

Anonymous Dissent in the Digital Age: A YouTube Dislikes Dataset

Sujan Dutta¹ *, Mallikarjuna Tupakula² *, Sumeet Kumar², and Ashiqur R. KhudaBukhsh¹ **

¹ Rochester Institute of Technology, Rochester NY, USA
`{sd2516, axkvse}@rit.edu`

² Indian School of Business, Hyderabad, India
`{mallikarjuna_tupakula, sumeet_kumar}@isb.edu`

Abstract. On December 13, 2021, YouTube made a major policy change with respect to the visibility of video dislikes. Citing the need to protect the well-being of individual content creators, YouTube stated that dislike information will only be visible to the video owners thus abolishing a decade-long tradition of publicly visible, anonymous outlet for user dissent. This paper makes two key contributions. First, it releases a valuable dataset of dislike information from 8.3 million videos gleaned from 159 popular news and debate outlets to characterize and chronicle the dislike behavior on YouTube. Second, it quantifies and investigates the information gap that this policy change leaves us with.

Keywords: Social Web · User Engagement · YouTube Policy Change

1 Introduction

In December 2021, YouTube introduced a platform-wide policy change that rendered dislike information publicly inaccessible. One month prior to rescinding API access to dislike information YouTube announced this change and gave web science practitioners an advance warning that they were on borrowed time. In our bid to preserve a snapshot of an anonymous outlet for user disapproval, this one-month window allowed us to collect likes-dislikes data for a large number of popular channels. After YouTube’s policy change took effect, the platform no longer displays the dislike count. This paper performs a longitudinal analysis of dislike behavior on YouTube on controversial issues and analyzes the possible information gaps YouTube’s policy change leaves us with.

Respecting YouTube’s rationale to protect small content creators, our analyses preclude small, individual contributors and primarily consider major news networks that are no strangers to audience reactions and TRP battles. Via a substantial dataset of 8.3 million videos that enjoyed overall 2.74 billion likes and endured 333 million dislikes, this paper characterizes and chronicles the like

* Joint first authors

** Corresponding author

and dislike behavior on YouTube. Central to the theme of inclusive AI, our paper considers (1) news and opinion outlets presented in a wide range of languages (see Table 1); (2) news outlets with minorities as the target audience; and (3) news outlets having a substantial geographical coverage (135 countries).

To summarize, our contributions are the following:

- **Resource:** We release a dataset of 8.3 million YouTube videos with like, dislike, video title, and video description information. Beyond the immediate goal of characterizing dislike behavior, our dataset can be valuable to social scientists to understand political events of import in the last decade.
- **Social:** We present a detailed characterization of dislike behavior of a global audience. Our analyses reveal that studying dislike behavior may present valuable insights into anonymous dissent. Especially, in countries with questioned democracy. While prior literature investigated potential reasons for YouTube engagement [13] and post-hoc effect of hiding dislikes [11], to our knowledge, ours is the first detailed characterization of dislike behavior at this scale.

Table 1: List of all languages in the dataset.

Language
Albanian, Amharic, Arabic, Azerbaijani, Bengali, Burmese, Chinese, English, Filipino, French, German, Greek, Gujarati, Hausa, Hindi, Hungarian, Indonesian, Italian, Japanese, Kannada, Korean, Malay, Malayalam, Marathi, Odia, Pashto, Persian, Portuguese, Russian, Sinhalese, Spanish, Swahili, Tamil, Thai, Turkish, Ukrainian, Urdu, Vietnamese.

2 Dataset

We collect a dataset³ that spans 159 YouTube channels with an overall subscriber count of 727,756,097 (4.43 ± 0.70 million subscribers). Examples include (1) official handles of prominent news networks (e.g., TVC News Nigeria (Africa), BBC News (Europe), CNN (North America), Sky News Australia (Oceania), and Aaj Tak (Asia)); (2) general and parliamentary debates (e.g., Oxford Union, UK Parliament, C-Span); (3) YouTube channels of prominent print media (e.g., New York Times (US), Daily Mail (UK), Bild (Germany), and Times of India (India)).

