



Disinformation on Youtube: A Dataset of Youtube Comments on Videos Related to Claims Made by Trump and Vance on Haitian Immigrants

RYAN M. OMIZO 

DISCUSSION PAPER

]u[ubiquity press

ABSTRACT

The *YouTube Disinformation Comments Corpus* provides 71,025 comments to 23 YouTube videos posted by US news organizations on the false claim that Haitian immigrants in Springfield, Ohio were stealing and eating neighborhood pets, which was picked up and amplified by the 2024 Republican presidential ticket of Donald Trump and J.D. Vance. Additionally, the *YouTube Disinformation Comments Corpus* features metadata associated with each YouTube video, including video titles, descriptions, transcripts, like counts, and YouTube content tags. This dataset is useful for people studying (1) natural language processing, (2) network textual analysis, (3) the circulation of dis- and misinformation on YouTube; (4) the reception of dis- and misinformation on YouTube; (5) news framing effects on YouTube and (6) public discussions of US immigration. TFIDF analysis is used to provide an overview of the variations across comment data.

CORRESPONDING AUTHOR:

Ryan M. Omizo

Department of English, Temple
University, Philadelphia,
United States

tuk35906@temple.edu

KEYWORDS:

YouTube comment data;
digital humanities; natural
language processing; political
polarization

TO CITE THIS ARTICLE:

Omizo, R. M. (2025).
Disinformation on Youtube: A
Dataset of Youtube Comments
on Videos Related to Claims
Made by Trump and Vance on
Haitian Immigrants. *Journal
of Open Humanities Data*, 11:
12, pp. 1–9. DOI: [https://doi.
org/10.5334/johd.283](https://doi.org/10.5334/johd.283)

(1) CONTEXT AND MOTIVATION

The *YouTube Disinformation Comments Corpus* dataset concerns a smear-campaign maligning Haitian immigrants to Springfield, Ohio amplified by the 2024 Republican nominees for President and Vice President, Donald Trump, and J.D. Vance, respectively. In early September 2024, Springfield, Ohio resident, Erika Lee published a Facebook post in which she claimed that her neighbor's daughter's cat had been stolen and butchered by a family of Haitian immigrants. This claim was then reposted on Twitter on September 5, 2024, where it was amplified across right-wing media ecosystems like Fox News (see Howard and Brewster, 2024 September 12). GOP vice presidential nominee, J.D. Vance then posted to X the following:

Months ago, I raised the issue of Haitian illegal immigrants draining social services and generally causing chaos all over Springfield, Ohio.

Reports now show that people have had their pets abducted and eaten by people who shouldn't be in this country. Where is our border czar? (@JDVance, 2024 September 9)

Vance's Tweet has garnered 11.4 million views and 27.9 thousand shares through Retweets and Quoted Tweets. Then, during the September 10, 2024, US presidential debate, Trump repeated the story, stating "They're [Haitian immigrants in Springfield, OH] eating the dogs. They're eating the cats. They're eating the pets of the people that live there," (qtd in Catalini et al., 2024). Such statements have been roundly debunked by state official and journalistic outlets (Orazco, 2024; Lozano, 2024; Olmstead, 2024; Timotija, 2024). Nevertheless, Trump, Vance, and GOP supporters have maintained the public lie, with Vance admitting to CNN's Dana Bash that he and Trump are knowingly spreading falsehoods about Haitian immigrants to raise awareness about real issues (Maher and Boyett, 2024; Olmstead, 2024). Consequently, narratives of Haitians abducting and eating pets in Springfield, OH can be considered political disinformation—deliberate deceit intended to agitate anti-immigrant animus for political gain.

This data was collected as part of an ongoing research project in the spread of political disinformation on YouTube.

