# Evaluating Platform Accountability: Terrorist Content on YouTube

***Dhiraj Murthy, School of Journalism and Media, Moody College of Communication and Department of Sociology, University of Texas at Austin, Austin, TX 78712-0109, USA.***

## ABSTRACT

YouTube has traditionally been singled out as particularly influential in the spreading of ISIS content. However, the platform along with Facebook, Twitter and Microsoft jointly created the *Global Internet Forum to Counter Terrorism* in 2017 as one mode to be more accountable and take measures towards combating extremist content online. Though extreme content on YouTube has been found to have decreased substantially due to this and other efforts (human and machine-based), it is valuable to historically review what role YouTube previously had in order to better understand the evolution of contemporary moves towards platform accountability in terms of extremist video content sharing. Therefore, this study explores what role YouTube's recommender algorithm had in directing users to ISIS-related content prior to large-scale pressure by citizens and governments to more aggressively moderate extremist content. To investigate this, a YouTube video network from 2016 consisting of 15,021 videos (nodes) and 190,087 recommendations between them (edges) was studied. Using Qualitative Comparative Analysis (QCA), this study evaluates 11 video attributes (such as genre, language, and radical keywords) and identifies sets of attributes that were found to potentially be involved in the outcomes of YouTube recommending extreme content. This historical review of YouTube at a unique point in platform accountability ultimately raises questions of how platforms might be able to be more pro-active rather than reactive regarding filtering and moderating extremist content.

**Keywords**: YouTube; algorithms; social media; social network analysis; terrorism; platform accountability


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

# Evaluating Platform Accountability: Terrorist Content on YouTube

## Introduction

Since its infancy, Islamic State (ISIS[1]) placed a strong emphasis on communication via video content on YouTube. ISIS' YouTube footprint became a campaign issue during the 2016 US presidential campaign. During 2015-2016, Hillary Clinton repeatedly argued that technology companies need to be held accountable for their role in disseminating terror-related content and should actively 'disrupt' ISIS' efforts on social media (Kosoff, 2016; Sanger, 2015). Former U.S. President Barack Obama suggested that technology companies need to work to silence ISIS online and make "alternative accounts" to ISIS' narrative available. Part of this call to action came from journalists documenting the importance of YouTube in the radicalization processes (Gates & Podder, 2015). For example, a Western ISIS defector stated, "A lot of people when they come [to ISIS], they have a lot of enthusiasm about *what they've seen online or what they've seen on YouTube*" (Freytas-Tamura, 2015; my emphasis). YouTube may have been singled out as "it remain[ed] easy to find content on YouTube that violate[d] the company's community guidelines against hate speech and/or explicitly promote[d] terrorism" (Neumann, 2013, p. 442).

In 2017, YouTube and other technology companies launched a major offensive through the *Global Internet Forum to Counter Terrorism* (GIFCT), whose stated purpose 'is to prevent terrorists and violent extremists from exploiting digital platforms' (Global Internet Forum to Counter Terrorism, 2019). This was seen as a collective mechanism to increase platform accountability by combating extreme content online. Through human moderation and machine-based detection, YouTube sought to be more accountable of the content it was hosting. Gauging the reality of ISIS content

[1] 'ISIS' is used synonymously with IS/DAESH/ISIL/Islamic State


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

on YouTube prior to aggressive moderation in 2017 is not only useful to understanding what motivates technology companies to be more accountable, but can provide insights for future platform accountability.

We found 11 official ISIS videos produced by their media wing, *Al Hayat* accessible on YouTube. Given these videos were identified by journalists as important to ISIS' recruitment efforts, we evaluated what type of videos were recommending them. Though we cannot evaluate what potential ISIS recruits are searching for, exploring the paths of recommendation is one avenue for evaluating how accessible these videos actually were prior to algorithmic upgrades as part of the accountability strategies implemented by YouTube after the formation of GIFCT in 2017. Moreover, our study is novel and its use of a pre-GIFCT YouTube data set, but with a retrospective post-GIFCT outlook. Using these 11 official ISIS videos as seeds, we derived a network consisting of 15,021 video nodes and 190,087 recommended edges. We use Social Network Analysis (SNA), a method of mapping networks based on graph theory, to visualize the video network as well as identify what types of videos recommend ISIS content. Videos in the recommender network do not have an equal likelihood of recommending an ISIS video and this study seeks to identify broad patterns of recommendation. To do this, we used YouTube's video genre labels, view count, comment count, and other metadata returned by the YouTube API. Lastly, we use Qualitative Comparative Analysis (QCA) (Rihoux & Ragin, 2008) to study features likely influencing YouTube's algorithmic decision making.

This study has three main conclusions: (1) Though rare, YouTube's recommender algorithm does recommend radical content from non-radical content, signaling that potential ISIS recruits had to


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

actively search for ISIS content in 2016 rather than stumbling upon it serendipitously; (2) That these non-ISIS videos recommending ISIS videos appear to be doing so partially due to similarities in title or description with seed videos and (3) ISIS videos do recommend other ISIS videos. Having any terrorist content on public platforms is an issue and the high-profile reaction by YouTube and others in 2017 with GIFCT reflects the fact that the broader public debate reached a particular tipping point that YouTube had to be accountable for any terrorist content that it was hosting, even if the volume of extremist content posted on YouTube was low.

## Critical Algorithm Studies

Algorithms are not neutral. Without us necessarily understanding the computations involved behind them or having input on how they are working, algorithms are increasingly shaping our lives (Kitchin, 2017). Algorithms do not necessarily exist in isolation, but interact with other algorithms, whether they are on one platform/site or are interacting with external algorithms (Slavin, 2011).

Some scholarship has placed an emphasis on understanding the role algorithms play in material reality. Fuller (2008, p. 16) as argued that "algorithms bear a crucial, if problematic, relationship to material reality", citing the Turing machine as an example. Feuz et al. (2011) add that algorithms “are of increased importance because of the role they play as parts of contemporary abstract infrastructure […] because, like water mains and roads [… algorithms also depend on the principle] that things in the world may be seen as computable data and information”. Others have argued that algorithms are ‘analytical procedures’ that can reconfigure particular social relations (Ruppert et al., 2013). These arguments have encouraged scholarship that “seeks to open the black


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

box of processors and arcane algorithms to understand how […] lines and routines of code […] work in the world by instructing various technologies how to act" (Kitchin & Dodge, 2011). These lines of inquiry have found that algorithms have a tremendous influence that shapes our online consumption and the people we interact with. Given the ubiquity of algorithm-driven suggestions of who to interact with and what to talk about, these types of perspectives are particularly prescient.

Gillespie (2014) draws attention to the default attitude of 'algorithmic objectivity', wherein there is a social perception of algorithmic impartiality. This work also renders visible a dissonance between what types of content we may 'expect' on online platforms and what content may actually be indexed under the purview of algorithms. Gillespie's (2014) work on inclusion/exclusion and the oxymoron of 'raw data' informs much of this line of work. Algorithms in social spaces are tasked with shaping our experiences quite profoundly. But, the desire to maximize helpfulness means that they go from 'raw data', helpfully trying to characterize it and recommend to those whom it thinks it will be valuable to – regardless of whether the content might be socially objectionable or not.

Moreover, the combination of fine-grained geolocation with sometimes long-standing search history can mean that the sort of metadata-driven portraits that some firms store about us can be surprisingly rich (Kliman-Silver et al., 2015). This can have profound implications on the operation of algorithms and ultimately on the construction of our material experience of social media. Additionally, the types of algorithms that are deployed on social media are not necessarily based on a linear model, but rather on a complex systems model, where networks can be dynamically incorporated into decision-making engines.

## YouTube and Content Discovery

There is a rich literature that studies how YouTube's content discovery systems operate in terms of developing models of user interests from browsing history, similarity, and popularity (e.g., Davidson et al., 2010; Li et al., 2013). Davidson et al.'s (2010) study was solely authored by Google authors and provides an authoritative starting point for understanding YouTube's content discovery systems by making clear that a user's recommended videos are "generated by using a user's personal activity (watched, favorited, liked videos) as seeds and expanding the set of videos by traversing a co-visitation based graph of videos" (p. 294). Perhaps a more candid peek at the operation of YouTube's content discovery features was posted on its 'Creator discovery handbook', a page taken down in 2015. Here, YouTube states, "we run hundreds of experiments each year to make this and every other discovery feature better [… and] current algorithms will change over time."[2] Indeed, Christos Goodrow, an engineering director at YouTube, noted that it takes over 1 million lines of code to figure out what videos to recommend (Computerphile, 2014).

