

TRANSPARENCY REPORT



TERRORIST CONTENT ANALYTICS PLATFORM

YEAR TWO: 1 DECEMBER 2021 – 30 NOVEMBER 2022



Terrorist Content
Analytics Platform





CONTENTS

Executive Summary	04
Overview.....	04
Statistical breakdown of Impact.....	05
1. Introduction.....	08
1.1 Why is transparency pivotal in guiding the TCAP?	08
2. Terrorist Content in Scope of the Terrorist Content Analytics Platform	09
2.1 Inclusion Policy	09
Expansion of the TCAP Inclusion Policy	10
3. Quantified Impact of the Terrorist Content Analytics Platform	12
3.1 Summary of the Key TCAP Metrics	12
3.2 Takedown Rates	13
3.3 TCAP Submissions and Alerts per Platform Type	14
3.4 TCAP Submissions and Alerts per Terrorist Entity	17
Focus: TCAP submissions and alerts per Islamist terrorist entity	18
Focus: TCAP submissions and alerts per far-right terrorist entity	18
3.5 Takedown Percentages per Terrorist Entity in Scope	19
3.6 Takedown Rates per Platform Type and Terrorist Entity	20
3.7 Takedown Rates per Extreme Content Flag	23
3.8 Takedown Rates per Personally Identifiable Information (PII) Flag	25
4. Annex.....	26
4.1 What is the Terrorist Content Analytics Platform?	26
4.1.1 Objectives.....	26
4.1.2 The TCAP process	27
4.1.3 TCAP application interface	30
4.1.4 Automated scraping – additional information	30
4.2 Policy Considerations.....	30
4.2.1 Key development principles.....	30
4.2.2 Content Classification and Verification Policy	33
4.2.3 Background: public consultation process	34
4.2.4 Legal consultation.....	34
4.2.5 Crisis Protocol Policy.....	35
4.3 Recognition.....	38
4.4 Global Engagement.....	39
4.5 What’s Next?	40



ABOUT TECH AGAINST TERRORISM

Tech Against Terrorism supports technology companies to counter the terrorist use of the internet. It is an independent public-private partnership initiated by the UN Security Council. Our research shows that terrorist groups - both jihadist and far-right terrorists - consistently exploit smaller tech platforms when disseminating propaganda. At Tech Against Terrorism, our mission is to support smaller tech companies in tackling this threat whilst respecting human rights and to provide companies with practical tools to facilitate this process. As a public-private partnership, the initiative works with the United Nations Counter Terrorism Executive Directorate (UN CTED) and has been supported by the Global Internet Forum to Counter Terrorism (GIFCT) and the governments of Spain, Switzerland, the Republic of Korea, and Canada.

techagainstterrorism.org
contact@techagainstterrorism.org



EXECUTIVE SUMMARY

OVERVIEW

- The Terrorist Content Analytics Platform (TCAP) is a secure and transparent online tool to detect and verify terrorist content and notify technology companies of the presence of such content on their platforms. In November 2020, with support from Public Safety Canada, Tech Against Terrorism launched the TCAP, and we are now building the world's largest database of verified terrorist content, collected in real time from verified terrorist channels online.
- The TCAP is developed using a transparency-by-design approach. This means that all the development work of the TCAP since its creation has ensured that we can be transparent about our actions and policies. In the online world, transparency is a vital pillar of trust between online service providers and their users. While remaining sensitive to operational security, we detail in this report the extent and scale of terrorist content discovered by the TCAP, and how the data gathered is utilised to disrupt and understand terrorist and violent exploitation of the internet. This is the second TCAP transparency report, which is one of several initiatives Tech Against Terrorism has taken to honour our founding principles. The report provides a detailed breakdown of the core metrics for the reporting period between **1 December 2021 and 30 November 2022**, and of key TCAP policies and processes.
- Since our last report, the TCAP has developed both policies and content collection practices to ensure we alert content relating to a range of terrorist entities and ideologies. Overall, most of the content we have alerted has been related to Islamist terrorist entities (92% of alerts in Year 2). This is compared to 98% of alerts containing Islamist terrorist content in Year 1 of the TCAP.
- Over the past 12 months, because of alerts from the TCAP, an average of 84% of Islamist terrorist content was removed, whilst an average 61% of far-right terrorist content was removed. Compared to our previous Transparency Report,³ the takedown rate for Islamist terrorist content has decreased from 94% but has increased from 50% for far-right terrorist content.
- Over the past 12 months, we have expanded the TCAP's Inclusion Policy⁴ to widen the scope of entities whose official content we alert to tech companies. Given the relative imbalance in the official designation of violent far-right compared to violent Islamist entities, we have focused on expanding our inclusion of violent far-right material as far as possible within the ambit of the law. We now alert promotional material (manifestos and livestreams) produced by individual attack perpetrators, based on their classification as "objectionable content" by the New Zealand Classification Office.⁵

¹ Tech Against Terrorism awarded grant by the Government of Canada to build Terrorist Content Analytics Platform, Tech Against Terrorism, <https://www.techagainstterrorism.org/2019/06/27/press-release-tech-against-terrorism-awarded-grant-by-the-government-of-canada-to-build-terrorist-content-analytics-platform/>

² Terrorist Content Analytics Platform, <https://www.terrorismanalytics.org/>

³ Terrorist Content Analytics Platform, Transparency Report, <https://terrorismanalytics.org/policies/transparency-report>

⁴ Inclusion Policy, Terrorist Content Analytics Platform, <https://www.terrorismanalytics.org/policies/inclusion-policy>

⁵ New Zealand Classification Office, <https://www.classificationoffice.govt.nz/>

- In September 2022, the Government of Canada awarded Tech Against Terrorism a second round of funding for the Terrorist Content Analytics Platform (TCAP) which will expand our alerting functionality to support smaller tech companies, as well as establish an archive of verified terrorist content.⁶ This was announced by Canadian Prime Minister Justin Trudeau at the Christchurch Call to Action Summit in 2022.

STATISTICAL BREAKDOWN OF IMPACT

During this reporting period

- Our open-source intelligence experts **submitted 18,995 URLs** containing terrorist content, and the TCAP sent **10,174 alerts to 57 tech companies**, **82%** of which is now offline. In total, 150 tech companies are registered and able to receive alerts as soon as we detect terrorist content on their platforms. We have increased the number of platforms we can alert from 114 in Year 1.
- **18,048 URLs containing Islamist terrorist content** were submitted to the TCAP, compared to **947 URLs containing far-right terrorist content**. 9,436 alerts containing Islamist terrorist content were sent, whilst 738 alerts containing far-right terrorist content have been sent to tech companies. The discrepancy in numbers is due to the different propaganda dissemination techniques employed by far-right and Islamist terrorist groups.⁷ However, we have begun to close the gap between Islamist and far-right terrorist content submissions and alerts and this difference is smaller than in Year 1.
- Tech platforms generally remove more Islamist terrorist content than far-right terrorist content because of our alerts. The average removal rate by tech companies following alerts of Islamist terrorist content is 84%, whereas the average removal rate of far-right terrorist content is 61%.
- Most Islamist terrorist content submitted to the TCAP, and made the subject of a **TCAP alert**, was produced by the **Islamic State** (44%), **al-Shabaab** (18%), and **al-Qaeda** (12%). We saw a marked decrease in the output of al-Qaeda in the Arabian Peninsula, who made up 22% of Islamist terrorist content submissions in Year 1.
- Most far-right terrorist content submitted to the TCAP, and alerted by the TCAP, was produced by **Atomwaffen Division** (19%), the **Christchurch attack perpetrator** (17%), and the **Buffalo attack perpetrator** (17%).
- The TCAP detected terrorist content on 14 different types of tech platforms. The three most exploited technology types in descending order were file sharing services, archiving services, and paste sites. We identified more content on paste sites compared to Year 1.

A TCAP alert is an email sent to tech platforms containing the URL of terrorist content on their services. Alerts also contain key metadata such as the related terrorist entity and if the content contains graphic material.

 Terrorist Content
Analytics Platform

⁶ Government of Canada announces up to \$1.9 million in funding to combat online terrorist and violent extremist content, Public Safety Canada, <https://www.canada.ca/en/public-safety-canada/news/2022/09/government-of-canada-announces-up-to-19-million-in-funding-to-combat-online-terrorist-and-violent-extremist-content.html>

⁷ Following the first TCAP Transparency Report, we published a blog post comparing the difference in our statistics of Islamist and far-right terrorist content online. Comparative Analysis of the TCAP Transparency Report Statistics on Content Collection and Removal Rates, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news/comparative-analysis-of-the-tcap-transparency-report>



- Platforms providing video hosting, link shortening, forum, and audio sharing services are most responsive and have **removed 100% of verified terrorist content** notified via the TCAP. Archiving platforms are the least responsive to our alerts, with 47% of alerted content being removed, this is a decrease from 59% of content removed on archiving platforms in Year 1.
- The TCAP notifies platforms if the alerted content contains graphic content. Some platforms prioritise removing content which depicts violence or is graphic in nature. Based on our data, book subscription, file sharing, and video sharing platforms are more likely to remove content if it contains extreme content.

Policy and Development during the reporting period

- Since February 2022, we **hash** all URLs containing terrorist content that are submitted to the TCAP.⁸ These unique hashes will be shared with the GIFCT's hash-sharing consortium,⁹ which forms a shared industry database of "perceptual hashes" of verified images and videos produced by terrorist entities or groups designated by the United Nations. This action will further achieve our mission to support smaller tech companies in removing terrorist content by allowing tech platforms to preemptively ban verified content without viewing user data.
- Over the report period, we added five new entities to our Inclusion Policy.¹⁰ Our Inclusion Policy is based on the legal designation of terrorist entities by democratic nation states and supranational institutions. We began alerting content created by the following entities:
 - o James Mason, designated by the Government of Canada
 - o During the reporting period, we also began alerting manifesto and livestream content created by attack perpetrators relating to attacks in:
 - Oslo and Utøya, Norway in 2011 - manifesto
 - Halle, Germany in 2019 - manifesto and livestream
 - Buffalo, New York, USA in 2022 - manifesto and livestream
 - Bratislava, Slovakia in 2022 - manifesto


A hash is a unique value assigned to a piece of data, like a digital fingerprint. Our developers have generated hashing software which creates a unique hash of each URL identified by the TCAP. The TCAP hashes verified terrorist content helping the tech sector, particularly smaller tech companies, with automated decision making when moderating terrorist content.



⁸ Announcement: The TCAP's hashing and hash-sharing capability, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news/hashing>

⁹ GIFCT's Hash-Sharing Database, Global Internet Forum to Counter Terrorism, <https://gifct.org/hsdb/>

¹⁰ Inclusion Policy, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/policies/inclusion-policy>

- 
- In February 2022, we published our Crisis Protocol Policy. The Crisis Protocol guides our actions when an emergency incident occurs, by ensuring we have provisions in place to alert the appropriate authorities and mitigate the threat posed by online violent extremist content. In the event of a potential threat to life, the Crisis Protocol Policy outlines the steps that TCAP staff take to evaluate the credibility and imminency of the threat to life and what proportionate actions should be taken. The policy is divided into three sections:
 - o **Pre-incident** – what we do when we encounter a potential threat to life,
 - o **During incident** – what we do in an active crisis event,
 - o **Post-incident** – how we respond to crisis events after an event has occurred.

¹¹ Crisis Protocol Policy, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/policies/crisis-protocol-policy>

1. INTRODUCTION

The Terrorist Content Analytics Platform (TCAP) was created by Tech Against Terrorism to support the tech sector in identifying terrorist content on their services so that such content can be reviewed and removed. In tackling terrorist exploitation of the internet, we believe it is essential to ensure we are transparent and accountable and set best practices for the tech sector. This is the second TCAP transparency report, which is one of several initiatives Tech Against Terrorism has taken to honour our founding principles. The report provides a detailed breakdown of the core metrics for the reporting period between 1 December 2021 and 30 November 2022, and explains some central TCAP policies and processes.

1.1 Why is transparency pivotal in guiding the TCAP?

Transparency is vital to ensure accountability towards the public and internet users. Since counterterrorism is often used as justification to disregard human rights and fundamental freedoms, including online freedoms, transparency reporting on counterterrorism efforts is crucial to understand the extent to which such abuse might occur. Transparency reporting is also an important means of increasing awareness of an organisation's internal decision-making processes. Tech Against Terrorism encourages tech companies and governments to be transparent about their online counterterrorism efforts. For tech platforms, regular transparency reports on online counterterrorism efforts, such as content moderation, provide significant insight into how a platform enforces its counterterrorism policies and responds to government and law enforcement requests.

