TRANSPARENCY REPORT

TERRORIST CONTENT ANALYTICS PLATFORM

YEAR ONE: 1 DECEMBER 2020 – 30 NOVEMBER 2021

tech against terrorism

Terrorist Content Analytics Platform

EXECUTIVE SUMMARY

Overview

- In November 2020, with support from Public Safety Canada,[1] Tech Against Terrorism launched the **Terrorist Content Analytics Platform (TCAP)**.[2] The world's largest database of verified terrorist content, collected in real time from verified terrorist channels on messaging platforms and apps, the 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.

- The TCAP is developed using a transparency-by-design approach. This is the first TCAP transparency report, which is one of several initiatives Tech Against Terrorism has taken in compliance with our core principles. The report provides a detailed breakdown of the core metrics for the reporting period between **1 December 2020 and 30 November 2021**, and of key TCAP policies and processes.

- The TCAP was commended by Prime Minister of Canada Justin Trudeau at the Christchurch Call to Action 2021 Summit. The TCAP was recognised in a number of reports by UN CTED and Human Rights Watch.[3]

Statistical breakdown of impact to date

During this reporting period:

- Our open-source intelligence experts submitted **18,958** URLs containing terrorist content, and the TCAP sent **11,074** alerts to **65** tech companies, **94%** of which is now offline. In total, **114** tech companies are registered and able to receive alerts as soon as we detect terrorist content on their platforms.

- **18,787** URLs containing Islamist terrorist content have been submitted to the TCAP, compared to **170** URLs containing far-right terrorist content. **10,959** alerts containing Islamist terrorist content were sent, whilst **115** 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.

[1] Tech Against Terrorism awarded grant by the Government of Canada to build Terrorist Content Analytics Platform
[2] Terrorist Content Analytics Platform
[3] 4.3 Recognition – page 29

- Tech platforms generally remove more Islamist terrorist content than far-right terrorist content as a result of our alerts. The average removal rate by tech companies following alerts of Islamist terrorist content is **94%**, whereas the average removal rate of far-right terrorist content is **50%**.

- Most Islamist terrorist content submitted to the TCAP, and made the subject of a TCAP alert, was produced by the Islamic State (40%), al-Qaeda in the Arabian Peninsula (22%), and al-Shabaab (21%).

- Most far-right terrorist content submitted to the TCAP, and alerted by the TCAP, was produced by the Christchurch attack perpetrator (51%) , Atomwaffen Division (AWD) (19%), and National Socialist Order (10%).

- The TCAP detected terrorist content on **13** different types of tech platforms. The three most exploited technology types in descending order were platforms providing file sharing, archiving, and link shortening services.

- Platforms providing link shortening, photo sharing, video hosting, and audio streaming services, as well as web hosting platforms, 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 **59%** of alerted content being removed.

The policies guiding our detection, verification, and alerts

- Before the development of the TCAP, Tech Against Terrorism conducted a consultation process during which we asked partners from tech companies, civil society, and academia what they thought we should take into consideration whilst building the TCAP. We produced a consultation report based on our findings.[4]

- From this consultation process, Tech Against Terrorism established the core principles which guide the TCAP's development, including the rule of law, transparency, accuracy, accountability, security, privacy, freedom of speech and platform autonomy.

- Before launching the TCAP, Tech Against Terrorism commissioned a review to establish the legal requirements to meet when developing the TCAP. Civil and criminal counterterrorism legislation in force in the United Kingdom, Europe, Canada, and the United States was considered. The top-level report was published in April.[5]

[4] Consultation Report
[5] Legal Review

- In addition to clarifying our principles and legal parameters, we also established robust policies to govern the TCAP in practice.
    - Our Inclusion Policy[6] is based on the legal designation of terrorist entities by democratic nation states and supranational institutions. During our reporting period, the TCAP notified material produced by the Islamic State, al-Qaeda, official affiliates, the Taliban, designated far-right groups, and the Christchurch attack perpetrator.

    - Our Content Classification and Verification Policy[7] ensures that all content subjected to TCAP processing is 'official' material produced by a designated entity clearly within the scope of counterterrorism efforts.

    - Our Crisis Protocol Policy[8] is used by our team to identify and assess a credible and imminent threat to life. This is based on the capability and intent of a potential attacker. The threats to life are given a score of low, medium, or high. All high threats to life are reported to the UK police and local authorities.

[6] Inclusion Policy
[7] Content Classification and Verification Policy
[8] Crisis Protocol Policy

TABLE OF CONTENTS

INTRODUCTION

In November 2020, with support from Public Safety Canada,[9] Tech Against Terrorism launched the **Terrorist Content Analytics Platform (TCAP)**,[10] a secure online tool that detects and verifies terrorist content and then alerts technology companies of the presence of such material on their platforms. The TCAP is one of the first technical tools designed to counter terrorist use of the internet while upholding human rights.

The TCAP was developed using a transparency-by-design approach.[11] This first transparency report, which provides a detailed breakdown of the core metrics for the reporting period between 1 December 2020 and 30 November 2021, is one step in ensuring the transparent and accountable development of the TCAP. It also details the core policies underpinning the TCAP to support tech platforms in their own efforts to action terrorist content swiftly whilst ensuring that human rights, including freedom of expression and the right to privacy, are protected.