More recent empirical work done by Google authors explains how the platform uses 'deep' neural methods for its content discovery systems (Covington et al., 2016). Covington et al. (2016, p. 193) emphasize that their content discovery systems are able to scale as their systems are '[i]nspired by continuous bag of words language models' and that their architecture learns 'high dimensional embeddings for each video in a fixed vocabulary and feed these embeddings into a feedforward neural network'. This scalable system also allows YouTube to 'assimilate many signals' and

[2] https://web.archive.org/web/20150329041618/https://support.google.com/youtube/answer/6060859?hl=en&ref_topic=6046759

integrate recently uploaded ('fresh') content into its content discovery systems (Covington et al., 2016). Academic work without access to YouTube's internal data has tended to focus on studying data sets accessible through the YouTube API. A noteworthy example is Rieder et al.'s (2018) study of the top 20 results using YouTube search over a period of 44 days, which found that the content that was discovered was "heavily influenced by both issue and platform vernaculars' and that YouTube's content discovery systems provide barriers that 'frustrate our attempts at causal explanation and are better served by strategies of 'descriptive assemblage'". Each of their daily searches was done for the following keywords – gamergate , islam australia, islam, trump, sanders, refugees and syria - and they found "YouTube's search mechanism is designed to pick up on what we have called 'newsy' moments, and whenever it does, results can change drastically from one day to the next". Rather than provide a causal explanation, their 'descriptive assemblage' follows Latour's notion of seeking to 'navigate' an assemblage using API-derived data of 7262 videos. Rieder et al. (2018, p. 63) importantly underscore that the watch time of videos is not available to external researchers via the API, making multivariate analysis problematic.

## Cyber-Extremism

Khader et al.'s *Combating violent extremism and radicalization in the digital era* (2016) provides a comprehensive discussion of cyber extremism and usefully singles out a working definition of online radicalization by the U.S. Community Oriented Policing Services which argues that the term is indicative of a shift by an individual from 'mainstream beliefs' to 'extreme views' and singles out Facebook, Twitter, and YouTube as spaces of 'online extremist content' (Office of Community Oriented Policing Services, 2014, p. 1). Khader et al. (2016, p. xxi) argue that ISIS's use of social media highlights that the unique affordances of these online platforms may help to "increas[e] an individual's susceptibility to persuasion by violent extremist ideologies". Empirical studies of cyber extremism have used both qualitative and quantitative approaches to better understand motivations for individuals to radicalize, whether there are patterns in posted text, and the demographics of posters where identifiable.

Torok (2020) uses a longitudinal study of Facebook and found that social media posts use 'narratives of grievance' to 'stimulate strong emotive responses to perceived injustices' (p. 39). She also notes that Facebook's group pages have been useful to 'attract like-minded radicals' and emphasizes that the mentorship, discursive engagement, and social backing seen in these spaces is instrumental as 'part of a virtual community of extremists with similar ideologies' (p. 40). A study of 100 Facebook pages and 50 twitter accounts (Awan, 2017) found that ISIS use social media successfully for recruitment and propaganda and that these social media posts were used to "glamorise 'extremism' and make it appear as though fighting with them is 'cool'". Awan (2017) found that 90% of identifiable posters sympathizing with ISIS or operating on behalf of ISIS were male and where country of residence could be identified, they found half of the list consisted of Western countries and the other half of predominantly Muslim countries. More computational work has focused on predicting the likelihood of individuals becoming radicalized online. Scanlon & Gerber (2015) used machine learning on a unique data set of posts from mouse double-click the Dark Web to study whether textual content from posts could be used to forecast future recruitment activity. They found some consistencies in posted text that enable machine learning methods to be used to automatically forecast violent extremist cyber-recruitment. Ahmed and Jacob (2020) studied over 50 million tweets containing the keywords 'ISIS' or 'ISIL', including the study of

Arabic language tweets. Their large corpora enabled them to empirically conclude that Twitter was actively used by both ISIS and Al-Qaeda for ongoing, intense feuds. Ahmed and Jacob (2020, pp. 1442-1443) also found evidence that ISIS' activity on Twitter was used for "pressuring other extremists to join ISIS".

Other work has focused on ISIS' strategies online. Nissen (2014, p. 2) found that ISIS uses a 'cross-media' perspective on a range of social media platforms to recruit and radicalize followers with high levels of success. ISIS appears to have also paid careful attention to producing content with emotive appeal that normalizes the lives of jihadis. For example, they used the Twitter account 'Islamic State of Cat' (@ISISCats) to post images of jihadists feeding cats milk and other cat-related content (Vincent, 2014). This approach of normalizing jihadi life has been an important avenue of radicalization as it can easily act as an unintended/accidental gateway to stronger, explicit ISIS-related content. Nissen (2014) argues that one of ISIS' most successful social media efforts was their app, 'The Dawn of Gold Tidings' – otherwise known as 'Dawn'. Though this app has since been shut down, the app was able to tweet from 40,000 user accounts at its prime (Berger, 2014), providing the illusion of a strong, pervasive message on Twitter.

**ISIS and YouTube**

Early online cyber extremism by Al-Qaeda and similar extremist groups used YouTube (Chen et al., 2008) and the platform has similarly been instrumental to ISIS not only in terms of recruitment videos, but also has been the prime platform for engaging with international media outlets. Islamic State content posted on YouTube has historically been quickly picked up by major news media. Journalists regularly report how Islamic State's content has 'spread widely through social media'


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

and includes 'slick propaganda videos' (Mazzetti & Schmitt, 2016). This is in distinction to counter-terrorist YouTube interventions by the U.S. State Department, which have been characterized as unsuccessful (Miller & Higham, 2015). First-generation ISIS videos including *The End of Sykes-Picot, The Flames of War, The Clanging of the Swords I-IV, and Upon the Prophetic Methodology* focused on narratives of recruitment and legitimacy. Using high-quality cinematic production, these videos make appeals that blend the cause of an Islamic caliphate with humanized depictions of ISIS fighters. From short, bite-sized videos to content featuring fighters feeding their cats, ISIS tailor-made content they believed would do well on YouTube.

**Existing studies on content moderation and algorithmic detection of radical content**

Computational approaches have been successful in the detection of clusters of extremist users on Twitter (e.g., Moussaoui et al., 2019) as well as content moderation and detection of radical content (e.g., in identifying jihadist text on Twitter (Ashcroft et al., 2015)). Vanguard work was actually conducted by hackers with Anonymous and a splinter group of theirs, Ghost Security Group, who used various machine learned and other methods to identify 100,000 social media accounts linked to ISIS (McElreath et al., 2018). Benigni et al. (2017) use Iterative Vertex Clustering and Classification (IVCC) to study a network of over 22,000 Twitter users sympathetic to ISIS. Their approach identifies ISIS-supporting subgraphs through iterated community detection and vertex classification. They use 5 seed account and collect account information 2° out (n= 119,156), which results in a sample with a "preponderance of accounts [… with] no visible affiliation with ISIS" (Benigni et al., 2017). They labeled some of these data and then evaluated various classifiers, such as Random Forest, SVM, and others. Benigni et al. (2017) were able to develop a classifier with an F1 Score of 75.8%, representing a good first attempt at this classification problem. Ferrara et al. (2016) use similar methods, though rather than a seed network, use a dataset of 25,000 users identified by Twitter as extremist, paired with their own random sample of 25,000 accounts. Though they have lower F-1 scores than Benigni et al., they do achieve high Area Under Curve (AUC) scores after evaluating Logistic Regression and Random Forests classifiers. Using machine-learned methods, Ferrara et al. (2016) were able to successfully detect extremist users, predict extremist content adoption, and achieved notable results in forecasting interactions between extremist users and non-extremist users.

