International Journal on Advances in Intelligent Systems, vol 14 no 1 & 2, year 2021, http://www.iariajournals.org/intelligent_systems/

A Topic Modeling Framework to Identify Online Social Media Deviance Patterns

Thomas Marcoux, Esther Mead, Nitin Agarwal
COSMOS Research Center
University of Arkansas at Little Rock
Little Rock, AR, USA
email: {txmarcoux, elmead, nxagarwal}@ualr.edu

***Abstract*—Following the COVID-19 pandemic and the subsequent vaccine related news, the information community has seen the emergence of unique misinformation narratives in a wide array of different online outlets, through social media, blogs, videos, etc. Taking inspiration from previous COVID-19 and misinformation detection related works, we expanded our topic modeling tool. We added filtering capabilities to the tool to adapt to more chaotic social media datasets and create a chronological representation of online text content. We curated a corpus of 543 misinformation pieces whittled down to 243 unique misinformation narratives, and collected two separate sets of 652,120 and 1,664,123 YouTube comments. From our corpus of misinformation stories, this tool has shown to accurately represent the ground truth of COVID misinformation stories. This highlights some of the misinformation narratives unique to the COVID-19 pandemic and provides a quick method to monitor and assess misinformation diffusion, enabling policy makers to identify themes to focus on for communication campaigns. To expand previous publications and further explore the potential of topic streams in understanding online misinformation, we propose a framework used as a filter to help whittle down big data corpora and identify latent misinformation within. This could be scaled and applied to very large social networks to highlight misinformation.**

Keywords-misinformation; disinformation; topic models; topic streams; COVID-19; misinfodemic; narratives.

I. Introduction

Social media is characterized as a powerful online interaction and information exchange medium. However, it has given rise to new forms of deviant behaviors, such as spreading fake news, misinformation, and disinformation. For this reason, we began this research in our previous publication [1] and are now introducing this extended version. Due to afforded anonymity and perceived diminished personal risk of connecting and acting online, deviant groups are becoming increasingly common. Online deviant groups have grown in parallel with Online Social Networks (OSNs), whether it is black hat hackers using Twitter to recruit and arm attackers, announce operational details, coordinate cyber-attacks [2], and post instructional or recruitment videos on YouTube targeting certain demographics; or state/non-state actors and extremist groups (such as the Islamic State of Iraq and Syria) savvy use of social communication platforms to conduct phishing operations, such as viral retweeting of messages containing harmful URLs leading to malware [3].

More recently, there is a surge in misinformation and scam cases pertaining to COVID-19. The problem of misinformation is actually worse than the pandemic itself. That is why it is called infodemic or more specifically, misinfodemic. Like the pandemic, misinformation cases are also rising exponentially. These cases are more difficult to track than the epidemic, as they can originate in the dark corners of the Internet. To make matters worse, we cannot enforce lockdown on the Internet to stop the spread of this infodemic. This is in part because, during crises, the Internet is usually the first mode of communication and source of information. Although there are some quarantine efforts, for instance form social media companies, such as Facebook, YouTube, and retail companies like Amazon are doing their best to block such content, by suspending bad actors or scammers who are spreading misinformation to further their political agenda or to try to profit off of this adversity. But such cases are simply too many and growing too fast. What makes this problem worse is the fact that the information spreads like a wildfire on the Internet, especially the false or misinformation. Many studies have concluded that misinformation travels faster than its corrective information, and the more questionable the misinformation is the faster it travels. This is simply because on social media people usually have a lot more virtual friends than they do in their real life. So, if they share or retweet some misinformation, wittingly or unwittingly, they expose all their virtual friends to the misinformation.

There are similarities between misinformation about COVID-19 and other misinformation cases that we have studied for NATO, US, EU, Singapore, and Canada, etc. Like in other cases, the motivation for spreading COVID-19 misinformation is monetization or to provoke hysteria. Bad actors or scammers are spreading misinformation to further their political agenda or simply trying to profit off of this adversity. For instance, there exists many cases of scammers selling fake masks, fake cures, using fake websites to ask for private/sensitive information from people by posing as government websites. However, there is a significant difference between COVID-19 and other misinformation campaigns that we have studied before. Being a global and rapidly evolving crisis, the nature of misinformation is also extremely diverse and super-fast. Other misinformation campaigns were specific to an entity, event, region, elections, military exercises.

However, misinformation about COVID-19 has both global as well as regional narratives. While fake masks, fake cures, etc., affect a global audience, the regional narratives include promoting medicines for bovine coronavirus as cure for human coronavirus affecting rural/agriculturalist regions. Moreover, the misinformation about COVID-19 ranges from health to policy to religion to geopolitical affairs, i.e., highly topically diverse. Given the volume, velocity, and variety of COVID-19 related misinformation, research is warranted to study such campaigns and their organization. As resources are stretched too thin, government and other regulatory bodies cannot afford to investigate all the misinformation campaigns and scams. Such research could help prioritize investigation of misinformation campaigns and scams.

Therefore, we propose a study of the themes and chronological dynamics of the spreading of misinformation about COVID-19. Our scope focuses on misinformation geographically relevant to us (Arkansas, USA), as well as some global stories, with our main corpus is a collection of unique misinformation stories manually curated by our team. In collaboration with the Arkansas Attorney General, we have shared our findings with their office and made all reports and misinformation stories publicly available online [4]. In addition, we have collected a variety of YouTube video titles and comments. This allows us to compare a curated corpus to a data set more chaotic and true to life. To highlight and visualize these misinformation themes, we use topic modeling, and introduce a tool to visualize the evolution of these themes chronologically.

In addition, to expand our previous work [1], we introduce a manual node-based design to filter very large datasets and identify information of interest within, while avoiding the bais that can come with artificial intelligence methods. This framework is tested with a set of 1,664,123 YouTube comments and is built to introduce further feature detection, such as commenting behavior, or even inorganic video engagement behavior, tackling the issue of multimedia misinformation.

The rest of this study is structured as follows. First, we will discuss the work done by other researchers in comparable research in Section 2, describe our methodology in Section 3, including data collection, processing, and topic modeling methodology. Then, in Section 4, we will discuss our results and the subjective findings of our misinformation team with the scientific topic streams visualizations that support them. Finally, we briefly introduce our free online resource where the misinformation stories used here can be found, before presenting our conclusions in Section 5.

II. Literature Review

In this section, we first argue of the importance of this field as it can directly relate to public safety, followed by the efforts of the research community to combat this issue. We then introduce the significance of the YouTube platform and argue our choice of using YouTube comments for this study, finishing this section with the relevant literature on our primary analysis technique: topic models.

A. The Significance of Misinformation

The information community has been tackling the issue of misinformation surrounding the COVID-19 pandemic since early in the outbreak. We base the claims found in this paper on the findings that misinformation spreads in a viral fashion and that consumers of misinformation tend to fail at recognizing it as such [5]. In addition to this, we believe this research is essential as rampant misinformation constitutes a danger to public safety [6]. We also believe this research is helpful in curbing misinformation since researchers have found that simply recognizing the existence of misinformation and improving our understanding of it can enhance the larger public's ability to recognize misinformation as such [5]. In order to better understand the misinformation surrounding the pandemic, we look at previous research that has leveraged topic models to understand online discussions surrounding this crisis. Research has shown the benefits of using this technique to understand fluctuating Twitter narratives [7] over time, and also in understanding the significance of media outlets in health communications [8]. Studies on information propagation [9] establish entire mathematical models around the diffusion of misinformation and emphasize that early detection is essential to allow a proper response.