Tech Against Terrorism's Transparency Reporting Guidelines on Online Counterterrorism Efforts¹² serve as a starting point for increased transparency, and it is our aim that all governments and companies will report on the baseline set out in the Guidelines. Whilst the TCAP is neither a tech company nor a government, we have adopted in this report the best practice identified by those guidelines.

Finally, given the growing role TCAP is playing in the removal of terrorist content online, it is also becoming increasingly important to provide as much as transparency as possible without compromising operational security. Therefore, we have added additional metrics to our statistical review, and incorporated a greater focus on the policies that we revised and augmented during this reporting period.

¹² Transparency Reporting Guidelines on Online Counterterrorism Efforts, Tech Against Terrorism, <https://transparency.techagainstterrorism.org/>

2. TERRORIST CONTENT IN SCOPE OF THE TERRORIST CONTENT ANALYTICS PLATFORM

2.1 Inclusion Policy

The TCAP includes official material produced by terrorist entities in scope of the TCAP's Inclusion Policy.¹³ Our Inclusion Policy is based on the designation lists produced by select democratic nation states and supranational organisations.¹⁴ At the time of writing, the TCAP includes content created by Islamist terrorist entities: Islamic State (and official provinces), al-Qaeda (and verified affiliates), the Taliban. The TCAP also included content created by designated far-right terrorist groups, such as Atomwaffen Division. The TCAP implements the Christchurch Call to Action¹⁵ by notifying tech companies of material produced by the Christchurch attack perpetrator. We also support the New Zealand Classification Office in alerting content created by the perpetrator of the Oslo and Utøya (2011), Christchurch (2019), Halle (2019), Buffalo (2022), and Bratislava (2022) attacks.

	UN	EU	US State	US Treasury	UK	Canada	Australia	New Zealand
Al-Qaeda	●	●	●	●	●	●	●	●
<i>Al-Qaeda in the Arabian Peninsula</i>	●	●	●			●	●	●
<i>Al-Qaeda in the Indian Subcontinent</i>		●	●			●	●	
<i>Al-Qaeda in the Islamic Maghreb</i>	●	●	●		●	●	●	●
<i>Al Shabaab</i>	●	●	●	●	●	●	●	●
<i>Hurras al-Din</i>		●					●	
<i>Jama'at Nusrat al-Islam was-Muslimin</i>	●	●	●		●	●	●	●
Islamic State	●	●	●	●	●	●	●	●
<i>Islamic State Algeria Province</i>		●						
<i>Islamic State Central Africa Province</i>		●	●	●		●		
<i>Islamic State East Asia Province</i>		●				●	●	
<i>Islamic State Greater Sahara Province</i>	●	●	●	●		●		●
<i>Islamic State India Province</i>		●						
<i>Islamic State Khorasan Province</i>		●	●	●		●	●	●
<i>Islamic State Libya Province</i>		●	●	●		●	●	●
<i>Islamic State Pakistan Province</i>		●						
<i>Islamic State Sinai Province</i>		●	●	●		●	●	●
<i>Islamic State Somalia Province</i>		●		●			●	
<i>Islamic State Tunisia Province</i>		●				●		
<i>Islamic State West Africa Province</i>	●	●	●	●		●	●	●
Taliban				●		●		

● Designated terrorist entity
 ● Designated under a synonym or umbrella group or by affiliation
 ● Content banned by the New Zealand Classification Office

Figure 1: Islamist terrorist groups in scope of the TCAP and where they are designated.

¹³ Inclusion Policy, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/policies/inclusion-policy>

¹⁴ In March 2023, Tech Against Terrorism published a landmark report on the designation practices of 12 countries and supranational institutions, including those used as the foundation for the Inclusion Policy. Who Designates Terrorism? The Need for Legal Clarity and Transparency to Moderate Terrorist Content Online, Tech Against Terrorism, <https://www.techagainstterrorism.org/2023/03/23/designatingterrorism2023/>

¹⁵ Christchurch Call, <https://www.christchurchcall.com/>

	UN	EU	US State	US Treasury	UK	Canada	Australia	New Zealand
Atomwaffen Division					●	●		
<i>National Socialist Order</i>					●	●	●	
Blood and Honour						●		
Combat 18						●		
Feuerkrieg Division					●			
National Action					●			
<i>National Socialist Anti-Capitalist Action</i>					●			
<i>Scottish Dawn</i>					●			
<i>System Resistance Network</i>					●			
Proud Boys						●		●
Russian Imperial Movement			●	●			●	
Sonnenkrieg Division					●		●	
The Base					●	●	●	●
James Mason						●		

● Designated terrorist entity
● Designated under a synonym or umbrella group or by affiliation
● Content banned by the New Zealand Classification Office

Figure 2: Far-right terrorist groups in scope of the TCAP and where they are designated.

	UN	EU	US State	US Treasury	UK	Canada	Australia	New Zealand
2011 Norway Attack Perpetrator								
<i>Manifesto</i>								●
2019 Christchurch Attack Perpetrator								●
<i>Livestream</i>								●
<i>Manifesto</i>								●
2019 Halle Christchurch Attack Perpetrator								●
<i>Livestream</i>								●
<i>Manifesto</i>								●
2022 Buffalo Attack Perpetrator								●
<i>Livestream</i>								●
<i>Manifesto</i>								●
2022 Bratislava Attack Perpetrator								●
<i>Manifesto</i>								●

● Designated terrorist entity
● Designated under a synonym or umbrella group or by affiliation
● Content banned by the New Zealand Classification Office

Figure 3: Content created by far-right terrorist attack perpetrators in scope of the TCAP.

Expansion of the TCAP Inclusion Policy

During the reporting period, we have expanded the Inclusion Policy with five new entities. All these entities are classified as far-right terrorist entities.

2011 Norway attack perpetrator – 23 December 2021

The New Zealand government banned the manifesto produced by a far-right terrorist who killed 77 people in bomb and gun attacks in Oslo and Utøya, Norway, on 22 July 2011. Given the material is now deemed objectionable, the viewing, making, and distributing of the Manifesto is illegal in New Zealand. Hosting such material is also illegal, and the Chief Censor (the Chief Executive of the Classification Office) can require tech companies to block access to the Manifesto in New Zealand. Failure to comply can be sanctioned with a fine. **During the reporting period, we submitted 140 URLs and made 69 alerts to 5 tech platforms.**

2022 Buffalo attack perpetrator – 17 May 2022

The New Zealand government banned the livestream and the manifesto of the far-right terrorist who killed 10 people in a gun attack in Buffalo, New York, on 14 May 2022. The New Zealand Classification Office initially classified the attacker's manifesto as objectionable on 15 May, followed by the 6 minute 52 second livestream video of the attack on 16 May. This criminalises the possession and distribution of both publications and meant that the TCAP could start alerting material produced by the perpetrator early on. **During the reporting period, we submitted 174 URLs and made 122 alerts to 10 tech platforms.** Tech Against Terrorism has also become a partner to the Christchurch Call, to allow better information sharing in crisis scenarios. To engage our crisis response mechanisms, the New Zealand Classification Office alerts us as soon as a decision is made to ban terrorist content which may fall within scope of the TCAP, such that we are able to respond more swiftly to crises.

James Mason – 01 September 2022.

James Mason is listed by the Canadian government as a terrorist entity.¹⁶ The listing emphasises Mason's operational connection to internationally designated neo-Nazi groups such as Atomwaffen Division (AWD) and the ideological influence of his book, *Siege*, on contemporary far-right terrorist movements. **During the reporting period, we submitted 110 URLs and made 105 alerts to 5 tech platforms.**

2019 Halle attack perpetrator – 19 October 2022

The livestream and manifesto created by the Halle, Germany attack perpetrator, published before and during the attack on 09 October 2019, are banned by the New Zealand government. **During the reporting period, we submitted 10 URLs and made 8 alerts to 3 tech platforms.**

2022 Bratislava attack perpetrator – 19 October 2022

The manifesto created and distributed by the Bratislava, Slovakia attack perpetrator, published before the attack on 12 October 2022, is banned by the New Zealand government. We were alerted to the banning of the material immediately, allowing us to alert URLs containing the manifesto as part of our ongoing crisis response. **During the reporting period, we submitted 8 URLs and made 5 alerts to 1 tech platform.** We published a blog post analysing the dissemination of the manifesto content in the immediate aftermath of the attack.¹⁷

¹⁶ James Mason, Public Safety Canada, <https://www.publicsafety.gc.ca/cnt/ntnl-scrtr/cntr-trrrsm/lstd-ntts/crmt-lstd-ntts-en.aspx#63>

¹⁷ Far-Right Lone-Actor Terrorist Attacks and Violent Extremist use of File-Sharing Platforms, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news/Bratislava-analysis>

3. QUANTIFIED IMPACT OF THE TERRORIST CONTENT ANALYTICS PLATFORM

3.1 Summary of the Key TCAP Metrics

This section contains a detailed breakdown of the TCAP performance metrics, all of which are calculated across the reporting period from 1 December 2021 to 30 November 2022.

Metric	Description	Total
TCAP submissions	The number of unique URLs containing terrorist content submitted to the TCAP (by TAT's OSINT analysts or automated scrapers).	18,995
Alerts sent to tech platforms	The number of automated alerts sent to tech companies notifying them of terrorist content on their platform. Alerts are only sent to tech companies registered for TCAP alerts.	10,174
Percentage of alerted URLs offline	The percentage of content alerted to tech companies which is no longer accessible.	82%
Tech platforms alerted	The total number of tech platforms the TCAP has sent automated alerts to.	57
Tech platforms registered	The total number of tech companies registered to the TCAP and able to receive alerts following detection of verified terrorist content.	150 ¹⁸

Figure 4: Key TCAP metrics, descriptions, and total values for the reporting period.

There is a disparity between submissions sent to the TCAP and alerts sent by the TCAP, given that not all content submitted to the TCAP is subsequently notified to platforms. There are four main reasons for this:

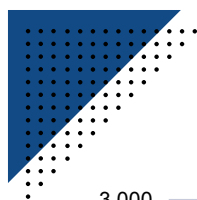
- 1) The content may have already been removed (no longer accessible);
- 2) We don't have a point of contact within the tech platform to send the alert to;
- 3) The platform where the content was identified is not subscribed to TCAP alerts;
- 4) The content may be hosted on a terrorist operated website (TOW).

Month	URL Submissions	Alerts Sent	Number of Tech Platforms Alerted
December 2021	1,516	937	36
January 2022	2,120	1,138	39
February 2022	1,800	936	30
March 2022	2,134	1,045	39
April 2022	2,481	1,271	32
May 2022	1,956	1,037	31
June 2022	1,129	655	29
July 2022	1,134	542	27
August 2022	1,369	662	23
September 2022	1,554	768	31
October 2022	877	565	25
November 2022	925	618	33
Total	18,995	10,174	57¹⁹

Figure 5: The TCAP metrics between 1 December 2021 and 30 November 2022.

¹⁸ Until the end of November 2022. At the time of writing, 195 tech companies are registered.

¹⁹ The total number of tech platforms alerted across Dec - Nov is not the sum of the individual months as each month there are several platforms consistently alerted.



