Tech Against Terrorism's response to Ofcom's Consultation: Protecting people from illegal harms online | February 2024

This document contains Tech Against Terrorism's response to the Ofcom Consultation: Protecting people from illegal harms online. Tech Against Terrorism and Ofcom met on 29 November 2023 to discuss the consultation. Ofcom requested Tech Against Terrorism's expertise on the terrorism chapter of the register of risk (Volume 2), the draft codes of practice (Volume 4), and the illegal contents judgement guidance (Volume 5). Tech Against Terrorism's prioritisation of concerns and a summary of our recommendations is included below.

Prioritisations of concerns

Type	Explanation	Risk level	Recommendation
Generative artificial intelligence (AI)	Generative AI services present multiple use cases for terrorist exploitation which pose significant risks in the immediate and long term. This includes use of generative AI for: • propaganda generation, recruitment, and radicalisation; • content dissemination and amplification; • circumvention of moderation; • operational planning and upskilling. Tech Against Terrorism has identified users exploiting generative AI tools to bolster the creation and dissemination of AI-generated propaganda in support of both violent Islamist and neo-Nazi ideologies. Whilst engagement with generative AI is likely to be in the experimental phase, these experiments indicate an emerging threat which should be reflected in Ofcom's risk registry.	High	Ofcom should consider incorporating the terrorist exploitation of generative AI into the terrorism register of risk and also consider the imposition of mitigation measures through the codes of practice.
Terrorist-operated websites	Terrorist-operated websites, or TOWs, pose a significant risk because they act as stable hosts of terrorist content and are easily discoverable on the surface web. Search engines can greatly increase the reach of terrorist	High	Incorporate terrorism content within the register of risk for search services, with specific reference to their role in increasing the discoverability of and access to terrorist-operated websites.

	content by directing users to TOWs, thereby posing a risk of harm to the public. TOWs and other pages hosting terrorist content are often indexed on search engine results, making them easily accessible by vulnerable individuals.		For mitigation, URL detection should be prioritised for the purpose of disrupting outlinks that lead users to TOWs.
Messaging platforms	Leadership messages and official statements from terrorist entities are frequently posted through one-way messaging channels which do not allow users to directly respond to messages or post in the channel, and which have a very large or unlimited audience. Channels like this are highly beneficial to terrorist entities as they allow complete control over the content in a channel and also constitute an uncomplicated means of providing clear direction to other platforms or channels. It is likely that terrorist entities are drawn to platforms that enable this type of one-way communication. Terrorist entities will almost certainly continue to exploit these channels to redirect users elsewhere throughout the short to long-term future, both to circumvent moderation and ensure content stability.	High	URL detection in the context of terrorism content should prioritise beacon and aggregator platforms, which include these messaging services.
File-sharing services and crisis situations	Tech Against Terrorism has identified an increased reliance on file sharing platforms in the immediate aftermath of crisis events. Outlinks from larger beacon platforms to smaller file-sharing sites to disseminate the Bratislava attack perpetrator's manifesto reflected a new typology of behaviour amongst a subset of far-right terrorist online networks seeking to exploit file-sharing platforms to host manifestos and livestreams.	Medium	Ofcom should highlight the specific exploitation of file-sharing platforms in crisis scenarios within its register of risk.

Recreations of terrorist attacks on gaming services	Gaming and gaming-adjacent platforms are increasingly being used by terrorist supporters to recreate gamified versions of terrorist attacks. In 2022, Tech Against Terrorism identified at least 40 computer-generated versions of offline attacks on one platform, including but not limited to Christchurch (2019), Buffalo (2022), Oslo and Utøya (2011), Bataclan, Paris (2015) and the Nairobi Westgate Mall attack (2013).[1] The computer-generated versions of offline attacks all allowed users to take part in a simulated game, most of which allowed users to play as the original attack perpetrators.	Medium	Ofcom should highlight this exploitation of gaming platforms within its register of risk.
Archiving and paste platforms	In 2023, Tech Against Terrorism found archiving platforms to be the second most heavily exploited platform type amongst those alerted via the Terrorist Content Analytics Platform (TCAP).[2] Of all platform types, only archiving platforms were exploited to a significant extent across the ideological spectrum, by both Islamist and far-right terrorist actors. Similarly, Tech Against Terrorism found a high concentration of terrorist content on a small number of paste sites. However, archiving and paste platforms are missing from Ofcom's identification of services capable of being used to commit or facilitate terrorist offences.	Medium	When identifying types of services that can be used to commit or facilitate offences related to terrorism, Ofcom should also include archiving platforms and paste platforms.
Platform features at risk of exploitation: minimal sign-up	Platforms that require minimal sign-up requirements are likely to attract terrorist actors because their personal security online is increased when their identities are less traceable.	Moderate	Ofcom could also include lack of registration and on-platform search as high-risk features.

[1] Tech Against Terrorism, State of Play (2022).
[2] Tech Against Terrorism, TCAP Insights: Patterns of Online Terrorist Exploitation (2023).

requirements, on-platform search functions, messaging channels	Platforms that are easy to navigate are typically preferred by all users, including terrorist actors. The availability of on-platform search functions makes a platform simple to navigate and content, groups, users, and other information easy to find. With on-platform search features, terrorist actors can more easily find terrorist content utilising keyword and phrase searches.		

Further summary of consultation feedback

Volume 4: The draft codes of practice:

- **Proposed mitigations for proscribed organisations.** Ofcom should consider reducing the burden on service providers by widening the "reasonable grounds" threshold for inferring an account's affiliation to a proscribed organisation.
- **URL detection for terrorism content.** This should be targeted towards beacons and aggregators and should account for the use of link shorteners to evade moderation. Tech platforms could utilise Tech Against Terrorism's TOW Database and Beacon Channels tracker for URL detection.
- **Keyword detection for terrorism content**. Tech platforms could utilise Tech Against Terrorism's Terminology Database for keyword detection. Recommendations for keyword detection should incorporate safeguards for human rights, including a balance with human review, and consideration of other contextual identifiers. Ofcom could also consider recommending image and symbol detection for terrorism content.
- **Hash matching for terrorism content**. There is evidence that hash-matching offers an effective solution for moderating terrorist content in many contexts, particularly where the content is stable and unlikely to be changed as it proliferates (e.g. a still image or document). This is especially so when used in tandem with access to a database of known terrorist material, such as that currently being built as part of the Terrorist Content Analytics Platform (TCAP). Video content poses a much more complex challenge for hash detection, especially where footage is edited or excerpted. Furthermore, as online actors seeking to disseminate terrorist propaganda begin adopting advanced generative AI tools to produce content, the opportunities for hash matching against known material are likely to correspondingly decrease.

Volume 5: Illegal content judgements guidance:

- While the guidance is accessible, it still places a significant burden of responsibility on services and their moderators to make appropriate judgements. This may overlook the

limited capacity of medium and smaller-sized service providers to hire and train moderators to make those informed decisions.

- Ofcom should consider providing additional practical resources to aide and complement these guidelines and simplify the priority content for platforms to target.
- Tech Against Terrorism's Knowledge Sharing Platform (KSP) contains a comprehensive and up-to-date resource that includes an image compendium and terminology database that covers 14 organisations proscribed by the UK, totalling over 160 images and 520 phrases. This would support tech companies in reaching illegal content judgements which relate to membership or support for proscribed organisations.