We consult each country’s Wikipedia page that documents major news outlets (cable news channels and print media). For these news outlets, we search and confirm if these outlets have a substantial YouTube presence (we set a threshold of at least 100K subscribers). For these YouTube channels, we scrape the videos using Selenium to obtain the video IDs. We use the publicly available YouTube API to collect the count of likes, dislikes, comments, and views of these videos. In addition, we also collect the video titles and video descriptions (if present). Our dataset is collected between November 13, 2021, and December 12, 2021.

We consider news channels due to the following reasons:

- **Global participation and audience diversity:** A broad and diverse audience can be achieved by geographically and linguistically diverse YouTube news channels. Our selected channels span over 130 countries and 38 languages. 64.7%

³ Available at <https://github.com/Suji04/YouTube-Dislike-Dataset>

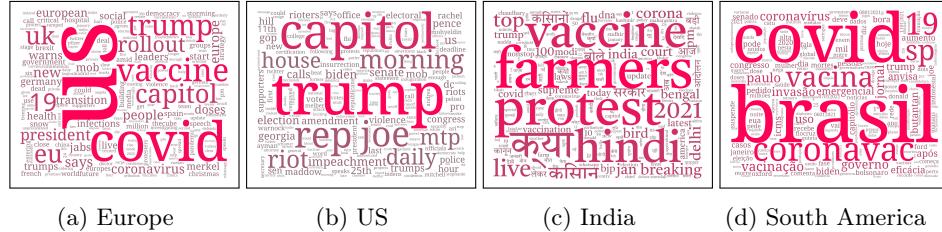


Fig. 1: Word cloud visualization of news headlines from prominent news channels around the globe for the week of Jan 6, 2021. EuroNews (Europe); CNN, Fox News, and MSNBC (US); Aaj Tak, Zee News, ABP News and NDTV (India); Band Jornalismo and C5N (South America).

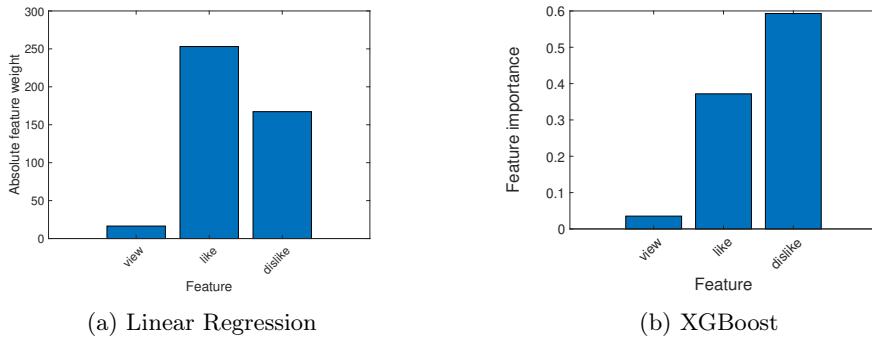


Fig. 2: Relative feature importance of view, like, and dislike in the task of predicting comment count on a given video.

of these channels are operated by official news networks, ensuring a reliable representation of global politics. We take special care in including news outlets from countries beyond the economic superpowers (e.g., TVC News Nigeria) or news outlets catering to a disadvantaged minority group (e.g., Black News Channel) or news outlets with a focus on countries under political instability (e.g., ToloNews from Afghanistan, Syria Al Khbaria for Syria). From Swahili to Pashto – Our linguistic diversity, again, ensures that we reach out to a truly global audience.

- ***Rich topical diversity:*** Within the narrow scope of US politics in 2020-2021, the country witnessed a raging pandemic, a bitterly fought election, a watershed moment in the history of racial justice, and a far from peaceful transfer of power. From climate change to migrant crises, from Superbowl to Wimbledon, and local street crimes to international disputes, the richness of our dataset offers a unique lens to understand public dissent.
 - ***News as parallel accounts of world events:*** Finally, this dataset captures multiple parallel accounts of world events as they happen. Figure 1 presents a world-cloud visualization of news titles from news outlets from different geographic regions for the week of Jan 6, 2021. We note that while the US was

grappling with an assault on democracy with an unprecedented riot at the Capitol, India was dealing with its own issue of farmers' dissent. At the same time, Brazil was dealing with the COVID-19 vaccine adoption challenges. Even within the same country, two networks can present competing views. For instance, Capitol insurrection is a phrase that seldom appeared on Fox while it was a regular fixture on MSNBC news titles. Together, these outlets paint a comprehensive picture of how the world works, and therefore, our dataset can be of great value to social scientists for a broad range of socio-economic and policy questions.