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 provide insight regarding to what extent such abuse is taking place. 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 Efforts 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 the Guidelines.[12]

[9] Press Release on the development of the TCAP
[10] Terrorist Content Analytics Platform website
[11] Tech Against Terrorism, The Terrorist Content Analytics Platform - Transparency by Design, Voxpol.
[12] Tech Against Terrorism supports platforms in introducing and improving their transparency reports as part of its Mentorship Programme. For more information, see: Tech Against Terrorism's Transparency Guidelines

2. TERRORIST CONTENT IN SCOPE OF THE TCAP

2.1 Inclusion Policy

The TCAP includes 'official' material produced by terrorist groups and 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. At the time of writing, this includes content created by Islamic State (and official provinces), al-Qaeda (and verified affiliates), the Taliban, and designated far-right terrorist groups, such as Atomwaffen Division. The TCAP also implements the Christchurch Call to Action by notifying tech companies of material produced by the Christchurch attack perpetrator.

	TCAP	UN	EU	US STATE	US TREASURY	UK	CANADA	AUSTRALIA
Islamic State (IS)	●	●	●	●	●	●	●	●
Al-Qaeda (Central)	●	●	●	●	●	●	●	●
Al-Qaeda in the Arabian Peninsula (AQAP)	●	●	●	●	●		●	●
Al-Qaeda in the Islamic Maghreb (AQIM)	●	●	●	●	●	●	●	●
Jama'at Nusrat al-Islam was-Muslimin (JNIM)	●	●	●	●		●	●	
Al-Shabab	●	●	●	●		●	●	●
Al-Qaeda in the Indian Subcontinent (AQIS)	●		●	●	●		●	●
Hurras al-Din	●							
Islamic State West Africa Province (ISWAP)	●	●		●				
Islamic State Sinai Province	●			●			●	●
Islamic State Somalia (ISS)	●							●
Islamic State Greater Sahara (ISGS)	●			●				
Islamic State Central Africa Province (ISCAP)	●							
Islamic State Libya Province	●			●				
Islamic State Algeria Province	●							
Islamic State Tunisia Province	●							
Islamic State Khorasan Province (ISKP)	●	●		●			●	●
Islamic State India Province	●							
Islamic State Pakistan Province	●							
Islamic State East Asia Province	●	●		●		●	●	●
Taliban	●		●		●		●	

● Designated Terrorist Group ● Designated under a synnoym or umbrella group or by affiliation

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

	TCAP	UN	EU	US STATE	US TREASURY	UK	CANADA	AUSTRALIA
Blood and Honour	●						●	
National Action	●					●		
Combat 18	●						●	
Sonnenkrieg Division (SKD)	●					●		●
Scottish Dawn	●					●		
National Socialist Anti-Capitalist Action (NS131)	●					●		
System Resistance Network (SRN)	●					●		
Feuerkrieg Division	●					●		
Atomwaffen Divison (AWD)	●					●	●	●
National Socialist Order (NSO)	●					●	●	●
Russian Imperial Movement	●			●			●	
The Base	●					●	●	●
Proud Boys	●						●	

● Designated Terrorist Group ● Designated under a synnoym or umbrella group or by affiliation

Figure 2: Infographic showing the far-right terrorist groups in scope of the TCAP and where they are currently designated

[13] See our Inclusion Policy

Islamist terrorist groups: considerations in creating our Inclusion Policy

IS and al-Qaeda affiliates

To determine which IS and al-Qaeda affiliates to include in the initial scoping, we examined the designation lists shown in figure 1 to first verify whether the group was designated as a terrorist organisation. Second, we assessed whether the group is an official affiliate of al-Qaeda or IS through conducting Open Source Intelligence (OSINT) analysis of IS and al-Qaeda propaganda outlets and their methods of dissemination. We also consulted with leading experts on terrorist groups to verify our findings.

Taliban

The Taliban was later included within the scope of the TCAP in August 2021, when the Taliban took over the government of Afghanistan. Our full press release detailing our reasons can be found here.

Far-right terrorist groups: considerations in creating our Inclusion Policy

In line with our Inclusion Policy, we have included far-right groups that have been designated as terrorist organisations by a democratic country or supranational organisation. We update our Inclusion Policy in accordance with evolving designations of far-right groups, an example of this being our inclusion of Atomwaffen Division when it was designated by Canada as a terrorist organisation. In May 2021, and in support of the Christchurch Call to Action, we began to notify registered tech platforms of instances of the Christchurch perpetrator's manifesto and livestream detected on their services.[14]

[14] In December 2021, we also decided to alert the Norway attacker's manifesto to tech companies. More can be found in our policy. Given this falls outside our reporting period, more on this will covered in the second TCAP Transparency report.

3. QUANTIFIED IMPACT OF THE TCAP

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 2020 to 30 November 2021.

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).	**18,958**
Alerts sent to tech companies	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.	**11,074**
Percentage of flagged URLs offline	The percentage of content alerted to tech companies which is no longer accessible.	**94%**
Tech platforms notified	The total number of tech platforms the TCAP has sent automated alerts to.	**65**
Tech platforms registered	The total number of tech companies registered to the TCAP and able to receive alerts following detection of verified terrorist content.	**114**[15]

Figure 3: 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 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 and Violent Extremist Operate Website (TVEOW), in which case such content will be dealt with through our OSINT workstream.[16]

[15] Until the end of November 2021, at the time of writing, 120 tech companies are registered.
[16] To read more on our work on Terrorist and Violent Extremist Operated Websites (TVEOWS), please see our latest report.