TCAP URL Submissions and Alerts Sent

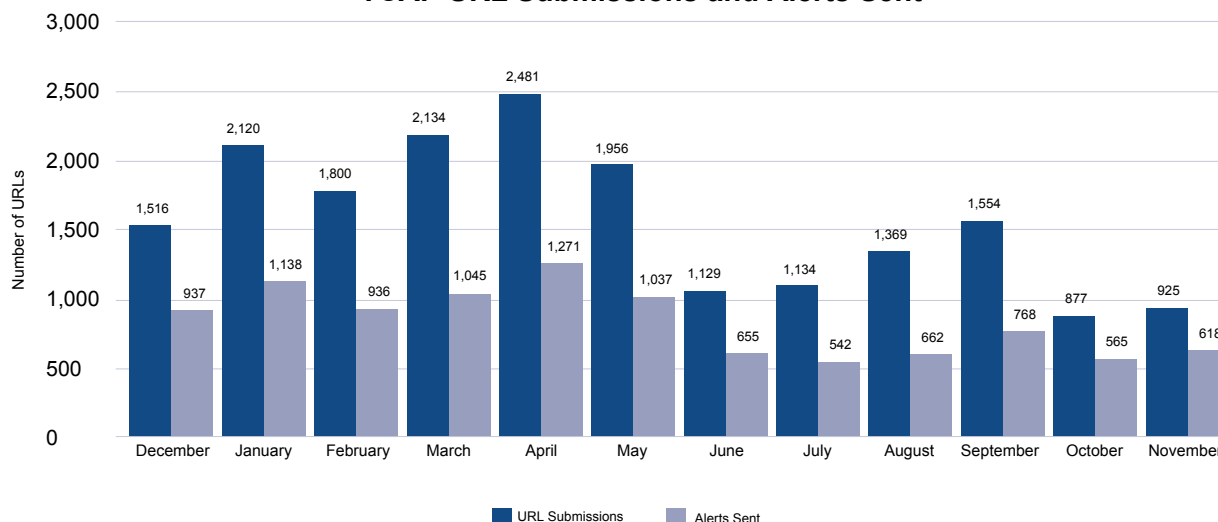


Figure 6: Alerts and submissions between 1 December 2021 and 30 November 2022

3.2 Takedown Rates

We record the percentage of flagged content which is no longer available after a TCAP alert is sent. We refer to content that is no longer available as being “offline.” For some URLs, the status is marked as “unknown” as we were unable to verify the status. For some URLs, platforms have restricted the content in certain locations, where we can determine this, we have classified the content as “geo-blocked.” As one of the TCAP’s key aims is to reduce the volume of terrorist propaganda on smaller tech platforms, the TCAP’s success may be measured in the high percentage of content recorded as offline after an alert is sent.

The percentage of URLs offline and online have been recorded per month. For the sake of this report, we checked all URLs after our reporting period, all URLs were checked in February 2023. The below table shows the monthly averages. In total, 82% of content is now offline.

Month	Alerts Sent	% URLs Offline	% URLs Online	% Geo-Blocked	% Status Unknown
December	937	75%	13%	11%	1%
January	1,138	89%	7%	2%	1%
February	936	84%	15%	1%	0%
March	1,045	77%	20%	0%	3%
April	1,271	86%	14%	0%	0%
May	1,037	82%	16%	0%	2%
June	655	91%	9%	0%	0%
July	542	78%	22%	0%	1%
August	662	90%	10%	0%	0%
September	768	74%	26%	0%	0%
October	565	84%	16%	0%	0%
November	618	71%	28%	0%	0%
Total	10,174	82%	16%	1%	1%

Figure 7: Breakdown of the key TCAP metrics across each month within the reporting period.



TCAP Alerted Content - Total Offline vs Online

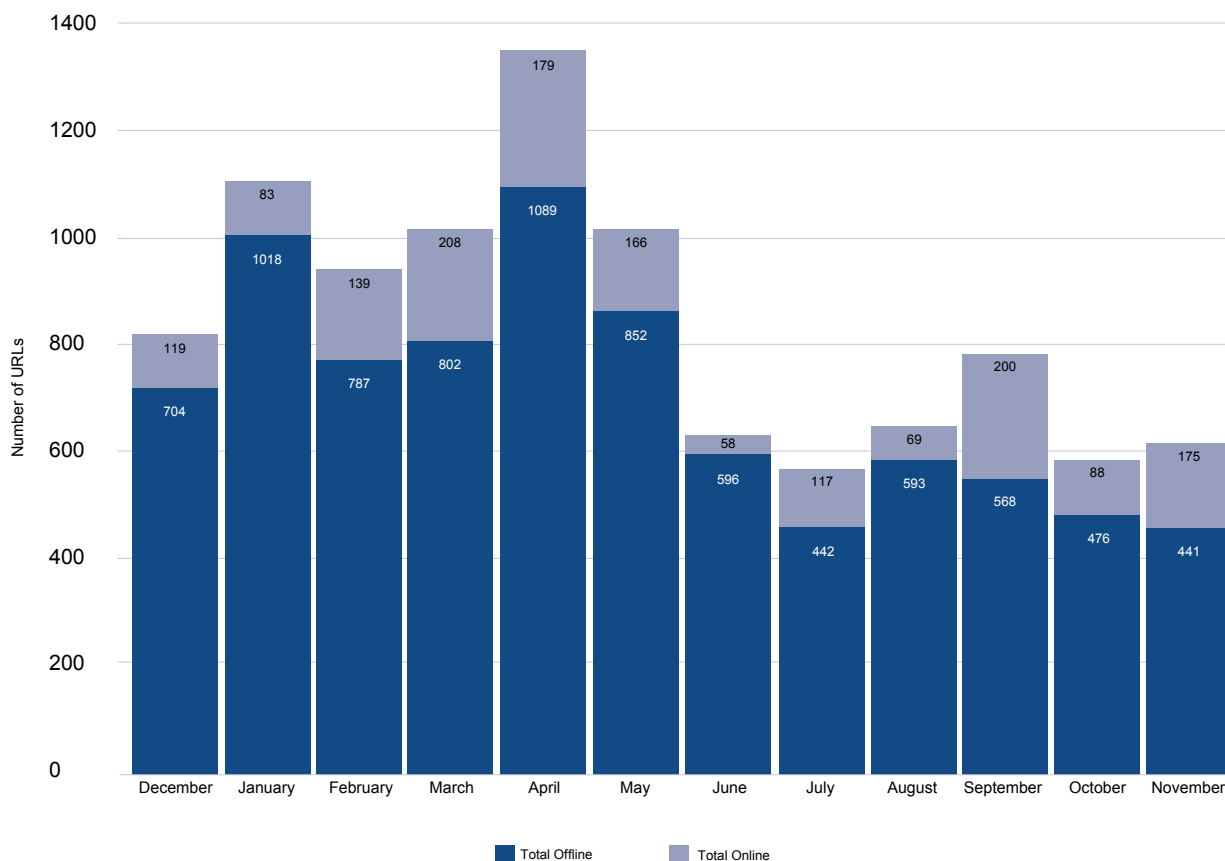



Figure 8: The number of URLs alerted to tech platforms which are online and the number which are offline (no longer available).

3.3 TCAP Submissions and Alerts per Platform Type

As outlined in the annex in section 4.1, the TCAP identifies and flags verified terrorist content found on various technology and internet platforms. These platforms vary in purpose and functionality. To date, the TCAP has identified terrorist content on 14 different types of platforms. The table below highlights these platform types and the core functionality of each; 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 used.



Platform Type	Functionality
Archiving	Storage of information on webpages or documentation from the past for anyone to view publicly.
Audio Sharing	Uploading, conversion, storage, and later consumption of audio content on the internet.
Book Subscription	Subscription to officially published and user-published books and documents.
File Sharing	Storage and public access of digital media online such as photos, videos, and documents, typically shared through a URL.
Forum	Online spaces for dedicated communities and chat rooms, typically consisting of specific conversation threads.
Link Shortener	Conversion of any URL into a shorter, more readable link to content hosted elsewhere online.
Messaging	Online chat in real time with individuals or larger groups and communities.
Paste Site	Uploading and sharing of text online, often used for sharing source code.
Photo Sharing	Uploading, conversion, storage, and later consumption of image content on the internet.
Search engine	Execution of web searches using key words and phrases and the indexing of webpages and websites.
Social Media	Creation and sharing of information through virtual communities and networks.
Video Hosting	Storage and access of digital video content online, typically accessed via a URL shared elsewhere online.
Video Sharing	Storage and access of digital video content online, with on-site search functions to navigate the platform.
Web Hosting	A static website or website available online, also including providers of such services.

The table below shows the total number of TCAP submissions and alerts within the reporting period, December 2021 – November 2022, categorised by the platform type on which the content was identified. The table also shows the percentage of the total number of TCAP alerts.

Platform Type	Number of URL Submissions	Number of Alerts Sent	Number of Tech Platforms	% of Total Alerts Sent
File Sharing	13,317	6,526	29	64%
Archiving	2,501	2,200	2	22%
Messaging	938	538	3	5%
Paste Site	1,127	489	3	5%
Video Sharing	181	158	5	2%
Social Media	120	99	5	1%
Video Hosting	115	70	2	1%
Book Subscription	32	29	1	0%
Photo Sharing	111	25	2	0%
Link Shortener	18	18	1	0%
Forum	21	12	1	0%
Audio Sharing	11	6	2	0%
Web Hosting	148	4	1	0%
Search engine	6	0	N/A	0%
Unknown	349	0	N/A	0%
Total	18,995	10,174	57	100%

Figure 9: The number of TCAP submissions and alerts per platform type.

Percentage of Total Alerts by Platform Type

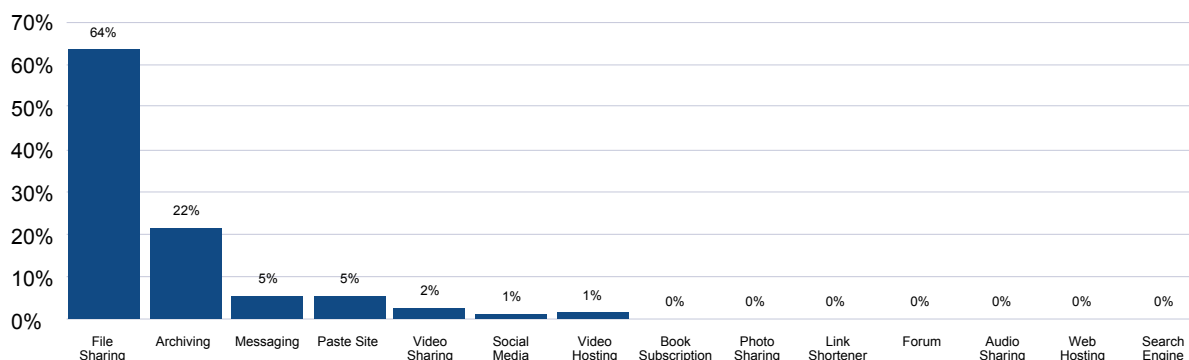


Figure 10: The percentage of TCAP alerts sent to different platform types.

The TCAP aims to counter terrorist use of the internet by supporting tech companies with the swift detection of terrorist content, after which they can take a decision on content moderation. The main goal is to ensure terrorist content can be removed before it gets the opportunity to spread further; the higher percentage of offline content after an alert is sent, the greater the success of the TCAP. Therefore, it is important to record the percentage of takedowns achieved by different types of platforms, to understand which type of platform best responds to our alerts, and which may need further support. The below table shows the takedown percentages per platform type.

Platform Type	Number of Alerts Sent	% Offline	% Online	% Geo-Blocked	% Status Unknown
File Sharing	6,526	94%	5%	0%	1%
Archiving	2,200	47%	47%	6%	0%
Messaging	538	68%	27%	0%	6%
Paste Site	489	99%	1%	0%	0%
Video Sharing	158	68%	32%	0%	0%
Social Media	99	96%	4%	0%	0%
Video Hosting	70	100%	0%	0%	0%
Book Subscription	29	76%	24%	0%	0%
Photo Sharing	25	96%	4%	0%	0%
Link Shortener	18	100%	0%	0%	0%
Forum	12	100%	0%	0%	0%
Audio Sharing	6	100%	0%	0%	0%
Web Hosting	4	25%	75%	0%	0%
Total	10,174	82%	16%	1%	1%

Figure 11: Takedown percentage per platform type.

3.4 TCAP Submissions and Alerts per Terrorist Entity

As outlined in the Inclusion Policy, the TCAP flags material produced by designated terrorist entities in scope. The below table summarises the number of URLs notified to platforms per entity type.

Terrorist Entity Ideology	URL Submissions	Alerts Sent
Islamist terrorism	18,048	9,436
Far-right terrorism	947	738
Total	18,995	10,174

Figure 12: Table showing the breakdown of TCAP submissions and alerts by the two terrorist entity ideologies in scope of the TCAP Inclusion Policy.

There are several explanations for the significant disparity between submissions and alerts for the two group types. Firstly, Islamist terrorist groups in scope of the TCAP often disseminate each piece of propaganda content (e.g., a video) with large lists of URLs that link to different file-sharing platforms. This dissemination technique makes the content easy to locate and to verify as official content as it is often disseminated from beacon channels.²⁰ This is very different from far-right terrorist groups, who often paste propaganda and material in-app, without sharing it in as an outlinked, URL version.

A second relevant factor is the verification of official content, which tends to be more difficult for far-right content. As mentioned, Islamist content is often disseminated through official beacon channels and can be verified due to the branding of official content with the associated media outlet. In contrast, a significant volume of far-right content is not branded but is supporter-generated, praising groups or individuals within scope of the TCAP through more subtle or coded messaging.

Third, the TCAP alerts tech companies that are willing to work with us and are not perceived as hostile, or a terrorist or extremist operated website. We often find far-right terrorist material on such platforms; in which case we cannot alert through the TCAP. In such scenarios, our OSINT team tackles this content in a different manner.

