Politicization During the 2024 United States Presidential Elections

Marcelo Sartori Locatelli[1,2⋆], Matheus Prado Miranda[1⋆], Wagner Meira Jr.[1], and Virgilio Almeida[1]

[1] Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, Brazil
Tel: +55 (31) 3409-5860
{locatellimarcelo,matheus.prado,meira,virgilio}@dcc.ufmg.br
[2] Max Planck Institute for Security and Privacy, 44799 Bochum, Germany

Abstract. Topic shifts occur naturally during conversations when a person changes the subject to one different from the original. This phenomenon may be extremely meaningful, being useful for modeling dialogues and measuring information distortion during a conversation, among other tasks. In this work, we explore topic shifts as a metric for the measurement of politicization during the 2024 U.S. presidential elections, using YouTube news as a case study. We find evidence of politicization during the studied period as over 69% of non-political videos contain at least one political comment. This politicization increases as we get closer to the date of the election, with videos from right-leaning channels having over 40% of comments being political in the week of the elections. We also identify topics relating to immigration to be the most politicized, as commenters discuss the government's stance on immigration, aggravated by displays of xenophobia, exemplifying the dangers that come with politicization.

Keywords: Politicization, Social Media, Election

1 Introduction

During the last decade, social media data has been a target of frequent studies by researchers due to its ability to serve as a window to people's behavior, which, when combined with the ease of acquisition and abundance of such data, has led to the research and characterization of interesting phenomena, such as polarization [9], echo chambers [8], and filter bubbles [22].

One such phenomenon, whose study has benefited greatly from the increased amounts of data, is politicization. Politicization occurs when a topic that is not inherently political receives a political tone [24]. This ideological charge can lead to manipulation and a loss of public trust in experts' opinions. Studies on

⋆ Equal contribution.

politicization typically address specific topics (e.g., climate change [19], COVID-19 [12], science [3], etc.) or rely on traditional media as their object of study, failing to fully leverage the abundance of data generated online.

Given the risks associated with politicization and the impacts it can have on society, this study attempts to measure and uncover how politicization takes shape during important political events in a country. In this case, we study the 2024 U.S. presidential elections through the lens of politicization. Unlike most previous studies, we seek to take those measurements with a topic-agnostic methodology to better understand the full landscape of politicization in this context. However, as previously mentioned, commonly adopted methods in literature fail to scale to the proposed scope.

To address the aforementioned issues and provide a text-processing focused alternative, this research explores a way to measure politicization, using so-called *topic shifts*, a methodology developed to measure politicization on social media networks [17]. Topic shifts occur when a person changes the subject mid-conversation, potentially moving to a tangentially related or entirely unrelated topic. This has been used in other tasks, such as identifying relevant sentences for summarization tasks [14].

The underlying intuition behind the usage of topic shifts as an indicator of politicization is simple. Imagine a topic that is becoming politicized. By definition, this will be a topic receiving ideological and political tones, even if it was not originally so. Thus, for a politicized topic, given a non-political text posted on any online platform, responses to this text will tend to be more political compared to those on non-politicized topics, leading to more topic shifts.

In this work, we aim to understand politicization in the context of the U.S. presidential elections through the usage of topic shifts. In particular, we seek to answer the following questions:

- **RQ1**: Are there evidences of politicization during the studied period?
- **RQ2**: Are there any temporal patterns related to politicization during the studied period?
- **RQ3**: How does the content creator's political leaning influence politicization and topic shifts?
- **RQ4**: Which topics were the most and least politicized?

We find evidence that videos from right and, to a lesser extent, from left-leaning news sources tended to be more susceptible to politicization than those from center-leaning channels (RQ1). We believe that this may suggest a link between polarization and politicization as users flock to either side. This is corroborated by our analysis of temporal dynamics of topic shifts (RQ2), where we find that videos from left and right-leaning channels tended to receive topic shifts faster than those from the center. Additionally, we note that the right-leaning group tended to be the most associated with politicization, especially after Biden dropped out of the election, as seen in the peaks in the proportion of political comments (RQ3).

Finally, we identify the most and least politicized topics (RQ4), finding that there is a drastic difference in the amount of political comments received depending on the subject being discussed, with the most political topics having over a quarter of their comments relating to American politics. Topics of the so called hard-news tended to be more politicized. We see that topic shifts are frequently accompanied by an increase in toxic behavior, such as racism, further highlighting the raising concern over the necessity of moderating such behavior.

2 Related Works

Topic shift. The concept of topic shift, although under-explored, was introduced quite some time ago as a way to find new information in a text and has been applied to various scenarios since then. The earliest attempt at automatic detection to the best of our knowledge was [21], who used this concept to identify new information in online news, enabling the analysis of larger amounts of data.