Month	URL submissions	Alerts sent	Number of tech platforms alerted
December 2020	625	219	19
January 2021	1110	526	20
February 2021	1579	1035	28
March 2021	1511	806	29
April 2021	1747	1044	31
May 2021	1630	905	32
June 2021	1208	811	37
July 2021	1709	1003	34
August 2021	1566	929	32
September 2021	2262	1290	34
October 2021	1771	1052	35
November 2021	2239	1454	43
Total	**18957**	**11074**	**65***

Figure 4: The TCAP metrics between 1 December 2020 and 30 November 2021

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

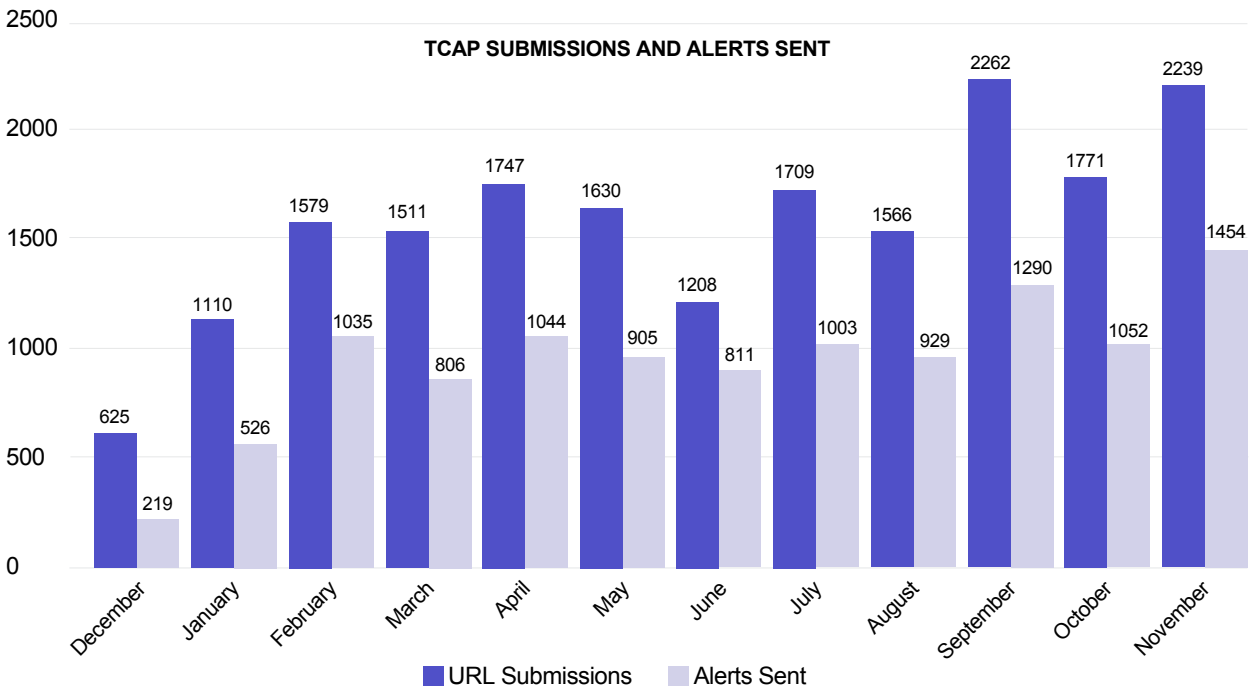

Figure 5: Alerts and submissions between 1 December 2020 and 30 November 2021.

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 unknown as we were unable to verify the status. 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 January 2022. The below table shows the monthly averages. In total, 94% of content is now offline.

Month	Alerts sent	% URLs offline	% URLs online	% Status unknown
December	219	96.80%	1.37%	1.83%
January	526	96.39%	2.85%	0.76%
February	1035	97.49%	1.55%	0.97%
March	806	98.26%	1.74%	0.00%
April	1044	98.85%	1.15%	0.00%
May	905	94.81%	4.86%	0.33%
June	811	98.64%	1.23%	0.12%
July	1003	95.61%	2.59%	1.79%
August	929	91.17%	7.00%	1.83%
September	1290	91.55%	7.05%	1.40%
October	1052	91.06%	8.75%	0.19%
November	1454	77.79%	22.08%	0.14%
Total	**11074**	**94.03%**	**6.40%**	**0.71%**

Figure 6: Breakdown of the key TCAP metrics across each month within the reporting period

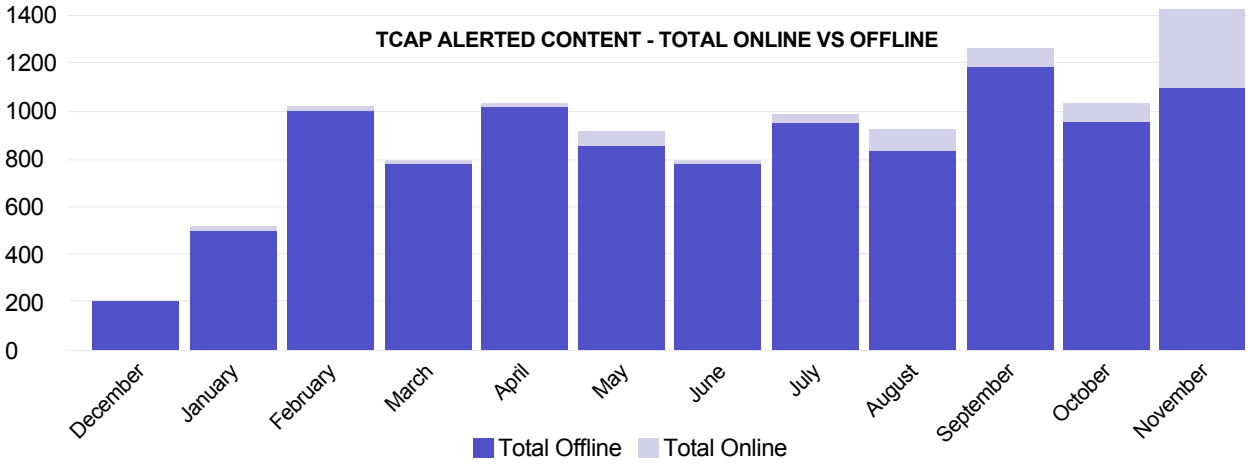

Figure 7: The number of URLs alerted to tech platforms which are online and the number which are offline (no longer available)

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 13 different types of platforms. The table below 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 used.