**Research questions**

The hypothesis of this study is that prior to its highly visible 2017 content moderation and accountability efforts, ISIS content was accessible through YouTube's search and recommender systems. The hypothesis is drawn from popular news reports as well as government reports that the YouTube website was part of the radicalization process for jihadis.

*Research question 1 (RQ1):* Given that ISIS content is regularly deleted from YouTube, were radical videos ISIS-related incorporated into the recommender algorithm?  Content discovery systems are good at dealing with new content by immediately looking for 'features'.  The purpose of this research question is to see if the features of radical videos are incorporated into the platform's recommender system, making the content more accessible.

*Research question 2 (RQ2):* Politicians and journalists have provided examples of how ISIS has used social media platforms to make their content highly accessible and have argued that these companies need to better moderate content (Kosoff, 2016; Malik et al., 2014). However, there is a dearth of academic work that studies the types of videos that might be recommending ISIS content.

Given that Google states the majority of YouTube video views are through the recommender system (Davidson et al., 2010), this research question asks if ISIS content is on YouTube and what types of videos are recommending ISIS content?

*Research question 3 (RQ3):* Part of platform accountability efforts seek to make more transparent the processes behind algorithms. However, given Rieder et al.'s (2018) findings that 'strong causality remains elusive' in YouTube API-derived data studies, can mixed methods be used to identify a set of attributes that might help explain the decision-making processes of YouTube's recommender algorithms? The purpose of this research question is to help inform future platform accountability debates and provide additional context to historical debates.

**Methods**

ISIS never had a stable, official YouTube channel with central hosting. Therefore, to identify accessible videos, we examined news reports for the titles of ISIS videos released as of May, 2016 and 15 videos were selected that were attributed to ISIS's media wing, *Al Hayat Media Center*. The anonymized web browser, Tor Browser, was used to search for each video title and then was closed and relaunched in order to not bias results based on cookies, location, or IP address (Google accounts were not signed into). This was done to make a concerted effort to avoid personalization, but we do acknowledge there are many strategies for personalizing results derived from the web and Application Programming Interfaces (APIs) (Hannak et al., 2013; Kliman-Silver et al., 2015) and that search results could have been biased towards a de-personalized Tor Browser. Of the 15 video titles identified, 11 were found on YouTube and used as seeds (see Table 1 for a list of videos that were found on YouTube; some were not readily found using English-language searches).

The YouTube API was queried in October 2016 via YouTube Data Tools (Rieder, 2015) to produce a list of 'recommended' videos. These 11 videos were used as seeds in order to build a network that represents videos: (a) recommended, (b) recommended by the recommended videos, and (c) videos recommended by recommended videos in (b). Data was stored as a directional network and these were visualized using the Social Network Analysis (SNA) package Gephi to determine what ISIS videos were being recommended. All seed videos were produced by *Al Hayat Media Center*. The graph network gathered through this process included 15,021 nodes and 190,087 edges. The collected data was cleaned by creating a network file which only included recommendations to ISIS videos (which we define as videos ISIS officially claimed to have released and must include branding of *Al Hayat* Media Center or other official ISIS media wings). Videos included in these networks had to be recommending ISIS videos. A known limitation of this method is that the sample network is derived from seeds rather than being a generic crawl of the YouTube video network.

**INSERT TABLE 1 HERE**

Using custom-developed Python scripts, we then read in the original network returned by YouTube data tools and inserted an additional column that indicated whether the node directly recommended one of the 11 seed videos (i.e. only videos that recommended an ISIS video). This produced a subset with 67 videos. We sent all non-English language video titles to the Google Translate API and inserted a new column with translations. We then added two additional columns and coded these manually – primary language of the video and whether the video was an ISIS video. The decision rule for primary language classification was purely based on Google Translate and for determining whether a video was an ISIS video or whether the video was officially acknowledged by ISIS and contained official branding by an ISIS media wing (e.g., *Al Hayat*, *Al-Furqan*, and *Ajnad Media*). We then read in the whole network list of 15,021 videos and used a Python script to add a column indicating if the video was an ISIS video[3]. The purpose of this level of data transformation was threefold: 1) to visualize the sub-network by metadata and hand-coded attributes; 2) to test for any association between YouTube metadata attributes and recommendation to ISIS content; and 3) to identify what metadata attributes are associated with official ISIS videos. Additionally, descriptive statistics were used to understand some metadata attributes.

By undertaking this level of data transformation, we are able to go from a very large network crawled to 2 degrees of YouTube videos from 11 seeds to visualizing just the recommending nodes around these videos. Correlations were performed in order to identify whether any association exists between node videos that YouTube had selected to recommend and seed videos. Kolmogorov-Smirnov tests were conducted in SPSS to determine if values followed a normal distribution. As with previous work using YouTube data (Huguenin et al., 2012), the YouTube metadata was not normally distributed. For this reason, the nonparametric Spearman's rho statistic was used for correlation analysis.

The explanatory power of these correlations is limited given that the collected data is not of YouTube as a whole. Rather, it is more appropriate for us to make an argument about attributes that *may be associated* with a particular outcome. We therefore use Qualitative Comparative

[3] If any API-derived field matched a list of keywords for official ISIS videos (e.g. mujatweets, ISIS/IS/ISIL + official video title), it was tentatively flagged and then human checked in a TorBrowser session before inclusion in the finalized data set.


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

Analysis (QCA) (Rihoux & Ragin, 2008), a combinatorial method which uses coding of content- either via 'crisp set' (where binary 0/1 values are coded) or 'fuzzy' (where greyer areas can also be coded, providing a spectrum of values). QCA was used to provide insights into patterns in order to better understand the types of attributes that may be influencing the decision-making of YouTube's algorithms when recommending ISIS videos. The crisp set method was selected as the video either had or did not have particular attributes (e.g., posted by an organization, published within two years, title contains a radical keyword[4], title includes Islamic State/ISIS/ISIL/Daesh). QCA operates by taking a coded matrix where each row represents a case (e.g., each row is one of the YouTube videos in the network) and values are entered for each of the binary variables being coded (0 = false and 1 = true). QCA's approach is driven by a 'truth table' algorithm (Rihoux & Ragin, 2008) which is iterated and qualitatively evaluated at each iteration for feasible explanations for particular outcomes.

QCA is not based on explicit causation, but rather data is iteratively examined comparatively in terms of particular relevance that is assisted by the truth table algorithm, which looks for some indication of relevance, but does not decide this relevance itself. Rather, potential options are presented to the researcher, who, with prior understanding of the data set, evaluates their feasibility. Systems of Boolean algebra used in comparative social science draw from the fact that outcomes (whether content is recommended by algorithms, for example) can be expressed as the presence and absence of conditions (this parallels the origins of Boolean algebra's roots in

[4] 'radical keyword' is defined as true if the video includes terms derived from the seed videos ("no respite", "Sykes picot", "Mujatweets", or "Al Hayat") in the title, description, or initial frames of the video being coded.

electronic circuits). Therefore, our goal is to develop 'recipes' of attributes behind causation related to YouTube's video recommendation.

We acknowledge that a limitation of our approach is that starting with seeds rather than random videos from YouTube to build a network introduces bias. However, given ISIS videos on YouTube are needles in a haystack, taking such an approach would have likely yielded no radical content. To mitigate bias, we selected: 50 videos from the initial seed network, excluding seeds, but including videos not recommending ISIS videos; 50 randomly selected videos from the derived seed network that did not recommend ISIS videos; and 50 videos randomly selected from the YouTube video network as a whole using the Random YouTube Videos platform (RandomVideoBot, 2016). A total of 188 videos were hand coded along a binary scale for the 11 attributes we selected.

**Results**

We find that, in 2016, one's likelihood of encountering ISIS content accidentally was rare. Such occurrences were likely due to a video sharing similar metadata attributes to official ISIS videos (particularly similarity in keywords in their titles). Therefore, in terms of RQ1, radical videos are incorporated into the recommender algorithm. This provides a background to GIFCT's mandate in combating terror-related content in digital platforms. ISIS content being recommended tended to be by other ISIS videos, particularly those that were non-English language, providing an answer to RQ2 which asks what types of videos are recommending ISIS content. Lastly, in terms of RQ3, which asks whether it is possible to identify a set of attributes that may help explain part of the YouTube algorithm's decision-making process, our QCA-derived results indicate that

YouTube's recommendations of radical content quite likely factor some weight to the presence of radical keywords in a video's title.

