tech against terrorism

Position paper
Content personalisation and the online dissemination of terrorist and violent extremist content
February 2021

Summary

We welcome the increased focus amongst policymakers on the role played by content personalisation and other algorithmic recommendation systems on online platforms.[1] Such scrutiny is warranted. Terrorist groups exploit platforms that make use of recommendation algorithms, and there are examples of individuals coming into contact with terrorist and violent extremist content via platforms using content personalisation. However, we are concerned that the current debate is, on a policy level, based on an incomplete understanding of terrorist use of the internet, and that a focus on content personalisation is a distraction from more important steps that should be taken to tackle terrorist use of the internet.

Recommendations

In our view, governments and policymakers should focus on priority challenges with regards to terrorist use of the internet. We recommend that policymakers:

- Improve designation of terrorist groups and provide legal clarity for tech companies with regards to what content should be actioned – this is the most important step governments can take to counter terrorist use of the internet[2]
- Focus on improving legal instruments. If policymakers wish to see content removed from tech platforms, they should criminalise such content in law. This content should be clearly defined, and any such legislation should have appropriate freedom of expression and human rights safeguards
- Provide evidence justifying action on content personalisation within the framework of online counterterrorism measures
- Avoid introducing any measures that would see removal of – or restricted access to – legal content as part of measures taken against content personalisation and algorithmic amplification
- Improve strategic leadership with regards to countering terrorist use of the internet – counterterrorism efforts should be led by democratically accountable institutions and not by private tech companies
- Improve support mechanisms for smaller tech platforms to help upscale counterterrorism responses across the tech industry
- Develop protocols that facilitate action on terrorist-operated websites in a manner consistent with international norms around freedom of expression

[1] See, for example, the EU's Counter-Terrorism Coordinator's paper "The role of algorithmic amplification in promoting violent and extremist content and its dissemination on platforms and social media" (December 2020): https://data.consilium.europa.eu/doc/document/ST-12735-2020-INIT/en/pdf

[2] In our work, we find that all (even occasionally hostile) platforms comply where there is legal certainty around a group's terrorist status

State of play: terrorist use of the internet and algorithms' role in radicalisation

It is by now well-documented that terrorists use an eco-system of predominantly smaller platforms to communicate and disseminate propaganda, the majority of which make no use of recommendation algorithms or other content amplification.[3] Terrorists instead rely primarily on simple web-based tools such as pasting, archiving, and file-mirroring sites, which are used to "swarm" the web with content to ensure longevity.[4] Increasingly, terrorists rely on terrorist-operated websites[5] and the use of alternative platforms and infrastructure providers to host, aggregate and disseminate propaganda, as well decentralised and encrypted platforms.[6] These platforms are used largely because terrorist content is more likely to be quickly removed on larger platforms, and demonstrate that terrorist use of the internet is a complex and multi-faceted threat that requires a systems-based approach across platforms.

Larger platforms that rely on content personalisation, like YouTube and Facebook, already remove almost all[7] of the terrorist and violent extremist content located on their platforms proactively. Whilst these companies still have work to do in terms of improving their response to terrorist content, they no longer constitute key propaganda dissemination platforms for terrorist groups.

There is very limited evidence that content personalisation leads to terrorism. There are cases in which an individual's radicalisation process has partly consisted of visiting platforms using content personalisation, but there is no conclusive evidence to suggest that algorithmic recommendation systems lead to terrorism. Even one of the studies most critical of YouTube (and the platform's role in radicalisation) does not highlight the platform's recommendation algorithm as the most significant reason behind the platform being exploited.[8] We encourage policymakers to look beyond the disproven 'conveyor-belt' theory of radicalisation, which suggests that people are radicalised by simply viewing terrorist content online. It is well-established that radicalisation is much more complex, and most academic research argues that individuals are largely radicalised offline and subsequently seek content online proactively.[9] In our view, it is crucial that any policy or regulation affecting content personalisation is underpinned by evidence and extensive research.

