Proceedings of the Workshop on Social Threats in Online Conversations: Understanding and Management (STOC-2020), pages 41–47
Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020

Corpus Development for Studying Online Disinformation Campaign: A Narrative + Stance Approach

Mack Blackburn, Ning Yu,
John Berrie, Brian Gordon, David Longfellow, William Tirrell, Mark Williams
Leidos Inc.
4001 Fairfax Dr. Arlington, VA 22203
{ mack.blackburn, ning.yu, john.w.berrie, brian.a.gordon, david.a.longfellow, william.c.tirrell, mark.b.williams}@abc.org

Abstract

Disinformation on social media is impacting our personal life and society. The outbreak of the new coronavirus is the most recent example for which a wealth of disinformation provoked fear, hate, and even social panic. While there are emerging interests in studying how disinformation campaigns form, spread, and influence target audiences, developing disinformation campaign corpora is challenging given the high volume, fast evolution, and wide variation of messages associated with each campaign. Disinformation cannot always be captured by simple factchecking, which makes it even more challenging to validate and create ground truth. This paper presents our approach to develop a corpus for studying disinformation campaigns targeting the White Helmets of Syria. We bypass directly classifying a piece of information as disinformation or not. Instead, we label the narrative and stance of tweets and YouTube comments about White Helmets. Narratives is defined as a recurring statement that is used to express a point of view. Stance is a high-level point of view on a topic. We demonstrate that narrative and stance together can provide a dynamic method for real world users, e.g., intelligence analysts, to quickly identify and counter disinformation campaigns based on their knowledge at the time.

Keywords: disinformation, narrative extraction, corpus development

1. Introduction

In this paper, we present our strategies for data collection and annotation to support studies of online disinformation spread within and across information platforms, mainly using a use case of disinformation campaigns towards the Syrian Civil Defense Force (a.k.a., White Helmets) on Twitter and YouTube. The biggest challenge in constructing such a corpus is how to annotate disinformation. Disinformation is not always easy to fact check; sometimes it can be a piece of true information used in a misleading context, or a fact that was once true but is now false. We propose to bypass directly labeling a piece of information as disinformation or not, instead labelling the narrative and stance of social media content about the White Helmets. In our context, *Narrative* is defined as a recurring statement that is used to express a point of view on a particular topic. Narratives may be explanations of events, interpretations of the motives of actors, statements which emphasize specific concepts, or other techniques to express a point of view. *Stance* is a point of view on a topic. These points of view should be very high level and should represent a user's attitude (usually for or against) a topic. While stance is the point of view itself, a narrative is a particular idea or claim which supports the stance. We show that narrative and stance together can provide a dynamic method for real world users, e.g., intelligence analysts, to quickly create their own data collection on disinformation campaigns with their context-specific knowledge. We first explain how we developed an on-topic corpus containing tweets, YouTube comments, and supporting data sources, then we discuss our initial efforts and lessons learned in defining, detecting and labeling narratives against a subset of this data. We developed an approach to extract individual narrative elements in a clearly interpretable form, drawing on work from information extraction and computational narratology. We also incorporated technologies such as semantic vector clustering in order to combine narrative elements with different structure but similar meaning. Finally, we briefly explain our ongoing efforts to refine and improve narrative and stance annotation guidelines.

2. Background

2.1 White Helmets disinformation campaign

Russia, in coordination with its allies, has orchestrated a large-scale online misinformation/disinformation campaign to discredit the White Helmets of Syria, who are potential witnesses to war crimes committed by the Assad Regime. Russia uses online social platforms like Twitter and YouTube to undermine the credibility and neutrality of the White Helmets by developing narratives about their association with terrorism (i.e., ISIS), Western governments, and even the black market organ trade. Studying such online disinformation operations may help forecast the impact of future disinformation campaigns and potentially allow early development of counter-message strategies.