These findings add to the literature by providing some empirical confirmation of Rieder et al.'s (2018) finding that YouTube content discovery systems are partially based on unique platform vernaculars and, very importantly, and by being part of the process of 'Auditing Radicalization Pathways on YouTube' that Ribeiro et al. (2020) recently framed. Our findings also indicate that by 2016, YouTube was not accessibly hosting a significant volume of ISIS content. Following Ribeiro et al.'s (2020) call for auditing YouTube, the mere presence of any extremist content on YouTube in 2016 is important to empirically substantiate.

From a qualitative perspective, the seeds included particularly extreme content including beheading videos and gaming-like shoot-out content alongside 'humanizing' content like guided tours of ISIS territory, kids smiling, and markets full of produce. An information seeker in 2016 could encounter these videos within diverse searchable parameters (e.g., 'life in Raqqah', 'inequality against Muslims' or 'Syrian Muslims', all terms we searched for). These findings do provide some backdrop to understanding political and public pressure for increased accountability by YouTube as such searches should not have been yielding ISIS content in the first place. Moreover, if recommendations were largely being made on keyword, title, and other metadata attributes, interceptions to halt such recommendations would likely have not been excessively difficult.

*Correlation Analysis*

We performed correlations of metadata derived from the YouTube API, including views, posting date, comment count, and like count. As these correlations are derived from a biased set of data (a seed network), these results are not meant to suggest association or causation, a similar position taken by Rieder et al. (2018). Rather, they provide a means to better understand aspects of the large network surrounding the seed videos. Other work has found so-called 'spill over' effects by items in proximity within a recommender network (Kummer, 2013). Moreover, these correlations provide a baseline understanding of our data set, which is itself a prerequisite to Qualitative Comparative Analysis (QCA), the primary method we used to determine the attributes of videos that recommend radical ISIS content.

As Table 2 illustrates, we found that the number of views of videos was negatively correlated with both whether a video recommended one of the 11 seed videos ($r_s = -.078$, $p <. 01$) and whether the video was an official ISIS *Al Hayat* video ($r_s = -.043$, $p <. 01$). Specifically, those videos with lower overall view counts were more likely to be videos that were recommending ISIS videos. Figure 1 illustrates this; view counts of videos not recommending ISIS content are normally distributed. Videos that recommend ISIS videos are multimodal and were not highly viewed. The number of comments of videos was negatively correlated with whether a video recommended a seed video ($r_s = -.066$, $p <. 01$) and whether the video was an ISIS video ($r_s = -.038$, $p <. 01$). Therefore, videos that received more comments were less likely to be recommending ISIS videos and, similarly, videos with more comments are less likely to be ISIS content. The number of likes was negatively correlated with whether a video recommended a seed video ($r_s = -.067$, $p <. 01$) and whether the video was an ISIS video ($r_s = -.040$, $p <. 01$). These results do indicate collinearity of likes and comments.


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

We also found that the video's YouTube category was associated with whether a video recommended a seed video ($r_s$ = -.072, p <. 01) and whether the video was an ISIS video ($r_s$ = -.052, p <. 01) (see Figure 2). The 16 categories were converted to integers and the test indicates that the lower value categories, particularly *People & Blogs* and *News & Politics*, are more likely to be the categories of videos recommending a seed video or whether a video in our data set was classified as an ISIS video.

**INSERT TABLE 2 HERE**

As Figure 2 illustrates, publication date was associated with whether a video recommended one of a seed video ($r_s$ = .058, p <. 01) and whether the video was an ISIS video ($r_s$ = .039, p <. 01). As YouTube data tools returns publication date in UNIX time, the number of seconds elapsed since January 1, 1970, this indicates that the newer the video, the more likely it is recommending a seed video or is itself an official ISIS video.

**INSERT FIGURE 1 HERE**

Categories of videos recommending official ISIS content have a small amount of diversity, concentrating in five categories, which are *News & Politics*, *People & Blogs*, *Music*, *Film & Animation*, and *Comedy*. However, videos classified as official ISIS videos are concentrated in two main categories – *People & Blogs* and *Music*. The *People & Blogs* category makes up 80.6% of ISIS videos coded in our sample.

**INSERT FIGURE 2 HERE**

ISIS videos are much less likely to be recommended by English-language videos (see Figure 3); only 11.94% of the videos recommending ISIS seed videos were English-language. This contrasts with the fact that 67.16% of videos recommending ISIS content are Arabic-language, an unexpected finding for RQ2.


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

Specifically, we found there is variation in recommendations based on the language of a video. Due to the small number of videos studied, future work is needed to thoroughly explore the generalizability of effects video language might have.

**INSERT FIGURE 3 HERE**

Figure 4 illustrates a case of YouTube recommending radical content from non-radical videos using the example of the ISIS video 'No Respite'. This video, which gained substantial press coverage as it labels George W. Bush as a liar and Bill Clinton as a fornicator, is recommended from some videos containing the words 'No Respite'. This surprisingly occurred with an Indian Telegu news program and a report from Al Jazeera English titled 'No respite for US from deadly tornadoes'. The latter disturbingly had the ISIS 'No Respite' video as the first recommended video. (This recommendation may not have appeared for every user as an anonymized Tor Browser session was used for this study). This provides empirical evidence to answer RQ1 in that we are able to conclude that radical content was assimilated by the recommender system. There seems to be no 'cold start' problem, wherein new radical content is disconnected from the recommender system.

**INSERT FIGURE 4 HERE**

Previous work argues that YouTube's algorithm makes some of its decisions based on other users' actions, preferences, and behavior (Baluja et al., 2008; Covington et al., 2016; Davidson et al., 2010). What the case of 'No Respite' suggests is that the title of a video mattered despite YouTube having a rich array of other data. It is computationally minimal to recommend based on metadata. However, other videos likely recommend radical content based on their position within the network. In the case of a Hindi video of two young girls singing, *Sadia Attaria and Maria Attari*, *No Respite* is recommended (see Figure 4). This seems to be based on network position or non-metadata attributes.


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

ISIS recommendations tended to come from ISIS videos. Though YouTube had already made it difficult for users to find these content in 2016, there were likely instances that once users discovered ISIS videos on YouTube, they were recommended other ISIS videos. It was also likely rare for users consuming non-ISIS videos to be directed to ISIS videos. For non-English videos, the likelihood was greater and this is a noteworthy finding regarding RQ2. Browser searches in 2016 confirmed that a *Mujatweets* video (minute-long ISIS videos that are relatively language neutral) recommended other videos in the *Mujatweets* series.

**INSERT FIGURE 5 HERE**

Figure 5 illustrates recommendations from *Mujatweets 2* to *Flames of War*, which our network data confirms then recommends other ISIS videos, including another *Mujatweets* video in Raqqah. As figure 5 illustrates, a YouTube user viewing *Mujatweets* 2 may have been recommended the *Flames of War* series or other ISIS videos, which, for example, constitute 32.84% of the network depicted in figure 6. Though beyond the remit of this study, the behavior of YouTube's recommender system potentially did make it easier to find ISIS content. Some of these videos were challenging to find manually (as confirmed via anonymized Tor Browser sessions).

**INSERT FIGURE 6 HERE**

*Qualitative Comparative Analysis (QCA)*

Qualitative Comparative Analysis (QCA) can be used to produce 'truth tables' (as in table 3), which are 'combinations of conditions' being evaluated with a 'selected outcome' (Ragin, 2018, p. 35), which in our study, is the recommendation of ISIS videos. Rather than evaluating causality, QCA follows how case-oriented researchers look for combinations of conditions that are "shared by the positive cases", "believed to be linked to the outcome", and "not displayed by negative cases" (Ragin, 2018, p. 45). The columns in truth tables indicate the explanatory power of various possible combinatorial solutions. The 'Attribute' column signifies "causal recipes", combinations of conditions believed to be linked to the outcome. "Raw coverage measures the proportion of memberships in the outcome explained by each term of the solution"; "Unique coverage measures the proportion of memberships in the outcome explained solely by each individual solution term (memberships that are not covered by other solution terms); and "Consistency measures the degree to which membership in each solution term is a subset of the outcome" (Ragin, 2018, p. 61). "Solution coverage measures the proportion of memberships in the outcome that is explained by the complete solution" and "Solution Consistency measures the degree to which membership in the solution (the set of solution terms) is a subset of membership in the outcome" (Ragin, 2018, p. 61).

