# A Multidimensional Analysis of YouTube Communities in the Indo-Pacific Region

Ugochukwu Onyepunuka, Thomas Marcoux, Mainuddin Shaik, Mayor Inna Gurung, Nitin Agarwal
*COSMOS Research Center*
*University of Arkansas at Little Rock*
Little Rock, AR, USA
email: {uponyepunuka, txmarcoux, mxshaik, mgurung, nxagarwal}@ualr.edu

***Abstract*—Although YouTube is the second most used social media platform in the world today, there is a need for more systematic research on video-based social network platforms, as few studies provide insights into the dynamics of their online discourse. Data, such as comments and other user engagement statistics, give insight into what information content creators are spreading and what moves communities. The analytical framework utilized in this research presents a multidimensional view into the geopolitical discourse on YouTube, with a focus on the Indo-Pacific region. We identified major YouTube channels that were discourse movers on polarizing topics, such as the treatment of Uyghur muslims, and COVID-19. We provided context on information consumption behaviors of these communities, as well as the engagement trends within these channels. This revealed segregated communities that often engaged in toxic behavior when clashing with opposite communities, as well as some inorganic engagement trends, allowing us to identify communities that may use automated means to push their narrative.**

***Keywords-indo-pacific; China; Uyghur; YouTube; misinformation; social media; information operations; cyber influence campaigns.***

## I. Introduction

In this study, we approach how YouTube videos can and have been used as vehicles of misinformation or political propaganda. We observe content popular in the Indo-Pacific region and explore three main topics: general discourse about the Chinese government, the Uyghur crisis, and COVID-19.

Compared to the significant body of work that studies the online discourse on social media platforms like Twitter, there is little research on discourse within video-based social networking platforms, such as YouTube, which is one of the most popular platforms in the Asia-Pacific region. Carrying out systematic research on video-sharing platforms like YouTube provides immense opportunities to gain situational awareness that could be pivotal in strategic policy making, especially in trade and defense. Data, such as video titles, comments, views, and likes can give insight into what information content creators and media houses are spreading along with the audience's engagement with those contents. Engagement statistics, such as likes and views, could also serve as potential sources to gain implicit knowledge about community interests on certain topics. Analyzing trending and influential content on social media platforms requires a rigorous, systematic approach and could yield many benefits, helping stakeholders to detect existing trends and content generators, track upcoming trends and accompanying narratives, and how discourse evolves.

Through this research, we aim to present a methodological pathway on how to identify suspicious and inorganic content in cyber-influence operations. We use the word "suspicious" when a message is being amplified through inorganic means: through commenter bots, inorganic growth in a channel's subscribers, or a video's views - driving traffic to the content.

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. Finally, in Section 4, we will discuss our findings before offering closing thoughts.

## II. Literature Review

Third party web traffic reports [1] tell us that YouTube is the second most popular website, and accounts for 20.4% of all search traffic. According to official YouTube sources [2], 1 billion hours of videos are watched each day. Another study [3] found that 60% of YouTube videos are watched at least 10 times on the day they are posted. The researchers also highlighted that, if a video does not attract viewership in the first few days after upload, it will likely not gain traction later on. YouTube provides an overwhelming amount of video data, with over 500 hours of content uploaded every minute on average. That number was 300 in 2013 [4]. In previous publications [5, 6], we identified YouTube as a potential vehicle of misinformation and proposed the use of YouTube metadata for understanding and visualizing these phenomena by observing data trends. Previous research [7] has studied 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 of polarizing content, this effect has been shown to be a contributing factor to user radicalization [8]. Recent research on the same subject leverages advanced Natural Language Processing (NLP) techniques on text entities, such as video comments [9] but we could find little work available on the video content itself.