---

[3] Tech Against Terrorism has identified more than 350 platforms exploited by terrorists, with more than two thirds being either smaller or "micro" platforms. See: https://www.techagainstterrorism.org/2019/04/29/analysis-isis-use-of-smaller-platforms-and-the-dweb-to-share-terrorist-content-april-2019/. Furthermore, Europol's Internet Referral Unit has also reported on this trend: https://www.europol.europa.eu/publications-documents/eu-iru-transparency-report-2018

[4] Ali Fisher, "Swarmcast: How Jihadist Networks Maintain a Persistent Online Presence", Perspectives on Terrorism, Vol 9. No. 3 (2015): http://www.terrorismanalysts.com/pt/index.php/pot/article/view/426/html; "Netwar in Cyberia" decoding the Media Mujahedeen", USC Center on Public Diplomacy (2018): https://uscpublicdiplomacy.org/sites/uscpublicdiplomacy.org/files/Netwar%20in%20Cyberia%20Web%20Ready_with%20disclosure%20page%2011.08.18.pdf.

[5] Tech Against Terrorism is currently aware of more than 50 such sites run by designated terrorist groups and violent extremist groups.

[6] Often collectively called 'alt-tech', these are companies that have a professed ideological commitment to keeping terrorist and/or violent extremist content online.

[7] 93.9% (YouTube) and 99.7% (Facebook) respectively, as per the transparency reports produced by YouTube and Facebook for Q3: https://transparencyreport.google.com/youtube-policy/removals?hl=en; https://transparency.facebook.com/community-standards-enforcement#dangerous-organizations.

[8] Becca Lewis, "Alternative Influence: Broadcasting the Reactionay Right on YouTube" (2018): https://datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf

9 See for example: Whittaker & Herath "Understanding the Online and Offline Dynamics of Terrorist Pathways" (2020) https://gnet-research.org/2020/07/13/understanding-the-online-and-offline-dynamics-of-terrorist-pathways/; Gill et al, "Terrorist Use of the Internet by the Numbers" (2017) https://onlinelibrary.wiley.com/doi/full/10.1111/1745-9133.12249; Gaudette, Scrivens & Vinkatesh, "The Role of the Internet in Facilitating Violent Extremism: Insights from Former Right-Wing Extremists" (2020) https://www.tandfonline.com/doi/abs/10.1080/09546553.2020.1784147; Gemmerli, "The Fight Against Online

Key considerations for policymakers

Rule of law

It is vital that the rule of law, human rights, and freedom of expression are respected in tackling terrorist use of the internet. In Tech Against Terrorism's Online Regulation Series,[10] we noted a global trend (including in democratic countries) where governments are introducing legislation aiming to counter terrorism online that risks undermining the rule of law. One example of this is governments compelling companies to remove content that is legal but "harmful". This has significant negative implications for freedom of expression and raises serious concerns around extra-legal norm-setting. It may also be difficult to operationalise. In the United Kingdom's Online Harms scheme[11] the UK Government has signalled that companies would need to take action on content that has significant "psychological and emotional impact". As experts have pointed out, such categories of harm are tied to subjective reactions of users and will be extremely difficult to operationalise consistently.[12] In some countries, companies are required by law to assess legality of content,[13] something which effectively makes tech companies, which unlike governments are democratically unaccountable, the *de facto* arbiters of speech legality online. It is vital that democratic countries lead by example and that core democratic principles, such as the rule of law, are respected. Norm-setting should occur via consensus-driven mechanisms led by democratically accountable institutions, not by private companies. As such, we encourage policymakers exploring introducing measures against content personalisation to ensure that such measures do not target legal speech.