B. Misinformation Detection

Because of the severity of the threat of misinformation campaigns and the need to quickly discover such efforts, we concern ourselves with detection models to help systematically recognize inorganic or concerted information operations. Because misinformation spreads so quickly and deals long lasting damage, we consider developing scalable models to quickly identifying misinformation a critically important endeavor. Of course, because of the severity of this public issue, there are a great many efforts within the information community striving to propose solutions. The state of the art in fake news detection could be roughly described as being divided between three main ideas. One is artificial intelligence models, where researchers will use traditional machine learning techniques [10, 11], multinomial Bayesian models [12, 13], or deep-learning [14, 15, 16]. Another school of thought in misinformation detection leveraging natural language processing processing technique. Some researchers, for example, focus on text features and experiment with natural language processing techniques, such as sentiment analysis [17]. The authors of this publication propose the use of this extra dimension as a source of auxiliary features. Finally, an emerging technique is the use of a combination of the previous two [18, 19].

While proponents of natural language processing point out that deep learning models tend to produce inexplicable black boxes that may lead to biased outputs [14], which is sometimes echoed by proponents of machine learning [18], the same researchers [18] rightfully point out that the bag-of-words nature of topic models impedes such methods from capturing features based on the sequential ordering of words. This is a weakness of note and why topic models should not be used alone when attempting to systematically detect misinformation, especially considering the more difficult to detect subtle pieces of misinformation. The authors also classified misinformation detection methods as belonging to either traditional machine learning models, topic models, or deep learning models.

Researchers agree that the fake news detection problem is a complex one and has not yet seen a perfectly appropriate solution.

Some approaches attempt to model claims as binary true or false and run into issues of representing further nuance and complexity. For this reason, we will steer our research to rely on a score and focus on detecting suspicious or inorganic behavior rather than real or fake claims. Other works [16] use multi-platform datasets and attempt to model complex information structures by classifying claims between specific categories (here: "fake news", "news bias", "rumors", and "clickbait"), and rely on annotations to build predictive models based on headline linguistic features, achieving an average effectiveness of 70.27%. Some researchers address the issue with classifying "realness" by representing both certainty and uncertainty [14] and accounting for user response and engagement. The authors found promising results and, as many other studies did [13], encouraged the use of wider arrays of features when attempting to detect social media misinformation. Researchers [14] also correctly point out many challenges of fake news detection. Such as multilingualism when relying on textual approaches, which has some researchers relying on meta-data or networking only approaches. Particularly challenging and effective misinformation also includes items which featured subtly inserted falsehood or half truths. The Multimedia nature of misinformation is another challenge.

Others use wide and deep models [18], relying on memorizing and generalizing information, which somewhat inspired our natural language processing based contribution, to advance interpretability and reduce unknown bias. These researchers also propose a framework model combining multiple design principles and detection methods. Although this particular study uses datasets of a slightly different nature: deceptive reviews and fraudulent emails

Using a self-constructed twitter dataset of 1,300 entries, researchers have been able to achieve an impressive near-real time 93% accuracy in detecting misinformation [15]. Twitter being a very prized source of data for such studies due to the wide array of metadata available [20]. One concern however is how scalability and ability to detect a very wide range of misinformation may become a hurdle for this model as it could detect merely dubious information. As opposed to our approach, these researchers ignored textual content and focused on networking and linguistic features. In contrast, other authors [21] found 49.2% accuracy with a much larger dataset of 34,918 claims. These claims were crawled from fact checking websites and include metadata, such as the creator of the misinformation, the checker, etc. This approach is more suited to predict performances for fact checking websites.

C. The Role of YouTube

From third party public resource and web traffic reports [22], we know that YouTube is the second most popular website, ceding the first spot to Google, and accounts for 20.4% of all search traffic. According to official YouTube sources [23], 1 billion hours of videos are watched each day. Another study by Cha et al. [24] found that 60% of YouTube videos are watched at least 10 times on the day they are posted. The authors also highlight that if a video does not attract viewership in the first few days after upload, it is unlikely to attract viewership later on. YouTube provides an overwhelming amount of streaming data: over 500 hours of videos are uploaded every minute on average. A number which was "only" 300 in 2013 [25]. In previous publications [26, 27] we identified YouTube as a potential vehicle of misinformation. We proposed the use of YouTube metadata for understanding and visualizing these phenomena by observing data trends. We also proposed the concept of movie barcodes as a tool for video summarization clustering [28]. In this publication, we present the movie barcode tool as a part of VTracker, as well as new video characterization tools. Previous research [29] has looked into engagement patterns of YouTube videos and highlighted the related videos engagement trends, later designated as the "rabbit hole effect" where users will be recommended increasingly relevant videos. In some cases, where the subject matter is a very polarizing one, this effect has been shown to be a contributing factor in user radicalization [30]. This last study takes the example of vaccine misinformation, which has attracted much interest from the information community. With some research highlighting that while users turn to YouTube for health information, many of the resources available failed to provide accurate information [31, 32], and public institutions should increase their online presence [33] to make reliable information more accessible. Recent research on the same subject leverages advanced NLP techniques on text entities, such as video comments [34] but we could find little work available on the video content itself.

D. Topic Modeling

To implement topic modeling, we use the Latent Dirichlet Allocation (LDA) model. Within the realm of Natural Language Processing (NLP), topic modeling is a statistical technique designed to categorize a set of documents within a number of abstract "topics" [35]. A "topic" is defined as a set of words outlining a general underlying theme. For each document, which in this case, is an individual item of misinformation in our data set, a probability is assigned that designates its "belongingness" to a certain topic. In this study, we use the popular LDA topic model due to its widespread use and proved performances [36]. One point of debate within the topic modeling community is the elimination of stop-words: i.e., analysts should filter common words from their corpus before training a model. Following recent research claiming that the use of custom stop-words adds few benefits [37], we followed the researchers' recommendation and removed common words **after** the model had been trained.

Our model choice has seen use in previous research using LDA for short texts, specifically for short social media texts, such as tweets [38, 39, 40]. Some other social media research using homogeneous social media sources, such as tweets or blog posts use associated hashtags to provide further context to topic models [41]. We expand this research on social media corpora by focusing one of the largest information propagator on the web: YouTube.

In this paper, we propose to leverage topic models to understand the main underlying themes of misinformation and their evolution over time using a manually curated corpus of known fake narratives.

As a secondary goal, we observe the performances of different topic models for understanding online discourse. To accomplish this, we repeated our methodology on a secondary data set using a Hierarchical Dirichlet Process (HDP) model [42]. For our purposes, the major difference between the two models is that LDA models require a number of topics prior to training and will actively attempt to fit that number to the corpus, potentially leading to biased results. On the other hand, the HDP model infers the number of topics present in the corpus during training.

III. Methodology

This study uses a two-step methodology to produce relevant topic streams. First, through a manual curating process, we aggregate different misinformation narratives for later processing. We consider misinformation narratives, any narrative pushed through a variety of outlets (social media, radio, physical mail, etc.) that has been or is later believably disproved by a third party. This corpus constitutes our input data. Secondly, we use this corpus to train an LDA topic model and to generate subsequent topic streams for analysis. We describe these two steps in more details in the next sections.

A. Collection of Misinformation Stories

This is the set referred to as **Dataset-1**. Initially, the misinformation stories in our data set were obtained from a publicly available database created by EUvsDisinfo in March of 2020 [43]. EUvsDisinfo's database, however, was primarily focused on "pro-Kremlin disinformation efforts on the novel coronavirus". Most of these items represented false narratives that were communicating political, military, and healthcare conspiracy theories in an attempt to sow confusion, distrust, and public discord. Subsequently, misinformation stories were continually gleaned from publicly available aggregators, such as POLITIFACT, Truth or Fiction, FactCheck.org, POLYGRAPH.info, Snopes, Full Fact, AP Fact Check, Poynter, and Hoax-Slayer. The following data points were collected for each misinformation item: title, summary, debunking date, debunking source, misinformation source(s), theme, and dissemination platform(s). The time period of our data set is from January 22, 2020 to July 22, 2020, which is the COVID-19 breakout period. The data set is comprised of 543 total stories and 243 unique misinformation narratives. For many of the items, multiple platforms were used to spread the misinformation. For example, oftentimes a misinformation item will be posted on Facebook, Twitter, YouTube, and as an article on a website. For our data set, the top platforms used for spreading misinformation were websites, Facebook, Twitter, YouTube, and Instagram, respectively. All the stories found by our team are made public through our partnership with the Arkansas Attorney General Office and can be found on our website.