²⁰ 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.

Focus: TCAP submissions and alerts per Islamist terrorist entity

The table below shows the breakdown of TCAP submissions and alerts across the TCAP's designated Islamist terrorist groups.

Terrorist Entity	URL Submissions	Alerts Sent
Islamic State (IS)	7055	4,177
Al-Shabaab	3574	1,686
Al-Qaeda (AQ)	2739	1,155
Al-Qaeda in the Arabian Peninsula (AQAP)	1799	903
Islamic State West Africa Province (ISWAP)	972	549
Al-Qaeda in the Indian Subcontinent (AQIS)	341	197
Islamic State Central Africa Province (ISCAP)	294	171
Al-Qaeda in the Islamic Maghreb (AQIM)	379	160
Islamic State Sinai Province (ISSP)	142	84
Islamic State Somalia (ISS)	170	77
Islamic State Khorasan Province (ISKP)	105	63
Islamic State Pakistan Province (ISPP)	132	60
Islamic State East Asia Province (ISEAP)	130	58
Islamic State Libya Province (ISLP)	114	36
Islamic State India Province (ISIP)	59	31
Islamic State Greater Sahara (ISGS)	39	26
Taliban	3	3
Jama'at Nusrat al-Islam wal Muslimin (JNIM)	1	0
Total	18,048	9,436

Figure 13: TCAP submissions and alerts per Islamist terrorist group in scope.

Focus: TCAP submissions and alerts per far-right terrorist entity

The table below shows the breakdown of TCAP submissions and alerts across the TCAP's designated far-right terrorist groups.

Terrorist Entity	URL Submissions	Alerts Sent
Atomwaffen Division (AWD)	143	137
2019 Christchurch attack perpetrator	188	125
2022 Buffalo attack perpetrator	174	122
James Mason	110	105
2011 Norway attack perpetrator	140	69
Feuerkrieg Division (FKD)	52	51
National Socialist Order (NSO)	47	43
The Base	40	40
Sonnenkrieg Division (SKD)	18	18
National Action (NA)	13	11
2019 Halle attack perpetrator	10	8
2022 Bratislava attack perpetrator	8	5
Blood and Honour (B&H)	3	3
Scottish Dawn	1	1
Total	947	738

Figure 14: TCAP submissions and alerts per far-right terrorist entity in scope.

3.5 Takedown Percentages per Terrorist Entity in Scope

Tech Against Terrorism has tracked the removal rates by tech companies following TCAP alerts, which we provide below as segmented by group. We have included separately material which is specifically marked as “geo-blocked”. Geo-blocking is the practice of restricting access to content in certain geographical areas. This allows platforms using it to comply with local and regional legislation.

Terrorist Entity	Alerts Sent	% Offline	% Online	% Geo-Blocked	% Status Unknown
Islamic State (IS)	4177	78%	19%	3%	0%
Al-Shabaab	1686	93%	7%	0%	0%
Al-Qaeda (AQ)	1155	90%	9%	0%	1%
Al-Qaeda in the Arabian Peninsula (AQAP)	903	91%	8%	0%	1%
Islamic State West Africa Province (ISWAP)	549	79%	18%	1%	2%
Al-Qaeda in the Indian Subcontinent (AQIS)	197	91%	8%	0%	1%
Islamic State Central Africa Province (ISCAP)	171	73%	27%	0%	0%
Al-Qaeda in the Islamic Maghreb (AQIM)	160	97%	3%	0%	0%
Islamic State Sinai Province (ISSP)	84	77%	23%	0%	0%
Islamic State Somalia (ISS)	77	94%	6%	0%	0%
Islamic State Khorasan Province (ISKP)	63	70%	29%	0%	2%
Islamic State Pakistan Province (ISPP)	60	67%	23%	0%	10%
Islamic State East Asia Province (ISEAP)	58	81%	19%	0%	0%
Islamic State Libya Province (ISLP)	36	69%	31%	0%	0%
Islamic State India Province (ISIP)	31	87%	13%	0%	0%
Islamic State Greater Sahara (ISGS)	26	69%	31%	0%	0%
Taliban	3	33%	67%	0%	0%
Total	9436	84%	14%	1%	1%

Figure 15: Takedown percentages per Islamist entity in scope.

Percentage of Content Offline per Islamist Territory Entity

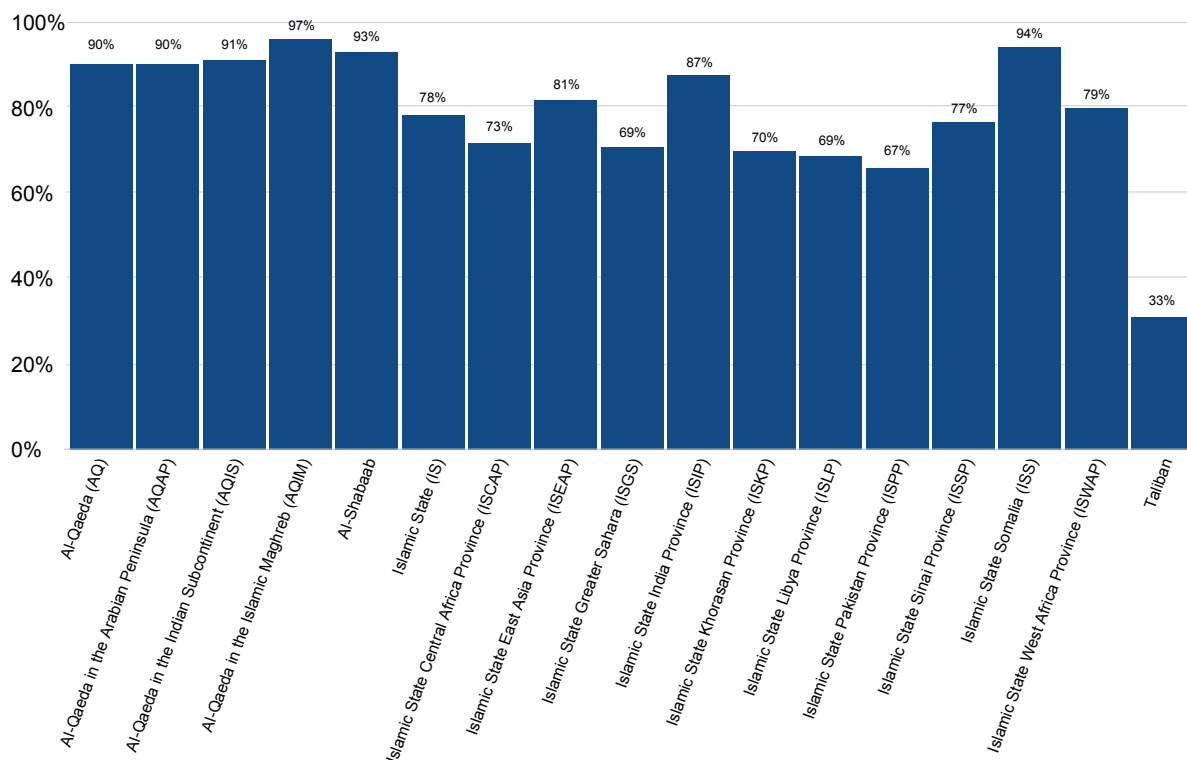


Figure 16: Takedown percentages per Islamist entity in scope.

Terrorist Entity	Alerts Sent	% Offline	% Online	% Geo-Blocked	% Status Unknown
Atomwaffen Division (AWD)	137	77%	18%	0%	5%
2019 Christchurch attack perpetrator	125	72%	22%	0%	6%
2022 Buffalo attack perpetrator	122	61%	31%	0%	7%
James Mason	105	8%	92%	0%	0%
2011 Norway attack perpetrator	69	36%	59%	0%	4%
Feuerkrieg Division (FKD)	51	92%	6%	0%	2%
National Socialist Order (NSO)	43	95%	5%	0%	0%
The Base	40	85%	10%	0%	5%
Sonnenkrieg Division (SKD)	18	89%	6%	0%	6%
National Action (NA)	11	18%	82%	0%	0%
2019 Halle attack perpetrator	8	38%	63%	0%	0%
2022 Bratislava attack perpetrator	5	40%	60%	0%	0%
Blood and Honour (B&H)	3	33%	67%	0%	0%
Scottish Dawn	1	0%	100%	0%	0%
Total	738	61%	35%	0%	4%

Figure 17: Takedown percentages per far-right entity in scope.



Percentage of Content Offline per Islamist Territory Entity

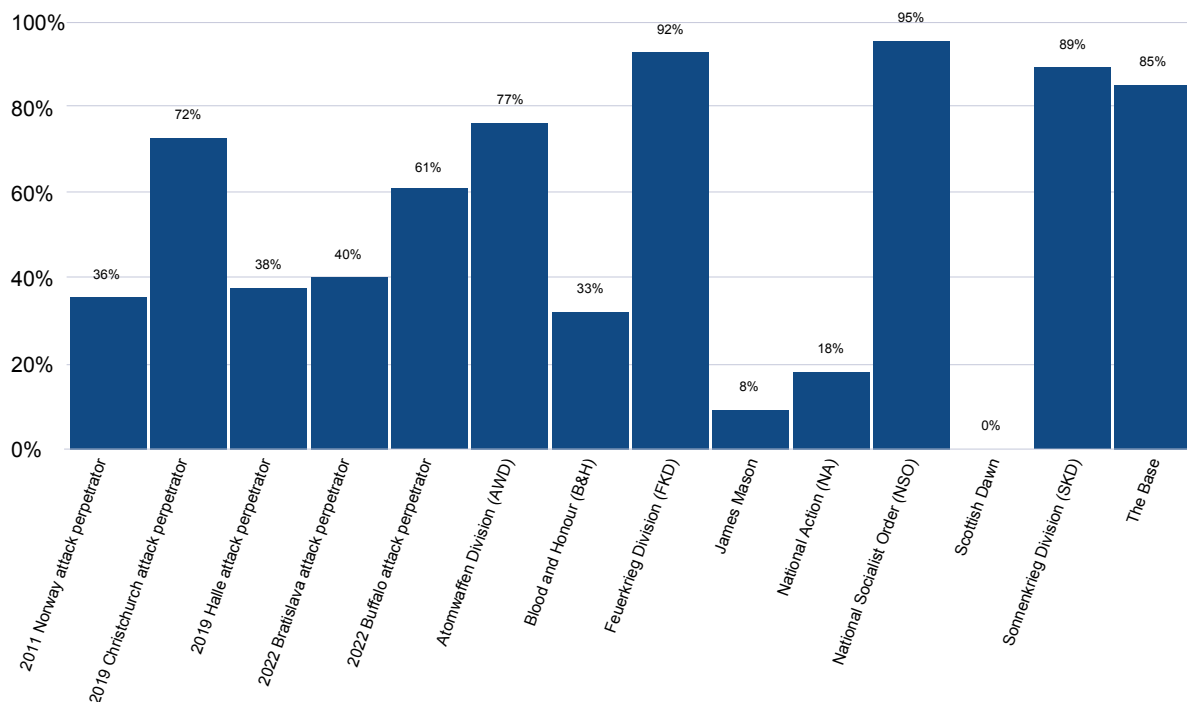


Figure 18: Takedown percentages per far-right entity in scope.

There is a significant difference between the removal percentage of Islamist terrorist content and far-right terrorist content. For Islamist content, this averages 84%, whilst for far-right terrorist content this averages a 61% takedown rate. Following on from our first Transparency Report we published a blog in March 2022 which explained the difference in takedown rates of content.²¹

²¹ Comparative Analysis of the TCAP Transparency Report Statistics on Content Collection and Removal Rates, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news/comparative-analysis-of-the-tcap-transparency-report>

3.6 Takedown Rates per Platform Type and Terrorist Entity

The below table segments takedown statistics for platform type by terrorist ideology.

Platform Type	Ideology	Alerts Sent	% Offline
File Sharing	Islamist	6,518	94%
	Far-right	8	100%
Archiving	Islamist	2,127	48%
	Far-right	73	12%
Messaging	Islamist	24	88%
	Far-right	514	67%
Paste Site	Islamist	487	99%
	Far-right	2	0%
Video Sharing	Islamist	81	91%
	Far-right	77	44%
Social Media	Islamist	60	95%
	Far-right	39	97%
Video Hosting	Islamist	70	100%
	Far-right	0	
Book Subscription	Islamist	17	82%
	Far-right	12	67%
Photo Sharing	Islamist	19	100%
	Far-right	6	83%
Link Shortener	Islamist	18	100%
	Far-right	0	
Forum	Islamist	9	100%
	Far-right	3	100%
Audio Sharing	Islamist	6	100%
	Far-right	0	
Web Hosting	Islamist	0	
	Far-right	4	25%
Total		10,174	82%

Figure 19: Alerts and takedown rates per platform type, separated by ideology.