It is important to provide some social and political context for this work. One major key idea of this study is China's desire to develop a 21st-century Maritime Silk Road coined

 ISBN: 978-1-68558-002-5 

as the belt and road initiative. One of the important hubs for the project Xinjiang, is home to Uyghurs Muslims [10]. The People's Republic of China has opened a number of what it calls "vocational education" camps, claiming to de-radicalize extremists individuals in the area. However, the West has called it cultural genocide and political brainwashing. The government of China is taking a range of countermeasures to mitigate those claims by utilizing the video-sharing platform YouTube. In these videos, content creators are calling the West hypocrites who are spreading fake news, and producing content mimicking trusted news channels like CNN [11]. While this type of content increases misinformation and gives raise to propaganda, it can have a severe impact on society. For instance, during COVID-19, 87 % of the users encountered relevant misinformation that suggested consuming ethanol or bleach as preventative measures for COVID-19 [12]. Similarly, politicizing YouTube content gave rise to anti-Asian hatred with the use of words like Chinese virus to refer to the pandemic [13]. Other research [14] suggests that this type of content is heavily concentrated among a small group that has high prior racial resentment. In contrast, YouTube content has been utilized to create public awareness as well. Studies [15] suggest that many content creators used the platform as a mean to convey information about diseases, interviewing subject matter experts, giving evidence and argument.

## III. Methodology

This study uses a three-step methodology to assess YouTube content's suspiciousness. First, we collected YouTube video metadata and comments, separated them according to the main narrative they pushed, and performed engagement and network analysis on each of the subsets identified.

### A. Narrative Segregation

We collected data from YouTube using a group of keywords relating to each narrative theme provided by subject matter experts from Arizona State University. The relevant keywords were identified by studying coverage in the Indo-Pacific region with further reviews to improve the inclusiveness. A data collection task was set up for each narrative using the group of keywords as a parameter to pull YouTube data relating to the narrative. To perform an analysis on a narrative, we queried our database using the keywords in the full-text search query. Table II is a snippet of the keywords for each narrative and Table I shows total collection statistics. YouTube data was collected using the official YouTube Data API and following the methodology described by Kready et al. [16].

TABLE I
Video collection statistics

| | |
|---|---|
| Videos | 25,673 |
| Channels | 6,806 |
| Comments | 5,538,730 |

### B. Engagement Analysis

To uncover any suspicious activity within the channels found in our research, we examined the engagement trends over certain periods. Our analysis includes time series data based on channels' activity trends, such as, daily view count, daily subscriber count, daily video count, total views, total subscribers, total comments, and total videos; following the methodology described by Kirdemir et al. [17].

Engagement trend analysis was executed on the channels we collected data for, to discover channels exhibiting suspicious behaviors. There are 5 key steps employed by the script in discovering suspicious behaviors in a channel:

1) Rolling window correlation analysis
2) Anomaly detection
3) Rule-based classification
4) Principal component analysis
5) Clustering

The first step is the rolling window correlation analysis that groups the data into a rolling window of 100 days and computes the pairwise correlation between the video production and engagement metrics for each window. It aims to capture channels inauthentic behaviors by analyzing the correlation between engagement metrics, e.g. channels with increasing views, but decreasing subscribers. The output from this step is the start and end date for each window and the value for correlation pairs between the metrics:

- Views and subscribers
- Views and videos
- Views and comments
- Subscribers and videos
- Subscribers and comments
- Videos and comments

The output from the rolling window correlation analysis goes into the anomaly detection step to train a Long Short-Term Memory (LSTM) model on the time series data and identify anomalies in the correlation between engagement metrics. LSTM models are artificial neural networks that, unlike traditional feedforward neural networks, use feedback connections. This lets the model process not only single data points, but also entire sequences of data by using a recurrent network. To capture anomalous periods, a threshold was set to average peak of the data for each channel's correlation pairs to capture all data points that were placed above that threshold.

To compute the anomaly threshold, we grouped the data into a rolling window of 100 days, then computed the pairwise correlation between the video engagement metrics. The output from this was passed into the anomaly detection step, where the data was trained on the LSTM model. It ran through each dataset (1000 data points approximately, where one data point represents 100 days) with a batch size of 32 and a lookback size of 1. The loss from the computation was represented as the anomaly confidence score. To capture the anomalous periods, the threshold was set to the anomaly confidence score for each channel's correlation pair. This resulted in a list containing the anomalous data points for each correlation pair.