B. Collection of YouTube Data

In order to observe results in uncontrolled, relevant social media environments, we also gathered YouTube data. We chose YouTube because it is a principal vector of information and communication between users and is heavily understudied. Using the official YouTube API, we performed separate searches for the following keywords on April 19th 2020: "Coronavirus, Corona, Virus, COVID19, COVID, Outbreak". The result is a set of the most popular videos at that time, as determined by YouTube's algorithm. From this search, we collected a total of 7,727 videos ranging mostly from January 1st to April 19th 2020. For this particular study, in order to focus on the most relevant videos possible, we selected only videos published between March 1st and March 31st (included). Like the previous set, this is a key month of the COVID-19 breakout period. This totals 444 videos, which is comparable to the number of narratives studied. For the purposes of this study, we will only look at the video titles. After selecting this corpus, we used the same API to collect comments posted in these videos and gathered a total of 652,120 comments. This is **Dataset-2**.

Based on a manual qualitative analysis of known alt-right public figures active on social media, a set of specific actors was identified and selected as seeds for preliminary data collection. YouTube data for our set of key actors was collected using the YouTube Data API according to the methodology described by Kready et al. [44]. During post-processing, the dataset was filtered to focus in on the two months prior and post the January 6, 2021 U.S. Capitol riot event, resulting in a timeframe of analysis of November 1, 2020 to March 1, 2021. We chose this period because that is where most discussion revolving around vaccines can be found. This is **Dataset-3**. In order to comply with YouTube's terms of service, this data cannot be made public.

C. Topic Modeling

In order to derive lexical meaning from this corpus, we built a pipeline executing the following steps. First, we processed each document in our text corpus. All that is needed is a text field identified by a date. Because in most cases of word of mouth or social media it is impossible to pinpoint the exact date the idea first emerged, we use the date of publication of the corresponding third party "debunk piece". We trained our LDA model using the Python tool Gensim, with the methodology and pre-processing best practices as described by its author [45] as well as best stop words practices as described earlier [37]. In this study, we found that generating

20 different topics best matched the ground truth as reported by the researchers curating the misinformation stories.

Still using Gensim, we also trained an alternative topic model using HDP [42]. The process is the same except for the number of topics. HDP infers the number of topics in a corpus (with a default threshold of 150). Therefore, we only select the first 20 topics, ordered by α, the weight of each document to topic distribution.

Once the models have been trained, we ordered the documents by date and created a numpy matrix where each document is given a score for each topic produced by the model. This score describes the probability that the given document is categorized as being part of a topic, i.e., if a probability score is high enough (more details below), the document is considered to be part of the topic. Through manual observations, we noticed that many documents retain ”noise probability”, giving them a probability to be in every topic of around 1% to 5%. For this reason, we set the probability threshold to a comfortable 10% and noticed consistent results. This allowed us to leverage the Python Pandas library to plot a chronological graph for each individual topic. We averaged topic distribution per day and used a moving average window size of 20 unless otherwise specified. This helped in highlighting the overarching patterns of the different narratives. Note, however, that this process hides some early and late data in our set as there are less data points around that time.

IV. Results

In this section, we discuss the thoughts of our data collection team and the ground truth as they were observed, and compare these with the results obtained through our topic modeling visualization tool.

A. Prominent Misinformation Themes Over Time

Although a variety of misinformation themes were identified, particular dominant themes stood out, changing over time. These themes were considered as dominant based on a simple sum of their frequency of occurrence in our data set. During the month of March, the prominent misinformation theme was the promotion of remedies and techniques to supposedly prevent, treat, or kill the novel coronavirus. During the month of April, the prominent themes still included the promotion of remedies and techniques, but additional prominent themes began to stand out. For example, several misinformation stories attempted to downplay the seriousness of the novel coronavirus. Others discussed the anti-malaria drug hydroxychloroquine. Others promoted the idea that the virus was a hoax meant to defeat President Donald Trump. Others consisted of various attempts to attribute false claims to high-profile people, such as politicians and representatives of health organizations. Also in April, although first signs of these were seen in March, the idea that 5G caused the novel coronavirus began to become more prevalent. During the month of May, the prominent themes shifted to predominantly false claims made by high-profile people, followed by attempts to convince citizens that face masks are either more harmful than not wearing one, or are ineffective at preventing COVID-19, and how to avoid rules that required their use. The number and variety of identity theft phishing scams also increased during May. Misinformation items attempting to attribute false claims to high-profile people continued throughout May. Also becoming prominent in May were misinformation items attempting to spread fear about a potential COVID-19 vaccine, and items promoting the use of hydroxychloroquine. During the month of June, the prominent theme shifted significantly to attempts to convince citizens that face masks are either more harmful than not wearing one, and how to avoid rules that required their use. Phishing scams also remained prominent during June. During the month of July, the dominant themes of the misinformation items shifted back to attempts to downplay the deadliness of the novel coronavirus. Another prominent theme in July was the proliferation of attempts to convince the public that COVID-19 testing is inflating the results.

B. Topic Streams

After using the tool described in Section III-C, we generated the graphs and tables described and discussed in this section. Our data for this step contained 243 unique misinformation narratives spanning from January 2020 to June 2020, when we stopped data collection. The data was curated by our research team through the process described in the methodology. Each entry contains, among other fields, a “date” used as a chronological identifier, a “title” describing the general idea the misinformation is attempting to convey, and a “theme” field putting the story in a concisely described category. For example, a story given the title *“US Department of Defense has a secret biological laboratory in Georgia”* is categorized in the following theme: *“Western countries are likely to be purposeful creators of the new virus.”* Each topic was represented by an identification number up to 20 and a set of 10 words. We picked the three most relevant words that best represented the general idea of each topic. Notably, obvious words, such as *covid* or *coronavirus* were removed from the topic descriptions since they are common for every topic.

In Tables I and II, we described some of the twenty topics found by each of our LDA models. These topics were chosen because they each described a precise narrative and have a low topic distribution (or proportion within the corpus). A low proportion is desirable because this indicates the detection of a unique narrative within the corpus; as opposed to an overarching topic including general words, such as “world”, “outbreak”, or “pandemic”. Do note that topic inclusiveness is not exclusive and documents can be part of multiple topics. This becomes apparent in Table I: from our topic model, we found a dominant topic encompassing 68% of narratives. It includes words such as “Trump”, “outbreak”, “president”, etc. Some other narratives also included words such as “flu”, “news”, or “fake”. Because the evolution of these narratives are consistent across the corpus and show little temporal fluctuation, we chose not to report on them further. For these reasons, the narratives we focused on below show a low percentage of distribution (Tables I & II).

TABLE I
MOST FREQUENT DOMINANT TOPICS FROM TITLES.

Topic ID	Word 1	Word 2	Word 3	Proportion
10	china	chinese	spread	2%
12	scam	hydroxy...	health	2%
17	state	donald	trump	2%
18	vaccine	gates	bill	5%

TABLE II
MOST FREQUENT DOMINANT TOPICS FROM THEMES.

Topic ID	Word 1	Word 2	Word 3	Proportion
3	fear	spread	western	2%
9	predicted	pandemic	vaccine	2%
16	phishing	hydroxy...	vaccine	2%

*1) Using narrative titles as a corpus - **Dataset-1**:* The general narratives described by the topics were thus:

- Topic 10 described the narratives related to the Chinese government and its responsibility in the spread of the virus. These stories represented an estimated 2% of the 243 stories collected.
- Topic 12 described the narratives related to personal health and scams or misinformation, such as the benefits of hydroxychloroquine. These stories represented an estimated 2% of the 243 stories collected.
- Topic 17 described the narratives related to the response of Donald Trump and his administration. These stories represented an estimated 2% of the 243 stories collected.
- Topic 18 described the narratives related to the involvement of Bill Gates in various conspiracies, mostly linked to vaccines. These stories represented an estimated 4% of the 243 stories collected.