(1.1) LIMITATIONS AND CAVEATS

This corpus is not meant to be a representative sample of mainstream news coverage of the Haitian disinformation campaign. Videos with less than 500 comments were ignored. No consideration was made for the type of news format contained in the video (for instance, on the scene reporting versus interview), national versus local reporting agency, or replication in the corpus. By extension, the results of the Term Frequency Inverse Document Frequency (TFIDF) analysis used in this study should only be taken as relevant to the comments in the corpus, not the mainstream media or comment landscape on YouTube as a whole.

(2) DATASET DESCRIPTION

REPOSITORY LOCATION

<https://zenodo.org/records/14166028>

REPOSITORY NAME

Zenodo

OBJECT NAME

YouTube Disinformation Comments Corpus

FORMAT NAMES AND VERSIONS

CSV

CREATION DATES

2024-09-13–2024-11-14

LANGUAGE

English

LICENSE

Creative Commons Attribution 4.0 International

PUBLICATION DATE

2024-11-14

The *YouTube Disinformation Comments Corpus* captures 71,025 comments on and metadata from 23 YouTube videos posted by the official YouTube accounts of ABC News, CBS News, CNN, Fox News, MSNBC, New York Post, and Forbes news outlets that reported on the debunking of J.D. Vance's claim that Haitian immigrants were eating neighborhood pets in Springfield, Ohio soon after the first US presidential debate between Kamala Harris and Donald Trump. The corpus contains three files. First, the *youtube_haitian_disinformation_videos_meta.csv* file includes comments and YouTube video metadata. Data are organized around per video information. The columnar values are:

- *video_id*
- *date* (video publication date)
- *title*
- *description*
- *channel_title*
- *transcript*
- *transcript_str* (Video transcript without timestamps)
- *views*
- *likes*
- *comments* (all comments per video)

Second, the *youtube_haitian_disinformation_comment_reply_metadata.csv* file includes comments, replies, and comment metadata. Each comment occupies its own row in the spreadsheet. The columnar fields are:

- *video_id*
- *comment*
- *comment_date*
- *comment_like_count*
- *author*
- *comment_id*
- *in_reply_to*
- *neg, neu, pos, compound* (VADER polarity scores)
- *named_entities* (spaCy named entity tags with tokens)
- *emoji* (spaCy Emojis and token spans)

Comments and associated metadata are represented in individual rows.

The third file, *youtube_disinfo_comments_tfidf_results_per_video.csv*, contains the results of the TFIDF analysis described herein. The TFIDF analysis features the top 5000 terms weights for the comments to each video in a .csv file. For space considerations, I only display the top-10 results of the TFIDF analysis in this paper (Table 1) and direct readers to the complete tabulations on the dataset.

(3) METHOD

Video metadata and comments were downloaded using YouTube's Academic API between September 10, 2024 and October 2, 2024. Videos were selected if they were uploaded by a major

news outlet (such as ABC or CNN) and featured comment counts greater than 500. Metadata for the corpus features the YouTube video id, publication date, video description, transcripts with and without timestamps, view count, like counts, comment counts, and YouTube tags. Comments, replies to comments, and comment metadata were also collected using the YouTube Academic API. Comment and reply metadata include the date of publication, like count, comment author name, comment id, and the account name of the comment receiving replies.

To enrich the data, I extracted named entities and emojis from comments using spaCy (Honnibal et al., 2020). I also computed comment sentiment polarity using VADER (Hutto and Gilbert, 2014). I include negative (*neg*), neutral (*neu*), positive (*pos*), and compound scores for each sample.

To describe the comments in the corpus, I employ TFIDF analysis to surface the most definitive words in comments to each video (see Gallagher et al, 2023a; Gallagher et al., 2023b). TFIDF is a well-established method in information retrieval, which represents documents as a series of term weights. Weights for terms are a function of a term’s frequency in a document compared to the frequency of the term across the entire corpus. Terms common across the corpus are downweighted. Terms that are specific to a document are elevated (Manning and Schutze, 1999). For this study, I vectorize the comments and label each comment by its associated video_id. I use the video_ids as class labels and calculate the most informative features for each class using the chi-square test of significance. Key terms such as “harris,” “trump,” “donald,” “vance,” “j.d.,” “kamala,” and “president” were removed because they refer to content already manifest in the general description of videos.