 ISBN: 978-1-68558-002-5 

TABLE II
NARRATIVE EXTRACTION SAMPLE

| S/N | Narrative | Date | Keywords |
|---|---|---|---|
| 1 | China | 2018 - 2021 | 'Komunis Cina—China pengaruh Indonesia', 'Menguasai Cina—China—Tiongkok— Tionghoa ekonomi Indonesia', ... |
| 2 | Uyghur | 2018 - 2021 | 'Uighur—Uyghur Indonesia', 'Penindasan Uighur—Uyghur bebaskan', 'Kejam Uighur—Uyghur', ... |

A rule-based classification algorithm was used to determine a suspicion score ranging from 0 (least suspicious) to 1 (most suspicious) for each engagement metric pair. The suspicion scores for all indicators are then aggregated to create a single suspicion score for each observation.

### C. Network Analysis

The YouTube co-commenter network represents the connections between commenters on YouTube videos, where the edge weight indicates the number of videos they commented on together. To identify the communities in the network, we used the ForceAtlas2 layout in Gephi, a force-directed layout similar to other algorithms used in network spatialization. This also showed the top node in each community, and the type of content the community engaged with. The modularity measure was used to partition the network into clusters or communities. As the degree centrality measure was used to find out the top nodes, it also shows how many ties/edges a node has. Once the node information was extracted for all nodes in a community, our dataset is queried to extract samples of the content some of the nodes engaged with.

## 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. The results show YouTube co-commenters networks for the different Indo-Pacific narratives where edges between commenters indicate that the commenters were active in 10 or more of the same videos. The results highlight the top nodes and different communities in the network. It also shows the type of content that the members of the various communities engaged with or shared. We notice that the communities tend to converge towards news channels, and then segregate based on their preferred narrative.

### A. Network Analysis - China Narrative (February – August 2021)

TABLE III
CHINESE NARRATIVE COMMUNITIES

| S/N | Top Commenter Name | Community | Degree |
|---|---|---|---|
| 1 | Nilesh Bhattacharya | 0 (Blue) | 108 |
| 2 | thndrngest | 0 (Blue) | 72 |
| 3 | Beware of the Leaven of the USA | 0 (Blue) | 66 |
| 4 | Discover China 探索中 | 1 (Purple) | 173 |
| 5 | True North Strong and Free | 1 (Purple) | 95 |
| 6 | Colchicum autumn crocus | 1 (Purple) | 66 |
| 7 | Pub Comrad | 4 (Green) | 113 |
| 8 | Olympic - 2022 | 4 (Green) | 76 |
| 9 | Last Chang | 4 (Green) | 72 |

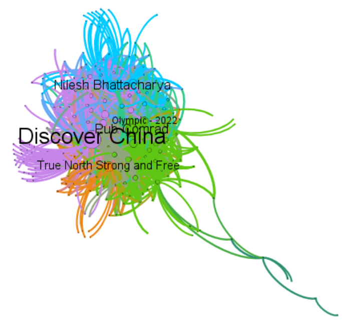


Fig. 1. China narrative (February – August 2021).

Figure 1 illustrates some of the data shown in Table III. It highlights three communities. The majority of the channels from community 0 publish videos relating to China. It is labelled as such due to the presence of high viewership videos, such as “Xi Jinping: China’s president and his quest for world power”, and “Global brands face backlash in China for rejecting Xinjiang cotton”. The comments from the top nodes in this community showed support for China, but showed some robotic or translated patterns, e.g., “CPC is indeed the Chinese people’s party. Long may it stay strong”, “Long live the people’s Republic of China”.

Videos in community 1 discussed China centenary celebration. It is labelled as such due to the presence of high viewership videos, such as “Xi Jinping leads celebrations marking centenary of China’s ruling Communist Party”, and “China’s largest military parade marks National Day”. The comments from two of the top nodes in this community were pro-China with comments on Chinas’ centenary celebration and others. While the comments from the other node were anti-China, and were critical of the decision to hold the celebration while citizens’ grievances went unaddressed.