3 Characterizing Dislikes

Research Question: *How are likes, dislikes, comments, and views correlated?* Table 2 summarizes the pairwise correlations between views, likes, dislikes, and comments. It is intuitive that the more views a video gets, the more likely it would get an audience who are willing to engage through likes, dislikes, and comments. We further note that likes (dislikes) are more correlated with comments than dislikes (likes). This result is also intuitive as not all videos will elicit polarization reactions. On the other hand, a polarized video is likely to receive more comments than a non-polarized one. This intuition is reflected as comments are more correlated with likes and dislikes than views.

Table 2: Pairwise correlation between video views, likes, dislikes, and comments.

	view	like	dislike	comment
view	-	0.66	0.50	0.40
like	0.66	-	0.45	0.58
dislike	0.50	0.45	-	0.50
comment	0.40	0.58	0.50	-

Research Question: *Does dislike present any additional informational value beyond video views and likes?* We design a prediction task where the inputs (features) are views, likes, and dislikes received by a video and the output is the predicted comment count of the video. Figure 2 summarizes the feature weights of individual features upon using two well-known regression algorithms (XGBoost [4] and linear regression [16]). We observe that dislike information is an important feature in predicting comment counts.

3.1 Quantifying Disagreement

To estimate viewership disagreement over a set of videos, we follow a measure introduced by KhudaBukhsh *et al.* [14]. This measure has been applied to quantify the shift in the popularity of fringe conservative news outlets preceding the Capitol riot [15].

For a given video v , let v_{like} and $v_{dislike}$ represent the number of likes and dislikes received by it. The disagreement of v is measured as $\frac{v_{dislike}}{v_{like} + v_{dislike}}$. Let

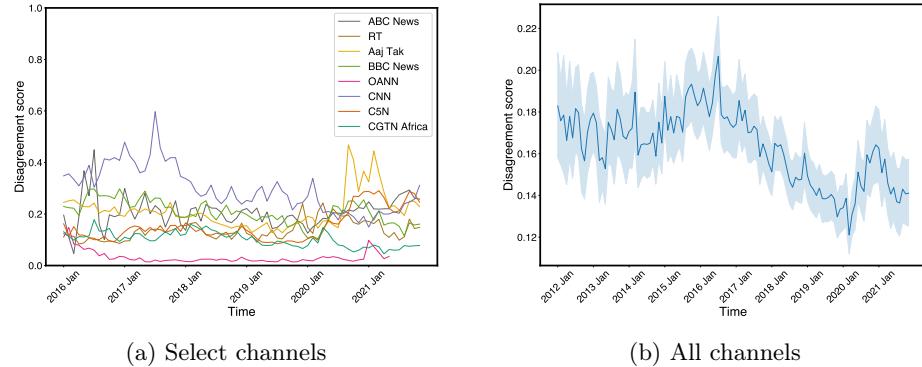


Fig. 3: Evolution of disagreement scores over time. Figure 3a considers six channels from six continents: CGTN (Africa); BBC (Europe); Aaj Tak (Asia); ABC News (Oceania); CNN (N. America); and C5N (S. America). OANN and RT are included for contrastive purposes. Figure 3b considers all channels.

$\mathbb{I}(v^i, \mathcal{F})$ be an indicator function that returns 1 if video v^i satisfies a set of filtering criteria denoted by \mathcal{F} . KhudaBukhsh *et al.* [14] mainly fix \mathcal{F} to a specific channel and upload time. For example, \mathcal{F} could be all videos uploaded on CNN in the month of January 2014. The *disagreement score* of the set of videos satisfying the filtering criteria \mathcal{F} is computed as $\frac{\sum_i \mathbb{I}(v^i, \mathcal{F})}{\sum_i \mathbb{I}(v^i, \mathcal{F})}$. A low *disagreement score* indicates that the majority of the viewers liked the content. Whereas, a high value signifies a rising proportion of disapproving viewership. In our work, we vary the filtering conditions \mathcal{F} to characterize like and dislike behavior along different aspects.

