.[38]
- Web-hosting platforms are one of the least engaged platform types, as we currently lack the capability to alert web domains via the TCAP. This year, Tech Against Terrorism is working on a strategy for tackling terrorist operated websites (TOW) and supporting web-hosting infrastructure providers. As part of the expansion of the TCAP, we aim to submit website domains via the TCAP, alerting the relevant infrastructure providers to the terrorist operated website.

Monitoring and evaluation

- Improve evaluation of the TCAP by using broader metrics to gauge success beyond the statistics of submissions and alerts.
- A more nuanced measurement for success is content removal rate; namely, the proportion of TCAP alerts actioned by platforms. However, removal rate is also insufficient as it does not measure how quickly that content is removed and how many times it is viewed. Given the purpose of propaganda is to reach as large an audience as possible as quickly as possible, speed of removal of content may be the best indicator for successful disruption. To measure this, we are improving the capability of our URL status checker to provide real-time data on the uptime of terrorist content submitted to the TCAP.
- As of April 2023, we will begin alerting tech platform at three different intervals in the day. In future, even more regular alerts could incentivise platforms to respond more quickly and remove content closer to the time of publication.
- To improve our monitoring and understanding of terrorist exploitation by platform size, we are including additional criteria for tech platforms added to the TCAP which will record platform size and stage. This will enable real-time statistical analysis of terrorist exploitation by platform size, guiding our own understanding and response to this threat.

Combating adversarial shift in terrorist use of the internet – Areas of further study

- The overall data for TCAP submissions and alerts points to a downward trend over the last half a year in Islamist terrorist URLs. Our analysis suggests this is due to a reduced use of outlinking via long lists of URLs. Instead, Islamist terrorist actors are increasingly sharing propaganda in-app, where it cannot be alerted, in private channels where it cannot be accessed, or on terrorist operated websites (TOWs), where content can be downloaded from directly. Tackling this content shared in-app requires careful engagement with messaging applications. More quantitative research is needed to better understand the ongoing shift in propaganda dissemination techniques by Islamist terrorist actors and especially the use of TOWs for hosting propaganda content.

38 Tech Against Terrorism's Policy Advisory and Response (PAR) Team provides bespoke support to tech platforms in strengthening their online counterterrorism efforts. Th team supports platforms through the Mentorship Programme (TAT's principal knowledge sharing and bespoke capacity-building support programme), the Knowledge Sharing Platform, counterterrorism and human rights compliance, transparency support, and regulatory analysis and impact assessment.

- Tech Against Terrorism will take into careful consideration the fact that the over-removal of terrorist content may lead to an adversarial shift, and a shift of terrorist actors to platforms and spaces (chats) where monitoring and moderation is more difficult. More empirical research into the effects of deplatforming terrorist actors is an important area of further study.
- As the dataset of the TCAP grows, the value of data-driven analysis will only increase. Tech Against Terrorism will utilise this unique dataset for research to inform our understanding of patterns of terrorist use of the internet and drive strategies for mitigation and interventions. The TCAP Insights research series will do this by providing actionable policy recommendations for at-risk tech platforms.
- The first year of the TCAP expansion will cover the archiving and classification of terrorist content to provide academics with a centralised knowledge base to inform counter-terrorism research. By providing a secure and ethical manner of sharing online terrorist content, the TCAP will ensure that academic institutions and expert researchers can contribute in a protected manner to an increased evidence-based understanding of terrorist use of the internet, which will drive actionable research recommendations.

METHODOLOGY

Research questions

TCAP Insights is a series of research and policy analyses into patterns of terrorist use of the internet that harnesses the analytical power of the TCAP's unique dataset. When considering patterns of terrorist use of the internet, our analysis focuses on three core areas: tech platform geographic region, size, and type. When considering patterns of terrorist use of tech platforms, we based this assessment primarily on patterns of terrorist content identified and submitted to the TCAP. Whether terrorists consider these platform features when choosing platforms for propaganda dissemination is difficult to verify and, therefore, our assessments are based on the likely utility these features would serve these actors. To assess these areas, our research has focused on the following questions:

- How do terrorists exploit tech platforms based on their geographic region, type, and size?
- Are there patterns in how terrorists exploit tech platforms within different geographic regions, or based on their type or size? If so, why might that be?
- How do terrorists of different ideologies exploit tech platforms based on the type of platform?
- At what rate do tech platforms of different geographic regions, types, and sizes remove terrorist content that is alerted to them via the TCAP? What reasons could there be for these removal rates?
- How can tech platforms from different geographic regions, and of different types and sizes – particularly those most heavily exploited - change their policies and processes to disrupt terrorist use of their services?
- What can be done to support tech platforms from different geographic regions, and of different types and sizes – particularly those most heavily exploited - in countering terrorist use of their services?

Data collection

The analysis in this report is based on data relating to terrorist content online collected by the TCAP between 26 November 2020 and 19 January 2023. This data set includes 39,964 URLs of terrorist content submitted to the TCAP (TCAP Submissions), including 22,615 of those sent as alerts to 95 different tech companies (TCAP Alerts). This is official terrorist content produced by the 37 different terrorist entities within scope of the TCAP, as defined by our Inclusion Policy. This data set also includes the removal rate of terrorist content alerted to tech companies – namely, whether a platform removes terrorist content after it has been alerted to them via the TCAP. The status of URLs used for the report was last checked on 19 January 2023.