Top videos in community 3 also captured polarizing videos from other communities, which tended to communicate anti-China sentiment, e.g., “Taiwan: ’China preparing for final military assault”, and “Global brands face backlash in China for rejecting Xinjiang cotton”. The comments made by the top nodes in this community were vehemently anti-China, with comments such as: “Of the 14 countries bordering China, it has conflict with 13 of them including Russia”, and “Without constant lying (on top of intimidation) the Chinese regime can’t exist”.

TABLE IV
UYGHUR NARRATIVE COMMUNITIES

| S/N | Top Commenter Name | Community | Degree |
|---|---|---|---|
| 1 | Discover China | 0 (Green) | 145 |
| 2 | Nilesh Bhattacharya | 0 (Green) | 102 |
| 3 | thndrngest | 0 (Green) | 95 |
| 4 | Beware of the Leaven of the USA | 1 (Purple) | 210 |
| 5 | Arthur Lincoln | 1 (Purple) | 85 |
| 6 | Aaron Baldwin | 1 (Purple) | 63 |
| 7 | Hüseme Erbolat | 2 (Blue) | 67 |
| 8 | John Francisco | 2 (Blue) | 55 |
| 9 | lin hai | 2 (Blue) | 35 |
| 10 | Yuni Sukawana | 3 (Orange) | 71 |
| 11 | Siti Nurjannah | 3 (Orange) | 71 |
| 12 | Camay Chayo | 3 (Orange) | 68 |

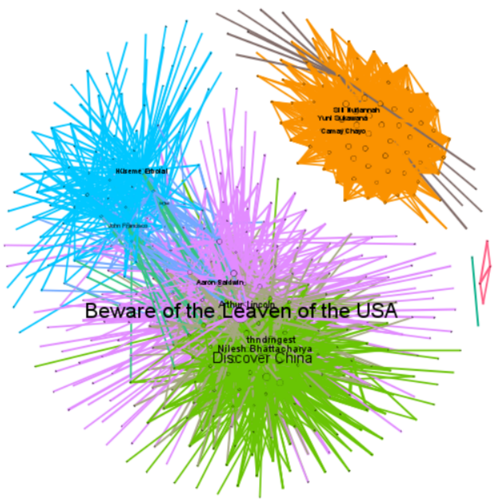


Fig. 2. Uyghur narrative (February – August 2021).

*B. Network Analysis - Uyghur Narrative (February – August 2021)*

Figure 2 illustrates some of the data shown in Table IV. It highlights four communities. In community 0, we noticed the top videos focused on Chinas' forced labor camps and the oppression of the Uyghurs, e.g., "Uyghurs Who Fled China Now Face Repression in Pakistan", "What's happening with China's Uighurs? — Start Here", and "GLOBALink — Chinese scholars debunk "forced labor" claims with own investigation". The comments made by top nodes in this community showed support to China and referred to China's oppression of the Uyghur Muslims as accusations. They also showed dislike for the western media.

The top videos in the nodes community 1 engaged with tended to be educational content, with titles such as "How Xinjiang Became Muslim ft. Let's Talk Religion", and "What's China's 're-education camp' in Xinjiang really about?". The comments from the top nodes in this community were mixed. One showed support for China and talked about how western media frames China as negative in their news, with another top node being anti-China.

Top videos in community 2 included content from popular politainment channels, such as "Last Week Tonight with John Oliver" and "VICE News", and also featured engaging titles, such as "China's Vanishing Muslims: Undercover In The Most Dystopian Place In The World". The comments made by the top nodes in this community were very polarized and argumentative, with both sides of the issues present.

The channels in community 3 share educational content on Islam, with top video in this community pertaining to the teachings of Islam. The title of the videos, translated to English, includes: "Extraordinary Ustadz Abdul Somad answered all the questions of this Malaysian congregation", and "2 ways to ask for God's help — Palembang". The comments from the top nodes were positive, and praised the speakers and their material.