Research Question: Are all news networks uniformly loved (or hated)?

Figure 3a considers a time duration of six years to account for elections, pandemics, protests for justice, and other factors that may disrupt news cycles temporarily with sudden shifts in popularity. We observe that not all news networks enjoy a similar disagreement score profile. For instance, for the most part of the time, CNN content was more disagreed than any other news channel. CNN also experiences a sharp spike starting in 2017, which marks the beginning of the long tussle between CNN and former President Trump, who repeatedly questioned CNN's reporting quality. Elections also created a spike in ABC News Australia at the beginning of 2016. The spike in disagreement scores in India's news channel Aaj Tak coincided with India's second wave COVID-19 crisis, which triggered a nationwide healthcare collapse. In contrast, One America News Network (OANN) has documented evidence of peddling election misinformation and COVID-19 misinformation [8]. We notice a disagreement score profile of being an echo chamber. Finally, RT Russia is a state-sponsored channel with a long history of accusations of being a propaganda news network. However, we see more disagreement on RT Russia than CGTN Africa. A deeper exploration of RT Russia follows in the later sections.

Table 3: Most disagreed trigrams across channels. Aaj Tak uses code-switched Hindi and English in their news titles. We translate the Hindi portions into English. For Russian, we present a translation performed by Google and verified by a native Russian speaker.

Aaj Tak	BBC	CNN	DW	Fox News	RT
abhishek banerjee wife	syria air strikes	assault weapons ban	germany focus europe	pelosi house dems	russian presidential elections
pm modis mann	iran nuclear deal	black lives matter	covid 19 vaccine	speaker pelosi holds	mishustin holds a meeting
opinion poll 2020	brexit trade deal	clinton donald trump	quadriga international talk	jen psaki holds	government members live
monsoon session 2020	leader jeremy corbyn	gop health bill	international talk quadriga	julie rogincky claps	fight covid 19
Gaurav Bhatia बोले [speaks]	labour leader jeremy	religious freedom bill	world max euromaxx	biden delivers remarks	Patriarch kirill congratulated
Bihar opinion poll	pm boris johnson	gop tax bill	chancellor angela merkel	schumer senate dems	head coach of the team
Rajasthan priest burnt	explained 60 seconds	white house trump	hillary clinton wins	covid 19 response	coordinating council fighting
100 Sep 2020	death toll rises	trump travel ban	world stories week	ted cru wins	putin holds a meeting
priest burnt alive	george floyd death	rep adam schiff	stories week reports	trump cruz feud	appeal vladimir putin
bengal jp nadda	aung san suu	sotu john mccain	young global leaders	gov cuomo holds	holds a meeting by members

Figure 3b considers our entire dataset and tracks the overall disagreement score. Starting from 2017 to the onset of COVID-19, the overall disagreement score was on a steady decline. Of course, the sharp spike around 2020 February coincides with a global pandemic requiring little explanation, we however wonder what could be the reason for the steady decline that started in 2017. Is it because the hyper-partisan media has polarized and segregated the news audience into different camps so that everyone has found their own cozy echo chamber?

3.2 Censorship

We now investigate the nature of news that attracts disagreement. We focus on CCTV China, a state-sponsored news outlet. China has documented evidence of social web censorship [1]. In fact, most videos on CCTV China have comments disabled. This implies that the only remaining outlet to express disagreement for this particular channel was through dislikes. Table 4 lists the titles of the top 10 most disagreed videos on CCTV. Research on the popularity of the Chinese Communist Party (CCP) indicates that state-controlled media play an important role in it [12,19]. Nonetheless, we observe that there exists considerable anonymous dissent against CCP that has little documentation.

3.3 Disagreement of Phrases

Research Question: Are there specific trigger words or phrases that are associated with greater disagreement?