Since then, the landscape of online data and computing has changed drastically, with massive amounts of data and increased computational power leading to new ways of detecting and interpreting topic shifts. Topal et al. [23] study the effect of this phenomenon in social media comments, specifically on Twitter, suggesting that abrupt topic changes in response to posts, combined with inflammatory comments, are detrimental to the user experience. They leverage the knowledge of this phenomenon to propose a detection methodology based on the discussion topic and the emotions expressed by users through comments, which could be used to filter out undesirable comments. Their results suggest that certain topics and emotions lead to higher occurrences of topic shifts.

Ermakova et al. [11] use topic shifts to demonstrate how true information about COVID-19 gets distorted throughout a sequence of comments. They find that as the topic shifts from medicine to politics and business, the information tends to become distorted, potentially leading to misinformation.

Politicization. The vast majority of studies on politicization are conducted qualitatively and focus on a single topic. Morten Bay [2] manually investigates tweets related to the movie "Star Wars: The Last Jedi", finding evidence of politicization and Russian bot activity. Peterson and Muñoz [20] use surveys to study the politicization of sports media, such as ESPN, over the years.

Quantitative studies, on the other hand, often limit themselves to analyzing keywords or hashtags, capturing mentions of political entities in news or social media. Hart et al.[12], and Chinn et al.[7] use this type of technique to study, respectively, the politicization of COVID-19-related posts on social media and climate change news from 1985 to 2017, finding signs of politicization correlated with an increase in mentions of Democratic and Republican politicians in the United States. A similar approach was used by Brown and Midberry [4] to study the politicization of drug addiction through keyword analysis.

Some researchers have explored alternative methods to measure politicization quantitatively. In [10], the researchers use topic modeling techniques to identify subjects related to partisan ideology. They find that topics related to ivermectin

Leaning	Videos	Comments	Channels	Unique Commenters
Any	21,075	9,324,489	68	3,648,739
Left	9,141	4,187,888	31	2,247,024
Center	4,229	2,495,698	12	1,130,910
Right	7,705	2,640,903	25	1,150,463

Table 1: Summary of the collected video statistics.

and other alternative COVID-19 treatments displayed more negative sentiment in Democratic states, providing evidence of politicization.

We extend previous research by focusing on the very recent 2024 American elections and by analyzing the characteristics of topic-shifting comments, deepening the understanding of this phenomenon.

3 Dataset

As politicization may manifest itself in how people react to news and other important events, we collect all videos from reliable English language news sources using the YouTube data API V3. In total, our dataset consists of 21,075 videos posted by 68 different channels, separated by political leaning ⋆⋆⋆, and over 9 million comments published during the period ranging from January 1 2023 to February 23 2025. Table 1 shows a summary of the video statistics. The structure of YouTube data is helpful to our purposes as the videos set a main topic to which the comments should respond to, making it easier to identify when users stray away from that topic.

We employed three annotators to label independent instances of 3000 videos and comments, which we shard into 5 different partitions to serve as train and test data for our models. Annotation guidelines were designed to identify political content, defined as any explicit or implicit reference to political candidates, parties, public offices, or ideologically charged entities. We assessed inter-rater reliability using Fleiss' Kappa [15], obtaining a score of 0.748. According to the interpretation scale proposed by Landis and Koch, this value indicates "substantial agreement" among annotators.

4 Detecting Topic Shifts

In order to measure how topics change in the direction of politics, we must first determine what is political. For that, we build a text based classifier that, given a comment or a video, is able to output how likely it is to be political. As we rely on the classifier to detect topic shifts, its performance on both categories is pivotal. We measure this performance in terms of the video and comments'

⋆⋆⋆ https://www.allsides.com/media-bias/media-bias-chart

F1, which equates to the harmonic mean between precision and recall, as we are using a binary classifier.

$$F_1 = 2\frac{precision \cdot recall}{precision + recall}$$

The performance metric F1 was chosen because we want to avoid detecting false topic shifts to politics on non-political videos (comment misclassified as political or video misclassified as non-political), as well as detecting topic shifts to non-politics on political videos (comment misclassified as non-political or video misclassified as political), thus, both precision and recall are important.

We first test the two-step Positive and Unlabeled(PU) learning algorithm proposed in [17]. The idea behind this algorithm is that it is fairly easy to have an initial set of posts and comments that are reliably political by simply using a set of keywords, while it is much harder to determine an initial set of non-political text. However, by leveraging the spies PU learning technique [16], using a weaker classifier such as naive bayes, we can obtain a set of reliably non-political pieces of text, which we can finally combine with our initial set of positive examples in order to train a stronger classifier (in the original paper, a XGboost [6] classifier).