2.2 Narrative

Narrative plays an important role in both online and offline environments and has been studied in the fields of literature, communication, marketing, and more recently, computational social science (e.g., Chambers and Jurafsky, 2008; Huhn 2019; Yarlott and Finlayson, 2016). Finlayson and Corman (2013) coined two levels of narratives. Level I narrative is related to event discourse: "a report of a sequence of actions or events that are locally coherent and connected, with clear chains of cause and effect concerning a set of agents and their goals and motivations." Most computational work on narrative focuses on Level I, and so does our narrative annotation. Level II narrative is related to action discourse and follows comprehensive narrative structure that adds things narratologists are concerned with such as use of metaphors and cultural tropes. This is an area of interest for future research.

Chambers and Jurafsky (2008) made early attempts to automatically extract, associate, and order narrative event chains from news articles. They parsed the raw text to extract narrative events tuples about a central actor. For each document, verb pairs linked by common entities are narrative events that make up a narrative chain, and each story can contain multiple narrative chains. Chambers and Jurafsky also ordered and clustered events in the same narrative chain. Miller (2018) discussed computational approaches, e.g., event extraction, to narrative detection. Our approach is closest to these computational narrative analyses.

Past research has often been conducted using lengthy documents such as news articles. There are fewer studies of narratives in the microblog space, where narratives can be generated by groups of users via communication with short messages and/or multi-media. The latter format is sometimes called "small stories" (Georgakopoulou, 2014). Stories created in this way may be contained entirely within one tweet, collaboratively constructed by multiple participants, or sequentially created by a single user across multiple tweets (Dayter, 2015; Georgakopoulou, 2014, 2016). Our work aims to extract elements of narratives and the narratives themselves from tweets and YouTube comments.

3. Data collection

3.1 Keyword Driven Data Gathering

Working with subject matter experts in information operation, we first created a list of keywords (e.g., "syria civil defense"), Hashtags (e.g., "#SyriaHoax") and Twitter accounts (e.g., @RT) that are related to online discussions of the White Helmets and/or from disinformation sources, in both English and Arabic. Querying this list through Gnip Historical PowerTrack API[1] against the period of April 2018 to April 2019 returns a total of 1.2 million tweets. The same keywords were also used to query YouTube Search API[2] and gathered information of 1,461 related YouTube videos and 631 channels. We downloaded basic video information such as title, as well as statistics composed of view, likes, dislikes, favorite and comment counts, comments, replies, and captions. To facilitate research of cross-platform information spread, we also get all tweets that refers to YouTube videos.

One drawback of keyword-based data gathering are the false positives due to use of keywords in a context different from the target one. For example, occasionally, White Helmets may be used in a sports context. We took a semi-automatic approach to address this challenge. On one hand, in search queries we reinforce the correct context word and add negative rules for known false context words, e.g., -(scooter OR bike OR bicycle OR football); on the other, we run topic modeling to identify clusters of false positive messages.

3.2 Privacy Protection

We identify personally identifiable information fields in our data (e.g., user ids, emails) and either remove or anonymize such information. Both our data gathering and anonymization strategies have been approved by our corporate security office and, in some cases, by online service providers (e.g., Twitter). For example, mentioning of a twitter user name in a Tweet "RT @SyriaCivilDef" will be anonymized as "RT @ iAo-MokhyIPkTNyhXbuJmQ."

While protecting personal privacy, we also try not to void data of analytic value by enabling researchers to link anonymized information. For example, URLs are anonymized by sections to allow matching at different levels: youtube.com/anonymizedA/anonymizedB will still partially match youtube.com/anonymizedA/anonymizedC while this similarity will vanish if URLs are anonymized as one single string.

3.3 Data Enrichment

We extend the data fields returned by data APIs to include information that may facilitate understanding of disinformation spread. Some enrichment examples are as follows.

Named entities are extracted using tools developed specifically for Twitter data (Ritter, 2011). It can help researchers focus on the mention of particular type of entity, e.g., location or person.