(4) RESULTS AND DISCUSSION

I present the top-10 ranked terms for each YouTube video in Table 1 below as an overview of comments. To contextualize the top bigrams/trigrams, I present example comments for select video in which the bigram appears. Readers are encouraged to peruse the data repository to view additional comments, video data citations and more extensive TIFDF information.

id ID	TITLE	SOURCE/DATE	EXAMPLE COMMENT	TOP-10 TFIDF TERMS
0VSV99PXbbw	Why Trump keeps pushing debunked claim about Haitian immigrants	CNN, 2024b, Sep 14	Hunter Biden Laptop 2.0 Russia Collusion 2.0 Its all pre-election spin to manipulate voters, and then sometime in January the NYT will put out an article saying that the illegals are in fact eating local wildlife and pets. (@Zio-e5t, 2024, Sep 16)	(‘fake news’, 0.47), (‘elon musk’, 0.09), (‘laura loomer’, 0.08), (‘cnn joke’, 0.08), (‘russia russia’, 0.08), (‘proven true’, 0.07), (‘neo nazi’, 0.06), (‘white supremacist’, 0.05), (‘comment section’, 0.05), (‘blood libel’, 0.04), (‘clown news’, 0.04), (‘cnn communist’, 0.04), (‘news channel’, 0.04), (‘bon appetit’, 0.04), (‘freedom speech’, 0.03)
39voRIZciVI	Springfield mayor responds to JD Vance’s claims about Haitian immigrants	CNN, 2024e, Sep 15	Another corrupt worthless mayor that is lying out his asssss (@gregc2069, 2024, Oct 17)	(‘dana bash’, 0.33), (‘fake news’, 0.21), (‘rob rue’, 0.13), (‘mayor democrat’, 0.09), (‘bright light’, 0.08), (‘god bless’, 0.06), (‘liar liar’, 0.06), (‘lie lie’, 0.05), (‘free speech’, 0.05), (‘bon appetit’, 0.05)
4IGd0BrrXol	Trump, Harris face off at the ABC News Presidential Debate	Fox News, 2024a, Sep 10	Blame Game the Great Narcissist has spoken! The Great Twist & Shout has FAILED ! Our eyes have been opened across the country to see how much America 🇺🇸 needs President Trump ! Jesus Christ has preserved Trump for a time such as this! (@sandraschwenn414, 2024, Oct 17)	(‘abc news’, 0.17), (‘taylor swift’, 0.14), (‘david muir’, 0.12), (‘god bless’, 0.1), (‘blah blah’, 0.1), (‘fake news’, 0.09), (‘word salad’, 0.08), (‘lie lie’, 0.08), (‘old man’, 0.08), (‘good job’, 0.07)
6 W0 mzgcDNv	JD Vance claims on Ohio’s Haitian immigrants eating pets debunked	CBS News, 2024a Sep 10	Ah yes the same people who said they wouldn’t fact check during the debate and then proceeded to do just that, and lie. Why should Americans trust you? (@boodledemic6430, 2024, Oct 16)	(‘fake news’, 0.36), (‘american woman’, 0.34), (‘drug addict’, 0.08), (‘matter fact’, 0.06), (‘russia russia’, 0.06), (‘true watch’, 0.05), (‘front lawn’, 0.05), (‘lie saw’, 0.05), (‘true though’, 0.05), (‘show proof’, 0.04)
AuItvjZ1Tro	Missing cat at center of Springfield conspiracy theory found hiding in basement	CNN, 2024f, Sep 20	What about the other cats that have gone missing. What about the cell phone footage that showed the man grabbing the cat (@JamesBrian-r2b, 2024, Sep 28)	(‘van lathan’, 0.18), (‘fear fear’, 0.15), (‘one cat’, 0.12), (‘neo nazi’, 0.1), (‘fake news’, 0.08), (‘hate crime’, 0.08), (‘white woman’, 0.06), (‘lost cat’, 0.06), (‘shame shame’, 0.06), (‘bottom line’, 0.06)