For this work we use a set of keywords relevant to the U.S. presidential elections to define the initial positive dataset. They range from specific American political figures to general political terms[†]. The list was curated through manual inspection of common political terminology in U.S. discourse, supported by prior work [17]. Since this method does not rely on our human-labeled training data, we use the training partitions of our manually labeled sample solely to define the model hyperparameters in a random search setting, with the test partition being used for evaluation, we repeat this 5 times with different test sets and aggregate the results when reporting.

As LLMs have proven to have great performances for most text based tasks, we also compare the results of the PU classifier with a more modern approach using a fine-tuned RoBERTa model leveraging the textual elements of videos and comments. For video classification, we simply use the title and description as inputs for the model, while for comments we attempt two approaches: the first uses solely the content of the comment for the classification, while the second incorporates the context of the video into the prediction, truncating when necessary. Figure 1 gives an example of the classification pipeline. The idea here is to leverage the fact that our dataset consists of video and comment pairs, which allows us to make use of the video title and video description when classifying a comment. This is beneficial in the detection of more nuanced comments. For example, the comment "I agree with this person" would be political in the context of a video titled "Watch as the president gives his discourse" and would not if it was posted under a video titled "We need to stop mistreating dogs". The PU classifier described above would treat these two comments as the same, although the context from the video completely changes the topics to which the comments

[†] The full list of keywords is: "trump", "biden", "kamala", "vance", "republican", "democrat", "party", "election", "debate", "president".

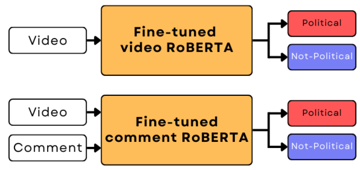


Fig. 1: We use two fine-tuned RoBERTa classifiers. The first (top) is responsible for classifying the video as political or non-political using its title and description as input. The second (bottom) classifies the comment using the comment text and the textual elements of the video as inputs, contextualizing the comment.

refer, an aspect that we wish to leverage on the RoBERTa classifier. We use a 5-fold cross validation approach and train the models for 5 epochs.

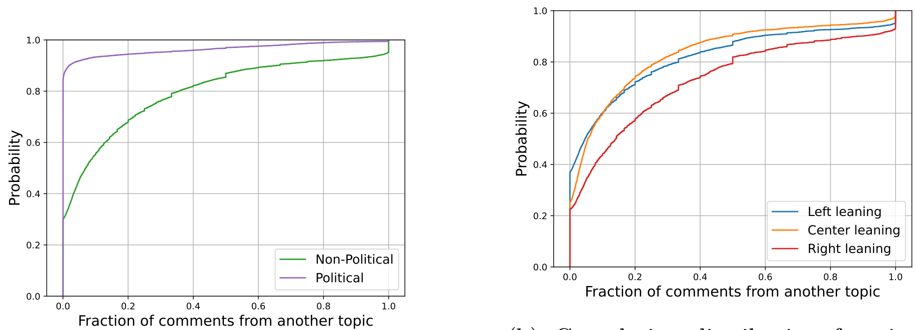
Table 2 shows the performance of either models on the political classification task, with a 99% confidence interval. We can see that the RoBERTa classifier outperforms the PU model when we incorporate the video context to the comment prediction, while the performance without this modification is even worse than the two-step PU classifier. To make the results more robust, we also conducted an unpaired t-test aiming to ascertain if there is any statistical difference between the two methods. We obtain the CIs -0.100 ± 0.02 and -0.197 ± 0.07 for the difference of means of the video and comment classifiers when comparing the PU method with the LLM based method, therefore the test passes with $p < 0.01$, demonstrating the superiority of the proposed approach for the task. We believe that, despite the errors in model classification, the relatively high classifier performance would still lead to tendencies being visible when we analyze large amounts of data.

5 Characterizing Topic Shifts

We use the best performing RoBERTa model to attribute a political or non-political label to all comments and videos on our dataset, classifying 65% of the comments and 60% of the videos as political. This suggests that most of the

	Video F1	**Comment F1**
Two-step PU learning	0.812+-0.029	0.668+-0.059
RoBERTa	0.918+-0.011	0.775+-0.089
+ **Video Context**	N/A	0.865+-0.031

Table 2: Performance of the two-step PU learning and fine-tuned RoBERTa model on the test set. The $\pm$ term represents the 99% confidence interval for the mean. The RoBERTa model far outperforms the two-step PU approach for the English YouTube news dataset.



(a) The cumulative distribution function (CDF) shows the proportion of topic-shifting comments posted on YouTube videos. We observe that approximately 80% of political videos have no topic-shifting comments, while around 80% of non-political videos contain topic-shifting comments in more than 80% of their total comments.