**INSERT TABLE 3 HERE**

Correlation analysis provided the requisite context needed for QCA, which we used to distill the attributes of YouTube videos recommending ISIS content. The fs/QCA software package (Ragin & Davey, 2014) was iteratively used and, at each phase, produced three possible solutions - complex, parsimonious, and simple. All sets were evaluated. After eight iterations, a final solution set (see table 3) was identified that was greater than chance for QCA (i.e. ≥.5), but also was in tandem with the results found during previous analysis of the content. Network and correlation analysis indicated that keyword play a role and QCA validates this. Unlike the correlation analysis, the QCA dataset includes a random sample of non-ISIS YouTube videos. Table 3 (per raw coverage) illustrates how 34.69% of recommendations of ISIS content are attributable to the presence of radical keywords in a title and a further 28.57% by the video being: a newscast, posted by an organization, primarily English language, and recent. (The remaining 3 Boolean sets in Table 3 provide a further explanation of ~ 16% of recommendations). The QCA results provide some attribute cues to understand why radical ISIS content was being recommended despite low view and comment counts.

**Limitations**

A clean, anonymized Tor Browser was used for searching and average users or those with YouTube accounts with viewing history might not have received the same recommendations. Moreover, recommendation results could have been biased by access via the YouTube API. However, using the YouTube app on a phone of one of the authors in October 2016 which had no previous searches for ISIS content, but did have a VPN enabled, yielded autocomplete recommendations that were similar to Tor Browser-derived results (see Figure 7). It can therefore be safely assumed that at least some users were receiving similar recommendations in 2016.

**INSERT FIGURE 7 HERE**

We have no way to identify how tagging, user attributes, global characteristics of user searching, or location posted from plays into our sampled network. Though we do know that "click through rate from a video to its related videos is high" (Zhou et al., 2010) and that "recommendations account for about 60% of all video clicks from the homepage" (Davidson et al., 2010, p. 296), we have no way of knowing the click through rate or watch time of recommended ISIS videos. Though future work could also use attributes such as YouTube channel popularity to measure aspects of engagement.

## DISCUSSION

Social media commentators have remarked on ISIS' successful social media strategy (Berger, 2014). Part of these conclusions involved a perception that ISIS videos were accessible on YouTube. Therefore, tracing some of ISIS' historic footprint on YouTube is important to provide empirical evidence to retroactively evaluate some of the claims made at the time.

We found ISIS-related videos within a variety of genres on YouTube. The attribute most likely to influence ISIS videos being recommended from non-ISIS videos was having a similarity of keywords in the title or some metadata similarities. Second, there were instances where ISIS videos were recommending other ISIS videos beyond our seed videos. Though not large in volume, ISIS content was searchable and being recommended. Moreover, any presence of extremist content is relevant to platform accountability debates.

YouTube does seem to have been effective in identifying and taking down ISIS content by *Al Hayat*, *Al Furqan*, and *Ajnad Media* at the time. Searches on TorBrowser confirm that the *Mujatweets* series - instantly found in 2016 - were not found at the time of writing. ISIS was reported to have published "about 90,000 pages on social networking sites, especially Facebook and Twitter" (al arab, 2015), so it is possible that these content were more responsible for driving traffic to ISIS videos and the rare recommendations to ISIS videos might not have been a major cause for concern in 2016.

Our findings also suggest that ISIS recommendations were not driven by any particularly complex neural process, despite the addition of deep learning to YouTube's content discovery systems during the time of our study (Covington et al., 2016). Given any incidence of extremist content is problematic, platforms like YouTube can potentially make simple pro-active algorithmic changes to be more accountable. Therefore, future pressure on platforms need not have wide ranging demands as small changes in decision rules can also vastly increase platform accountability.

## CONCLUSIONS

In 2016-2017 there was much public interest in understanding ISIS' footprint on social media platforms such as YouTube and this has influenced the searchability and accessibility of this content today. Though YouTube had been engaged in efforts to combat terror-related content prior to 2017, a major step change took place then and in the intervening years. In 2020, terror-related content on YouTube is not readily accessible. However, in 2016, a different picture emerges. This study finds that ISIS content was present on YouTube and incorporated within recommendations. Though the volume of ISIS content found was low, the fact that it was searchable and being recommended provides some insight into the public and governmental pressure for YouTube to become more accountable and audit its own algorithms. This was not an issue restricted to YouTube as several technology companies came together in 2017 to found the Global Internet Forum to Combat Terrorism (GIFCT).

This study used 11 ISIS videos as seeds to build a 2-degree network of 15,021 videos with 190,087 edges. First, we found evidence that despite YouTube's efforts to remove terror-related content, official ISIS content was searchable and being recommended. Second, we found that ISIS videos tended to recommend other ISIS videos. Third, we found that some non—ISIS videos did recommend ISIS videos, but this was rare and seemingly due to shared metadata attributes with ISIS videos (particularly having a similar title).

Our findings also indicate that users who were actively searching for ISIS videos may have been assisted by YouTube's recommender algorithm, particularly once they found ISIS videos. However, our findings also suggest that in 2016 it was not easy finding ISIS videos, but it was possible. As Table 1 indicates, ISIS videos accessible on YouTube in 2016 are noticeably absent several years later despite extensive searching. In 2020, the only seed video that could be found (and not uncut) was by searching for 'ISIS發布最新影片 驚見中華民國國旗' (translated as 'ISIS releases the latest film, surprised to see the flag of the Republic of China')[5] and the audio is dubbed. In 2016, a search for 'No Respite' yielded an uncut version with 'ISIS' in the title. In 2020, the only readily findable ISIS content is embedded within news or academic commentary. Clearly, YouTube has taken seriously efforts to combat ISIS videos.

The pressure mounted on YouTube after our data were collected mark a shift towards the platform acting somewhat like a publisher, rather than merely as a video sharing platform. It is also clear that for some of the recommendations being made (e.g., non-ISIS videos recommending ISIS videos based on similar titles), relatively small changes would have likely eliminated these

[5] The Republic of China (ROC) flag is included alongside other flags, where the video claims these countries are part of a 'Global Coalition Against the Islamic State'. Chinese and Taiwanese journalists reported on the ROC flag's inclusion in *No Respite*.


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

recommendations. We hope these findings help encourage future platform accountability efforts by social media companies.


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

**REFERENCES**

Ahmed, A.-R., & Jacob, G. (2020). Jihadist Propaganda on Social Media: An Examination of ISIS Related Content on Twitter. In Information Resources Management Association (Ed.), *Cyber Warfare and Terrorism: Concepts, Methodologies, Tools, and Applications* (pp. 1442-1457). IGI Global. https://doi.org/10.4018/978-1-7998-2466-4.ch085

al arab. (2015). ترسانة إعلامية ضخمة لتنظيم الدولة الإسلامية. Retrieved October 24, 2016 from http://alarab.co.uk/?id=47271

Ashcroft, M., Fisher, A., Kaati, L., Omer, E., & Prucha, N. (2015, 7-9 Sept. 2015). Detecting Jihadist Messages on Twitter. 2015 European Intelligence and Security Informatics Conference,

Awan, I. (2017, 2017/04/01). Cyber-Extremism: Isis and the Power of Social Media. *Society, 54*(2), 138-149. https://doi.org/10.1007/s12115-017-0114-0

Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D., & Aly, M. (2008). Video suggestion and discovery for youtube: taking random walks through the view graph. Proceedings of the 17th international conference on World Wide Web,

Benigni, M. C., Joseph, K., & Carley, K. M. (2017). Online extremism and the communities that sustain it: Detecting the ISIS supporting community on Twitter. *PLOS ONE, 12*(12), e0181405. https://doi.org/10.1371/journal.pone.0181405

Berger, J. M. (2014). How ISIS Games Twitter. *The Atlantic*. http://www.theatlantic.com/international/archive/2014/06/isis-iraq-twitter-social-media-strategy/372856/

Chen, H., Thoms, S., & Fu, T. (2008, 17-20 June 2008). Cyber extremism in Web 2.0: An exploratory study of international Jihadist groups. 2008 IEEE International Conference on Intelligence and Security Informatics,

Computerphile. (2014). *YouTube's Secret Algorithm - Computerphile*, YouTube. https://www.youtube.com/watch?v=BsCeNCVb-d8

Covington, P., Adams, J., & Sargin, E. (2016). *Deep Neural Networks for YouTube Recommendations* Proceedings of the 10th ACM Conference on Recommender Systems, Boston, Massachusetts, USA. https://doi.org/10.1145/2959100.2959190

Davidson, J., Liebald, B., Liu, J., Nandy, P., Vleet, T. V., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B., & Sampath, D. (2010). *The YouTube video recommendation system* Proceedings of the fourth ACM conference on Recommender systems, Barcelona, Spain.