Table 4: Video titles and disagreement scores of the most disagreed videos on CCTV. Translations are obtained using Google Translate and verified by a proficient Chinese speaker.

CCTV	Disagreement score
“Ping Talk” is close to people-Xi Jinping’s favorite allusion: “I will be selfless and live up to the people”	0.78
The song “Without the Communist Party, there would be no new China” “Celebrate the 100th anniversary of the founding of the Communist Party of China Theatrical performance “The Great Journey”	0.60
Theme song “Liancheng Family” Singing: Miriam Yeung, Long Shijie, Zheng Qiyuan, etc. [Celebrating the 20th anniversary of Macao’s return to the motherland]	0.49
Situational chorus “Roar, the Yellow River” “Celebrating the 100th anniversary of the founding of the Communist Party of China Theatrical performance “The Great Journey”	0.48
The Central Committee of the Communist Party of China and the State Council held a Spring Festival group meeting, Xi Jinping delivered a speech 2020-01-23	0.48
Chinese Ministry of Foreign Affairs: U.S. remarks on Hong Kong appear to be fair, but in fact reveal ulterior motives	0.43
“News Network” races against time and keeps pace with history - General Secretary Xi Jinping’s speech at the 2020 Spring Festival group gathering sparked enthusiastic responses 2020-01-24	0.37
VR micro-record “Happiness Coordinates”	0.36
“News Network” and “Seeking Truth” magazines published an important article by General Secretary Xi Jinping “Speech at the Discussion at the Dunhuang Academy” 2020-01-31	0.36
Wang Kai pays tribute to the heroes of the Anti-Japanese War “March of the Volunteers” shows the power of national rejuvenation “Heroes in the Poster”	0.32

We define disagreement of a word (phrase) as the disagreement score of all videos containing that word (phrase) in their titles. While we note that many such phrases (words) would not necessarily capture the stance of the video, this simple approach may lead to a deeper understanding of trigger words or phrases. On our entire dataset, we find that `climate change` (disagreement score: 0.23) is a much more divisive phrase than `weather forecast` (disagreement score: 0.10). Of the words `citizen` (disagreement score: 0.16), `immigrant` (disagreement score: 0.20), and `refugee` (disagreement score: 0.25), `refugee` is more disagreed than `citizen`. The disagreement score gap between refugee and citizen substantially widens when we restrict our focus to EuroNews videos.

Table 3 lists most disagreed trigrams from six prominent news channels: Aaj Tak (India); BBC (UK); CNN (US); DW (Germany); Fox News (US); and RT Russia (Russia). We observe that religious polarization and hate crimes [3], corruption issues surface among the most disagreed trigrams in India [17]. BBC News, being a world news channel with an emphasis on British news, shows both international political controversies (e.g., the Iran nuclear deal or the Syrian war) and internal political crises (e.g., Brexit) among its most disagreed trigrams. CNN and Fox News paint an interesting contrast together. We observe that Fox is primarily fixated on entities (e.g., US President Joe Biden, Former US House Speaker Nancy Pelosi, Senate Majority Leader Chuck Schumer) while CNN disagrees more on issues (e.g., assault weapons ban or black lives matter). We further observe that the most disagreed trigrams in RT Russia indicate anonymous disapproval of Russian President Vladimir Putin.

Table 5: Predicting disagreement. Each model is trained on a balanced dataset of 16,000 examples with an equal number of *disagreed* and *not-disagreed* examples. The evaluations are conducted on a held-out set of 4,000 examples.

	\mathcal{D}_{cnn}	\mathcal{D}_{fox}	\mathcal{D}_{msnbc}
\mathcal{M}_{cnn}	$80.9 \pm 0.5\%$	$57.3 \pm 0.2\%$	$68.1 \pm 0.5\%$
\mathcal{M}_{fox}	$59.1 \pm 0.7\%$	$83.5 \pm 0.8\%$	$56.4 \pm 0.4\%$
\mathcal{M}_{msnbc}	$63.3 \pm 0.6\%$	$55.7 \pm 0.2\%$	$81.8 \pm 0.4\%$

3.4 Is Disagreement Learnable?

Table 3 indicates that there could be untold stories of anonymous negative reactions to world leaders and dislike information offer vital insights. We next investigate if it is possible to predict controversiality of a news item from video titles and descriptions. Prior literature has considered predicting controversial news using Facebook reactions [2] and early prediction of controversiality on Reddit threads [9].