Segmented hashtags are hashtags separated into individual words, e.g., #SupportWH to "Support" and "WH" (Maddela, Xu, & Preotiuc, 2019). Hashtags are important in spreading information and in carrying crucial information across social networking and microblog platforms. Segmenting and analyzing hashtags reveal information contained in each and thus enable accurate hashtag alignment.

Sentiment is labeled at the message level using TweetMotif[3], which provides means for researchers to investigate the impact of sentiment on information spread.

User alignment provides a probability score in terms of how likely two accounts on different platforms belong to the same person. At this point, this is simply calculated by the string similarity of username (before they are anonymized) using the Levenshtein distance. This enrichment enables researchers to not only track information across platform, but also across multiple usernames belonging to the same user.

External references are pages linked from tweets. They either complete the information in the tweet or provide context for the tweet.

For Arabic messages, we also provide English translations using Google translate. For the rest of the paper focusing on narrative labelling, we are going to consider English data only as it is easier to interpret the results than Arabic data when it comes to narrative labeling.

4. Narrative Labelling

In our White Helmets data there are many narratives related to White Helmets, e.g., they are related to terrorist groups,

[1]https://developer.twitter.com/en/docs/tweets/batch-historical/overview

[2]https://developers.google.com/youtube/v3/docs/search/list

[3]https://github.com/ntietz/tweetment

they staged an attack at a particular location, or that they are saving lives. Many of them are not easily verified. Others rely on misleading information, have logical leaps, or are purely statements of opinion. Although researchers may rely on information sources as one factor to judge if a piece of information is disinformation or not, we cannot simply assume certain information sources will always spread disinformation about White Helmets because even propaganda sites share a mix of true and false information. Practically, we cannot manually label millions of messages either. In the rest of this section, we will present our data exploration with LDA to gain a sense of the topic space, then present alternative approaches to test to what extent automatic approaches can help us with narrative labeling.

4.1 Data Exploration with LDA

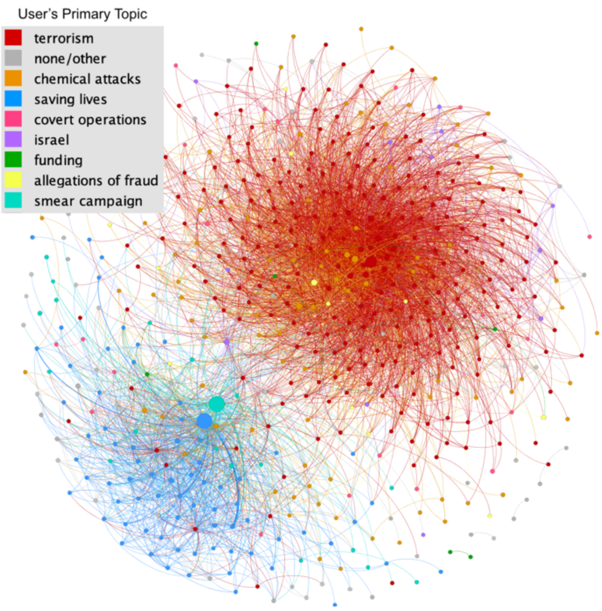

Figure 1: Retweet and quote network for the top users ranked by PageRank. Users are colored by the predominant keyword group in their tweets.

We ran LDA topic modeling (Blei, Ng, & Jordan 2003) on our data and expected to use topics to approximate narratives. However, sometimes LDA topics are less coherent, combine multiple distinct elements, or represent a semantic unit that cannot be interpreted as a narrative. For example, one LDA topic is represented by the words: *russia*, *assad*, *russian*, *regime*, *disinformation*, *claims*, *kremlin*, *target*, *crimes*, and *conspiracy*. This topic is useful in that it provides some insights on distributionally related content in the dataset. However, it would be a mistake to assume that all texts assigned high confidence for this topic share the same narrative. For example, below are two tweets that are assigned high confidence for this topic:

- "The White Helmets are eyewitnesses 193 to war crimes carried out by the Assad regime (which is backed by Russia)..."
- "How Twitter #disinformation is spread by a combination of Assad apologists, Kremlin bots, dupes and paid propagandists"

Although both tweets mention Assad and Russia aligned keywords, they have a distinctly different meaning.