Related studies have found that finger-pointing narratives usually lead to negative sentiment and toxicity in online communities [38, 46, 39].

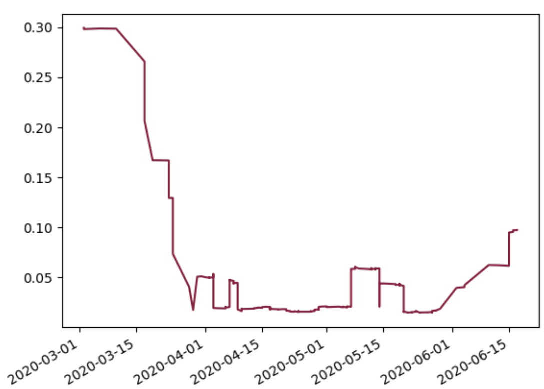

Fig. 1. Topic's probability distribution of titles for topic 10 (keywords: china, chinese, spread) over time (LDA model)

Figure 1 shows the evolution of Topic 10, the topic describing China-related narratives. It shows that these narratives were already in full force from the beginning of our corpus and slowly came to a near halt during the month of April. We notice a short spike again towards the end of the corpus during the month of June. This is consistent with the ground truth of online narratives that focused on the provenance of the virus during the early stages.

Figure 2 shows the evolution of Topic 12, the topic describing narratives related to health, home remedies, and general hoaxes and scams stemming from the panic. We can see it was consistent with the rise of cases in the United States and panic increased as with the spread of the virus. It is interesting to note that this figure roughly coincides with the daily number of confirmed cases for this time period [47].

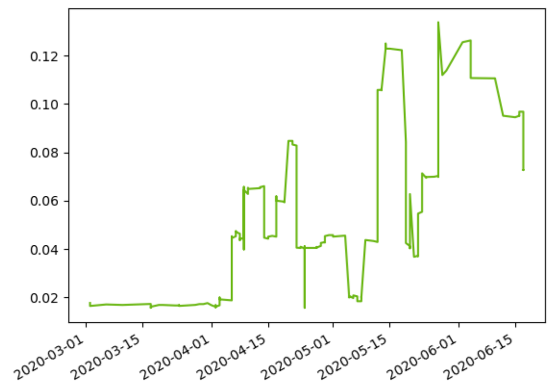

Fig. 2. Topic's probability distribution of titles for topic 12 (keywords: hydroxychloroquine, health, scam) over time (LDA model)

Figure 3 shows the evolution of Topic 17. This topic described stories related to Donald Trump and his administration. These stories generally referred to claims that the virus was manufactured as a political strategy, or claims that various public figures were speaking out against the response of the Trump administration.

Figure 4 shows the evolution of Topic 18. This topic described stories such as Bill Gates and his perceived involvement with a hypothetical vaccine, and other theories describing the virus' appearance and spread as an orchestrated effort. As with Figure 1, these narratives were especially strong early on (albeit this narrative remained active for a slightly longer time), before coming to a near halt.

We notice that, as theories about the origins of the virus slowed down, hoaxes and scams increased - as shown on Figure 2. This includes attempts at identity theft, especially toward senior citizens, and attempts to sell miracle cures and miracle personal protection items.

2) Using narrative themes as a corpus: For this section, we inputted narrative themes as the corpus. Note that the topic IDs are independent from the previous set of topics using titles. Similarly to Section IV-B1, we found a dominant topic

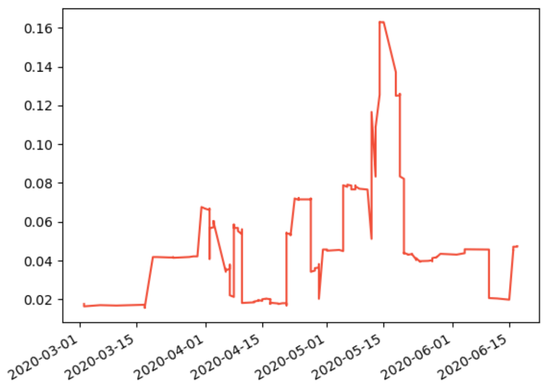

Fig. 3. Topic's probability distribution of titles for topic 17 (keywords: donald, trump, state) over time (LDA model)

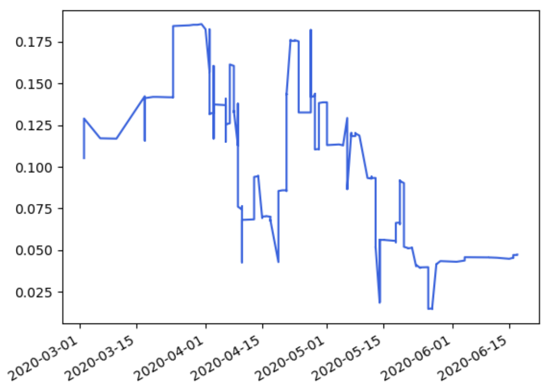

Fig. 4. Topic's probability distribution of titles for topic 18 (keywords: bill, gates, vaccine) over time (LDA model)

encompassing 68% of narratives as well. This time including words such as "attempt", "countries", and "purposeful". As for section IV-B1, we chose not to report on that topic as well as other smaller but general topics showing little fluctuation. Therefore, the narratives we focused on below show a low percentage of distribution. The general narratives described by the topics are thus:

- Topic 3 described the narratives related to the speculations on the spread of the virus, especially in an international relations context. These stories represented an estimated 2% of the 243 stories collected.
- Topic 9 described the narratives related to stories claiming the creation and propagation of the virus were either designed or predicted, along with voices claiming a vaccine already exists. These stories represented an estimated 3% of the 243 stories collected.
- Topic 16 described the narratives related to personal health and scams or misinformation such as the benefits of hydroxychloroquine. These stories represented an estimated 2% of the 243 stories collected.

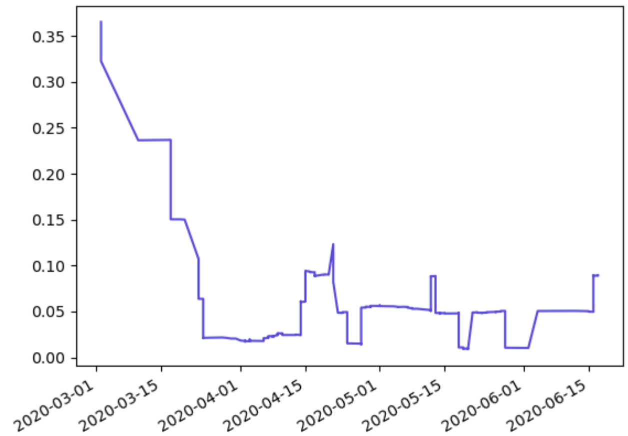

Fig. 5. Topic's probability distribution of themes for topic 3 (keywords: fear, spread, western) over time (LDA model)

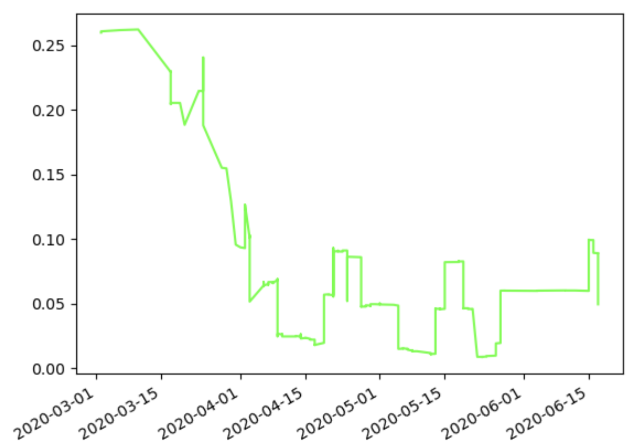

Fig. 6. Topic's probability distribution of themes for topic 9 (keywords: predicted, pandemic, vaccine) over time (LDA model)

Figure 5 shows the evolution of Topic 3. It is linked to early fear of the virus and presented narratives as opposing the western block with the East, notably China. It matched closely with Figure 1 and its China-related narratives. In both cases, we see an early dominance of the topic followed by a near halt as the virus touched the United States.