Platform type	Functionality provided
File Sharing	Access to digital media such as photos, videos, and documents.
Archiving	Storage of information on webpages or documentation from the past for anyone to view publicly.
Video Sharing	Uploading, conversion, storage and later consumption of video content on the internet.
Paste Site	Uploading and sharing of text online, often used for sharing source code.
Link Shortener	Conversion of any URL into a shorter, more readable link.
Instant Messenger	Online chat in real time with individuals or larger groups and communities.
Search	Execution of web searches using keywords or phrases.
TOW	Terrorist operated websites are sites which are controlled by terrorist groups.
Social Media	Creation and sharing of information through virtual communities and networks.
Book Subscription	Subscription to officially published and user-published books and documents.
Audio Sharing	Uploading, conversion, storage and later consumption of audio content on the internet.
Photo Sharing	Uploading, conversion, storage and later consumption of photo content on the internet.
Web Hosting	Posting a website or webpage onto the internet.

Figure 8: Types of platforms exploited by terrorists and alerted by the TCAP

The table below shows the total number of TCAP submissions and alerts within the reporting period, December 2020 – November 2021, 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 submissions	Number of platforms	Number of alerts	% of total
File Sharing	13257	34	8634	78%
Archiving	1829	2	1286	12%
Link Shortener	512	2	514	5%
Paste Site	734	1	250	2%
Video Sharing	252	8	183	2%
Messaging	589	5	98	1%
Social Media	47	7	38	0%
Book Subscription	26	1	23	0%
Photo Sharing	21	1	20	0%
Video Hosting	558	1	19	0%
Audio Streaming	11	1	7	0%
Web Hosting	2	1	2	0%
Unknown	1119	n/a	n/a	n/a
Total	**18957**	**65**	**11074**	**100%**

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

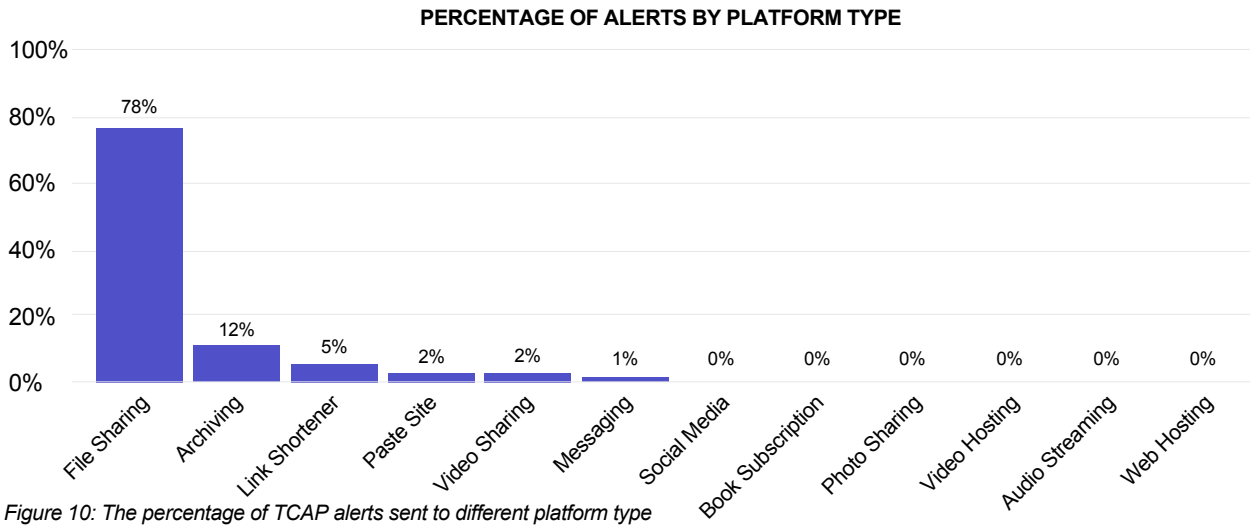

Figure 10: The percentage of TCAP alerts sent to different platform type

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 platform, 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	% URLs offline	% URLs online	% Status unknown
File Sharing	863	97.78%	1.42%	0.80%
Archiving	1286	59.02%	40.98%	0.00%
Link Shortener	514	100.00%	0.00%	0.00%
Paste Site	250	94.40%	5.60%	0.00%
Video Sharing	183	88.52%	9.29%	2.19%
Messaging	98	72.45%	22.45%	5.10%
Social Media	38	89.47%	7.89%	2.63%
Book Subscription	23	86.96%	13.04%	0.00%
Photo Sharing	20	100.00%	0.00%	0.00%
Video Hosting	19	100.00%	0.00%	0.00%
Audio Streaming	7	100.00%	0.00%	0.00%
Web Hosting	2	100.00%	0.00%	0.00%

Figure 11: Takedown percentage per platform type

3.4 TCAP submissions and alerts per designated terrorist group

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

Terrorist group type	TCAP submissions	Alerts sent
Islamist terrorist	18,787	10,959
Far-right terrorist	170	115
Total	**18,957**	**11,074**

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

There are a number of 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.[17] 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 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 oftentimes 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.