To clean up the LDA results, we hand-selected narrative related keywords from both LDA output and common terms used in top tweets for each topic. Then we manually grouped these keywords into semantically similar sets. As a result, we got 57 sets (narratives). When mapping such keyword sets onto a retweet network of the most influential users, Figure 1 shows a visibly polarized network where users on each side talk about similar narratives, whether supporting or opposing the White Helmets. This suggests that there are possible disinformation campaigns (upper right) and counter campaigns (lower left) visible in our data.

Overall, several patterns became evident during exploration of the dataset: 1) Many authors and sources can collaboratively construct a narrative that is distributed across disconnected texts; 2) The bias or stance of the narrative toward some issue or entity is often the most important component. As a matter of fact, Lehnert (1981) stated that emotional states are the building blocks for a narrative text; and 3) A narrative can often be summarized with a single statement of fact or opinion, e.g., "rescuers."

4.2 Narrative Extraction Experiment

Next, we developed several narrative extraction methods, ran them against a 50,000 tweets subset of the White Helmets dataset, and tested them against the 500 most retweeted tweets manually annotated with one of the 57 narratives identified as described in section 4.1 (See Appendix A).

4.2.1 Model Selection

While event extraction systems such as McClosky et al., 2011; Reschke, et al., 2014; and Wang, 2018 may be effective at extracting narrative events, one additional consideration is that understanding of a narrative requires extracting elements which cannot be classified as an event in the sense of a change of state. For example, relationships between characters, or attributes assigned to characters in a story may be essential to understanding the narrative as intended. However, these kinds of narrative elements may be extracted using methods from open information extraction (OpenIE).

OpenIE systems are designed to extract relations between noun phrases (Mausam, 2016). Many OpenIE systems use a combination of dependency parsing and learned patterns (Mausam and Etzioni, 2012; Wu and Weld, 2010). While some IE systems only extract binary relations, expressions in natural language may also involve more than two noun phrases, or exactly one. Some OpenIE systems have already explored n-ary relations (Christensen and Etzioni, 2011; Pal and Mausam, 2016). Others have also utilized clustering of both noun phrases and relations in order to reduce semantic redundancy (Vashishth and Talukdar, 2018). We incorporate ideas from OpenIE into our verb phrase clustering algorithm, most notably n-ary relations and clustering of embeddings.

Verb Phrase Clustering (VerbPC): vectors are generated for each of the unique verb phrases extracted using dependency parsing, and those are clustered into 100 groups using agglomerative clustering. The number of clusters was fixed at 100 in order to have a fair comparison with LDA and NMF, which also had 100 clusters. An example of an extracted verb phrase: {'verb': ['stage'], 'nsubjpass': ['chemical', 'attack'], 'agent': ['militant']} for "Chemical attack staged by militants."

Ngram Clustering (NgramC1, NgramC2): Scikit-Learn is used to extract 1-3grams from all texts. The FastText vectors of each ngram are clustered using agglomerative clustering. We evaluated with two separate versions of this model: 1) number of clusters fixed at 100 (NgramC2), and 2) distance parameter of agglomerative clustering was set at 1.5 and the number of clusters was induced (NgramC1).

For comparison, we also tested the topic modeling algorithms **LDA** and **NMF** (non-negative matrix factorization) with tf/idf using 100 clusters. For both methods, each text is represented by a binary vector showing the topic with the highest confidence. Additionally, we used the naive approach of bag of words (**BoW**) vectors, limited to the top 500 most common 1-3 grams.