Figure 6 describes the evolution of narratives claiming the virus was predicted or even designed. This figure is consistent with the results shown by Figure 4 which shows claims regarding Bill Gates, early vaccines, etc. They both showed stories of early knowledge of the virus and peaked early,

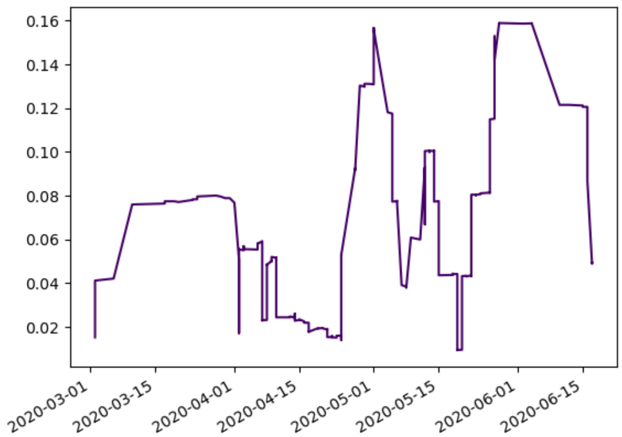

Fig. 7. Topic's probability distribution of themes for topic 16 (keywords: hydroxychloroquine, vaccine, phishing) over time (LDA model)

appearing more or less sporadically as time goes on and as cases increased.

Figure 7 is parallel to Figure 2. Both showed hoax stories promoting scams and health-related misinformation. We noticed an early rise in Figure 7, most likely due to the inclusion of the keyword "vaccines" in the topic, which caused some overlap with Topic 9 as shown in Figure 6.

C. YouTube Data

In this section, we explore how different topic models affect our YouTube data set. We focus on a subset of data published during the month of March to limit the number of comments to process.

1) YouTube videos - ***Dataset-2****:* The first observation for this set is that our HDP model did not perform as well as the LDA model. Our HDP model identified one dominant topic present in 87% of videos, with seemingly unrelated identifying keywords ("cases", "hindi", "nyc", "italy"). While the rest of the topics are present in around 1% of the videos. The second most dominant topic (1.8% of documents) also features contradicting words such as "plandemic" and "hospitals". One would expect language connected to the plandemic narrative in this topic, such as mentions of "Bill Gates" like we saw in the previous sets, but it is missing. There are two possible explanations for this. One is that performance may be due to the size of the set (more in the next section) as there were only 444 video titles processed. The other is that the set features numerous multilingual titles, which may skew results.

Our LDA model, however, behaved as expected and was able to identify major topics, mostly news videos (Topics 0 & 17), as well as what we suspect to be a vehicle of misinformation (Topic 6). As described in Table III and visualized in Figure 8. Figure 8 has been smoothed with a moving average equal to 15% of the total data set size (67) in order to improve legibility and reveal patterns. Due to most of the videos being published late in March, this has removed some granularity towards early March from the plot. However, we notice news topics staying fairly consistent while Topic 6 sees a decline, possibly as the number of covid cases makes maintaining the "fake pandemic" narrative more difficult and other misinformation narratives take over, such as various scams and hoaxes as seen in section IV-B1.

TABLE III
RELEVANT TOPICS FROM VIDEO TITLES (LDA MODEL)

Topic ID	Word 1	Word 2	Word 3	Proportion
0	news	update	live	12.4%
17	outbreak	doctor	cases	7.6%
6	plandemic	dempanic	dem	2.7%

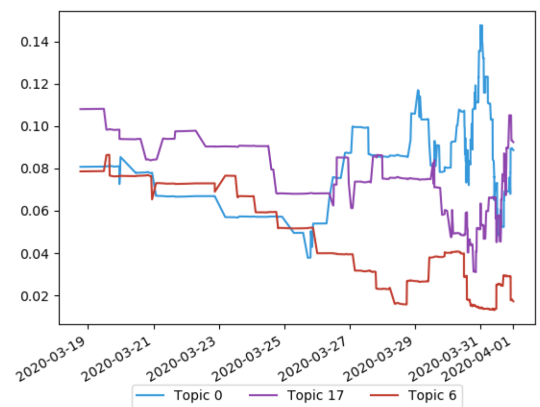

Fig. 8. Topic's probability distribution of topics 0, 17 & 6 over time (LDA model)

2) YouTube comments - ***Dataset-2****:* Contrary to the previous section, this is a much larger data set of 652,120 comments. This led to better performances, but still inferior to the LDA model. Our HDP model was able to identify non-English comments (11.4% German, 4.5% Spanish, 1.6% French). More importantly, the HDP model identified a topic that could be described as polarizing discourse, some of the most frequent terms including "Trump", "China', and "virus". This topic accounts for 6.6% of the corpus. The evolution of this topic is shown by Figure 9 where we notice that topic is on an upward trend. A moving average equal to 3% of the set size is applied to better identify patterns.

On this very large set, our HDP model somewhat outperformed LDA for our purposes as it was able to identify a probable topic for misinformation. When applied to our comments set, our LDA model mostly found general terms while also successfully isolating non-English comments. The model did identify a topic with some toxic language and some that could be used in a hostile way or communicate sinophobic sentiments (Topic 7 & 17). See Table IV. While discussion of China has so far been on a downward trend since the start

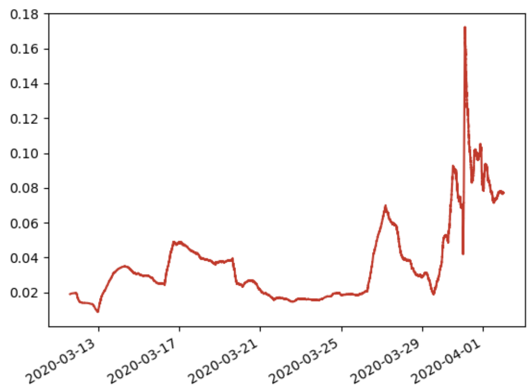

Fig. 9. Topic's probability distribution of Topic 4 over time (HDP model)

of the pandemic, the mention of the term "virus" along with "china" suggests toxic behavior. See Figure 10.

TABLE IV
RELEVANT TOPICS FROM FROM DATASET-2 COMMENTS (LDA MODEL).

Topic ID	**Word 1**	**Word 2**	**Word 3**	**Proportion**
7	china	virus	made	3.5%
17	trump	dumb	bats	3.3%

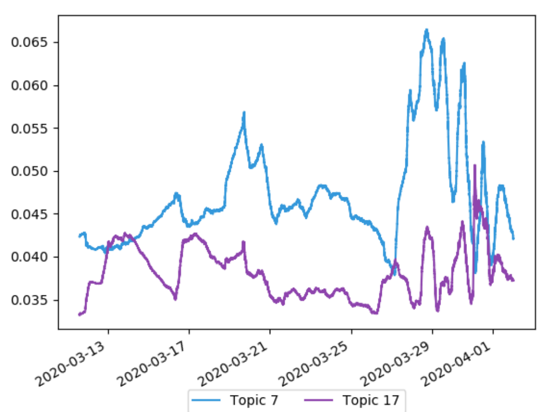

Fig. 10. Topic's probability distribution of Topic 7 & 17 over time (LDA model)

*3) YouTube comments - **Dataset-3**:* This larger set of 1,664,123 comments comes from efforts relating to content liked with the January 6, 2021 U.S. Capitol riot [48]. Due to its larger size, this set is our test bed for our new Pipeline Framework.