*C. Network Analysis - COVID-19 Narrative (All)*

TABLE V
COVID-19 NARRATIVE COMMUNITIES

| S/N | Commenter Name | Community | Degree |
|---|---|---|---|
| 1 | True North Strong and Free | 0 (Orange) | 16 |
| 2 | Lau Billy | 0 (Orange) | 3 |
| 3 | Frederic Chen | 0 (Orange) | 1 |
| 4 | Nilesh Bhattacharya | 1 (Green) | 12 |
| 5 | Jef Chen | 1 (Green) | 11 |
| 6 | Beware of the Leaven of the USA | 1 (Green) | 11 |
| 7 | Colchicum autumn crocus | 2 (Purple) | 13 |
| 8 | thndrngest | 2 (Purple) | 8 |
| 9 | LVPN 1 | 2 (Purple) | 6 |

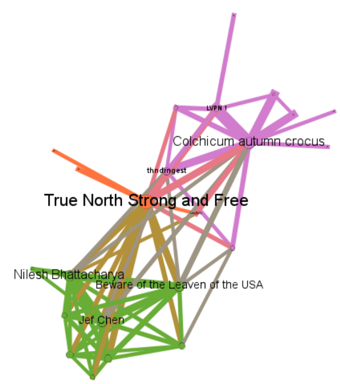


Fig. 3. COVID-19 narrative (all dates).

Figure 3 illustrates some of the data shown in Table V. It highlights three communities. In community 0, we once again see mostly News channels, including "CGTN", "New China TV", "球Global Times', "DW News", and "China Daily 中日". Most of the top videos in this community include content from the "China Global Television Network" (CGTN) channel and include "Fighting Together: China sends aid to Colombia to combat virus", and "Ambassador Lin Songtian discusses China-Africa pandemic aid". Most videos seemed to discuss Chinese international aid in vaccine distribution. The comments made by the top nodes in this community showed high toxicity and attacked the video content, as well as the other commenters. The top nodes showed polarized attitudes in regards to China, as well as inorganic behavior. Notable, one of the top nodes posted emojis exclusively.

The top nodes in community 1 engaged with the videos listed below, with a majority of videos making reference to China, e.g., "The Truth About The COVID Origin and the Lab Leak Theory", and "China says it has not received COVID-19 aid from U.S. government". We noticed obviously inorganic behavior, with some of the comments made by one of the top nodes, with repeating comments for every video commented on. The comments from the other two nodes hinted towards the commenters being pro-China: "China is the future", "China is a good example for the world", etc.

Videos the top nodes from community 2 engaged with only include two news channels: CGTN, and News China TV. Top videos from these channels discussed COVID-19 aid, e.g., "Medical teams return home with Wuhan in their hearts", and "Why did Sichuan experts volunteer to help Italy?". Comments were consistent with this pro-China sentiment.

### D. Engagement Trends Analysis - Indo-Pacific Channels

Engagement trend analysis is used to uncover channel-level suspicious activity on YouTube by processing metrics of video production and engagement through a multi-step analytical pipeline including rolling window correlation analysis, anomaly detection, peak detection, rule-based classification, Principal Component Analysis (PCA), and unsupervised clustering. From the list of 5,000 Indo-Pacific channel ids collected, we were able to get daily data on 3,517 channels from Social Blade, an online tool tracking social media statistics and analytics, commonly used for traffic and earnings estimations.

Figure 4 shows the output of the anomaly detection model, which returned anomalous periods with a large feature set. PCA was then used to reduce the dimensions of the dataset and a scatterplot was created using the first two principal components and the suspicion score was used to color the data point, as reported in Table VI, which shows the highest scores. Once the suspicion score was generated for every observation of the 3,517 Indo-Pacific channels, the channel with the highest suspicion score was "Breaking News TV" with a suspicion score of 0.72. The anomalous period for this data point occurred between 2019-03-24 and 2019-10-08. The channel has now been taken down.