Fourth, the TCAP includes more Islamist groups than far-right groups, as there are less far-right groups designated by democratic countries and supranational institutions. At Tech Against Terrorism, we encourage governments to designate more far-right terrorist groups and ban material, to ensure that tech companies have legal grounding and clarity on what types of content they should moderate. Throughout the TCAP's reporting year, we have included far-right groups as soon as democratic nation states have designated far-right terrorist groups, Atomwaffen Division being one example.

[17] 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 officials channels on them that signify their central communications.

Focus: TCAP submissions and alerts per designated Islamist terrorist groups

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

Terrorist group	Submissions	Alerts
Islamic State	7146	4364
al-Shabab	4521	2325
al-Qaeda in the Arabian Peninsula (AQAP)	4186	2450
al-Qaeda	1269	781
Islamic State West Africa Province (ISWAP)	410	246
Jama'at Nusrat al-Islam wal Muslimin (JNIM)	281	134
Islamic State Sinai Province	144	164
Islamic State Khorasan Province (ISKP)	161	112
al-Qaeda in the Indian Subcontinent (AQIS)	272	114
Islamic State Central Africa Province (ISCAP)	102	61
al-Qaeda in the Islamic Maghreb (AQIM)	69	37
Hurras al-Din	87	56
Islamic State Libya Province	38	34
Islamic State Pakistan Province	38	31
Islamic State Somalia (ISS)	41	28
Taliban	13	13
Islamic State India Province	9	9
Total	**18787**	**10959**

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

Focus: TCAP submissions and alerts per designated 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 group / entity	Submissions	Alerts
Christchurch Attack Perpetrator	97	59
Atomwaffen Division (AWD)	29	22
National Socialist Order (NSO)	12	11
Feuerkrieg Division	11	9
The Base	6	4
Combat 18	5	4
Blood and Honor	4	3
National Action	3	1
National Socialist Anti-Capitalist Action (NS131)	1	1
Russian Imperial Movement (RIM)	2	1
Total	**170**	**115**

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

The above table shows that most far-right content we submitted and alerted was created by the Christchurch attack perpetrator. This is partly because we often see this video being shared as a URL across multiple channels, whereas much of the content produced by far-right terrorist groups included in the TCAP's scope is often shared directly within a chat room or channel without outlinks often meaning it takes longer to detect and is not suitable to upload in URL format.

3.5 Takedown percentages per designated 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.

Terrorist group	Alerts	Alerts %	Inactive	% Offline	% Online
Islamic State (IS)	4364	40%	3847	88%	12%
Al-Shabab	2325	21%	2313	99%	1%
Al-Qaeda in the Arabian Peninsula (AQAP)	2450	22%	2404	98%	2%
Al-Qaeda (Central)	781	7%	733	94%	6%
Islamic State West Africa Province (ISWAP)	246	2%	213	87%	13%
Jama'at Nusrat al-Islam wal Muslimin (JNIM)	134	1%	131	98%	2%
Al-Qaeda in the Indian Subcontinent (AQIS)	164	1%	161	98%	2%
Islamic State Khorasan Province (ISKP)	112	1%	102	91%	9%
Islamic State Sinai Province	114	1%	108	95%	5%
Islamic State Central Africa Province (ISCAP)	61	1%	49	80%	20%
Hurras al-Din	3	0%	37	100%	0%
Al-Qaeda in the Islamic Maghreb (AQIM)	56	1%	55	98%	2%
Islamic State Libya Province	34	0%	34	100%	0%
Islamic State Pakistan Province	31	0%	31	100%	0%
Islamic State Somalia (ISS)	28	0%	2	82%	18%
Taliba	13	0%	12	92%	8%
Islamic State India Province	9	0%	9	100%	0%
Total	**10959**	**100%**	**10262**	**94%**	**6%**