As is illustrated in Figure 11, this architecture is a node-based system where the framework first reads raw data, then have each node ingest filtered or annotated data from the previous one. These nodes can be chained in any order but, in this study, we demonstrate what could be labelled as the data filtering layer. As was suggested in our previous publication [1], we are now using the more objective HDP model to divide a corpus into topics and then identify which topic to filter and send to our LDA model to identify latent narratives.

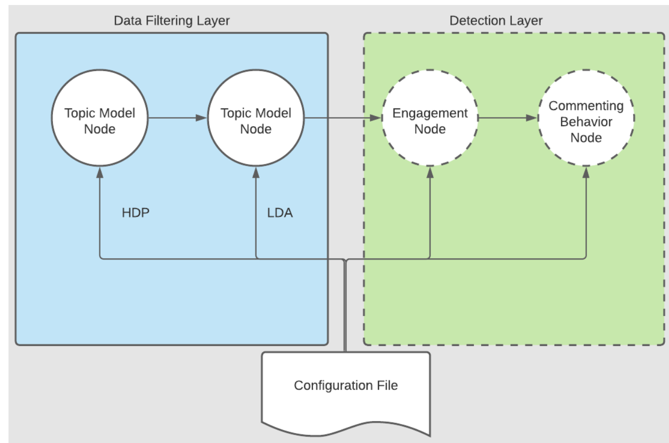

Fig. 11. Pipeline Framework

TABLE V
RELEVANT TOPICS FROM FROM DATASET-3 COMMENTS (HDP MODEL).

Topic ID	**Word 1**	**Word 2**	**Word 3**	**Word 4**	**Proportion**
3	gender	women	men	man	8.2%
2	covid	vaccine	even	know	6%
5	trump	ben	think	biden	3.1%

From Table V, which shows some of the most relevant words from the 20 topics we retained (in order of prominence within the dataset), we notice that Topic 2 is especially relevant to our subject at hand. For this reason, the comments belonging (where "belongingness" is characterized by a probability superior to 0.3 of belonging to a given topic) to that topic are sent to the next node where our LDA model is then retrained on these comments. The resulting main topics of interest and their descriptive keywords are described in Table VI.

TABLE VI
RELEVANT TOPICS FROM FROM DATASET-3 COMMENTS (LDA MODEL)

Topic ID	**Word 0**	**Word 1**	**Word 2**	**Word 3**	**Proportion**
15	leftist	welcome	tears	change	5%
8	rumble	back	joined	parler	4.3%
1	trump	address	back	party	2.9%

Table VI and its temporal visualizations tell give us the following insight: From the keywords described in Topic 15, there seems to be a celebration of some event perceived as a victory over the opposing party. This event is represented within the graph in Figure 12 by a very obvious peak.

Topic 8 shown on Figure 13 aggregates keywords discussing other apps focused on free speech and anonymity. Interestingly, this type of speech has seen a very big revival shortly

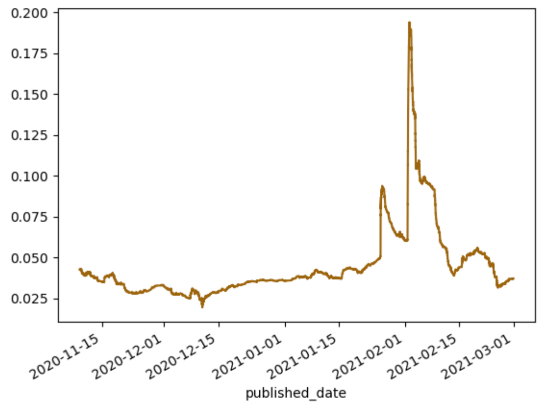

Fig. 12. Topic's probability distribution of Topic 15 over time (LDA model)

before the events on January 6th, and then another spike directly after with periodic movement following. This may suggest some level of organization or at least a desire to move away from mainstream platforms that could have been a factor in the Capitol riots.

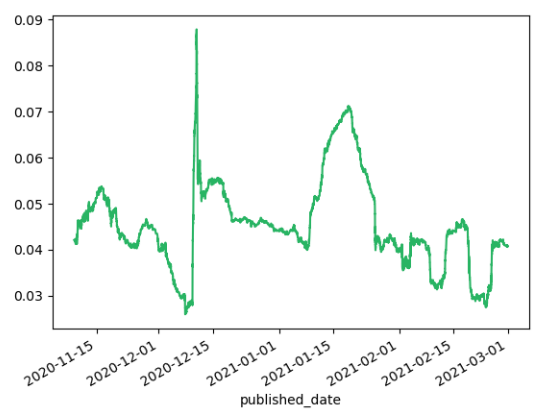

Fig. 13. Topic's probability distribution of Topic 8 over time (LDA model)

Finally, Topic 1 shown on Figure 14 shows discourse surrounding Donald Trump and his appearances. Unsurprisingly, the popularity of this topic has been on the decline since the 2020 presidential elections and then saw a revival around the January 6th riots. We also notice some periodicity.

Chaining topic models to help filter larger data sets has shown good results that are explainable by real world events and is a promising start to further enrich our framework for deviant behavior detection. Unlike deep learning networks, every node and features is strictly defined, reducing risk for bias. Of course, one limitation of such method becomes the bias of human experts designing features and also the risk of models becoming outdated. To address these weak points, we

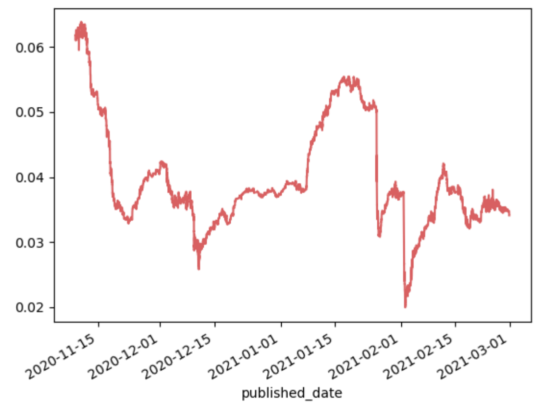

Fig. 14. Topic's probability distribution of Topic 1 over time (LDA model)

will further expand the pipeline to accept fully modular and interchangeable nodes.

D. Future Works

As shown in Figure 11, our framework will be appended with more nodes whose goal is to annotate and "detect" misinformation by providing score based on commenting behavior as well as engagement behavior in the source video of the comment. This is one way to tackle multimedia misinformation as video misinformation has presented a significant challenge, and threat, especially due to the popularity of such video content. The design of the framework aims to allow for chaining nodes in any order, and one other goal will be to automate this process to obtain and measure the most accurate results, but also to let researchers contribute their own nodes.

E. Public Website and Citizen Science

We have put together a website with known cases of misinformation about COVID-19. As of January 2021, we have documented close to 600 cases that we identified from numerous sources (social media - Facebook, YouTube, Twitter, blogs, fake websites, robocalls, text/SMS, WhatsApp, Telegram, and an array of such apps) - see Figure 15 [4]. The principal difference between our effort and other similar efforts by Google and social media companies is that we are paying special attention to cases of misinformation and scammers that are affecting our region, while also including global cases. We update the database periodically with newly detected cases. Moreover, we have put together a list of over 50 tips on the website for people to learn how to spot misinformation. We have also provided a feature for people to report fake websites or scams that are not currently in our database.