The final step was using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to identify clusters (groups of channels with similar engagement trends) from the PCA scatter plot - see Figure 5. Figure 5 matches Figure 4 and helps match communities to trends of suspicious engagement.

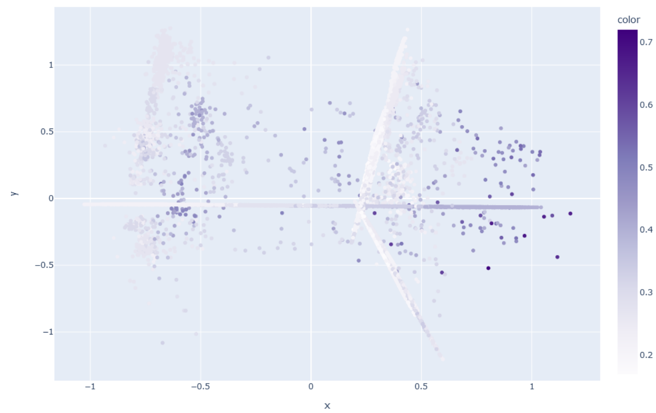


Fig. 4. Principal component analysis scatter plot showing suspicion scores.

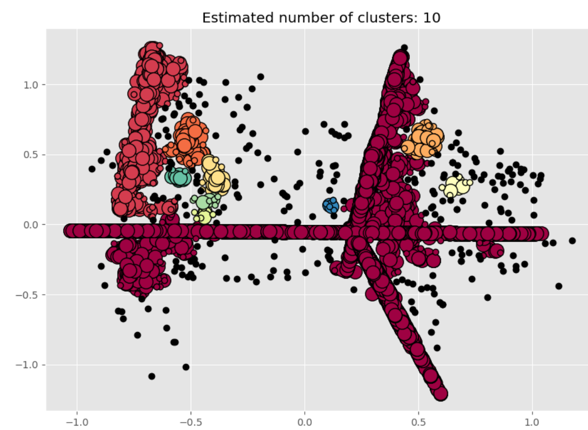


Fig. 5. Scatter plot showing clusters identified from DBSCAN.

## V. Conclusion and Future Work

In this study, we presented a multi-dimensional analysis of popular political and societal discourse in the Indo-Pacific region. We focused on YouTube, the second most used social media platform in the world today, and highly popular in the Indo-Pacific region, as a source of information and entertainment. To understand the patterns behind these online discourse dynamics, the interests of these communities, and the content preferences of the discourse movers (channels), we used network analysis and narrative segregation to identify various communities within YouTube networks, and provided context on their information consumption behaviors and engagement metrics. This revealed highly polarized user to user interactions, as well as some inorganic engagement trends, allowing us to identify "suspicious" communities. Through this study, we want to present the pathway on how to analyze video-based platforms, especially YouTube, to obtain situational awareness during any social cyber-influence operation. Future points of improvement for this research include considering bot accounts and their impact, and further automating the process of connecting community detection and clustering to engagement trends, creating a suspicion score at the community level instead of the channel level. This will allow analysts

 ISBN: 978-1-68558-002-5

TABLE VI
CHANNELS WITH MOST SUSPICIOUS DATA POINT

| Channel ID | Channel Name | Suspicious Score | Date Range |
|---|---|---|---|
| UCN_qIhm7BAq9Qxa6j9XMVXA | Breaking News TV | 0.72 | 2019-03-24 to 2019-10-08 |
| UCrGZO3wJ20CWiy36Fdu6vdw | Badminton Talk | 0.68 | 2019-04-24 to 2019-11-08 |
| UCkQSMH1vP1Kcx-SPArUYLLQ | viral_makkodak | 0.67 | 2020-02-07 to 2020-09-04 |
| UCmvtaFkiWOSHhnOwt4Fz68g | SAFA News | 0.64 | 2019-11-26 to 2020-06-22 |

to automatically identify communities that tend to participate in online information operations, purposefully or otherwise.