(b) Cumulative distribution function of the fraction of topic-shifting comments posted on non-political YouTube videos per leaning. Videos from center leaning channels seem to be the least susceptible to topic shifts, closely followed by left-leaning channels, while videos from right-leaning channels seem to suffer from it the most.

Fig. 2: Cumulative distribution functions (CDF) for topic shifts.

news published by reputable news sources in English on YouTube relate to U.S. politics. These 65% of political comments, however, are not restricted solely to political videos. We estimate that 64% of the non-political videos contain at least one political comment, which could be a sign of politicization as people start to bring American politics into most topics.

Figure 2a illustrates the fraction of topic-shifting comments for political and non-political videos. Political videos tend to maintain a higher topical coherence in their discussions, with a comparatively lower incidence of comments shifting to unrelated topics. This suggests that explicit political framing effectively directs audience discourse, maintaining focus on the intended political themes. Non-political videos, however, show a higher variability, with a considerable proportion containing moderate to high fractions of political comments.

When examining how the leaning of news sources affects topic shifts, an initial investigation reveals that videos from neutral news sources receive fewer topic-shifting comments in general - see Figure 2b. Although the distribution of left-leaning channels resembles that of center leaning channels more closely, both left-leaning and right-leaning sources appear to exhibit a higher incidence of topic-shifting comments. We hypothesize topic shifting towards politics might be related to polarization, and thus, in the more polarized right and left-leaning contexts, one would be more likely to observe an out of topic political comment.

Topic	Representative Words	Political Comments	Sample
Immigration	Migrant, Immigration, Mexican.	47% ± 0.3%	**News.**"A Growing number of Chinese migrants seek asylum at U.S. Mexico border" **Politicized Comment.**"Smart China, conquer US with voters instead of weapons."
Earthquakes	Earthquake, Syria, Turkey.	46% ± 0.9%	**News.**'Shellshocked: Southwest Turkey and Syria Rocked Again by Earthquake | Jerusalem Dateline" **Politicized Comment.** "Politicians don't have brains"
TikTok Regulation	Tiktok, Ban, Regulate.	35% ± 1.4%	**News.**'This is What TikTok Does to Your Child" **Politicized Comment.**"... people who make the decisions are in favor of it. Saying that Liberals and Conservatives can come together on this doesn't matter when one of them wants this to continue... "
Space	Spacex, Starship, Rocket.	31% ± 0.4%	**News.**"Las Vegas family that reported aliens stands by original story | Banfield" **Politicized Comment.**"The footage they took was already confiscated, and the threats they are getting are from the secret government organizations running this technology..."
Natural Disasters	Storm, Hurricane, Snowstorm.	28% ± 0.4%	**News.**"11 million people under flood alerts in California | WNT" **Politicized Comment.**"Meanwhile USA is under attack and Biden could care less! for all! Geoengineered?!"

Table 3: Most politicized non-political YouTube news topics. Disasters tend to get consistently politicized, possibly due to a mix of climate activism, conspiracies and government blaming. The ± term indicates the 99% proportion confidence interval.

5.1 Politicization of Topics

After finding evidence of politicization during the studied period, the next step is to discern which topics were more politicized. To do so, we separate the non-political YouTube news and use BERTopics, a topic modeling technique based on sentence embeddings with c-TF-IDF, to produce interpretable non-political topics. Having obtained the topics for the published videos, we then find the percentage of topic-shifting comments that appear on each of them to identify the most and least politicized topics. Since the topics are highly interpretable, at this stage we can manually exclude some uninteresting or likely to be misclassified topics. In our case, we exclude one topic related to daily, weekly, and monthly news, since videos from this topic tend to contain political content even if their title or description is unrelated. An example of such video title would be "Top U.S. & World Headlines — December 11, 2024". In total, we identified 28 different topics that made up the news between the years of 2023 and 2024, including technology, crime, travel, food, natural disasters, and many others.

Table 3 and 4 show the top-5 most and least politicized topics. We find that immigration is the most politicized topic, with upwards of 45% of the comments on those videos being classified as relating to politics much of this politicization traces back to foreign citizens - especially Latin American, Chinese and Islamic people - possibly due to the consistent clashes between the Republican and Democratic party on their stance around immigration and deportations, with racism possibly playing a part (§5.2). The politicization of disasters such as earthquakes, storms, floods and fires (24% political comments, not in the table) is interesting as it appears as an intersection from many factors, including attributing blame to governments, conspiracy theories, and climate activism, leading to the high rate of political comments.