3.7 Takedown Rates per Extreme Content Flag

All our alerts feature a specific flag for any content that is graphic or extreme in nature. This is to provide tech platform moderators with a warning before viewing content that they may find distressing. The flag can also be used by tech platforms for internal reporting purposes to classify content for transparency reporting. The below table summarises the submissions, alerts, and removal rates by the presence or absence of the extreme content flag.

Extreme Content	URL Submissions	Alerts Sent	% URLs Offline	% URLs Online	% Geo-Blocked	% Status Unknown
Yes	6,749	3,829	83%	14%	2%	1%
No	12,049	6,206	82%	16%	1%	1%
Unknown	197	139	76%	20%	0%	4%
March	18,995	10,174	82%	16%	1%	1%

Figure 20: Submissions, alerts, and takedown rates per extreme content flag.

Some platforms prioritise removing content which depicts violence or is graphic in nature. Based on our statistics, book subscription, file sharing, and video sharing platforms are more likely to remove content if it contains extreme content. The graph below summarises the removal rates of different platform types based on the extreme content flag. We have only included platform types which have received alerts containing both graphic and non-graphic content.

Removal Rates of Graphic vs Non-Graphic Content by Platform Type

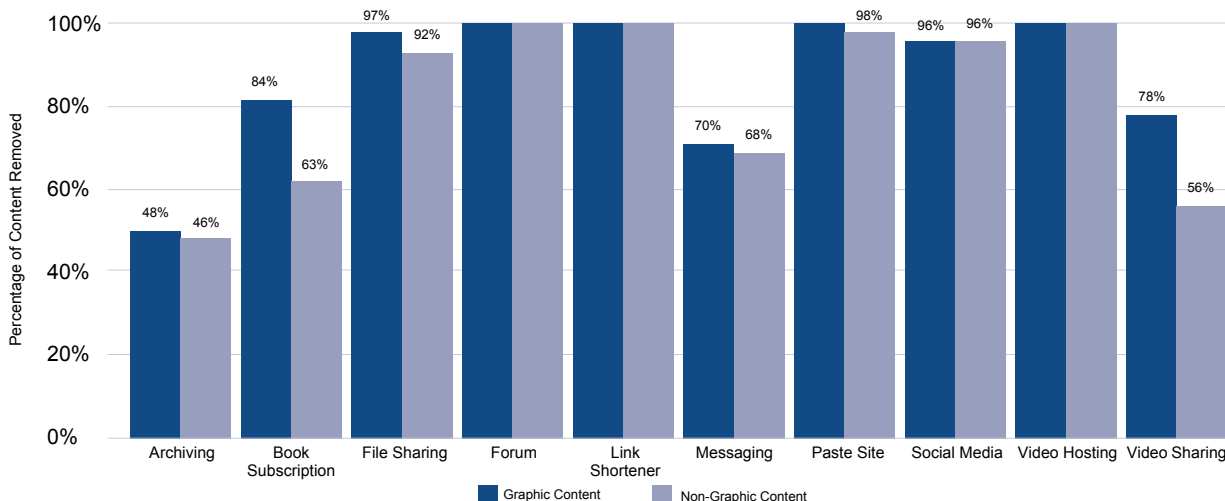


Figure 21: Takedown percentages per platform type, separated by extreme content flag.

The following graphs show the removal rates of content classified by terrorist entity based on the extreme content flag. Entities which we have not alerted with both extreme and non-extreme content have not been included in these graphs.

Removal Rates of Graphic vs Non-Graphic Content by Islamist Entity

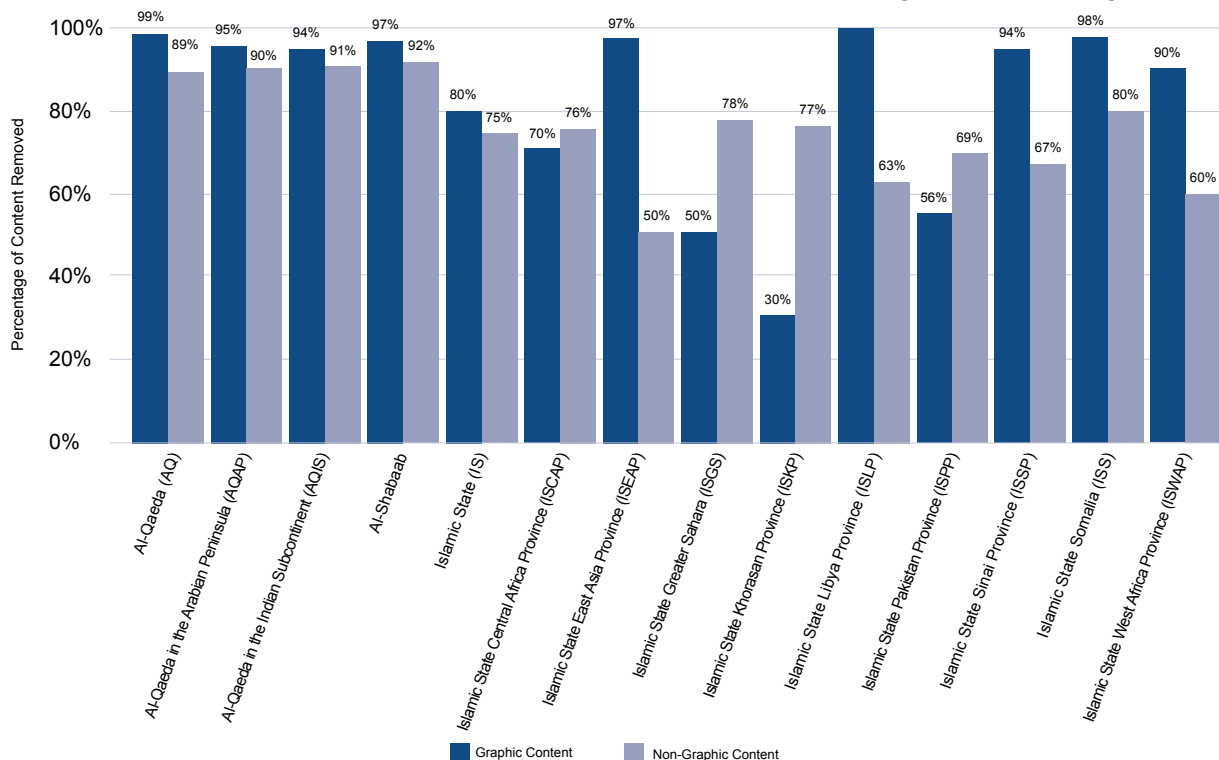


Figure 22: Takedown percentages per Islamist entity, separated by extreme content flag.

Removal Rates of Graphic vs Non-Graphic Content by Far-Right Entity

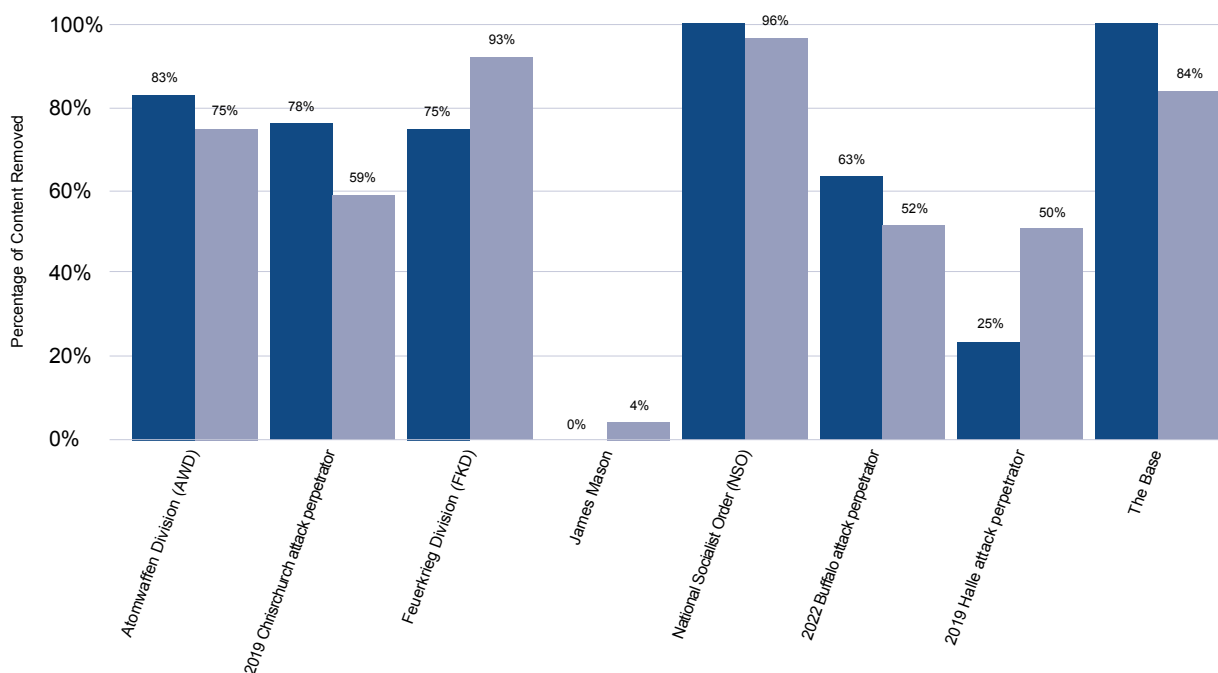


Figure 23: Takedown percentages per far-right entity, separated by extreme content flag.

3.8 Takedown Rates per Personally Identifiable Information (PII) Flag

All our alerts are flagged if the content contains Personally Identifiable Information (PII). Currently, most of our alerts are marked as “Unknown” as we cannot definitively determine if the content contains PII. There are multiple reasons for this, including content being published in languages we are unable to accurately translate and video content which is of low quality. The table below summarises the submissions, alerts, and takedown rates based on the PII classification.

PII	URL Submissions	Alerts Sent	% URLs Offline	% URLs Online	% Geo-Blocked	% Status Unknown
Yes	620	366	79%	19%	3%	0%
No	5,573	3,225	76%	22%	1%	1%
Unknown	12,802	6,583	85%	13%	1%	1%
March	18,995	10,174	82%	16%	1%	1%

Figure 24: Submissions, alerts, and takedown rates per Personally Identifiable Information (PII) flag.

As a UK-based NGO, we must abide by the legal requirements set out in the General Data Protection Regulation. Due to the nature of our alerts, the URL would have to contain PII, such as if a URL were to contain a name and date of birth. However, to ensure platforms have all relevant information in the alerts, the PII flag is used in the alert if PII is detected in the content itself. A condensed version of the full TCAP legal review may be requested on our website.²²

²² Legal Review, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news/tcaps-legal-review>

4. ANNEX

4.1 What is the Terrorist Content Analytics Platform?

4.1.1 Objectives

The key objectives of the Terrorist Content Analytics Platform²³ (TCAP) are as follows:

1. Support tech companies in detecting terrorist content on their platforms by alerting them to terrorist content, and by helping to inform and manage company moderation procedures by reference to the TCAP.
2. Facilitate affordable intelligence sharing for smaller internet platforms and help smaller tech companies to address terrorist use of their platforms expeditiously by means of an alert function.
3. Facilitate secure intelligence sharing between expert researchers and academics. By giving vetted academics and expert researchers access to the platform and a centralised dataset, the TCAP aims to improve the quantitative analysis of terrorist use of the internet and inform the development of accurate countermeasures.
4. Facilitate the coordination of data-driven solutions to counter terrorist use of the internet by making content on the platform available as a training dataset for the development of automated solutions.

The TCAP alerts tech companies to terrorist content found on their platforms. TCAP alerts are made on an advisory basis, and it is the sole decision of the tech platforms on how to proceed with content moderation decisions. In preparing its alerts, the TCAP marshals a large database of terrorist content collected in real time from verified terrorist channels on messaging platforms and apps. As a repository of verified terrorist content (imagery, video, PDFs, URLs, audio) collected from open-source platforms and existing datasets it also facilitates secure intelligence sharing between platforms.