Ferrara, E., Wang, W.-Q., Varol, O., Flammini, A., & Galstyan, A. (2016). Predicting Online Extremism, Content Adopters, and Interaction Reciprocity. In E. Spiro & Y.-Y. Ahn, *Social Informatics* Cham.

Feuz, M., Fuller, M., & Stalder, F. (2011). Personal Web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation [Digital Methods, Personalisation, Web Search, Google, Advertising]. *First Monday, 16*(2). https://doi.org/10.5210/fm.v16i2.3344

Freytas-Tamura, K. D. (2015, September 20, 2015). ISIS Defectors Reveal Disillusionment. *The New York Times*. http://www.nytimes.com/2015/09/21/world/europe/isis-defectors-reveal-disillusionment.html

Fuller, M. (2008). *Software studies : a lexicon*. MIT.

Gates, S., & Podder, S. (2015). *Social Media, Recruitment, Allegiance and the Islamic State* (Vol. 9) [Jihadism; ISIS; foreign fighters; recruitment]. http://www.terrorismanalysts.com/pt/index.php/pot/article/view/446

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, K. A. Foot, W. E. Bijker, W. B. Carlson, & T. Pinch (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167-194). MIT Press.

Global Internet Forum to Counter Terrorism. (2019). *Global Internet Forum to Counter Terrorism (GIFCT)*. Retrieved October 17 from https://gifct.org/

Hannak, A., Sapiezynski, P., Kakhki, A. M., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). *Measuring personalization of web search* Proceedings of the 22nd international conference on World Wide Web, Rio de Janeiro, Brazil.

Huguenin, K., Kermarrec, A.-M., Kloudas, K., & Taïani, F. (2012). Content and geographical locality in user-generated content sharing systems. Proceedings of the 22nd international workshop on Network and Operating System Support for Digital Audio and Video,

Khader, M., Neo, L. S., Ong, G., Eunice Tan, M., & Chin, J. (Eds.). (2016). *Combating Violent Extremism and Radicalization in the Digital Era*. IGI Global. https://doi.org/10.4018/978-1-5225-0156-5.

Kitchin, R. (2017, 2017/01/02). Thinking critically about and researching algorithms. *Information, Communication & Society, 20*(1), 14-29. https://doi.org/10.1080/1369118X.2016.1154087

Kitchin, R., & Dodge, M. (2011). *Code/sp recommender system,ace : software and everyday life*. MIT Press.

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., & Mislove, A. (2015). *Location, Location, Location: The Impact of Geolocation on Web Search Personalization* Proceedings of the 2015 Internet Measurement Conference, Tokyo, Japan.

Kosoff, M. (2016, June 14, 2016). Clinton Calls on Tech Companies to Fight ISIS Propaganda. *Vanity Fair*. https://www.vanityfair.com/news/2016/06/clinton-calls-on-tech-companies-to-fight-isis-propaganda

Li, H., Ma, X., Wang, F., Liu, J., & Xu, K. (2013). *On popularity prediction of videos shared in online social networks* Proceedings of the 22nd ACM International Conference on Information and Knowledge Management, San Francisco, California, USA.

Malik, S., Laville, S., Cresci, E., & Gani, A. (2014, Wed 24 Sep 2014). Isis in duel with Twitter and YouTube to spread extremist propaganda. *The Guardian*. https://www.theguardian.com/world/2014/sep/24/isis-twitter-youtube-message-social-media-jihadi

Mazzetti, M., & Schmitt, E. (2016, July 17, 2016). In the Age of ISIS, Who's a Terrorist, and Who's Simply Deranged? *The New York Times*. http://www.nytimes.com/2016/07/18/world/europe/in-the-age-of-isis-whos-a-terrorist-and-whos-simply-deranged.html?_r=1

McElreath, D. H., Doss, D. A., McElreath, L., Lindsley, A., Lusk, G., Skinner, J., & Wellman, A. (2018). The Communicating and Marketing of Radicalism: A Case Study of ISIS and Cyber Recruitment. *International Journal of Cyber Warfare and Terrorism (IJCWT), 8*(3), 26-45. https://doi.org/10.4018/IJCWT.2018070103

Miller, G., & Higham, S. (2015, May 8, 2015). In a Propaganda War against ISIS, the US Tried to Play by the Enemy's Rules. *The Washington Post*. https://www.washingtonpost.com/world/national-security/in-a-propaganda-war-us-tried-to-play-by-the-enemys-rules/2015/05/08/6eb6b732-e52f-11e4-81ea-0649268f729e_story.html

Moussaoui, M., Zaghdoud, M., & Akaichi, J. (2019). A possibilistic framework for the detection of terrorism-related Twitter communities in social media. *Concurrency and Computation: Practice and Experience, 31*(13), e5077. https://doi.org/10.1002/cpe.5077

Neumann, P. R. (2013, 06//). Options and Strategies for Countering Online Radicalization in the United States [Article]. *Studies in Conflict & Terrorism, 36*(6), 431-459. https://doi.org/10.1080/1057610X.2013.784568

Nissen, T. E. (2014). Terror. com: IS's Social Media Warfare in Syria and Iraq. *Contemporary Conflicts: Military Studies Magazine, 2*(2).

Office of Community Oriented Policing Services. (2014). *Online Radicalization to Violent Extremism*. https://www.theiacp.org/sites/default/files/2018-07/RadicalizationtoViolentExtremismAwarenessBrief.pdf

Ragin, C. (2015). From Fuzzy Sets to Crisp Truth Tables. University of Arizona, http://www.compasss.org/wpseries/Ragin2004.pdf

Ragin, C. (2018). *User's Guide to Fuzzy-Set/Qualitative Comparative Analysis 3.0*. University of California.

Ragin, C. C., & Davey, S. (2014). fs/QCA [Computer Programme], version 2.5. *Irvine, CA: University of California*.

RandomVideoBot. (2016). *Random YouTube Videos*. https://randomyoutube.net/watch

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W. (2020). *Auditing radicalization pathways on YouTube* Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain. https://doi.org/10.1145/3351095.3372879

Rieder, B. (2015). *YouTube Data Tools.* In Bernhard Rieder. https://tools.digitalmethods.net/netvizz/youtube/

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to 'ranking cultures':Investigating the modulation of visibility in YouTube search results. *Convergence, 24*(1), 50-68. https://doi.org/10.1177/1354856517736982

Rihoux, B., & Ragin, C. C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). Sage Publications.

Ruppert, E., Law, J., & Savage, M. (2013). Reassembling social science methods: The challenge of digital devices. *Theory, culture & society, 30*(4), 22-46.