We also see how politicization leads to fear and distrust, with many of the news being called fake news even when they come from reputable news sources

Topic	Representative Words	Political Comments	Sample
Relationships	Divorce, Marriage, Part.	2% ± 0.5%	**News.**“Secrets To A Successful Marriage" **Politicized Comment.**“Honestly. Some comments are just nonsense. And no I'm not leftist. I'm very much conservative. ..."
Oceangate	Titanic, Submarine, Shipwreck.	6% ± 1.2%	**News.**“I'm fearful': Submarine commander on missing Titanic vessel" **Politicized Comment.**“It's ALLL TRUMP'S Fault! Trump will be indicted for this soon!"
Weight	Diet, Obesity, Weightloss.	8% ± 1.3%	**News.**“Ozempic Underworld The Black Market Of Weight Loss Drugs" **Politicized Comment.** “Blame your government regulations not the pharmaceutical companies! They aren't breaking any laws"
Automation	AI, Ai-powered, Robots.	8% ± 0.3%	**News.**“How The Massive Power Draw Of Generative AI Is Overtaxing Our Grid" **Politicized Comment.**“Trump: buy Greenland, take Canada"
Economy	Inflation, CPI, Recession.	9% ± 0.5%	**News.**“How The U.S. Is Stalling A Recession" **Politicized Comment.**“Vote Democrats OUT!"

Table 4: Least politicized non-political YouTube news topics. Notice how the topics are more varied than the most politicized ones. The ± term indicates the 99% proportion confidence interval.

in addition to conspiracy theories claiming that crimes and reports are just a political tool to mislead the population. In fact, 14% ± 1% of non-political videos have at least one comment containing the term “fake news”.

Interestingly, some topics frequently associated to politics were not as politicized during the studied period, for instance, videos related to inflation, recession and other economic processes had only around 10% political comments. On a manual investigation, many of the comments are actually referencing corporate greed or suggesting solutions to these processes rather than focusing on State policies or politicians. Automation appearing with only 8% political comments is also surprising, especially as advances such as the Chinese deepseek have sparked new discussions around the China-U.S. competition, leading to pronunciations from important politicians such as D. Trump[‡], who said the model was a "wake-up call" to U.S. industry. This illustrates the difficulty in studying politicization, showing that different facets of a topic may be politicized, but they will not necessarily represent the topic as a whole.

Although we show samples of politicized comments on the least politicized topics, those are much harder to find in our actual dataset. However, they serve to illustrate that on a more politicized context, you can find political discussion even on unassuming topics such as relationship discussions or weight-loss tips.

As politicization is a process, it is important to understand how it varies with time. For this reason, we explore both the speed of topic shifts (i.e. how long it takes for a non-political video to receive its first political comment) as well as the overall proportion of topic-shifting comments on our database. Figure 3a shows the variation in proportion of political videos posted on the database throughout the year. It is possible to see that, especially for videos from right-leaning channels, the proportion of political comments drastically increases after Joe Biden drops out of the presidential race, peaking at the week of the elections.

[‡] https://www.reuters.com/world/us/trump-deepseeks-ai-should-be-wakeup-call-us-industry-2025-01-27/

This same tendency is not observed on the left-leaning videos, whose percentage of politic comments remained steady throughout the year, showing a small peak after the elections, possibly due to discussions arising after Kamala's defeat. We further discuss the characteristics of the political discourse outside political videos in Section 5.2.

Figure 3b shows the median time (in comments) before a topic-shifting comment is posted on a video. From this figure, we can see that videos from center leaning channels, in general, had longer periods without topic-shifting comments. This corroborates with our hypothesis that topic shifts might be related to polarization, as the news channels whose views more closely aligned with the U.S. parties, be it Democratic or Republican, tended to be topic shifted faster. Additionally, we note that throughout most of the studied period, right-leaning videos tended to receive topic-shifting comments faster, which, coupled with the higher ratio of political comments during most weeks and the higher fraction of videos with a high number of topic-shifting comments (see Figure 2b), suggests that right-leaning content is more associated with politicized discourse.

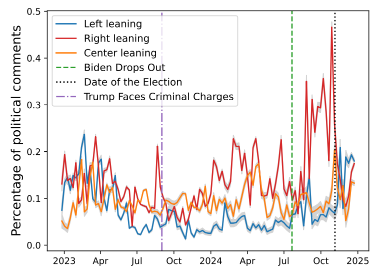
5.2 Characteristics of topic-shifting Comments

Finally, we conduct an investigation on what is driving those topic shifts. For this, we define Democratic leaning word set as "democrat", "Kamala", "Harris", "left-wing" and Republican leaning word set as "republican", "Donald", "Trump", "right-wing". Using the VADER [13] sentiment analyzer, we obtain the sentiments of topic-shifting comments containing each of those sets. The idea behind this experiment is to find which group is driving politicization: the republicans or democrats. If the right-leaning groups are responsible for the topic shifts, we expect the sentiments towards the Republican set to be positive and Democratic set to be negative. The opposite is true for the left-leaning group.

