SD Powering solutions to extremism and polarisation

Briefing Note The Management of Terrorist Content: How Al Qaeda Texts Continue to Evade Facebook and YouTube Detection

By Moustafa Ayad

July 23, 2019

Introduction

In June 2019, <u>ISD researchers identified Arabic language content on Facebook and YouTube</u> promoting influential extremists from AI Qaeda (AQ), Islamic State of Iraq and the Levant (ISIL) and affiliates for both organisations. This research highlighted the gaps in content moderation and the challenges for the companies enforcing terms of service across languages and markets. In a new briefing paper launched this week, ISD highlights additional AQ content that exposes the continued persistence of gaps in the companies' ability to identify and remove content, which they have committed to doing. Below we outline a selection of key findings, including the content that was identified, how it was identified, and our analysis about the tactics being used to evade the companies' content detection systems.

Key Findings:

• ISD researchers have identified Facebook and YouTube networks of users, channels and pages sharing al Qaeda and Islamic State of Iraq and the Levant (ISIL) "legacy terrorist content" across both platforms. This includes seminal texts that underpin the strategies and objectives of Salafi-Jihadist groups, including *The Management of Savagery: The Most Critical State Through which the Islamic Nation Shall Pass* and *The Global Islamic Resistance Call*. While a good deal of research has been done to understand these works in terms of their centrality to Salafi-Jihadist strategy and practice, including by Google's global lead on counterterrorism, Will McCants, there has been limited research into the prevalence of legacy terrorist content such as these seminal works on both Facebook and YouTube.¹ Much of both companies' focus over the past four years has been on removing and limiting the impact of ISIL-affiliated content on their platforms at the behest of governments around the world. Our research however identifies a gap in response to Arabic legacy al Qaeda content on both platforms. While much of the Islamist terrorist activity online has migrated to alternative platforms such as Telegram, ISD researchers have uncovered a significant amount of Arabic language terrorist content still available on both Facebook and YouTube simply by searching for their titles in Arabic.



¹ McCants is credited with being the first to translate The Management of Savagery from Arabic to English. He has previously highlighted the book's centrality to ISIL-governance in ransacked cities such as Mosul, and Raqqa, other Salafi-jihadist leaders, and plots in countries such as Saudi Arabia and Lebanon.

- Simple Arabic language searches surfaced five public Facebook pages, with more than 4,000 followers, dedicated to sharing the ideas contained in al Qaeda linked Global Islamic Resistance Call and The Management of Savagery. Researchers also found four dedicated YouTube profiles sharing a 10-part audiobook version of The Management of Savagery, with more than 25,000 views, and linked to large subscriber base channels. Researchers surfaced direct linkages to users, and online networks, propagating the ideas of each book, and in some instances providing users with easier modes of consuming the concepts in both texts. The largest Facebook page dedicated to The Management of Savagery seems to have linkages to Francophone audiences ("Sauvagerie de Gestion"). All of the content found has been on the platform for more than two years.
- It is evident that both YouTube and Facebook manual review processes, and automated detection systems for terrorist content, have clear gaps. In June 2019, ISD flagged a number of users and pages to the companies that had been circulating content related to AQ and ISIL. While they were largely removed, problematic content linked to these pages and accounts were not fully addressed. For example, YouTube banned a single video in a terrorist content playlist, but allowed other terrorist-affiliated videos in the same playlist to remain live. This suggests a clear operational issue for technology companies in both their manual and automated detection processes for non-English speaking markets.



- Users are taking advantage of these gaps and deploying certain tactics to evade both manual and automated detection by the companies. Some of the tactics observed include:
 - (1) Labelling content as 'educational'. Facebook users identified terrorist content or pages as "educational," tagging terrorist material as "books," 'public figures," or "companies," in order to evade detection. Educational content that might otherwise violate Community Guidelines is generally spared take downs by both YouTube and Facebook. Facebook's Community Guidelines outline this exception under the company's "Objectionable Content" section. YouTube similarly outlines this exception under its "Violent Criminal Organizations" policy.
 - (2) Sharing links rather than content. Facebook representatives have previously noted during GIFCT events that they do not track and monitor links shared on the platform and it appears that supporters of Salafi-Jihadism are taking advantage of this. ISD researchers found eight different direct

- links to the PDF version of *The Management of Savagery* on Facebook. The links were shared 107 times across the platform. One public page doubling as an "electronic library" with more than 137,000 followers posted a link to the PDF version of the terrorist strategy that has been shared 61 times on Facebook. The same link has been liked more than 160 times.
- (3) Linking to mainstream content. This approach, observed on YouTube, appears designed to "game" the platform search algorithms by linking mainstream content to terrorist content, and hence driving traffic back to problematic content. A number of the playlists, which were posting sections on The Management of Savagery, were also posting a smattering of mainstream content, including videos of popular Egyptian singers alongside mobile-phone shot videos of families.
- Facebook's "related pages" function created linkages to other public pages sharing extremist content. The "related pages" option functions as an algorithmic bridge between networks of terrorist content propagators. Researchers were linked from pages that were sharing explicitly terrorist content to other "softer" pages supportive of Salafi-jihadism, Salafi-jihadist principles, and Salafi-jihadist ideologues. Two clicks away from *The Management of Savagery* through related pages is a "Company" page title that reads "Jihad for the Sake of God." This page claims more than 42,000 followers, and is dedicated to countering "misguided sheiks."



• YouTube and Facebook should update their hash databases and direct their manual and automated Arabic language detection efforts to ensure capture of 'legacy terrorist content'. This should include the Arabic names of al Qaeda and ISIL ideologues, and the titles of terrorist strategies, pamphlets, books, and films. While the platforms have been diligent at removing some English-language terrorist content, the same approach has yet to be reproduced in Arabic. Until that basic level of detection is instituted, users will be able to evade detection on both sites. Moreover, as part of their efforts in the Global Internet Forum for Countering Terrorism (GIFCT), the companies should commit to greater transparency in sharing some of the on-going challenges that they are facing in removing terrorist-related content globally, across languages and global regions.

Media enquiries

For further information, please contact: Media@isdglobal.org +44 (0)20 749 3933