Sanger, D. E. (2015, December 6, 2015). Hillary Clinton Urges Silicon Valley to 'Disrupt' ISIS. *The New York Times*. http://www.nytimes.com/2015/12/07/us/politics/hillary-clinton-islamic-state-saban-forum.html?_r=0

Scanlon, J. R., & Gerber, M. S. (2015). Forecasting Violent Extremist Cyber Recruitment. *IEEE Transactions on Information Forensics and Security, 10*(11), 2461-2470. https://doi.org/10.1109/TIFS.2015.2464775

Slavin, K. (2011). *How algorithms shape our world*. http://www.ted.com/talks/kevin_slavin_how_algorithms_shape_our_world

Torok, R. (2020). Social Media and the Use of Discursive Markers of Online Extremism and Recruitment. In Information Resources Management Association (Ed.), *Cyber Warfare and Terrorism: Concepts, Methodologies, Tools, and Applications* (pp. 478-508). IGI Global. https://doi.org/10.4018/978-1-7998-2466-4.ch030

Vincent, J. (2014, August 21, 2014). 'I Can Haz Islamic State Plz': ISIS propaganda on Twitter turns to kittens and lolspeak. *The Independent*. http://www.independent.co.uk/life-style/gadgets-and-tech/isis-propaganda-on-twitter-turns-to-kittens-and-lolspeak-i-can-haz-islamic-state-plz-9683736.html

Zhou, R., Khemmarat, S., & Gao, L. (2010). The impact of YouTube recommendation system on video views. Proceedings of the 10th ACM SIGCOMM conference on Internet measurement,

## TABLES AND FIGURES

| ***Video ID* in 2016** | **Title** | **Available in 2020** | **Notes** |
|---|---|---|---|
| | | | |
| AQ9b6xmW1ZE | Mujatweets 6 | No | |
| KySHORzZu90 | Mujatweets 2 | No | |
| 9mKmnJotOTM | Islamic Caliphate Mujatweets Raqqah Market | No | |
| toZ6m9IftKI | Mujatweets 3 | No | |
| b1WWFw3zin0 | Mujatweets4 | No | |
| k2MupZqJBu4 | The Religion of Kufr Is One - AlHayat Media Center | No | |
| T2VjhJOQwnc | ISIS VIDEO Flames of War Trailer 720p | No | |
| MSNCPj7s6vI | ISIS - The End of the Sykes-Picot Agreement | No | |
| 5tYIjqWCTKY | No Respite FR MP4 360 | No | French Language |
| Ta75NNb6MQ8 | ISIS發布最新影片 驚見中華民國國旗! | Yes, but only by searching for ISIS 發布最新影片 驚見中華民國國旗 and audio is dubbed | Uncut English-language 'No Respite' Video posted by China Television Corporation; searched for 'No Respite' |
| OPl9gLlAQdY | ISIS - داعش تتحدى العالم | No | Arabic language 'No Respite' |
| | Deterring the Hirelings | No | Not found in 2016 |
| | Deterring the Hirelings 2 | No | Not found in 2016 |
| | Clanging of the swords | No | Not found in 2016 |
| | Upon the prophetic methodology | No | Not found in 2016 |

**Table 1: Seed Videos used for Data Collection; inclusion of Video ID indicates the video was used as a seed video**


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

Table 2: Correlations of seed videos and ISIS videos

| | | | | | dislikes | likes | comments | views | date | recommend | official | category |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Spearman's rho | recommendation | Correlation Coefficient | | | -.056** | -.067** | -.066** | -.078** | .058** | 1.000 | .577** | -.072** |
| | | Sig. (2-tailed) | | | .000 | .000 | .000 | .000 | .000 | . | .000 | .000 |
| | | N | | | 14249 | 14249 | 14249 | 14249 | 14249 | 14249 | 14249 | 14249 |
| | | Bootstrap[c] | Bias | | .000 | .000 | .000 | .000 | .000 | .000 | -.002 | .000 |
| | | | Std. Error | | .006 | .006 | .006 | .006 | .008 | .000 | .057 | .008 |
| | | | 95% Confidence Interval | Lower | -.067 | -.080 | -.077 | -.091 | .042 | 1.000 | .456 | -.089 |
| | | | | Upper | -.044 | -.055 | -.054 | -.065 | .073 | 1.000 | .686 | -.055 |
| | official | Correlation Coefficient | | | -.034** | -.040** | -.038** | -.043** | .039** | .577** | 1.000 | -.052** |
| | | Sig. (2-tailed) | | | .000 | .000 | .000 | .000 | .000 | .000 | . | .000 |
| | | N | | | 14249 | 14249 | 14249 | 14249 | 14249 | 14249 | 14249 | 14249 |
| | | Bootstrap[c] | Bias | | .000 | .000 | .000 | .000 | -.001 | -.002 | .000 | .000 |
| | | | Std. Error | | .006 | .006 | .006 | .006 | .009 | .057 | .000 | .007 |
| | | | 95% Confidence Interval | Lower | -.045 | -.050 | -.050 | -.055 | .019 | .456 | 1.000 | -.064 |
| | | | | Upper | -.022 | -.027 | -.026 | -.030 | .054 | .686 | 1.000 | -.035 |

**. Correlation is significant at the 0.01 level (2-tailed).

c. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

| **Attribute** | **Raw Coverage** | **Unique Coverage** | **Consistency** |
|---|---|---|---|
| | | | |
| radical_keyword | 0.346939 | 0.244898 | 1 |
| newscast*~organization*arabic*recent | 0.0408163 | 0.0204082 | 1 |
| newscast*organization*english*recent | 0.285714 | 0.244898 | 0.82352 |
| explicit_reference*~organization*~arabic*~recent | 0.0612245 | 0.0408164 | 1 |
| ~newscast*explicit_reference*~organization*english | 0.0612245 | 0.0408164 | 1 |
| | | | |
| **solution coverage: 0.693878** | | | |
| **solution consistency: 0.918919** | | | |

Table 3: Optimal QCA truth table; '~' signifies 'Logical Not', where the term is not present and '*' (asterisk) "is used to indicate set intersection"(Ragin, 2005), meaning that newscast*organization*english~arabic indicates videos that are newscasts by organizations in English and not Arabic language

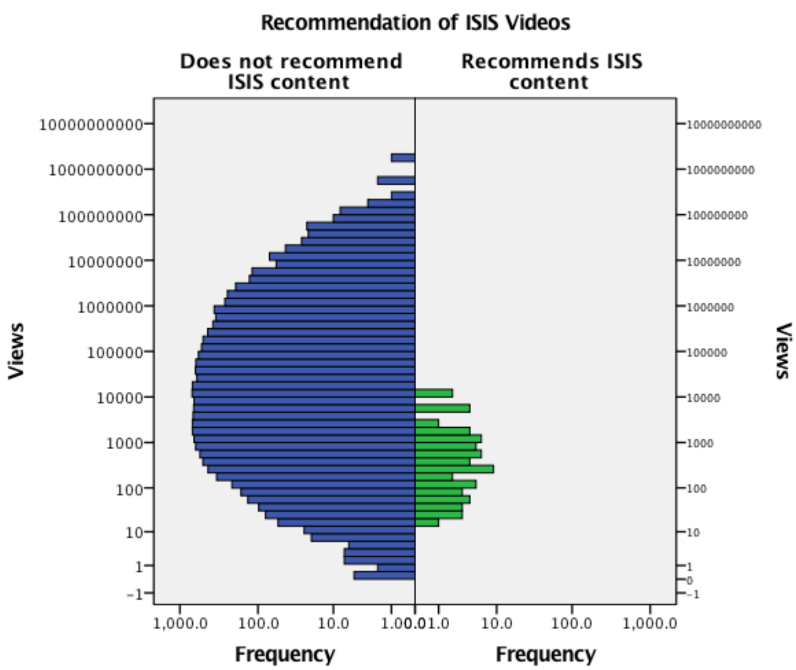


Figure 1: YouTube view count by ISIS recommendation; 'Frequency' indicates number of videos in a view count bin