Figure 15: The takedown percentages per Islamist group included

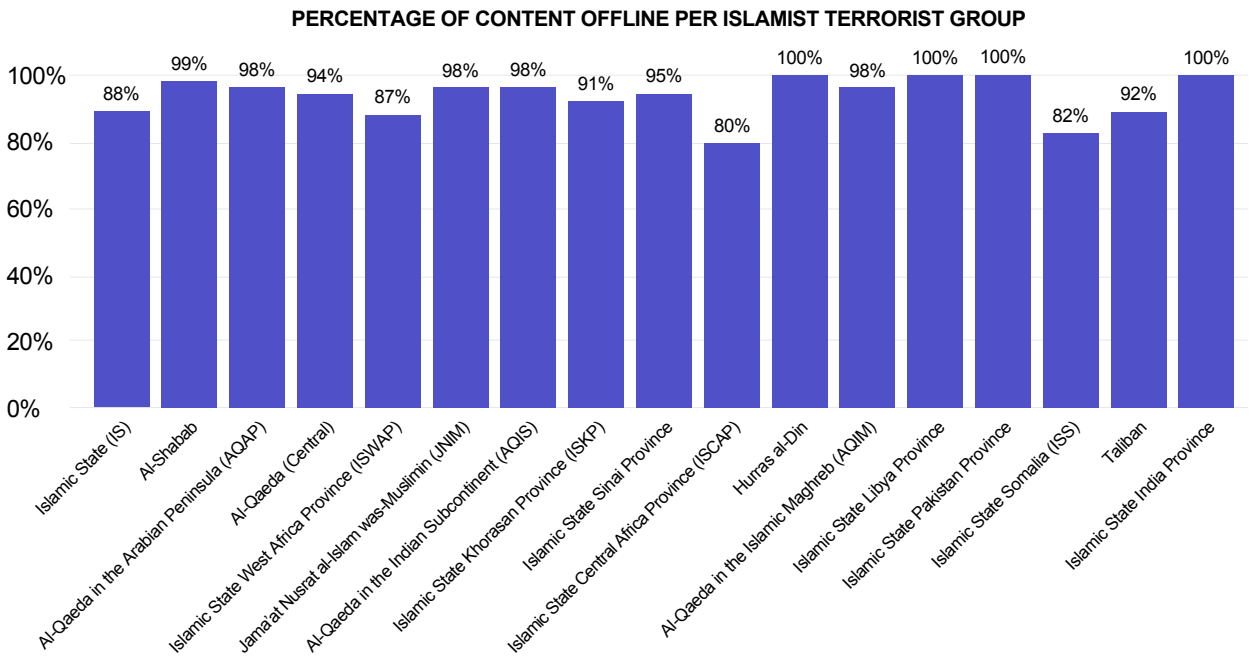

Figure 16: Takedown percentage per Islamist terrorist group in scope

Terrorist group	Alerts	Alerts %	Inactive	% Offline	% Online
Christchurch attack perpetrator	59	51%	26	44%	56%
Atomwaffen Division	22	19%	14	64%	36%
National Socialist Order (NSO)	11	10%	3	27%	73%
Feuerkrieg Division	9	8%	9	100%	0%
The Base	4	3%	2	50%	50%
Combat 18	4	3%	3	75%	25%
Blood and Honour	3	3%	0	0%	100%
National Action	1	1%	0	0%	100%
National Socialist Anti-Capitalist Action (NS131)	1	1%	1	100%	0%
Russian Imperial Movement (RIM)	1	1%	0	0%	100%
Total	**115**	**100%**	**58**	**50%**	**50%**

Figure 17: The takedown percentages per far-right terrorist group

There is a significant difference between the removal percentage of Islamist terrorist content versus far-right terrorist content. For Islamist content, this averages 94%, whilst for far-right terrorist content this averages a 50% takedown rate. For a more detailed explanation of why we assess this is, please keep an eye out on our upcoming article on the TCAP website.

3.6. Takedown rates per platform type and designated entity

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

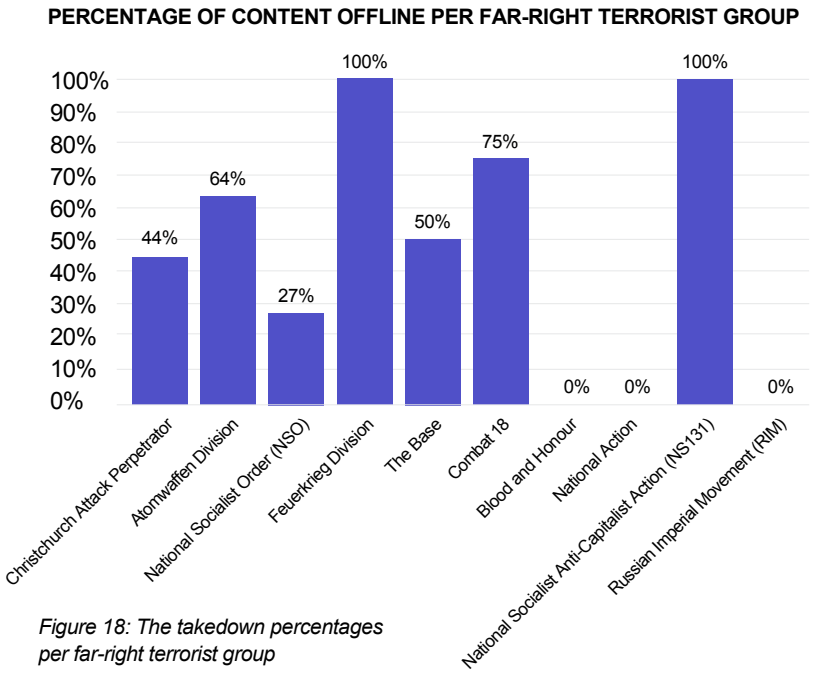

Figure 18: The takedown percentages per far-right terrorist group

Platform type	Alerts	Alerts %	Inactive	% Offline
File Sharing	Far-right	0		N/A
	Islamist	8634	8442	97.78%
Archiving	Far-right	43	9	20.93%
	Islamist	1243	750	60.34%
Link Shortener	Far-right	0		N/A
	Islamist	514	514	100.00%
Paste Site	Far-right	0		N/A
	Islamist	250	236	94.40%
Video Sharing	Far-right	49	34	69.39%
	Islamist	134	128	95.52%
Messaging	Far-right	18	14	77.78%
	Islamist	80	57	71.25%
Social Media	Far-right	2	1	50.00%
	Islamist	36	33	91.67%
Book Subscription	Far-right	3	0	0.00%
	Islamist	20	20	100.00%
Photo Sharing	Far-right	0		N/A
	Islamist	20	20	100.00%
Video Hosting	Far-right	0		N/A
	Islamist	19	19	100.00%
Audio Streaming	Far-right	0		N/A
	Islamist	7	7	100.00%
Web Hosting	Far-right	0		N/A
	Islamist	2	2	100.00%

Figure 19: Alerts and takedown rates per platform type and ideology of the terrorist entity creating the content

4. ANNEX

4.1 What is the TCAP?

4.1.1 Objectives

Launched in November 2020, with support from Public Safety Canada,[18] the key objectives of the Terrorist Content Analytics Platform[19] (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,[20] who struggle to action extremist content due to limitations of capacity, capability, and subject matter knowledge.[21] 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.[22] 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.