To evaluate each model, the output was fed to a K Nearest Neighbors classifier, and their precision and F1 were recorded in Table 1.

Method	F1	Precision
NgramC1	0.33	**0.71**
NgramC2	0.35	0.66
VerbPC	0.35	0.60
Baselines		
BoW	0.28	0.60
LDA	0.36	0.57
NMF	0.38	0.64

Table 1: KNN Classification Results on Narrative Extraction Methods

4.2.2 Results and Discussions

The most precise algorithm is clustering of ngrams. Verb phrase clustering was more effective than LDA and BoW, but was less effective than NMF and ngram clustering. This may suggest that one approach forward would be to extract text units from the documents that are smaller and more common across texts than verb phrases, but would still convey more of a coherent meaning than ngrams alone. Phrase mining systems such as (Liu et al., 2015), which can extract high-quality readable phrases, may be effective here. While all embedding algorithms here used FastText and cosine distance for agglomerative clustering, incorporating more sophisticated semantic distance measurements may be more effective in the future.

4.3 Supervised Approach

Given that none of the fully automatic narrative extraction approaches we examined in section 4.2 yield results that are good enough to be used as ground truth and there is still more research to be done on this topic. Hence, we are pursuing in parallel a supervised learning approach, which requires more training data.

Here are the steps we plan to take to create the annotation set. Starting with the full data:

Twitter:

- Remove all texts that have fewer than 200 retweets
- Sample of unique texts randomly, weighted by # of times occurring in corpus, random ordering
- Final annotation set is 10,000 tweets

YouTube:

- Randomly Sample of unique texts, weighted by # of times occurring in corpus, with random ordering
- In the annotation set, the number of texts from YouTube should be proportional to the number of relevant YouTube texts in all unique text values.
- Final annotation set is length 10,000 * ((number of YouTube texts matching relevance query)/(number of unique texts))

After generating the annotation candidates, we asked 9 annotators for two annotation tasks: stance and narrative. We assigned a few small batches (30-60 pieces of text) to all annotators in order to see their agreement scores and make changes to the annotation guidelines if necessary. Once all annotators had completed 120 messages, we split the rest of the data into separate batches. Each annotator annotated 100 messages by themselves and then 100 together. Periodically we calculated the inter-annotator reliability by Fleiss' kappa to determine if we need to give them more guidance or modify the guidelines.

Once we have all the training data annotated, it will be used to train several supervised multi-classification systems with text representations from simple tf/idf vector to multilingual BERT or FastText with pretrained Spanish embeddings.

5. Conclusion and Future Work

In this paper, we have demonstrated our end-to-end effort in developing a corpus for studying disinformation campaigns across platforms. We focused on the most challenging annotation tasks and discussed our early exploration of an automatic approach to extract elements of narratives on microblogs. While our approach shows promising results, we still have a long way to go in terms of accurately generating ground truth data. Our future plans are two-fold: First, we will continue our focus on optimizing narrative event extraction as well as linking these events into narratives by taking full advantage of microblog attributes. Secondly, we will continue to improve our annotation guidelines and processes and start to explore a supervised approach.

6. Ethical Considerations

There are some privacy concerns related to the work we discussed here: disclosing social media users' personal point of view without their explicit consent (Fiesler and Proferes, 2018), and the risk of wrongly associating users with disinformation spread activities during our manual or automatic labeling process. To mitigate those risks, we anonymize our data, reach agreement with each social company regarding out data collection and anonymization plan, and strictly follow IRB and private guidance provided by the research program. We also only allow researchers who have completed DARPA privacy training and meet all privacy compliance requirements to access data.

7. Acknowledgements

We thank our anonymous reviewers for their valuable feedback. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) and the Army Research Office (ARO) under Contract No. W911NF-17-C-0095. The content of the information in this document does not necessarily reflect the position or the policy of the U.S. Government, and no official endorsement should be inferred.

8. Bibliographical References

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), pp. 993-1022.