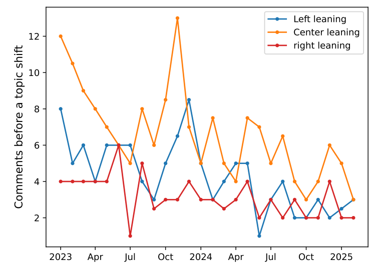
We find that on average, topic-shifting comments are slightly more negative than non-topic-shifting comments (t-test, $p<0.001$), however we find no practical differences between the sentiments of comments containing words from the Republican or Democrat word sets. We also do not find any evidences of the leaning of the video channel influencing the sentiment of the topic-shifting comments. Thus, the hypothesis that a specific group might be more responsible for topic shifts seems unlikely.

To further understand what issues may be fueling these changes in topic, we analyze the toxicity of topic-shifting comments, focusing on different categories that may be relevant for politicization. We detect toxicity using a RoBERTa fine-tuned model [1] chosen due to its ability to discern between the targets of hate speech, categorizing toxic comments into six different categories: "disability", "racism", "religion", "sexism", "sexual orientation" or "other".

We find that topic-shifting comments are considerably more likely to be toxic than average, with more than 3.5% of them being categorized as toxic compared to the only 1.5% of non topic shifts. Interestingly, 85% of the topic-shifting comments classified as toxic are racist. This is consistent with our topic analysis findings, which indicate that many politicized topics may be traced back to the



(a) Percentage of topic-shifting comments posted on YouTube non-political news videos per week. The gray area indicates the 99% proportion confidence interval.



(b) Median number of comments before a non-political video receives a topic-shifting comment per month.

Fig. 3: Temporal patterns for topic shifts.

presence of foreign citizens or cultures, serving as further evidence of the possible harms that come with politicization.

Figure 4 shows the proportion of comments belonging to each hate speech category, divided by different time frames: before Biden drops out (until July 24, 2024), after Kamala joins the race (before November 5) and after the day of the elections (remainder of the studied period). We find that among non-topic-shifting comments, the amount of toxicity remained roughly the same throughout the entire studied period. Additionally, for these comments, there was a similar proportion of sexism and racism. However, when we look at the topic-shifting comments we see that they are very skewed towards racism, with other kinds of hate-speech being less prevalent than their non-topic-shifting counterparts. Additionally, we see a considerable increase of racist comments after Kamala joins the presidential race, possibly a reflection of the inflammatory comments relating to her Black and Indian background.

Interestingly, besides racism and sexism, we also see a relatively high amount of hate speech related to religion. We hypothesize that these are related to the many ongoing conflicts related to religion, such as the Israel-Palestine war and the controversies regarding Muslim immigrants. Consequently, religious tensions appear to intersect notably with political discussions, potentially exacerbating polarization within online discourse.

5.3 Semantic Polarization

To complement our classifier-based approach to politicization, we adopt a lexical-semantic perspective using the Word Embedding Association Test (WEAT), originally introduced in Caliskan et al. [5]. WEAT quantifies differential associations between two sets of target concepts (e.g., ideological groups) and two sets

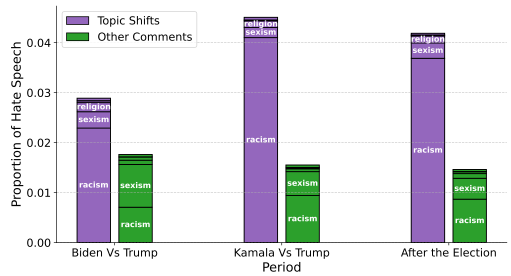


Fig. 4: Proportion of toxic comments per hate speech category, for different time periods of the US elections.

of attribute words (e.g., moral valence) by computing cosine similarity-based statistics in word embedding spaces. It provides an interpretable metric—effect size (Cohen's d)—and a permutation-based significance test, enabling the measurement of implicit semantic biases in text corpora.

This methodological choice is motivated by prior work utilizing WEAT to detect ideological framing in online discourse. In particular, Ottoni et al. [18] leveraged WEAT to measure implicit associations in YouTube content, revealing meaningful semantic biases in both comment and video data from right-wing channels. Adapting this framework to our setting, we apply WEAT to independently trained word embedding spaces, learned using Word2Vec models built from our data. It consists of topic-shifted and non-political comments, enabling a comparative analysis of their respective ideological associations.

We define multiple ideologically relevant WEAT tests reflecting thematic axes derived from our dataset topics. These axes, presented in table 5, were constructed by analyzing lexical patterns in comments, focusing on recurrent contrasts in group references and evaluative language across highlighted topics of our work. We also incorporate a pretrained Word2Vec model trained on the Google News corpus as a general-purpose semantic baseline, approximating a neutral Wikipedia-like embedding space for comparison.