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.

[18] Tech Against Terrorism awarded grant by the Government of Canada to build Terrorist Content Analytics Platform
[19] Terrorist Content Analytics Platform
[20] Q1-Q2 report
[21] Analysis: ISIS use of smaller platforms and the DWeb to share terrorist content
[22] GIFCT Technical Approaches Working Group

Figure 20: 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 ultimate 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:

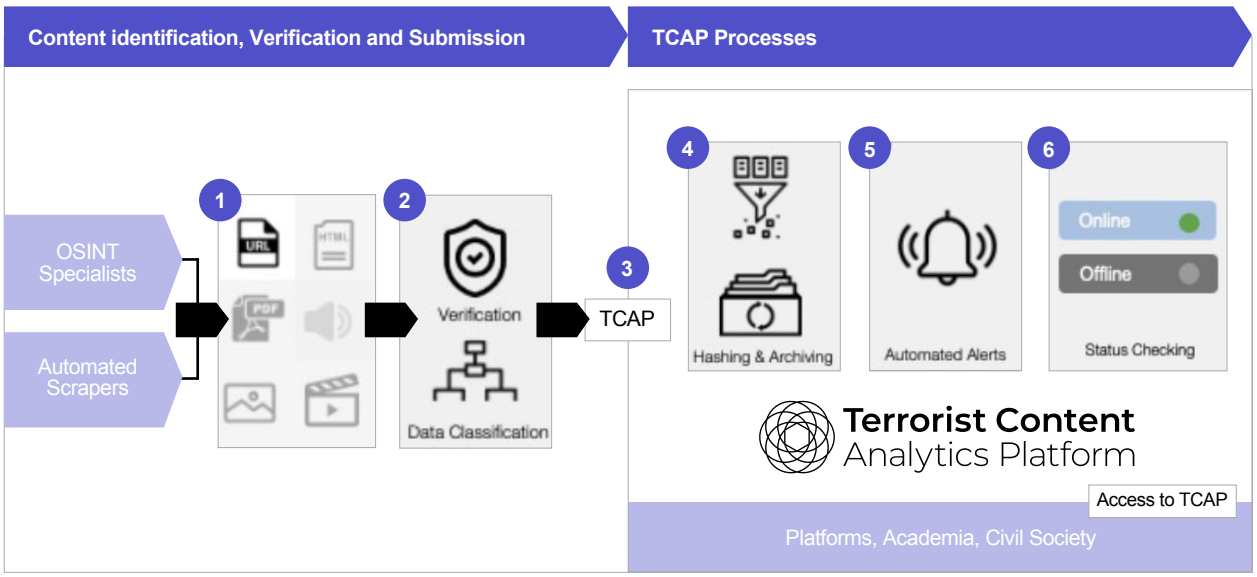

Figure 21: The TCAP's process of identifying, collecting, verifying, archiving, and alerting terrorist material

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 scrapers[23] 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 summaries the data attributes captured for each content item:

[23] Automated scrapers

Data attribute	Description
Associated terrorist group	The terrorist group 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
Channel URL	A link to the channel where the content was identified
PII warning	Content containing personally identifiable information
Extreme content warning	Content containing violent graphics
Data & 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

Figure 22: Data attributes stored for each content item on the TCAP

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.[24] For more detail on our hashing of URLs and hash-sharing, please see this blogpost here. For more details on GIFCT's hash-sharing database, please see here.

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 available for access; however, 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 that was collected on tech platforms registered for TCAP alerts. This content will be notified to the platform in question via an automated email alert. Emails 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.

[24] GIFCT hashing sharing consortium

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.

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. To date, we have had no content disputed by tech companies.

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[25] across multiple platforms.[26] The TCAP utilises the Selenium framework 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 key data from the web or mobile site:

- Channel MetaData (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 chat room 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.

[25] 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
[26] Please find more on this in our content classification and verification policy section in this report.

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 undue norm setting of speech, with its inherent risks to human rights, 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.	• 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. • To date, we have only included official content, using our Content Classification and Verification Policy.
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.[27]	• We are developing the TCAP through "transparency-by-design", ensuring we are transparent in all phases of the process. • All platform policies are available on request. • We launched a public consultation process, the findings of which can be found in our report. • 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.[28] • Anyone with TCAP access can share their views on classification. They can contest whether a generated alert concerns terrorist content.

[27] 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.
[28] More information on TCAP office hours

Principle	Justification	Implementation
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.[29]	• 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. • When we start sharing hashes with tech companies, we will ensure to build a "look-up" function, that allows tech companies to un-hash the material and examine the actual content • We implement a rigorous verification process using in-house terrorism experts to verify the content as terrorist in nature - for more information see above in our policy section. • 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. • 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. • 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.

[29] One database to rule them all (VoxPol 2020)

Principle	Justification	Implementation
Security 	Given that TCAP archives content and its location, it is imperative that we build TCAP securely, so that terrorist/violent extremist (T/VE) entities don't gain access to the platform. We also need to ensure that T/VE entities do not become aware of our operations to the extent that it inhibits our mission or risks our operational security (OpSec).	• We follow strict OpSec protocols when conducting our open-source intelligence monitoring. • 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. • Our development team executes frequent penetration testing so that the TCAP as a platform can resist any attack.
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 T/VE entities. It is therefore critical to enforce the right to privacy.	• Alerts to tech platforms come with a tag to show whether the content contains personal identifiable information (PII). • A record of captured PII will be kept to preserve its potential function as digital evidence in war crimes trials or the prosecution of other human rights abuses.[30] Using Amazon Web Services infrastructure, all data will be kept in a highly secure, controlled environment. • 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 (see below).