The TCAP is also concerned with the method by which terrorists and violent extremists spread their content on the internet. Tech Against Terrorism assesses that terrorist and violent extremist use of the internet is increasingly concentrated on smaller platforms,²⁴ who struggle to action extremist content due to limitations of capacity, capability, and subject matter knowledge.²⁵ Our analysis suggests that smaller tech companies struggle with the technical requirements of moderating terrorist content and with implementing the solutions that are available to them.²⁶ Given that terrorist content will remain accessible if just one smaller tech company keeps this content online, we conclude that all smaller tech companies need to be supported in order to counter terrorist use of the internet effectively.

²³ Terrorist Content Analytics Platform, <https://www.terrorismanalytics.org/>

²⁴ State of Play: Trends in Terrorist and Violent Extremist Use of the Internet 2022, Tech Against Terrorism, <https://www.techagainstterrorism.org/2023/01/19/state-of-play-trends-in-terrorist-and-violent-extremist-use-of-the-internet-2022/>

²⁵ Analysis: ISIS use of smaller platforms and the DWeb to share terrorist content – April 2019, Tech Against Terrorism, <https://www.techagainstterrorism.org/2019/04/29/analysis-isis-use-of-smaller-platforms-and-the-dweb-to-share-terrorist-content-april-2019/>

²⁶ GIFCT Technical Approaches Working Group, Global Internet Forum to Counter Terrorism, <https://gifct.org/wp-content/uploads/2021/07/GIFCT-TAWG-2021.pdf>

To date, we have accomplished aims 1 and 2 of TCAP development. We are currently working on aims 3 and 4, further detail will be provided below.

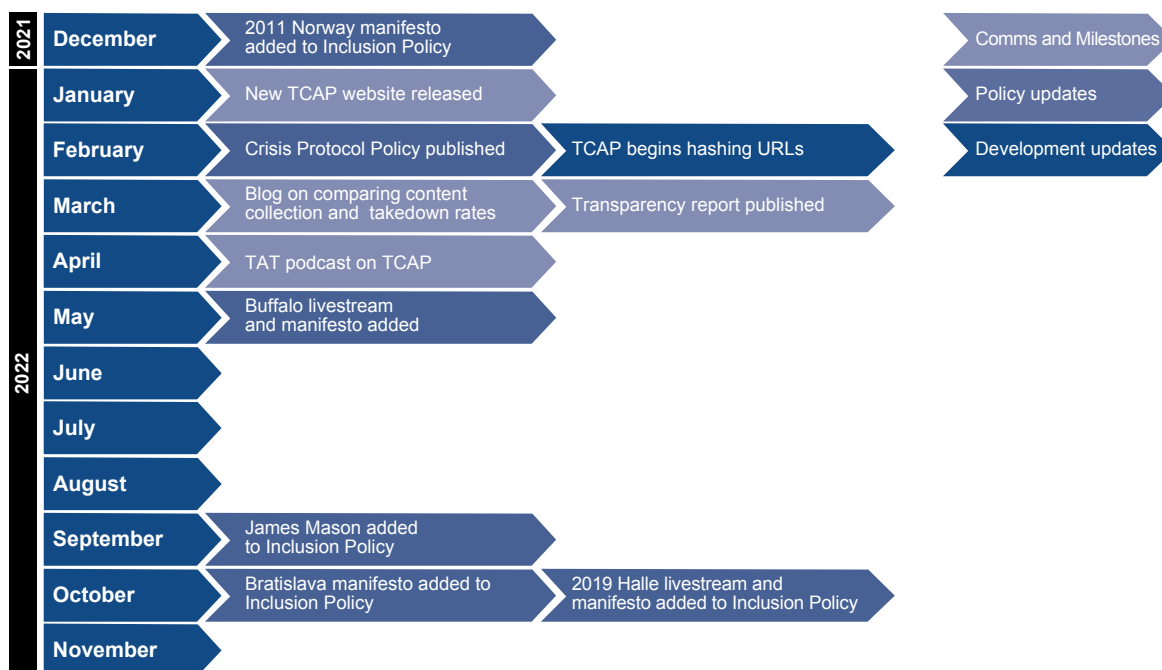


Figure 25: Top-level timeline of development, policy, and communication updates within the reporting period.

4.1.2 The TCAP process

This section details the end-to-end process of the TCAP, from identification of terrorist content on tech platforms to sending automated alerts.

The TCAP interferes with the dissemination of terrorist content on multiple levels. First, our OSINT experts trace terrorist groups to their preferred beacon platform, on which terrorists disseminate outlinks that direct users to smaller content stores, terrorist operated websites, and social media platforms on which the content is hosted. By means of beacon platforms, terrorists can spread propaganda exponentially. The TCAP aims to identify and alert platforms to the existence of these outlinks with the aim of the link being removed; in turn, content goes offline just as exponentially as it spreads, and as a result the terrorist content is harder to find. The TCAP therefore disrupts the entire ecosystem of tech platforms exploited by terrorists to disseminate their propaganda.

The below visualisation presents a top-level view of the end-to-end process used by the TCAP in collecting, classifying, and flagging terrorist content:

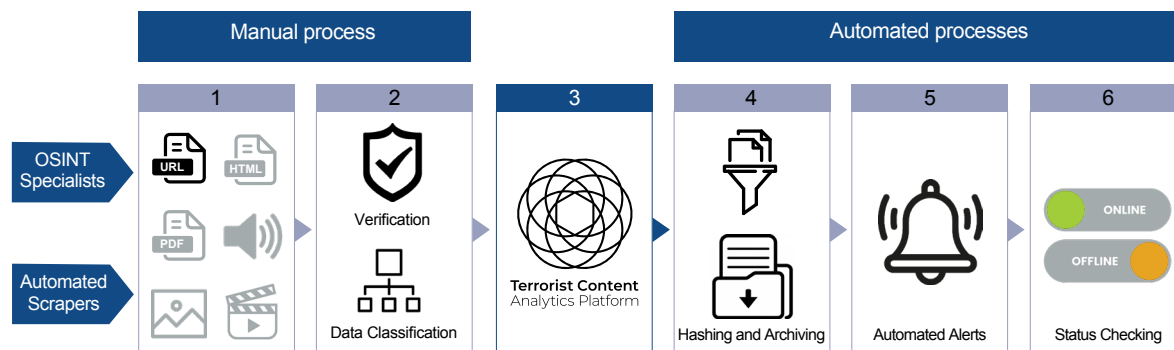


Figure 26: The TCAP’s process of identifying, collecting, verifying, archiving, and alerting terrorist material.

Step 1: Content discovery

The first step of the TCAP is the discovery of terrorist content, in line with our Inclusion Policy, across tech platforms. As of October 2021, the TCAP has two approaches to identifying terrorist content:

- Open-source intelligence (OSINT) analysis: Tech Against Terrorism’s OSINT team proactively traces terrorist groups to their preferred beacon platform. Terrorists use beacon platforms to post links to content stored on smaller platforms and terrorist-operated websites. These links are identified by the OSINT team.
- Automated web and mobile scrapers: The TCAP engineering team has built several automated scrapers²⁷ to extract data from those beacon platforms, comprising channels and chat rooms, which are known to host terrorist content. Once the scraper has exported the chat, an automated script scans the export to extract outlinks.

Step 2: Content verification & classification

After content has been identified, it is verified by the open-source intelligence team to ensure it is within scope of the TCAP’s Inclusion Policy. Any content identified which cannot be attributed to a designated group within the Inclusion Policy will not be uploaded to the TCAP. Content in scope will be classified and each content item assigned several different data attributes. The table below summarizes the data attributes captured for each content item:

Data Attribute	Description
Terrorist entity	The terrorist entity responsible for creating the content.
Tech Platform	The platform where the content was identified.
Channel Name	The specific channel on the tech platform where the content was identified, if applicable.
Channel URL	A link to the channel where the content was identified, if applicable.
PII Warning	Content containing Personally Identifiable Information.
Graphic Content Warning	Content containing violent graphics.
Date and Time of Collection	The date and time the content was submitted to the TCAP.
Outlink to Collected Content	The direct URL link to the content item.
Content Description	Top-level description of the content, such as the name of a video.

Figure 27: Data attributes stored for each content item on the TCAP.

²⁷ Web Scraping for OSINT: Techniques and Best Practices, Be4Sec, [https://be4sec.com/2023/03/14/web-scraping-for-osint-techniques-and-best-practices/#:~:text=Open%20Source%20Intelligence%20\(OSINT\)%20is,automatically%20extracting%20data%20from%20websites.](https://be4sec.com/2023/03/14/web-scraping-for-osint-techniques-and-best-practices/#:~:text=Open%20Source%20Intelligence%20(OSINT)%20is,automatically%20extracting%20data%20from%20websites.)

Step 3: Submission to TCAP database

After content has been verified and classified it is submitted to the TCAP to be processed for storage and informing notifications.

Step 4: Hashing and archiving content

Immediately after submission, the TCAP generates a hash of each content item. A hash is a distinct algebraic record of the content, which can be used to identify duplicated content. TCAP will soon begin sharing these hashes with GIFCT for inclusion in their hash-sharing database to support their work to prevent terrorist and violent extremist exploitation of digital platforms.²⁸ For more detail on our hashing of URLs and hash-sharing, we released a blog post.²⁹

The content, its associated metadata, and the hash is then added to the TCAP archive to ensure a record of the content is available for human rights and academic research purposes. The TCAP archive is currently not publicly accessible, but in later phases of development Tech Against Terrorism will look to grant access to verified academics and researchers.

Step 5: TCAP automated alerts

The TCAP then automatically identifies content collected from tech platforms which are registered for TCAP alerts. This content will be notified to the platform concerned via an automated email alert. Email alerts contain a link to where the content can be found on the platform in question, information about which designated terrorist group produced the content, and a warning for graphic content or material that contains PII. TCAP alerts are made on an advisory basis, and it is at the exclusive discretion of the alerted platforms to decide how to proceed with content moderation decisions.

For content identified on platforms not registered to the TCAP, the team will identify a contact email for the platform and share a preliminary notification to the content as well as explaining how the TCAP alerts operate. The tech platform can then register with the TCAP or ask to discontinue receiving notifications.

Step 6: Content status checking

After content has been flagged the TCAP runs an automated process to continuously validate the online status of each content item. This is used to determine whether the content has been taken down. Content which is tagged as 'online' is still publicly available via the source submitted to the TCAP and content tagged as 'offline' is no longer available.

²⁸ GIFCT's Hash-Sharing Database, Global Internet Forum to Counter Terrorism, <https://gifct.org/hsdb/>

²⁹ Announcement: The TCAP's hashing and hash-sharing capability, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news/hashing>

4.1.3 TCAP application interface

Registered tech platforms can log in to the TCAP interface to view and assess all terrorist content discovered on their platform to inform and support their content moderation decisions. The TCAP has a feature to allow tech platforms to dispute content they do not deem to be terrorist affiliated; the TCAP team reviews each content dispute and responds to the tech platform within 7 working days. During the reporting period, we received 36 content disputes. All these disputes were due to incorrect outputs of our automated takedown monitor. Following these content disputes, our team manually investigated the status of the alerted content. Following these disputes, we manually altered the status of 35 URLs and resolved the content disputes. We did not change the status of one URL as our analysts considered that the content remained active, but the content dispute was nonetheless resolved.

4.1.4 Automated scraping – additional information

Web and mobile scraping is the process of extracting data from websites and mobile applications. The TCAP has deployed several web and mobile scrapers to extract data on an ongoing basis from known terrorist channels across multiple platforms.³⁰ The TCAP utilises the Selenium framework³¹ for scraping platforms and a Celery framework, an open-source Python task queue which focuses on real time operations, for handling scraping requests.

The Selenium framework allows retrieval of essential data from the web or mobile site:

- Channel Meta Data (Channel Name, Share Link, Subscriber Count, Subscriber Names, Channel Description, Post Count)
- Channel Posts (Post Content, Post Number, Date Posted)

This data extracted by the scrapers is stored within a secure local AWS database of the TCAP's web framework application. The OSINT team analyses all content extracted by scrapers to ensure it continues to comply with our Inclusion Policy.

We define a channel as a specific location within a tech platform. For example, on a messaging platform, a channel is a specific chatroom where individuals are communicating.

4.2 Policy Considerations

4.2.1 Key development principles

At Tech Against Terrorism, our aim is to counter terrorist use of the internet while respecting human rights. This naturally extends to our development projects and includes (amongst other measures) building in safeguards to protect freedom of speech and the right to privacy. In developing the TCAP, there are eight principles that are crucial to our work. Below is a summary of these principles and how we implement them in practice.