## 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, N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), 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] *Youtube.com Traffic Analytics and Market Share — Similarweb*, URL: https://www.similarweb.com/website/youtube.com/#overview (visited on 2022-07-15).

[2] *How YouTube Works - Product Features, Responsibility, & Impact*, URL: https://www.youtube.com/intl/en-GB/howyoutubeworks/ (visited on 2022-07-15).

[3] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon, "Analyzing the Video Popularity Characteristics of Large-Scale User Generated Content Systems", in: *IEEE/ACM Transactions on Networking* 17.5 (2009-10), pp. 1357–1370, ISSN: 1558-2566, DOI: 10.1109/TNET.2008.2011358.

[4] J. Hale, *More Than 500 Hours Of Content Are Now Being Uploaded To YouTube Every Minute*, 2019-05, URL: https://www.tubefilter.com/2019/05/07/number-hours-video-uploaded-to-youtube-per-minute/ (visited on 2022-07-15).

[5] M. N. Hussain, S. Tokdemir, N. Agarwal, and S. Al-khateeb, "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.

[6] 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.

[7] 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.

[8] L. Tang et al., ""Down the Rabbit Hole" of Vaccine Misinformation on YouTube: Network Exposure Study", in: *J Med Internet Res* 23.1 (2021-01), e23262, ISSN: 1438-8871, DOI: 10.2196/23262, URL: http://www.ncbi.nlm.nih.gov/pubmed/33399543.

[9] 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, 2020-07, URL: https://www.aclweb.org/anthology/2020.nlpcovid19-acl.17.

[10] M. Clarke, "The Belt and Road Initiative: Exploring Beijing's Motivations and Challenges for its New Silk Road", in: *Strategic Analysis* 42.2 (2018), Publisher: Routledge _eprint: https://doi.org/10.1080/09700161.2018.1439326, pp. 84–102, DOI: 10.1080/09700161.2018.1439326, URL: https://doi.org/10.1080/09700161.2018.1439326.

[11] B. Alpermann, "'In other news': China's international media strategy on Xinjiang –CGTN and Xinhua on YouTube", in: 2020-10.

[12] S. Zhang, W. Pian, F. Ma, Z. Ni, and Y. Liu, "Characterizing the COVID-19 Infodemic on Chinese Social Media: Exploratory Study", in: *JMIR Public Health Surveill* 7.2 (2021-02), e26090, ISSN: 2369-2960, DOI: 10.2196/26090, URL: http://www.ncbi.nlm.nih.gov/pubmed/33460391.

[13] Y. Yang, C. Noonark, and C. Donghwa, "Do YouTubers Hate Asians? An Analysis of YouTube Users' Anti-Asian Hatred on Major U.S. News Channels during the COVID-19 Pandemic", in: *Global Media Journal - German Edition* 11.1 (2021-07), URL: https://www.globalmediajournal.de/index.php/gmj/article/view/198.

[14] A. Y. Chen, B. Nyhan, J. Reifler, R. E. Robertson, and C. Wilson, *Subscriptions and external links help drive resentful users to alternative and extremist YouTube videos*, 2022, DOI: 10.48550/ARXIV.2204.10921, URL: https://arxiv.org/abs/2204.10921.

[15] F. A. Sofian, "YouTubers Creativity in Creating Public Awareness of COVID-19 in Indonesia: A YouTube Content Analysis", in: *2020 International Conference on Information Management and Technology (ICIMTech)*, 2020, pp. 881–886, DOI: 10.1109/ICIMTech50083.2020.9211149.

[16] J. Kready, S. A. Shimray, M. N. Hussain, and N. Agarwal, "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.

[17] B. Kirdemir, O. Adeliyi, and N. Agarwal, "Towards Characterizing Coordinated Inauthentic Behaviors on YouTube", en, in: *The 2nd Workshop on Reducing Online Misinformation through Credible Information Retrieval (ROMCIR 2022) held with the 44th European Conference on Information Retrieval (ECIR 2022)*, Stavanger, Norway, 2022-04.