Our website uses a three-pronged approach:

- We identify new cases of fake websites, misinformation content, and bad actors. We use social network analysis and cyber forensic methodologies to identify such cases.
- We believe in educating people to be self-reliant because we might not be able to detect all possible cases of

Fig. 15. COVID-19 Website Front page - Showing the latest misinformation stories

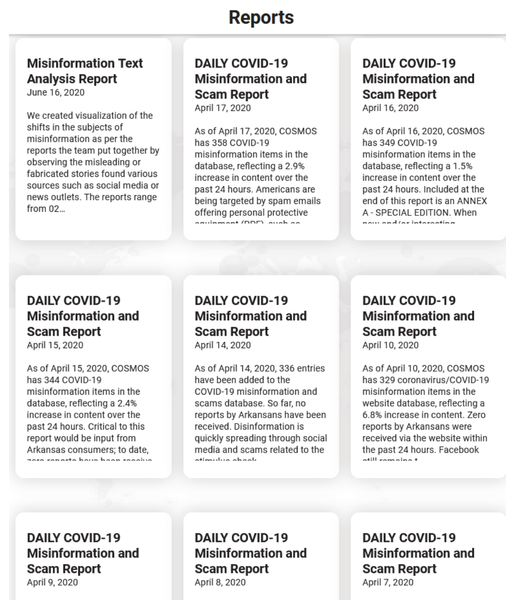

Fig. 16. COVID-19 Website Reports page - Showing all reports made to the Arkansas Attorney General Office

misinformation. Therefore, we go through identified cases and prepare a list of common telltale signs to detect whether a piece of information is genuine or not.

- For the cases that are not in our database and people cannot distinguish, we provide a way for people to submit cases of misinformation that we have not captured in our database.

The database of known misinformation cases and scams is publicly available for the research community to use [4]. We envision a tremendous value of this research database to various disciplines. The website is available for regulatory bodies (Arkansas Office of the Attorney General) and any citizen, which serves as an invaluable resource to not only educate people of the misinformation and scams about COVID-19 but also assisting legal authorities in taking action against malicious actors and groups. We are assisting the Arkansas' Attorney General's office by providing reports on cyber forensic evidence about scam/fake websites reported by people - see Figure 16 . The study presented in this paper will be developed into the system as a real-time campaign tracking feature. We will continue to work with Arkansas' Attorney General's office to assist in their effort to combat COVID-19 misinformation and scams to protect Arkansans.

V. Conclusion

In this study which expands our last publication [1], we have highlighted some of the narratives that surfaced during the COVID-19 pandemic. From January 2020 to July 2020, we collected 243 unique misinformation narratives and proposed a tool to observe their evolution. We have shown the potential of using topic modeling visualization to get a bird's eye view of the fluctuating narratives and an ability to quickly gain a better understanding of the evolution of individual stories. We have seen that the tool is efficient to chronologically represent actual narratives pushed to various outlets, as confirmed by the ground truth observed by our misinformation curating team and independent international organizations. Working with the Arkansas Office of the Attorney General, this study illustrates a relatively quick technique for allowing policy makers to monitor and assess the diffusion of misinformation on online social networks in real-time, which will enable them to take a proactive approach in crafting important theme-based communication campaigns to their respective citizen constituents. We have made most of our findings available online to support this effort.

In addition to these results, we have introduced much larget datasets, one of 652,120 YouTube comments, and another of 1,664,123 more comments. To accomodate these sets, we introduce a new node-based framework which functions as a pipeline where nodes can be intercheangeably used to filter and annotate documents. At this current stage, the framework supports topic model nodes based on the LDA and HDP model. By feeding into our LDA model documents belonging to specific topics as identified by our HDP model, we are able to focus on specific communities of interest and reveal latent patterns and events within those communities. The future of this tool is in the addition of more nodes that will examine wider features, such as commenting behavior and engagement behavior with videos and channels where comments are posted to detect suspicious behavior.

Acknowledgement

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920, IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2675, N00014-17-1-2605, N68335-19-C-0359, N00014-19-1-2336, N68335-20-C-0540), U.S. Air Force Research Lab, U.S. Army Research Office (W911NF-20-1-0262, W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059), Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

References

[1] T. Marcoux, E. Mead, and N. Agarwal. "Studying the Dynamics of COVID-19 Misinformation Themes". en. In: *2021 Seventh International Conference on Human and Social Analytics (HUSO)*. Nice, France, July 2021.

[2] S. Al-khateeb et al. *Exploring Deviant Hacker Networks (DHM) on Social Media Platforms*. Vol. 11. Journal of Digital Forensics, Security and Law, 2016, pp. 7–20. DOI: 10.15394/jdfsl.2016.1375. URL: https://commons.erau.edu/jdfsl/vol11/iss2/1.

[3] M. Calabresi. *Inside Russia's Social Media War on America.* 2017. URL: https://time.com/4783932/inside-russia-social-media-war-america/. (accessed: 01.19.2021).

[4] COSMOS. *COSMOS - COVID-19 Misinformation Tracker*. 2021. URL: https://cosmos.ualr.edu/covid-19. (accessed: 06.15.2021).

[5] G. Pennycook et al. "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention". In: vol. 31. 7. _eprint: https://doi.org/10.1177/0956797620939054. 2020, pp. 770–780. DOI: 10.1177/0956797620939054. URL: https://doi.org/10.1177/0956797620939054.

[6] R. Kouzy et al. "Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter". eng. In: vol. 12. 3. Publisher: Cureus. Mar. 2020, e7255–e7255. DOI: 10.7759/cureus.7255. URL: https://pubmed.ncbi.nlm.nih.gov/32292669.

[7] H. Sha et al. *Dynamic topic modeling of the COVID-19 Twitter narrative among U.S. governors and cabinet executives*. 2020. arXiv: 2004.11692 [cs.SI].

[8] Q. Liu et al. "Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach". In: vol. 22. 4. Apr. 2020, e19118. DOI: 10.2196/19118. URL: http://www.jmir.org/2020/4/e19118/.

[9] H. Zhang et al. "Detecting misinformation in online social networks before it is too late". In: *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 2016, pp. 541–548. DOI: 10.1109/ASONAM.2016.7752288.

[10] B. Al Asaad and M. Erascu. "A Tool for Fake News Detection". In: *2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*. 2018, pp. 379–386. DOI: 10.1109/SYNASC.2018.00064.

[11] S. Lyu and D. C.-T. Lo. "Fake News Detection by Decision Tree". In: *2020 SoutheastCon*. 2020, pp. 1–2. DOI: 10.1109/SoutheastCon44009.2020.9249688.

[12] A. Jain and A. Kasbe. "Fake News Detection". In: *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. 2018, pp. 1–5. DOI: 10.1109/SCEECS.2018.8546944.

[13] S. Yu and D. Lo. "Disinformation Detection Using Passive Aggressive Algorithms". In: *Proceedings of the 2020 ACM Southeast Conference*. ACM SE '20. event-place: Tampa, FL, USA. New York, NY, USA: Association for Computing Machinery, 2020, pp. 324–325. ISBN: 978-1-4503-7105-6. DOI: 10.1145/3374135.3385324. URL: https://doi.org/10.1145/3374135.3385324.

[14] Q. Zhang et al. "Reply-Aided Detection of Misinformation via Bayesian Deep Learning". In: Feb. 2019. DOI: 10.1145/3308558.3313718.

[15] L. van de Guchte et al. "Near Real-Time Detection of Misinformation on Online Social Networks". In: *Disinformation in Open Online Media*. Ed. by M. van Duijn et al. Cham: Springer International Publishing, 2020, pp. 246–260. ISBN: 978-3-030-61841-4.

[16] N. Lee et al. "On Unifying Misinformation Detection". In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Online: Association for Computational Linguistics, June 2021, pp. 5479–5485. DOI: 10.18653/v1/2021.naacl-main.432. URL: https://aclanthology.org/2021.naacl-main.432.

[17] M. Alonso Pardo et al. "Sentiment Analysis for Fake News Detection". In: *Electronics* 10 (June 2021), p. 1348. DOI: 10.3390/electronics10111348.

[18] Y. Chai et al. "Disinformation Detection in Online Social Media: An Interpretable Wide and Deep Model". en. In: *SSRN Electronic Journal* (2021). ISSN: 1556-5068. DOI: 10.2139/ssrn.3879632. URL: https://www.ssrn.com/abstract=3879632 (visited on 09/14/2021).