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

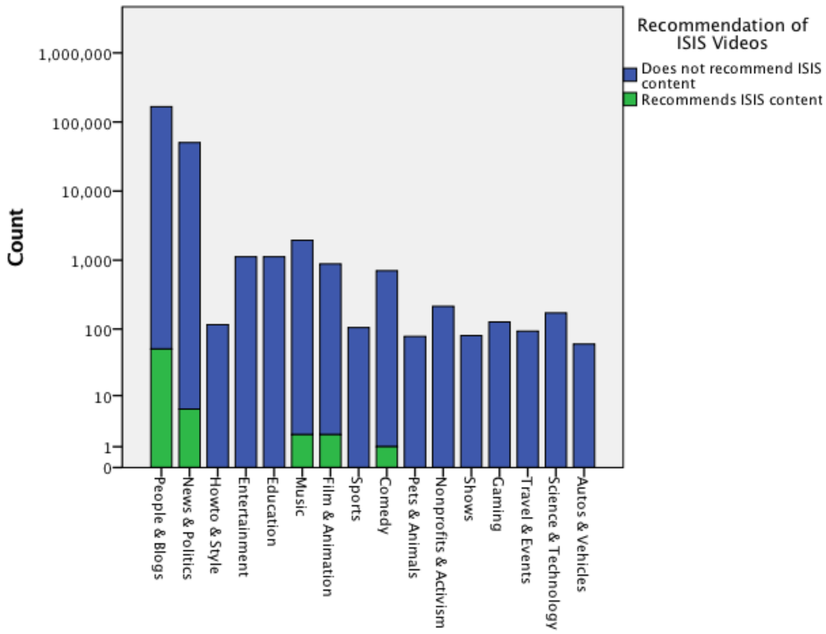


Figure 2: Frequency of videos in our sample by YouTube category; 'Count' indicates number of videos labeled as belonging to each YouTube category

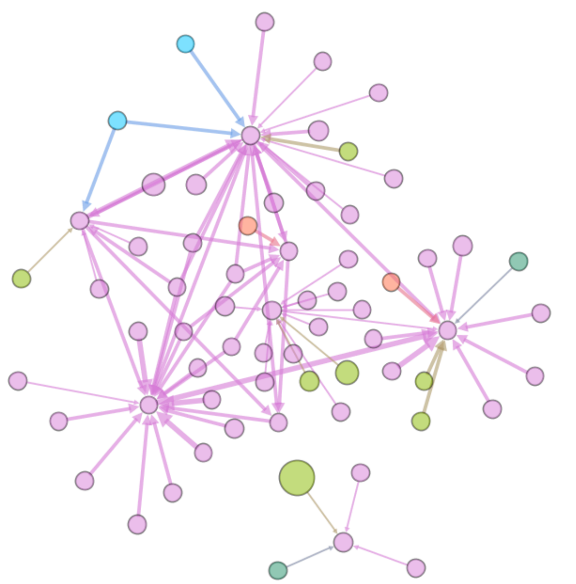

Figure 3: Videos recommending one of the 11 seed videos; node size scaled by views; colors represent language of video - pink = Arabic, green = primarily visual content (mostly *Mujatweets* videos), blue = English, orange = Hindi/Bengali/Telegu, dark green = French; arrows illustrate direction of the recommendations, and thickness of edges indicates edge weight (i.e., a thicker line indicates a greater proportional number of recommendations between the two nodes)


Figure 4: 'No Respite' ISIS video network; black nodes represent ISIS videos and blue nodes do not; node size scaled by views; Arabic titles translated using Google translate; arrows illustrate direction of the recommendations; and thickness of edges indicates edge weight (i.e., a thicker line indicates a greater proportional number of recommendations between the two nodes)

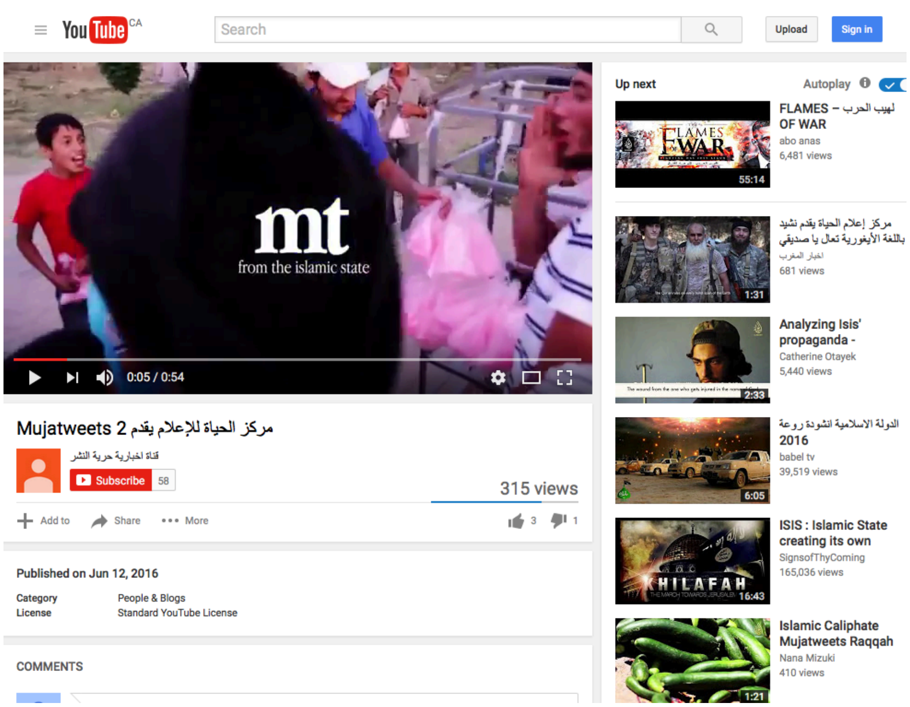


Figure 5: Mujatweets 2 video recommendations

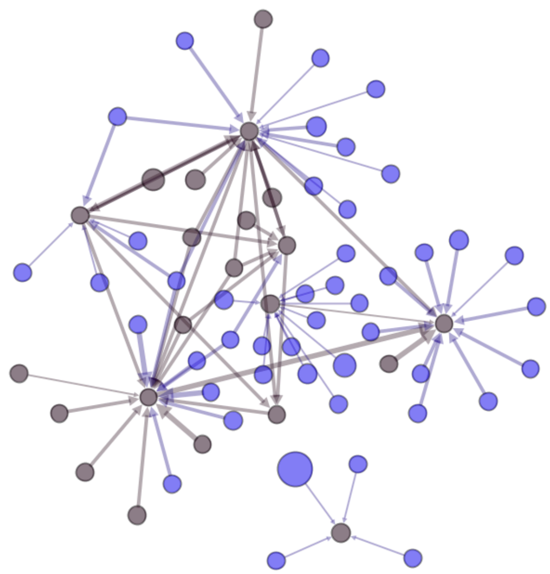

Figure 6: Videos recommending ISIS videos (i.e., videos officially acknowledged by ISIS that contained official branding by an ISIS media wing (e.g., Al Hayat, Al-Furqan, and Ajnad Media); black nodes represent ISIS videos and blue nodes do not; node size scaled by views and arrows illustrate direction of the recommendations; thickness of edges indicates edge weight (i.e., a thicker line indicates a greater proportional number of recommendations between the two nodes)


Version of record published in *American Behavioral Scientist* at https://journals.sagepub.com/doi/full/10.1177/0002764221989774

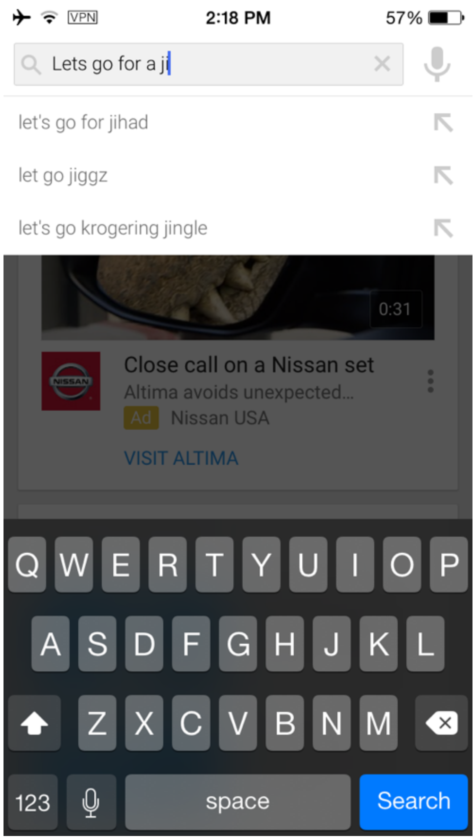


Figure 7: Screenshot illustrating autocomplete suggestions.