TCAP submissions and alerts[39]

- **Submissions:** Tech Against Terrorism's open-source intelligence team monitor and identify terrorist content daily. Each piece of content is verified against the TCAP Inclusion Policy and attributed to the corresponding terrorist organisation. Once content has been verified and classified, it is uploaded to the TCAP (submission).

- **Alerts:** Once content is submitted, the TCAP emails the platform in question with the link to where the content can be found, the associated terrorist entity, and a warning for content that is graphic in nature or contains personally identifiable information. TCAP alerts are made on an advisory basis, meaning it is the platform's decision to proceed with content moderation.

39 For further information on the methodology of the TCAP, please see Terrorist Content Analytics Platform, How it works.

Analysis

To analyse the dataset collected from the TCAP, we allocated three core data tags to each platform: its region, its type, and its size (divided into user base and resources/capacity).

Platform region

To assess a platform's geographic region, we used publicly available information accessible directly on a platform's website, or from other sources such as SimilarWeb and CrunchBase. We labelled each platform with a country, as well as a broader geographic region to facilitate the assessment of patterns in the data. The geographic regions were split into: Asia, Africa, Europe, Middle East, North America, Oceania, South America, and Unknown.

Platform type

To divide tech platforms by their type, we used the typology operated by the TCAP, published in the TCAP Transparency Report for December 2020 – November 2021.[40] This typology can be seen in Figure 22, which highlights platform types and their respective core functionality.[41] Where a platform had more than one functionality in practice, we examined the platform's own branding, as well as the main purpose for which it is exploited by terrorists.

Platform Type	Functionality Provided
File-sharing	Access to digital media such as photos, videos, and documents.
Archiving	Storage of information from defunct webpages or documents for anyone to view publicly.
Forum	Discussion site for conversations in the form of posted messages.
Video-hosting	Posting videos online.
Video-sharing	Uploading, conversion, storage, and later consumption of video content.
Link Shortener	Conversion of any URL into a shorter, more readable link.
Social Media	Creation and sharing of information through virtual communities and networks.
Messaging	Online chat in real time with individuals or larger groups and communities.
Photo-sharing	Uploading, conversion, storage, and later consumption of photo content on the internet.
Audio Streaming	Uploading, conversion, storage, and later consumption of audio content on the internet.
Paste Site	Uploading and sharing of text online, often used for sharing source code.

40 Tech Against Terrorism (2022), Transparency Report: Terrorist Content Analytics Platform, Year One: 1 December 2020 – 30 November 2021.

41 For the purpose of this report, we removed 'file-hosting' as a platform type, and incorporated platforms with this label under the 'file-sharing' platform type, due to file-hosting and file-sharing services' similar functionalities.

Search Engine	Performing web searches using key words or phrases.
Book Subscription	Subscription to officially published and user-published books and documents.

Figure 22: Types of platforms exploited by terrorists and alerted by the TCAP.

When assessing how terrorists of different ideologies exploit platforms based on their type and functionality, we split the ideology of terrorist content alerted by the TCAP between 'Islamist' and 'far-right' ideology, based on the TCAP Inclusion Policy.

Platform size

To assess terrorist exploitation of platforms of different sizes, we categorised the platforms in our sample into the following categories: micro, small, medium, large. Such classifications were based on publicly available information regarding a platform's average monthly user base, obtained through SimilarWeb.

Categorisation of a tech platform's size should also include consideration of its resources, particularly the size and therefore capacity of its team. Where publicly available, we incorporated information on a platform's 'stage' – separated into 'early', 'mid', and 'enterprise', based on the eSafety Commissioner's categorisation of tech company size – determined by number of employees.[42] We added 'very early' as a category to this classification, to effectively analyse the exploitation of the smallest platforms, and to differentiate between many of the platforms we alert via the TCAP which have a smaller number of employees.

Micro	< 100,000 average users per month
Small	> 100,000 average users per month
Medium	> 10 million average users per month
Large	> 1 billion average users per month

Figure 23: Classification of platform size by average user base.

Very Early	0 – 10 employees
Early	11– 49 employees
Mid	50 – 249 employees
Enterprise	250+ employees

Figure 24: Classification of platform stage, based on the eSafety Commissioner's categorisation of tech company size.[43]

42 eSafety Commissioner (2021), Development of industry codes under the Online Safety Act.

43 In this report, we used the labels for a platform's 'stage' ('very early', 'early', 'mid' and 'enterprise') to classify the number of employees at a given platform, thus giving an indication of its resources and capacity to moderate content. We used these labels to differentiate from those used to classify a platform's size as per its user base. A platform's 'stage', therefore, is not in reference to how many years a platform has been established.

ABOUT TECH AGAINST TERRORISM

Tech Against Terrorism enables technology companies in countering terrorist use of the internet whilst respecting human rights. Tech Against Terrorism is an independent public-private partnership initiated by the United Nations Security Council.

We work across the tech sector and are supported by the UN and other international bodies as well the governments of Spain, Switzerland, the Republic of Korea, and Canada.

techagainstterrorism.org
contact@techagainstterrorism.org

tech
against
terrorism