[30] More on this can be read in a Human Rights Watch in a September 2020 on removal of terrorist content and war crime evidence

Principle	Justification	Implementation
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"[31] and retain the right to free expression. We are aware that we, as a non-governmental organisation, should not set global norms for online speech.	• We base our Inclusion Policy on provisions of law, ensuring that we do not contribute to undue norm speech-setting of online content. • We alert tech companies with the URLs containing the terrorist content so they can review the content and avoid a dependence on automated removals compromising freedom of speech. • Civil society participation ensures that relevant concerns can be raised and addressed. We ensure this participation through regular feedback sessions in office hours and our consultation report (see above). • All alerts are made on an advisory basis.
Tech Platform Autonomy 	To avoid content 'cartelisation',[32] the TCAP alerts companies on an advisory basis only.	• 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. • This is supported through our Knowledge Sharing Platform[33] and Online Regulation Series[34] that makes tech platforms aware of their duties in certain jurisdictions when notified of terrorist content on their platform.

Figure 23: Core principles of the TCAP

[31] 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. Evelyn Douek. The Rise of Content Cartels (Colombia University, 2020).
[32] Ibid
[33] Knowledge Sharing Platform
[34] Online Regulation Series

4.2.2 Content Classification and Verification Policy

To include only official material from the above terrorist groups in scope, we have created a Content Classification and Verification Policy[35] 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. In order to assess the content, our team conducts an intelligence assessment that evaluates attributes of the content associated with a high level of probability that the material was produced by a designated terrorist organisation in scope of the TCAP.

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 also solicited feedback on the fundamental principles.

In August 2020, we published a report[36] 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.

Key findings and observations included:

- 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 with regard to 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.

[35] Our full Content Classification and Verification Policy can be found here. As this is a public report, we cannot go into the operational detail for security reasons. Please reach out to us if you would like the report or join our monthly Office Hours for a more detailed explanation.
[36] Consultation Report

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 available (for a select number of stakeholders) this legal analysis. Whilst the full document is legally privileged, you can request the condensed, top-level version of the legal review here.

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.

4.2.5 Threat to Life Protocol

Beyond the principles and policies discussed, we have also developed a Threat to Life Protocol that guides our Open Source Intelligence team should a potential threat to life situation be revealed as part of the content discovery process. The threat to life protocol is integrated into our Crisis Protocol.[37]

A threat to life is determined as the deliberate intention to cause:

- A real and immediate threat to life (real and immediate defined as a risk that is reasonably assessed to be real, and the potential assailant has the intention and capability to carry out the threat)
- Threat to cause serious harm
- Threat of injury
- Threat of serious sexual assault and/or rape

How we identify a threat to life:

We assess as much information available, including:

- The potential attacker and his/her capability
- The intent of the attacker
- The victim(s)
- The location
- The timescale
- The information we are missing

[37] Crisis Protocol Policy

Based on the above information, we then assess the information and give it a score. These can be:

- **Low:** No "real and immediate" threat, indicating that the perpetrator has no intention or capability to follow through with a threat.
- **Medium:** The alleged threat is likely to occur if the perpetrator has the right resources; the threat is therefore conditional. The intention will need to be assessed as "highly certain" to justify raising the threat level.
- **High:** The threat is credible, immediate, and specific. Suspect, victim, and location of the threat is identifiable. However, there can be a high threat to life but without all this information – leading to an unspecific high threat to life.

When the TCAP team encounters a **High Threat to Life**, Tech Against Terrorism will alert United Kingdom police and relevant authorities if the location is known.

4.3 RECOGNITION

TCAP has also received significant recognition from stakeholders.

- The TCAP was commended by Prime Minister of Canada Justin Trudeau at the Christchurch Call to Action 2021 Summit.
- The TCAP was mentioned in a report by Human Rights Watch, which states that the considerations taken by Tech Against Terrorism in building the TCAP will inform Human Rights Watch's mechanism for archiving removed material for evidence of war crimes.
- The TCAP was also mentioned in the Digital Lockers Human Rights Report which discusses 'Voluntary Partnership Models' in archiving media evidence of 'Atrocity Crimes'. The report was published by UC Berkeley, and is accessible here.
- The TCAP's contribution to countering Islamic State's propaganda was highlighted by the United Nations Counter-Terrorism Directorate (UNCTED) in their Twelfth report on the Threat posed by ISIL, Daesh to international peace and security.

4.4 WHAT'S NEXT?

- **Expanding Inclusion Policy:** We will seek to update our Inclusion Policy to include more designated terrorist groups in line with evolving and existing designation. In this process we will take into account the particular threat that a group poses as well as the amount of online content a given group disseminates. As we expand, designation will be the cornerstone of inclusion. 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.

● **Including supporter-generated material:** Until now, the TCAP has only flagged to tech companies official material from designated terrorist groups in scope. When we expand our policy, we will also include supporter-generated material that supports, incites, or glorifies the groups in scope of the TCAP. For Islamist groups, this includes material produced by supporter-run media outlets. For far-right groups and entities, this can include material that explicitly supports or glorifies terrorist groups or entities, as well as videogame versions of the Christchurch attack. This will ensure our submissions and alerts for groups and actors that adhere to a far-right terrorist ideology will become more equal to the groups that fall under the Islamist terrorist umbrella. We hope this will ensure that the TCAP becomes as successful in countering far-right terrorist content as with Islamist terrorist material.

● **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 this material to tech companies. If not, we will assess whether the material violates any other laws and notify this to 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.

● **Database for academics:** 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, as far-right terrorist groups oftentimes 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.