Chambers, N., & Jurafsky, D. (2008). Unsupervised learning of narrative event chains. In Proceedings of ACL-08: HLT. Association for Computational Linguistics, pp. 789–797.

Christensen, J., Soderland, S., & Etzioni, O. (2011). An analysis of open information extraction based on semantic role labeling. In Proceedings of the Sixth International Conference on Knowledge Capture, pp. 113-120.

Dayter, D. (2015). Small stories and extended narratives on twitter. Discourse, Context and Media, 10, pp. 19–26.

Fiesler, C., & Proferes, N. (2018). "Participant" perceptions of twitter research ethics. SocialMedia+ Society, 4(1):2056305118763366

Finlayson, M. A., & Corman, S. R. (2013). The military interest in narrative. Sprache und Datenverarbeitung, 37(1-2), pp. 173-191.

Georgakopoulou, A. (2014). Between narrative analysis and narrative inquiry: The long story of small stories research. Urban Language and Literacies, 131, pp. 1-17.

Hühn, P., et al. (eds.): The living handbook of narratology. Hamburg: Hamburg University. http://www.lhn.uni-hamburg.de/

Pal, H., & Mausam (2016). Demonyms and compound relational nouns in nominal open IE. In Proceedings of the 5th Workshop on Automated Knowledge Base Construction, pp. 35-39.

Lehnert, W. G. (1981). Plot units and narrative summarization. Cognitive science, 5(4), pp. 293-331.

Liu, J., Shang, J., Wang, C., Ren, X., & Han, J. (2015). Mining quality phrases from massive text corpora. In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, pp. 1729-1744.

Maddela, M., Xu, W., & Preoţiuc-Pietro, D. (2019). Multi-task pairwise neural ranking for hashtag segmentation. arXiv preprint arXiv:1906.00790.

McClosky, D., Surdeanu, M., & Manning, C. D. (2011). Event extraction as dependency parsing for bionlp 2011. In Proceedings of the BioNLP Shared Task 2011 Workshop. Association for Computational Linguistics, pp. 41-45.

Miller, B. (2018). Cross-document narrative alignment of environmental news: A position paper on the challenge of using event chains to proxy narrative features. In Proceedings of the Workshop Events and Stories in the News 2018, pp. 18-24.

Ritter, A., Clark, S., & Etzioni, O. (2011). Named entity recognition in tweets: an experimental study. In Proceedings of the conference on empirical methods in natural language processing. Association for Computational Linguistics, pp. 1524-1534.

Reschke, K., Jankowiak, M., Surdeanu, M., Manning, C. D., & Jurafsky, D. (2014). Event extraction using distant supervision. In Proceedings of the International Conference on Language Resources and Evaluation Conference (LREC), pp4527-4531.

Vashishth, S., Jain, P., & Talukdar, P. (2018). Cesi: Canonicalizing open knowledge bases using embeddings and side information. In Proceedings of the 2018 World Wide Web Conference, pp. 1317-1327.

Wang, W. (2018). Event detection and extraction from news articles (Doctoral dissertation, Virginia Tech).

Wu, F., & Weld, D. S. (2010). Open information extraction using Wikipedia. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, pp. 118-127.

Yarlott, W. V. H., & Finlayson, M. A. (2016). Learning a better motif index: Toward automated motif extraction. In 7th Workshop on Computational Models of Narrative (CMN 2016).