Following the results presented in figure 5, we find that topic-shifted comments tend to exhibit higher semantic polarization in politically charged associations. Among the tested axes, Racism and Natural Disaster reveals a clear distinction between topic-shifted and other groups. The higher effect size observed for the shifted group suggests that political content introduced into non-political contexts tends to reinforce more extreme or stereotyped associations along these semantic axes. Furthermore, the evaluative language in these associations tends to disproportionately target the discriminated class with negative attributions.

In contrast, Immigration emerges as a persistently polarized axis across all groups, including non-political and caption-based comments. In this sense, immigration-related associations are deeply embedded in online discourse, independent of whether the surrounding content is explicitly political. Wikipedia, serving as a baseline, exhibits the expected lower effect sizes compared to our

Group	Immigration	Natural Disaster	Racism
Class 1 *(Discriminated)*	immigrant, refugee, mexican, border, migrant	flood, fire, earthquake, hurricane, tornado	black, african, minority, brown, mexican
Class 2 *(Dominant)*	citizen, native, patriot, local, resident	government, biden, trump, state, politicians	white, european, american, caucasian, majority
Attributes 1	criminal, dangerous, invader, illegal, enemy	unavoidable, natural, random, accident, fate	criminal, dangerous, gang, violent, illegal
Attributes 2	peaceful, kind, honest, innocent, hardworking	avoidable, accountable, corrupt, neglect, failure	peaceful, honest, friendly, worker, neighbor

Table 5: Word sets used for each class and attribute in our Word Embedding Association Tests (WEATs).

dataset for the Immigration and Natural Disaster axes. Its relatively high score on the Racism axis, however, suggests that certain biased associations — particularly those involving race — may be more deeply embedded in the general language space. This aligns with our findings of racism as an extreme and prominent category within hate speech discourse, exceeding merely frequent political targeted figures like Kamala Harris.

6 Discussion

In this work we explored the politicization of YouTube news videos during the period leading up to the 2024 American presidential elections. Focusing on a set of reputable English speaking news sources,we analyze over 8 million comments published between the start of 2023 and the end of 2024. To allow for our analysis, we develop a political classifier leveraging the capabilities of a fine-tuned RoBERTa model, achieving 0.9 and 0.86 F1 for classifying videos and comments, respectively and use its outputs to detect political comments on non-political videos which serve as examples of topic shift.

While YouTube is the focus of this study, it is important to consider that the platform has unique characteristics—such as video-based content, recommendation algorithms, and a specific user demographic—that may influence politicization dynamics. Therefore, while the observed trends are strong on YouTube, future work should aim to validate whether similar patterns emerge on other social media platforms, such as Twitter, Reddit, or TikTok, which have different content formats, audience behaviors, and moderation approaches.

We find that non-political videos frequently contain topic-shifting comments, with over 69% of our dataset containing at least one political comment. This suggests a widespread tendency of online discourse to introduce politics to unrelated topics, often increasing in response to important political events, such as Biden dropping out of the presidential race. These tendencies of politicization were especially prevalent among right-leaning channels, with center channels being the

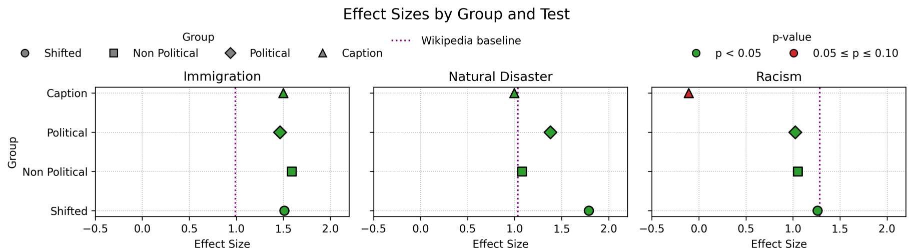


Fig. 5: Effect sizes (Cohen's d) for WEAT tests across content groups. Shapes indicate source type, colors denote significance levels (p-values), and the dashed line shows the Wikipedia baseline.

least affected. This suggests that media polarization might play a role in the phenomenon of politicization, although further studies are required to confirm this hypothesis.

We also find that topics relating to hard news tended to be more politicized. On our dataset, this is predominantly seen in the form of immigration related videos, which were the most politicized topic. On manual investigation, topic-shifting comments on those videos commonly contained a mix of xenophobia, racism and conspiracy theories, further highlighting the dangers of politicization as it amplifies social divisions and fuels misinformation. Complementing these findings, we observe that topic-shifted comments encode stronger biased associations than other groups, particularly in relation to race and institutional trust, highlighting the semantic intensity of politicization.