[19] K. Pelrine, J. Danovitch, and R. Rabbany. "The Surprising Performance of Simple Baselines for Misinformation Detection". In: *Proceedings of the Web Conference 2021*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 3432–3441. ISBN: 978-1-4503-8312-7. URL: https://doi.org/10.1145/3442381.3450111.

[20] F. Pierri, C. Piccardi, and S. Ceri. "A multi-layer approach to disinformation detection in US and Italian news spreading on Twitter". In: *EPJ Data Science* 9.1 (Nov. 2020), p. 35. ISSN: 2193-1127. DOI: 10.1140/epjds/s13688-020-00253-8. URL: https://doi.org/10.1140/epjds/s13688-020-00253-8.

[21] I. Augenstein et al. "MultiFC: A Real-World Multi-Domain Dataset for Evidence-Based Fact Checking of Claims". In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 4685–4697. DOI: 10.18653/v1/D19-1475. URL: https://aclanthology.org/D19-1475.

[22] *Youtube.com Traffic, Demographics and Competitors - Alexa*. URL: https://www.alexa.com/siteinfo/youtube.com (visited on 09/15/2021).

[23] *How YouTube Works - Product Features, Responsibility, & Impact*. URL: https://www.youtube.com/intl/en-GB/howyoutubeworks/ (visited on 09/15/2021).

[24] M. Cha et al. "Analyzing the Video Popularity Characteristics of Large-Scale User Generated Content Systems". In: *IEEE/ACM Transactions on Networking* 17.5 (Oct. 2009), pp. 1357–1370. ISSN: 1558-2566. DOI: 10.1109/TNET.2008.2011358.

[25] J. Hale. *More Than 500 Hours Of Content Are Now Being Uploaded To YouTube Every Minute*. May 2019. URL: https:

//www.tubefilter.com/2019/05/07/number-hours-video-uploaded-to-youtube-per-minute/ (visited on 06/13/2021).

[26] M. N. Hussain et al. "Analyzing Disinformation and Crowd Manipulation Tactics on YouTube". In: *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. ASONAM '18. event-place: Barcelona, Spain. IEEE Press, 2018, pp. 1092–1095. ISBN: 978-1-5386-6051-5.

[27] T. Marcoux et al. "Understanding Information Operations Using YouTubeTracker". In: *IEEE/WIC/ACM International Conference on Web Intelligence - Companion Volume*. WI '19 Companion. Thessaloniki, Greece: Association for Computing Machinery, 2019, pp. 309–313. ISBN: 9781450369886. DOI: 10.1145/3358695.3360917. URL: https://doi.org/10.1145/3358695.3360917.

[28] R. Erol et al. "YouTube Video Categorization Using Moviebarcode". en. In: *The Sixth International Conference on Human and Social Analytics (HUSO 2020)*. Porto, Portugal, 2020.

[29] X. Cheng, C. Dale, and J. Liu. "Statistics and Social Network of YouTube Videos". In: *2008 16th Interntional Workshop on Quality of Service*. 2008, pp. 229–238. DOI: 10.1109/IWQOS.2008.32.

[30] L. Tang et al. ""Down the Rabbit Hole" of Vaccine Misinformation on YouTube: Network Exposure Study". In: *J Med Internet Res* 23.1 (Jan. 2021), e23262. ISSN: 1438-8871. DOI: 10.2196/23262. URL: http://www.ncbi.nlm.nih.gov/pubmed/33399543.

[31] C. H. Basch et al. "Preventive Behaviors Conveyed on YouTube to Mitigate Transmission of COVID-19: Cross-Sectional Study". In: *JMIR Public Health Surveill* 6.2 (Apr. 2020), e18807. ISSN: 2369-2960. DOI: 10.2196/18807. URL: http://www.ncbi.nlm.nih.gov/pubmed/32240096.

[32] H. O.-Y. Li et al. "YouTube as a source of information on COVID-19: a pandemic of misinformation?" In: *BMJ Global Health* 5.5 (May 2020), e002604. DOI: 10.1136/bmjgh-2020-002604. URL: http://gh.bmj.com/content/5/5/e002604.abstract.

[33] G. Donzelli et al. "Misinformation on vaccination: A quantitative analysis of YouTube videos". In: *Human Vaccines & Immunotherapeutics* 14.7 (2018). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/21645515.2018.1454572, pp. 1654–1659. DOI: 10.1080/21645515.2018.1454572. URL: https://doi.org/10.1080/21645515.2018.1454572.

[34] J. C. Medina Serrano, O. Papakyriakopoulos, and S. Hegelich. "NLP-based Feature Extraction for the Detection of COVID-19 Misinformation Videos on YouTube". In: *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*. Online: Association for Computational Linguistics, July 2020. URL: https://www.aclweb.org/anthology/2020.nlpcovid19-acl.17.

[35] D. M. Blei, J. D. Lafferty, and A. N. Srivastava. *Text Mining: Classification, Clustering, and Applications*. CRC Press, 2009, pp. 71–88.

[36] D. M. Blei, A. Y. Ng, and M. I. Jordan. "Latent dirichlet allocation". In: vol. 3. Journal of Machine Learning Research, 2003, pp. 993–1022.

[37] A. Schofield, M. Magnusson, and D. Mimno. "Pulling Out the Stops: Rethinking Stopword Removal for Topic Models". In: *15th Conference of the European Chapter of the Association for Computational Linguistics*. Vol. 2. Association for Computational Linguistics. 2017, pp. 432–436.

[38] A. Abd-Alrazaq et al. "Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study". In: vol. 22. 4. 2020, e19016. DOI: 10.2196/19016. URL: http://www.ncbi.nlm.nih.gov/pubmed/32287039.

[39] R. Chandrasekaran et al. "Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study". In: vol. 22. 10. Oct. 2020, e22624. DOI: 10.2196/22624. URL: http://www.jmir.org/2020/10/e22624/.

[40] Y. Zhang, W. Mao, and J. Lin. "Modeling Topic Evolution in Social Media Short Texts". In: *2017 IEEE International Conference on Big Knowledge (ICBK)*. 2017, pp. 315–319.

[41] M. H. Alam, W.-J. Ryu, and S. Lee. "Hashtag-based topic evolution in social media". In: vol. 20. 6. Nov. 2017, pp. 1527–1549. DOI: 10.1007/s11280-017-0451-3. URL: https://doi.org/10.1007/s11280-017-0451-3.

[42] Y. W. Teh et al. "Hierarchical Dirichlet Processes". In: vol. 101. 476. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/016214506000000302. 2006, pp. 1566–1581. DOI: 10.1198/016214506000000302. URL: https://doi.org/10.1198/016214506000000302.

[43] EUvsDisinfo. *EUvsDisinfo. March 16, 2020. The Kremlin and Disinformation About Coronavirus.* 2020. URL: https://euvsdisinfo.eu/the-kremlin-and-disinformation-about-coronavirus/. (accessed: 01.19.2021).

[44] J. Kready et al. "YouTube Data Collection Using Parallel Processing". In: *2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*. 2020, pp. 1119–1122. DOI: 10.1109/IPDPSW50202.2020.00185.

[45] R. Řehůřek and P. Sojka. "Software Framework for Topic Modelling with Large Corpora". In: May 2010, pp. 45–50. DOI: 10.13140/2.1.2393.1847.

[46] H. Budhwani and R. Sun. "Creating COVID-19 Stigma by Referencing the Novel Coronavirus as the "Chinese virus" on Twitter: Quantitative Analysis of Social Media Data". In: vol. 22. 5. May 2020, e19301. DOI: 10.2196/19301. URL: http://www.ncbi.nlm.nih.gov/pubmed/32343669.

[47] H. Ritchie et al. *United States: Coronavirus Pandemic - Our World in Data*. 2020. URL: https://ourworldindata.org/coronavirus/country/united-states?country=~USA. (accessed: 07.29.2020).

[48] *U.S. Capitol Riot - The New York Times*. URL: https://www.nytimes.com/spotlight/us-capitol-riots-investigations (visited on 09/15/2021).