We cast the problem of predicting disagreement as a binary classification task. We consider the videos in the top and bottom 25 percentile for a given news channel based on the disagreement score. Since disagreement lies on a continuum, we perform this quantization step to segregate the subsets of most-disagreed and least-disagreed videos. All videos featuring in the top 25 percentile and the bottom 25 percentile have the same labels *disagreed* and *not-disagreed*, respectively. Considering the extensive literature on political polarization in hyper-partisan news media [10,14,7], we investigate three YouTube channels: CNN, Fox News, and MSNBC. We use a BERT-based [6] text classifier for this prediction task.

Table 5 shows that for a given YouTube channel, predicting disagreement is a learnable problem. All models perform quite well on unseen examples from their own channels. However, when we try to predict disagreement on different news channels we observe political alignments coming into play. We observe both \mathcal{M}_{cnn} and \mathcal{M}_{msnbc} predict disagreement on each other’s videos better than that on \mathcal{D}_{fox} . Interestingly, as a side note, it is worth mentioning that despite being a far-right news channel from a different continent, \mathcal{M}_{fox} predicts disagreement better on ABC News Australia ($62.6 \pm 0.8\%$) than on CNN, and MSNBC. We conduct an additional experiment with RT Russia (\mathcal{M}_{rt}) and BBC Russia ($\mathcal{M}_{bbc-russia}$) (see Table 6). Again, we observe that model predictions

Table 6: Predicting disagreement on BBC Russia and RT Russia.

	$\mathcal{D}_{bbc-russia}$	\mathcal{D}_{rt}
$\mathcal{M}_{bbc-russia}$	$78.9 \pm 2.0\%$	$46.4 \pm 0.8\%$
\mathcal{M}_{rt}	$44.0 \pm 0.9\%$	$76.6 \pm 1.6\%$

on unseen data from one’s own channel are reliable. However, both \mathcal{M}_{rt} and $\mathcal{M}_{bbc-russia}$ perform worse than chance when predicting for each other.

The ability to reliably predict disagreements on a given channel suggests our dataset can predict signals in newer videos after YouTube’s policy change. The poor cross-dataset generalization shows the global audience’s diversity and conflicting opinions, which our dataset can help explore further.

4 Discussions and Ethics Statements

Despite a growing call for incorporating disagreements in machine learning models [18,5] and moving toward more transparent AI systems in academia, big-tech social web companies are tightening their grip on publicly available information. We see our dataset as an important contribution to capture a decade’s worth of disagreements from a global audience. This paper starts a dialogue on web preservation and presents a compelling argument that the whims of platforms can create information gaps that are difficult to bridge. We rest our case with a compelling concluding example.

Figure 4 presents a screenshot of a YouTube news video hosted by CNN in which the former president’s both-sidesism drew severe scrutiny. The video now only shows that it has 3K likes. However, what we do not see now is that more than one year back, the video already had 4K dislikes. With the video title and description as input, \mathcal{M}_{cnn} correctly predicts this video as a highly disagreed one. The disagreement score of the word **Charlottesville**, a city in Virginia that seldomly appeared on news without the connection of the car attack, is a staggering 0.35. Nearly 5.9% of the videos in our dataset have more dislikes than likes. In those videos, the disapproval is effectively the majority opinion. Figure 4 presents a compelling example where the dislikes raised a voice against hate which got lost due to YouTube’s policy change on dislike visibility.