9. Appendix A: 57 White Helmets Narratives

Narrative Tag	Description
israel_evac_wh	discusses event of Israel evacuating WH
wh_save_lives	author believes WH save peoples lives
wh_in_danger	WH are under threat or are deliberately targeted by Assad military, WH need to be rescued
wh_fake_evidence	WH stage videos or photos, or otherwise provide fake evidence
us_funding_freeze	discussion of event of US freezing WH funding
wh_terrorists	WH are linked to terrorists, help

	terrorists, or share facilities/resources with terrorists
uk_asylum	Discussion of event of UK providing asylum to WH members
uk_connection	WH connection to the UK by funding, policy, or official statements
us_connection	WH connection to the US by funding, policy, or official statements
russia_assad_connection	Russia and Assad act as a coordinated axis (negative)
civilian_casualties	Deaths of civilians during military actions by Russia or Assad gov
anti_wh_smear_campaign	WH are being targeted by misinfo or smear campaign
wh_propaganda	WH are propaganda tools or make propaganda
western_connection	WH are connected to "the west", NATO, or the EU, or are favored by "western" entities
wh_used_to_promote_regime_change	WH are a tool used to promote a "regime change" agenda
netherlands_funding_freeze	Discussion of Netherlands freezing funding for WH
russia_opposes_wh	Russia opposes the WH either in general or in a way distinct from a smear campaign
israel_connection	WH connection to Israel or "zionism" by funding, policy, or official statements
wh_participate_in_execution	WH participate in, are present during, or clean up after executions
wh_committed_war_crimes	WH committed mass murder or otherwise broke international law in violent ways
wh_not_legitimate	WH are a "fake" or "illegitimate" group or are contrasted with Assad government affiliated groups

media_favor_wh	Media outlets are biased in favor of wh, are complicit in falsifying evidence, or refuse to convey anti-WH information
wh_win_oscar	Discusses WH winning Oscar or refers to them as "oscar-winning"
wh_evac	Discussion of wh evacuation in general without mentioning Israel
wh_asylum	wh will be provided asylum or resettled in nonspecific country
wh_not_helpful	WH do not help civilians or do not accomplish what they claim
assad_war_crimes	Assad military actions are mass murder or other very violent acts
canada_asylum	WH are provided with asylum in Canada
covert_ops	WH are involved in covert operations or are secretly affiliated with foreign military or intelligence agencies
wh_foreign_influence	WH are affiliated with governments or organizations foreign to Syria, which makes them illegitimate.
germany_asylum	WH will be provided asylum in Germany
roger_waters_emails	Discussion of emails sent to Roger Waters requesting he endorse the WH, and his statements after that
wh_illegal_acts	WH engage in other heinous acts such as kidnapping, drugging people, or mishandling dead bodies
wh_organs	WH are organ traffickers or harvest organs of dead or living people
wh_document_crimes	WH provide video or photo evidence of war crimes by Russia or Assad
misinformation	Discussion of mis/disinformation

	or fake news as opposed to "smear campaign" which makes no claims of misinformation
official_hearings	Official hearings on the WH at the UN or at the Hague
france_connection	WH connected to French government
chemical_weapons	Discusses use of chemical weapons by Assad or Russia
censorship	Claims of censorship by YouTube, twitter for anti-WH statements
wh_weapons	Claims WH have weapons such as guns or bombs
elie_wiesel_award	WH win Elie Wiesel award
exposing_truth	Vague general statements about exposing lies or truth
wh_member_deaths	Statements paying respect to dead members of the WH
nobel_prize	WH nomination for nobel peace prize
events_pro_assad	General descriptions of events from a pro-Assad stance
anti_wh_campaign_interests_conspiracists	States that the anti-WH campaign is generally aligned with other conspiracy theories
critique_israel	Criticizes other Israeli actions in Gaza, etc.
general_anti_wh	Generally negative toward WH without clarification
jo_cox	Discussion of politician Jo Cox, who supported WH
james_le_mesurier	Discussion of WH founder with ties to UK
wh_threat_to_host	WH are a threat to host countries where they will be relocated
russia_wants_peace	Russia is faced with NATO aggression and is attempting to promote peace
canada_connection	WH is connected to Canadian government
wh_misc_positive	Miscellaneous positive statements or positive discussion of secondary WH programs
qanon	QAnon US politics (deep state, conspiracies, etc)
unrelated	False positive in data collection (e.g. Football team white helmets)