In conclusion, this research provides valuable insights into the mechanisms and consequences of politicization in online discussions. As digital platforms continue to serve as primary venues for political discourse, understanding how and why non-political topics become politicized is crucial for mitigating polarization and fostering healthier online interactions.

Acknowledgement. This work was partially funded by CNPq, CAPES, FAPEMIG, and IAIA - INCT on AI.

References

1. Antypas, D., Camacho-Collados, J.: Robust hate speech detection in social media: A cross-dataset empirical evaluation. In: The 7th Workshop on Online Abuse and Harms (WOAH). pp. 231–242 (2023)
2. Bay, M.: Weaponizing the haters: The last jedi and the strategic politicization of pop culture through social media manipulation. First Monday (2018)
3. Bolsen, T., Druckman, J.N.: Counteracting the politicization of science. Journal of Communication **65**(5), 745–769 (2015)
4. Brown, D.K., Midberry, J.: Social media news production, emotional facebook reactions, and the politicization of drug addiction. Health communication **37**(3), 375–383 (2022)

5. Caliskan, A., Bryson, J.J., Narayanan, A.: Semantics derived automatically from language corpora contain human-like biases. Science **356**(6334), 183–186 (Apr 2017). https://doi.org/10.1126/science.aal4230, http://dx.doi.org/10.1126/science.aal4230
6. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proc. of the KDD. pp. 785–794. ACM (2016). https://doi.org/10.1145/2939672.2939785, http://doi.acm.org/10.1145/2939672.2939785
7. Chinn, S., Hart, P.S., Soroka, S.: Politicization and polarization in climate change news content, 1985-2017. Science Communication **42**(1), 112–129 (2020)
8. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., Starnini, M.: The echo chamber effect on social media. Proceedings of the National Academy of Sciences **118**(9), e2023301118 (2021)
9. Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., Flammini, A.: Political polarization on twitter. In: Proc. of the ICWSM. vol. 5, pp. 89–96 (2011)
10. Diaz, M.I., Hanna, J.J., Hughes, A.E., Lehmann, C.U., Medford, R.J.: The politicization of ivermectin tweets during the covid-19 pandemic. In: Open Forum Infectious Diseases. vol. 9, p. ofac263. Oxford University Press (2022)
11. Ermakova, L., Nurbakova, D., Ovchinnikova, I.: Covid or not covid? topic shift in information cascades on twitter. In: 3rd RDSM. pp. 32–37 (2020)
12. Hart, P.S., Chinn, S., Soroka, S.: Politicization and polarization in covid-19 news coverage. Science communication **42**(5), 679–697 (2020)
13. Hutto, C., Gilbert, E.: Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Proc. of the ICWSM. vol. 8, pp. 216–225 (2014)
14. Konigari, R., Ramola, S., Alluri, V.V., Shrivastava, M.: Topic shift detection for mixed initiative response. In: Proc. of the SIGDIAL. pp. 161–166 (2021)
15. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics **33**(1), 159–174 (1977), http://www.jstor.org/stable/2529310
16. Liu, B., Lee, W.S., Yu, P.S., Li, X.: Partially supervised classification of text documents. In: ICML. vol. 2, pp. 387–394. Sydney, NSW (2002)
17. Locatelli, M.S., Calais, P., Miranda, M.P., Junho, J.P., Muniz, T.L., Meira Jr, W., Almeida, V.: Topic shifts as a proxy for assessing politicization in social media. In: Proc. of the ICWSM. vol. 18, pp. 972–984 (2024)
18. Ottoni, R., Cunha, E., Magno, G., Bernardina, P., Meira Jr., W., Almeida, V.: Analyzing right-wing youtube channels: Hate, violence and discrimination. In: Proc. of the ACM Websci. p. 323–332. WebSci '18, ACM (2018). https://doi.org/10.1145/3201064.3201081
19. Pepermans, Y., Maeseele, P.: The politicization of climate change: problem or solution? Wiley Interdisciplinary Reviews: Climate Change **7**(4), 478–485 (2016)
20. Peterson, E., Muñoz, M.: "stick to sports": Evidence from sports media on the origins and consequences of newly politicized attitudes. Political Communication **39**(4), 454–474 (2022)
21. Radev, D.R.: Topic shift detection-finding new information in threaded news (1999)
22. Ross Arguedas, A., Robertson, C., Fletcher, R., Nielsen, R.: Echo chambers, filter bubbles, and polarisation: A literature review (2022)
23. Topal, K., Koyuturk, M., Ozsoyoglu, G.: Emotion-and area-driven topic shift analysis in social media discussions. In: 2016 IEEE/ACM international conference on advances in social networks analysis and mining. pp. 510–518. IEEE (2016)
24. Wiesner, C.: Introduction: rethinking politicisation in politics, sociology and international relations. Springer (2021)