³⁰ Tech Against Terrorism's OSINT team have undertaken extensive analysis of the internet to identify platforms and chat rooms used by terrorists to disseminate propaganda.

³¹ Selenium Framework, [https://en.wikipedia.org/wiki/Selenium_\(software\)](https://en.wikipedia.org/wiki/Selenium_(software))

Principle	Justification	Implementation
Rule of Law	Abiding by the rule of law provides democratic accountability and helps protect fundamental human rights. As the TCAP helps tech companies take content offline, it is essential that it is grounded in the rule of law to preserve these freedoms. To prevent setting speech norms unduly, with its inherent risks to human rights and especially freedom of expression, accuracy and accountability are vital for our work. Without this grounding, the TCAP risks establishing parallel and democratically unaccountable online speech norms.	<p>Our Inclusion Policy is based on designation lists of democratic nation states and supranational organisations' designation lists – this provides tech companies with the legal grounding to remove terrorist content from their platforms and protects freedom of expression.</p> <p>To date, we have only included official content, using our Content Classification and Verification Policy.</p>
Transparency	We want to ensure that the TCAP can be held accountable for the role it plays in countering terrorist use of the internet, which we can only do through transparency. We want to ensure that stakeholders have insight into the TCAP and the policies that guide it, as well the ability to give feedback on this process. ³²	<p>We are developing the TCAP through “transparency-by-design”, ensuring we are transparent in all phases of the process.</p> <p>All platform policies are available on request.</p> <p>We launched a public consultation process, the findings of which can be found in our report.</p> <p>We hold monthly Office Hours in which we provide an update on the development of the TCAP and stakeholders can ask questions and provide feedback.</p> <p>Anyone with TCAP access can share their views on classification. They can contest whether a generated alert concerns terrorist content.</p>
Accuracy and Accountability	We are aware that civil society groups have cautioned that a reliance on automated tools risks resulting in the wrongful removal of content and breaches of freedom of expression. ³³	<p>We only notified tech companies of verified content from targeted groups. These alerts are contained in email alerts which provide a URL to the content so the tech company in question can review the actual content.</p> <p>When we start sharing hashes with tech companies, we will build a “lookup” function, that allows tech companies to un-hash the material and examine the actual content.</p> <p>We implement a rigorous verification process using in-house terrorism experts to verify that the content is terrorist in nature - for more information see above in our policy section.</p> <p>Tech companies can dispute content when they think an alert is based on incorrect classification, and our team will review such content and keep a record for our transparency report.</p> <p>At the time of writing, we are setting up an Academic Advisory Board which will oversee our alerts, archive, and appeal process. The Board will superintend the accuracy of our alerts and their compliance with our Inclusion Policy and will also adjudicate any appeals made by TCAP's users.</p> <p>At all stages of development, we include civil society organisations, such as Human Rights Watch and Witness, to ensure we mitigate risks to human rights.</p>

³² At Tech Against Terrorism, we advise governments and tech companies to conduct regular transparency reports, to substantiate their transparency processes. We have launched our Transparency guidelines which considers how entities can do the same. Guidelines on transparency reporting on online counterterrorism efforts, Tech Against Terrorism, <https://transparency.techagainstterrorism.org/>

³³ One Database to Rule Them All, VoxPol, <https://www.voxpol.eu/one-database-to-rule-them-all/>

Principle	Justification	Implementation
Security	Given that TCAP archives content and its location, it is imperative that we build TCAP securely, so that terrorist entities don't gain access to the platform. We also need to ensure that terrorist entities do not become aware of our operations to the extent that it inhibits our mission or risks our operational security (OpSec).	<p>We follow strict OpSec protocols when conducting our open-source intelligence monitoring.</p> <p>Some of our policies and our office hours recordings are made available upon request, following a strict vetting process to ensure hostile actors won't be granted access.</p> <p>Our development team executes frequent penetration testing so that the TCAP as a platform can resist any attack.</p>
Privacy	Given the often sensitive nature of our alerts and the content we archive, the right to privacy is protected in the TCAP. This is also to prevent data ending up in the wrong hands, which could lead to individuals being targeted by retaliatory attacks from terrorist entities. It is therefore critical to enforce the right to privacy.	<p>Alerts to tech platforms come with a tag to show whether the content contains personal identifiable information (PII).</p> <p>A record of captured PII will be kept to preserve its potential to be used as digital evidence in war crimes trials or the prosecution of other human rights abuses.³⁴ Using Amazon Web Services infrastructure, all data will be kept in a highly secure, controlled environment.</p> <p>PII will only be shared when we come across an immediate and credible threat to life in line with our emergency Threat to Life Protocol.</p>
Freedom of Speech	We are very aware that the TCAP could pose risks to freedom of expression in content moderation without sufficient safeguards in place. When tackling terrorist use of the internet it is vital that this right is respected and not undermined by extra-legal mechanisms. We aim to safeguard against "content cartels" ³⁵ and uphold the right to free expression. We are aware that we, as a non-governmental organisation, should not set global norms for online speech.	<p>We base our Inclusion Policy on provisions of law, ensuring that we do not set speech norms online.</p> <p>We alert tech companies with the URLs containing the terrorist content so they can review the content and thereby avoid a dependence on automated removals compromising freedom of speech.</p> <p>Civil society participation ensures that relevant concerns can be raised and addressed. We support this participation through regular feedback sessions in office hours and our consultation report.</p> <p>All alerts are made on an advisory basis.</p>
Tech Platform Autonomy	To avoid content 'cartelisation', the TCAP alerts companies on an advisory basis only.	<p>All alerts are made on an advisory basis and will explain the reason for submission as well as the relevant designation guidelines relating to the groups in question.</p> <p>This is supported through our Knowledge Sharing Platform³⁶ and Online Regulation Series³⁷ that makes tech platforms aware of their duties in certain jurisdictions when notified of terrorist content on their platform.</p>

Figure 28: Core principles of the TCAP.

³⁴ "Video Unavailable" Social Media Platforms Remove Evidence of War Crimes, Human Rights Watch, <https://www.hrw.org/report/2020/09/10/video-unavailable/social-media-platforms-remove-evidence-war-crimes>

³⁵ Content cartel is a term coined by Evelyn Douek, who describes it as tech companies working together and taking content moderation decisions together without oversight. The Rise of Content Cartels, Colombia University, <https://knightcolumbia.org/content/the-rise-of-content-cartels>

³⁶ Knowledge Sharing Platform, Tech Against Terrorism, <https://ksp.techagainstterrorism.org/>

³⁷ Online Regulation Series, Tech Against Terrorism, <https://www.techagainstterrorism.org/wp-content/uploads/2021/07/Tech-Against-Terrorism---The-Online-Regulation-Series---The-Handbook-2021.pdf>

4.2.2 Content Classification and Verification Policy

To include only official material from the above terrorist entities in scope, we have created a Content Classification and Verification Policy³⁸ which we unveiled at the beginning of 2021. Our full policy is accessible on our website with registration required for security reasons.

Our Content Classification and Verification Policy operates in tandem with the Inclusion Policy to ensure that only official content is submitted to the TCAP. Official content is the material produced by a terrorist group or their media agency and differs from supporter-generated material, which is material published in support of a terrorist organisation. Our Content Classification and Verification Policy guides the analysis of content in the TCAP. Both the source and the material itself are assessed by our open-source intelligence experts. To verify the source, our experts identify core beacon channels through which a terrorist groups' messaging and propaganda is shared. To assess the content, our team conducts an intelligence assessment to determine whether the content has attributes associated with a high level of probability that the material was produced by a designated terrorist organisation in scope of the TCAP.

³⁸ Content Classification and Verification Policy, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/policies/content-classification-and-verification>

4.2.3 Background: public consultation process

Before commencing development of TCAP in 2019, Tech Against Terrorism opened a public consultation process by which tech companies, academics and members of civil society could provide feedback on what Tech Against Terrorism would need to consider when building the TCAP. Questions included the scope of TCAP and what type of tools would be most useful and solicited feedback on the fundamental principles.

In August 2020, we published a report³⁹ detailing the findings from this process as part of our commitment to ensuring that the platform is developed both transparently and in full observance of human rights and fundamental freedoms, including freedom of speech. The main findings and observations were as follows:

- Researchers and tech companies stressed that the TCAP should feature tools to facilitate analysis of terrorist content, in addition to an archive of terrorist content.
- Researchers emphasised the need to include content spanning multiple ideologies, with a particular focus on the global violent far-right.
- The TCAP should be transparent, and the platform should remain independent. Respondents also underlined the importance of respecting tech platform autonomy regarding moderation policy and enforcement decisions. As such, our alerts are given on an advisory basis only.
- Respondents from every sector stressed the importance of safeguarding the mental health and welfare of researchers and content moderators.

4.2.4 Legal consultation

In early 2021, Tech Against Terrorism commissioned a legal review to inform us about the legal considerations involved in building a platform of the TCAP's breadth. The legal review went on to be published in April 2021.

To uphold our principle of transparency and share best practice in the field, we want to make this legal analysis available for a select number of stakeholders. Whilst the full document is legally privileged, you can request the condensed, top-level version of the legal review on our website.⁴⁰

The legal review is divided into two sections: 1) civil actions, including offences such as defamation, malicious falsehood, misuse of private information 2) terrorism offences under relevant terrorism legislation. It also sets out some of the legal risks facing a publisher of terrorist material based in England – where Tech Against Terrorism is based – including some practical steps that can be taken to mitigate the risk of liability. The review also references relevant legislation from the European Union, Canada, the United States, and the United Kingdom.

³⁹ Consultation Report, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/policies/consultation-report>

⁴⁰ Legal Review, Terrorist Content Analytics Platform,

4.2.5 Crisis Protocol Policy

Data collection for the Terrorist Content Analytics Platform (TCAP) requires a wide-range of open-source intelligence (OSINT) across a variety of tech platforms. This data collection is targeted towards areas where terrorist and violent extremists spread propaganda, communicate, and recruit. Throughout our investigations, there is a possibility of finding data which gives information about an ongoing or future attack. As such, we have developed a Crisis Protocol Policy which covers three key areas of emergency incident management. These areas are pre-incident, during incident, and post-incident. Our Crisis Protocol Policy aims to be flexible to ensure that we can handle critical incidents in the most effective way possible. This Crisis Protocol Policy guides our actions when an emergency incident occurs, by ensuring we have provisions in place to alert the appropriate authorities and mitigate the threat posed by online violent extremist content.

Our Crisis Protocol Policy is based on similar policies created by the UK Police and Home Office. We aim to keep our Crisis Protocol Policy updated based on the development of the TCAP and aim to enhance the function of the TCAP as part of our crisis response workflow.

Pre-Incident

In the event of a potential threat to life, the Crisis Protocol Policy outlines the steps that TCAP staff take to evaluate the credibility and imminency of the threat to life and what proportionate actions should be taken.

A threat to life can be considered as:

- Real and immediate threat to a loss of life
- Threat to cause serious harm
- Threat of injury to another
- A threat to life also includes:
 - o serious sexual assault
 - o rape

Our assessment is based on considering the intent and capability of a potential attacker and collating intelligence to share with the appropriate law enforcement agencies. Each threat to life will be assessed as low, medium, or high, and is monitored for status change. We consider our ethical responsibility of reporting a threat to life as overriding the entities within the TCAP Inclusion Policy. While the Inclusion Policy may be used to support our report of a threat to life, association with a listed entity is not necessary for us to report a credible threat to life to authorities.

In the event of a potential, credible threat to life, we will inform the UK and local authorities, any relevant intelligence agencies, and continue to monitor the event. We will also ensure we keep an accurate archive of all relevant data, should it be needed.

In the event of a threat to life which cannot be verified as credible or immediate, such as in the event of doxing of a public figure, we will inform the relevant authorities and intelligence agencies. We will also continue to monitor the situation and escalate when necessary.

You can see our full threat to life protocol below, showing the workflow progression and the principal decisions involved in our assessments.