4.1 Ethics Statement

We collected our dataset when dislike information was publicly available. We further note that the internet archive has been tracking more than 25% of these videos. Hence, some of these videos’ dislike information is still publicly viewable. Furthermore, respecting YouTube’s goal to protect small-content creators, our

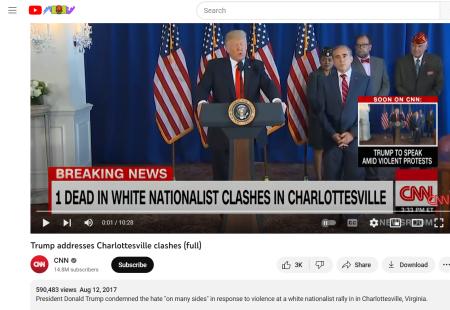


Fig. 4: A screenshot of a CNN video captured on March 1, 2023. This video was uploaded in the wake of the Charlottesville car attack in which a white supremacist ran over and killed a 35 year old woman protesting in favor of racial justice.

dataset precludes any individual content contributor and only considers popular news and opinion outlets that are no strangers to public dissent and TRP battles.

References

1. Bamman, D., O'Connor, B., Smith, N.: Censorship and deletion practices in chinese social media. *First Monday* (2012)
2. Basile, A., Caselli, T., Nissim, M.: Predicting Controversial News Using Facebook Reactions. In: CLiC-it (2017)
3. Basu, D.: Majoritarian politics and hate crimes against religious minorities: Evidence from India, 2009–2018. *World Development* **146**, 105540 (2021)
4. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., et al.: Xgboost: extreme gradient boosting. R package version 0.4-2 **1**(4), 1–4 (2015)
5. Davani, A.M., Dfaz, M., Prabhakaran, V.: Dealing with disagreements: Looking beyond the majority vote in subjective annotations. *TACL* **10**, 92–110 (2022)
6. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. In: NAACL-HLT 2019. pp. 4171–4186. Association for Computational Linguistics (2019)
7. Dutta, S., Li, B., Nagin, D.S., KhudaBukhsh, A.R.: A Murder and Protests, the Capitol Riot, and the Chauvin Trial: Estimating Disparate News Media Stance. In: IJCAI. pp. 5059–5065 (2022)
8. Garcia-Hodges, A.: YouTube suspends OANN for violating its Covid-19 policy (2020), <https://www.nbcnews.com/news/all/youtube-suspends-oann-violating-its-covid-19-policy-n1248845>, NBC News
9. Hessel, J., Lee, L.: Something's brewing! early prediction of controversy-causing posts from discussion features. In: Burstein, J., Doran, C., Solorio, T. (eds.) NAACL-HLT 2019. pp. 1648–1659 (2019)
10. Hyun, K.D., Moon, S.J.: Agenda setting in the partisan TV news context: Attribute agenda setting and polarized evaluation of presidential candidates among viewers of NBC, CNN, and Fox News. *Journalism & Mass Communication Quarterly* **93**(3), 509–529 (2016)
11. Jung, S.G., Salminen, J., J. Jansen, B.: The effect of hiding dislikes on the use of youtube's like and dislike features. In: ACM WebSci 2022. pp. 202–207 (2022)
12. Kennedy, J.J.: Maintaining popular support for the Chinese Communist Party: The influence of education and the state-controlled media. *Political Studies* **57**(3), 517–536 (2009)
13. Khan, M.L.: Social media engagement: What motivates user participation and consumption on YouTube? *Computers in human behavior* **66**, 236–247 (2017)
14. KhudaBukhsh, A.R., Sarkar, R., Kamlet, M.S., Mitchell, T.: We don't speak the same language: Interpreting polarization through machine translation. In: Proceedings of the AAAI Conference on Artificial Intelligence. pp. 14893–14901 (2021)
15. KhudaBukhsh, A.R., Sarkar, R., Kamlet, M.S., Mitchell, T.M.: Fringe news networks: Dynamics of us news viewership following the 2020 presidential election. In: 14th ACM Web Science Conference 2022. pp. 269–278 (2022)
16. Mitchell, T.M., et al.: Machine learning, vol. 1. McGraw-hill New York (2007)
17. Parry, J.: The crisis of corruption and the idea of India. *Morals of legitimacy: between agency and system* **12**, 27 (2000)

18. Pavlick, E., Kwiatkowski, T.: Inherent disagreements in human textual inferences. *TACL* **7**, 677–694 (2019)
19. Stockmann, D., Gallagher, M.E.: Remote control: How the media sustain authoritarian rule in China. *Comparative political studies* **44**(4), 436–467 (2011)