Securitisation of the internet and freedom of expression

We note that several separate challenges are often cited in relation to perceived risks with content personalisation. These include the amplification of hate speech, disinformation, digital literacy, and the role played by the "attention economy". These are significant challenges; however, we caution against securitising them under the rubric of counterterrorism. Whilst there is some evidence of a connection between hate speech, disinformation, and terrorism, it is evident that applying counterterrorism measures to other areas carries serious risks to freedom of expression and human rights.[14] We remind policymakers that several of the above-mentioned content categories constitute legal speech in several jurisdictions. If action on legal speech from tech companies is desired, governments should focus on anchoring such speech in criminal code with appropriate freedom of speech safeguards.

Smaller platforms

Radicalisation Starts Offline", (2015) https://www.jstor.org/stable/resrep13160?seq=1#metadata_info_tab_contents; Babuta "Online Radicalisation: The Need for an Offline Response" (2017) https://rusi.org/commentary/online-radicalisation-need-offline-response

[10] For more information, see: https://www.techagainstterrorism.org/2020/12/22/the-online-regulation-series-summary/

[11] See the UK Online Harms White Paper here: https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper. To see our analysis of the UK's regulatory response to online harms, see: https://www.techagainstterrorism.org/2020/10/22/online-regulation-series-the-united-kingdom/

[12] Graham Smith, "The Online Harms Edifice Takes Shape" (December 2020): https://www.cyberleagle.com/2020/12/the-online-harms-edifice-takes-shape.html

[13] Germany is one such example. See: https://www.techagainstterrorism.org/2020/10/21/the-online-regulation-series-germany/.

[14] https://www.techagainstterrorism.org/2020/10/05/the-online-regulation-series-singapore/

Terrorist use of smaller platforms constitute the most important strategic threat with regards to terrorist use of the internet. Almost all of the companies we engage with want to contribute to countering terrorist activity, but many smaller platforms are managed by just one person and require support in identifying and mitigating terrorist use of their services. Tech Against Terrorism has supported the tech sector in this endeavour since 2017 via its Mentorship Programme.[15] In addition, we are currently building the Terrorist Content Analytics Platform (TCAP), which will be the world's largest data set of verified terrorist content. The TCAP is already supporting smaller tech companies in detecting terrorist use of their platforms through an automated alerting function. The platform will also provide training data for algorithmic solutions to identify terrorist content online.[16]

Going forward: focussing on priority areas to ensure positive impact

Tech companies should play a role in combating terrorist use of the internet, and several platforms need to improve their response. Platforms that use content personalisation should also take steps to ensure that such systems are not exploited by nefarious actors, including terrorist groups, or conspiracy movements with potential for offline violence, for example. Importantly, such efforts are not confined to content removal but can also include other forms of disengagement methods. It is therefore welcome that the discussion around content personalisation provides an opportunity to explore ways in which platforms can work to reduce the spread of content as an alternative to removal.

Policymaking should focus on measures that will have tangible positive impact and contribute to reduce the threat of terrorism. At the moment there is not enough evidence to suggest that content personalisation is a key threat or the most important priority with regards to tackling terrorist use of the internet. Instead of focussing on a technical feature used by a few platforms, emphasis should be on supporting smaller tech platforms in tackling terrorist use of their services and identifying systems-based solutions that address the eco-system aspect of the threat in order to have tangible positive impact.

We also encourage policymakers to distinguish more clearly between content types and clarify which would be in scope for any potential regulatory mechanism or industry code. It is vital that any regulation only covers speech that is criminal and does not lead to takedown of legal content. Further, we should not expect tech platforms to reduce the spread of content that does not violate the law or their content policies.

In our view, governments can help prevent terrorism by taking measures that target the (overwhelmingly offline) root causes of terrorism and by creating definitional clarity around terrorist groups, for example via improved designation. Such measures would improve strategic leadership in the online counterterrorism space, and avoid policies that risk putting the cart before the horse when tackling terrorist use of the internet.

[15] As part of our Mentorship Programme, Tech Against Terrorism supports platforms at no cost in introducing effective counterterrorism policies and enforcement mechanisms. See: https://www.techagainstterrorism.org/membership/tech-against-terrorism-mentorship/

[16] See the dedicated TCAP website: https://www.terrorismanalytics.org/