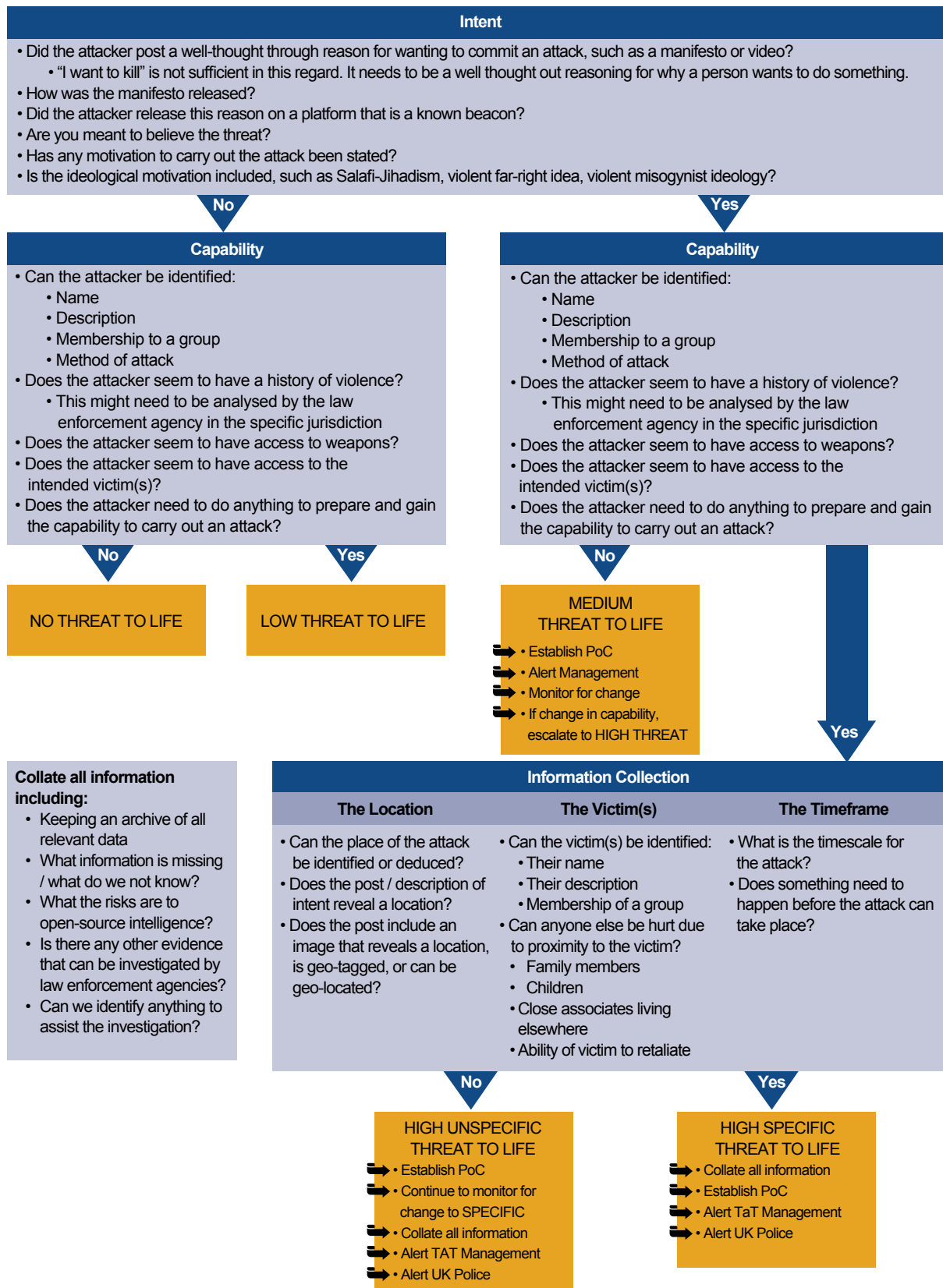


Figure 29: Threat to Life workflow for the Terrorist Content Analytics Platform.



During Incident

As seen with the Christchurch attack in 2019, there is an increasing threat of terrorist and violent extremist attackers utilising tech platforms to livestream and document attacks. In the event of an attack which is being livestreamed, the priority of the TCAP is to limit the spread of the content by flagging it to content moderators across a wide range of platforms. While large tech platforms are most likely to have the capability to immediately flag and remove duplicate versions of a livestream, it is equally likely that small tech platforms do not.

As with the pre-incident protocol, the potential threat-to-life involved in an ongoing crisis incident overrides the TCAP Inclusion Policy when the safety and security of the public is at stake. In the event of a livestreamed attack, we will provide the UK police and any other relevant authorities with all available data.

Currently, our alerting system sends alerts at 18:00 GMT daily, these alerts collate all URLs from the past 24 hours in one email to send to tech platforms. In the future, we will develop the TCAP to function as an immediate alerting system for all tech platforms to flag content from an attacker, whether it is an original livestream or a duplicate version. This will allow TCAP staff to override the regular alert function to send immediate alerts to tech companies, with the ability to add information about the event and content.

As a second priority, we also archive livestreams and footage of ongoing incidents. This archive has multiple purposes. The archive may be used to support prosecutions of terrorist and violent extremist actors by ensuring evidence is reliable and from an original source. The archive may also be used in the future to support expansion of the TCAP Inclusion Policy if the attacker is designated as a terrorist entity by a democratic nation state or supranational organisation. Finally, the archive may also be used to train artificial intelligence to assist in automated content moderation by training algorithms to identify potentially harmful content which can be flagged to human moderators more quickly for further review.

Post-Incident

As part of our regular TCAP data collection, we alert content which depicts attacks claimed by terrorist entities within the TCAP Inclusion Policy. By monitoring the designation lists of democratic nation states and supranational organisations and keeping our Inclusion Policy under review, we are not limited in our ability to flag terrorist content from a wide range of entities.

Our post-incident response to a crisis may also involve securely transferring intelligence data (such as livestream footage or other open-source data) to the relevant authorities.



4.3 Recognition

During the reporting period, the TCAP was widely acclaimed by multiple stakeholders:

- On 20 September 2022, the Government of Canada announced that it had awarded Tech Against Terrorism funding over three years for Phase II of the TCAP. At the Christchurch Call 2022 Leaders' Summit, the Prime Minister of Canada, the Right Honourable Justin Trudeau MP, announced the renewed funding.⁴¹
- Jonathan Hall KC, the Independent Reviewer of Terrorism Legislation in the UK, published his Annual Report on "The Terrorism Acts in 2020", in which he referenced Tech Against Terrorism's 'impressive focus on transparency and detailed analysis, going beyond mere research, which attempts to identify, through inclusion in a Terrorist Content Analytics Platform, content whose removal is justified.'^{42 43}
- Tech Against Terrorism was recognised in the Delhi Declaration issued by the UN Security Council Counter-Terrorism Committee on 29 October 2022.⁴⁴ Tech Against Terrorism attended the special meeting organised by UN CTED, highlighting emerging trends in terrorist use of the internet and the work of the TCAP in tackling this threat.

⁴¹ Government of Canada announces up to \$1.9 million in funding to combat online terrorist and violent extremist content, Public Safety Canada, <https://www.canada.ca/en/public-safety-canada/news/2022/09/government-of-canada-announces-up-to-19-million-in-funding-to-combat-online-terrorist-and-violent-extremist-content.html>

^{42,43} The Terrorism Acts in 2020, Jonathan Hall Q.C., https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1071570/IRTL_Report_Terrorism_Acts_in_2020.pdf

⁴⁴ Delhi Declaration on countering the use of new and emerging technologies for terrorist purposes, The Counter-Terrorism Committee, https://www.un.org/securitycouncil/ctc/sites/www.un.org/securitycouncil.ctc/files/outcome_document_ctc_special_mtg_final_e.pdf



4.4 Global Engagement

Over the reporting period, we briefed policymakers around the world on the TCAP:

- In January 2022, Anne Craanen, the Policy Lead of the TCAP, spoke on the Radicalisation Awareness Network's (RAN) podcast, discussing the impact of technology and the TCAP in preventing and countering violent extremism.⁴⁵
- In March 2022, Anne Craanen presented at the Club of Venice conference where she gave a demonstration of the TCAP and explained how we plan to develop the TCAP in the future.
- In May 2022, Anne Craanen presented at the Global Network on Extremism and Technology (GNET) Second Annual Conference.⁴⁶ She showcased the TCAP and our first Transparency Report.
- In June 2022, Anne Craanen presented at the Terrorism and Social Media conference in Swansea, showcasing the successes of the TCAP and explaining how we aim to develop the TCAP in the future.⁴⁷
- In June 2022, Anne Craanen presented on the TCAP at the Radicalisation Awareness Network's Strategic Communications Meeting on "Exploitation of technology by radicalising forces: developing an agile response."⁴⁸
- In July 2022, our Executive Director Adam Hadley attended the Global Internet Forum to Counter Terrorism's (GIFCT) Global Summit,⁴⁹ at which he emphasised the tremendous impact (see stats) of the TCAP in removing terrorist content.
- In July 2022, Anne Craanen appeared on the European Observatory of Online Hate's (EOOH) podcast 'Zooming in on Hate' discussing the future of the TCAP.⁵⁰
- In September 2022, Charley Gleeson, Open-Source Intelligence Analyst, presented at Tech Against Terrorism and the Global Internet Forum to Counter Terrorism's West Africa Conference on the role of the TCAP in countering terrorist propaganda in West Africa.⁵¹
- Throughout the reporting period, we hosted 12 monthly Office Hours sessions, giving us an opportunity to give our stakeholders regular updates on the development of the TCAP.⁵²

⁴⁵ RAN Podcasts, Radicalisation Awareness Network, https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran/ran-media/ran-podcasts_en

⁴⁶ The Second Annual GNET Conference, Global Network on Extremism and Technology, <https://gnet-research.org/resources/the-second-annual-gnet-conference/>

⁴⁷ Terrorism and Social Media, Cyber Threats Research Centre, <https://www.swansea.ac.uk/law/cytrec/projects/tasm/>

⁴⁸ Radicalisation Awareness Network, https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran_en

⁴⁹ Global Summit, Global Internet Forum to Counter Terrorism, <https://gifct.org/global-summit-2022/>

⁵⁰ Episode 9: Disrupting Online Terrorism, European Observatory of Online Hate, <https://eoooh.eu/podcasts/he2rnyb04b5no6q6k1xbdpieiy0kd0-nf5e8-z4879-alw4e-r6r2y-lsyth-jd8yx>

⁵¹ West Africa Workshop on Countering Terrorism and Violent Extremism Online, Global Internet Forum to Counter Terrorism, <https://gifct.org/events/west-africa-workshop-countering-terrorism-violent-extremism-online/>

⁵² Office Hours, Terrorist Content Analytics Platform, <https://terrorismanalytics.org/project-news>

4.5 What's Next?

Expanding Inclusion Policy

We will seek to update our Inclusion Policy to include more designated terrorist entities in line with evolving and existing designation. In this process we will consider the threat that an entity poses as well as the amount of online content a given entity disseminates. However, given many different groups are under consideration for inclusion, we will consider factors such as offline threat and quantity of online material disseminated when prioritising groups for inclusion. We continue to monitor the threat of other ideological forms of terrorism and may expand the scope of TCAP to include material produced by groups affiliated with other violent extremist ideologies when we have a legal basis to do so.

Tiered Alerts System

There has been a growing recognition in the field of online counterterrorism of the need to move beyond a purely group-based approach to understanding and defining terrorist content online. Through the TCAP tiered system, we will move beyond a reliance on terrorist designation to reflect and counter the post-organisational nature of the global terrorist threat. Meanwhile, we are committed to grounding our approach in the rule of law by providing legal bases for our policies and providing strict criteria for the inclusion of terrorist content to avoid setting undue speech norms and infringing on the right to freedom of speech.

Trusted Flagger Mechanism

We are working on a trusted flagger mechanism that allows practitioners and academics encountering terrorist content on the internet to alert this material to us. We will then verify the material to assess whether it is in scope of the TCAP. If it is, we will notify tech companies of this material. If not, we will assess whether the material violates any other laws and notify the authorities if legally required to do so. We hope that this mechanism will allow for practitioners and academics to flag more content for removal and thereby uphold the duty to report terrorist content.

The TCAP Archive

The TCAP will support academic research on terrorist content by providing a highly secure database of TCAP content accessible to verified academics. This will also allow us to include more far-right terrorist material since, as discussed, far-right terrorist groups frequently paste the material in-app, rather than through URL-sharing.

Development Features

- **Real-time scrapers:** We will develop additional real-time web and mobile scrapers capable of automatically detecting more terrorist content on a larger number of platforms. This in turn will increase TCAP submissions and alerts to tech platforms.
- **Application Programming Interface (API):** We are developing a TCAP API to allow tech companies to receive TCAP alerts directly within their platforms.
- **Content moderation workflow tool:** We will develop the technical infrastructure for a content moderation workflow tool in the TCAP to help tech companies prioritise content moderation queues and decisions.
- **Content analysis algorithms:** Subject to funding, we will look to design and develop content analysis algorithms to automate content moderation.