id ID	TITLE	SOURCE/DATE	EXAMPLE COMMENT	TOP-10 TFIDF TERMS
bPbVUxQ7btE	Residents in Springfield, Ohio react to accusations of immigrants abducting and eating pets	WBNS 10TV, 2024, Sep 12	But ppl still think it's fake (jessica-db4is, 2024, Sep 18)	('go home', 0.33), ('back haiti', 0.18), ('blah blah', 0.1), ('wrong wrong', 0.07), ('voo doo', 0.06), ('teach children', 0.05), ('might eat', 0.05), ('born america', 0.05), ('learn drive', 0.05), ('back came', 0.04)
DPCXxNXraKw	'NO DOGS OR CATS' Haitian immigrant pleads with Trump and Springfield, Ohio residents	LiveNow from FOX, 2024, Sep 18	Of coarse they deny it. (@saulovernow, 2024, Oct 12)	('david muir', 0.44), ('know haitian', 0.22), ('fake news', 0.22), ('last night', 0.12), ('yeah right', 0.07), ('abc news', 0.06), ('bird flu', 0.06), ('said fake', 0.06), ('game warden', 0.06), ('make fake', 0.05)
eytLX2OTKPs	Trump falsely claims Haitian migrants are eating pets in Ohio	CBS News, 2024b, Sep 10	Haitians commonly eat cats in Haiti and can even buy them at the store URL (@Lar-invxib, 2024, Oct 9)	('deport deport', 0.31), ('send back', 0.14), ('fake news', 0.13), ('god bless', 0.1), ('vote red', 0.09), ('follow money', 0.07), ('wake america', 0.06), ('close border', 0.06), ('fight fight', 0.05), ('martha vineyard', 0.05)
f1RF5OQrURI	Trump amplifies false claims about Haitian immigrants eating pets	ABC News, 2024b, Sep 19	Isnt it more possible that african immigrants are eating the pets since they cant stop eating bush meat and have been flooding the boarder too? (@bannedintheusa5714, 2024, Sep 25)	('fake news', 0.11), ('abc news', 0.1), ('lie lie', 0.09), ('last woman', 0.04), ('blah blah', 0.04), ('go home', 0.04), ('true right', 0.04), ('free speech', 0.04), ('take video', 0.03), ('david muir', 0.03)
GTmdOGut4Gw	Springfield, Ohio, man reported Haitian migrants snatching geese out of park just 2 weeks ago	New York Post, 2024, Sep 11	how the guy know they are Haitians? they can be local residents too... at least they are not dogs and cats (@AnythingJW, 2024, Sep 26)	('fake news', 0.13), ('springfield sue', 0.1), ('still support', 0.09), ('russia russia', 0.09), ('hate crime', 0.07), ('big fat', 0.06), ('dumb dumber', 0.06), ('could vote', 0.06), ('shame shame', 0.06), ('send bill', 0.06)
i0IvIekwz3c	Criminal charges filed by Haitian nonprofit against Trump and JD Vance	WSYX ABC 6, 2024a, Sep 11	These idiots are going to wind up proving that they in fact did eat geese and cats 😊 but of course by that time the election will be over (@uarbor70, 2024, Sep 29)	('send bill', 0.15), ('hate crime', 0.14), ('fake news', 0.12), ('god bless', 0.09), ('make pay', 0.09), ('springfield sue', 0.06), ('roy cohn', 0.06), ('drink bleach', 0.06), ('full stop', 0.06), ('old man', 0.06)
L5Gq1Tc63UE	Springfield resident describes impact of Haitian migrants on community	Fox News, 2024d, Sep 12	Unbelievable!! I went to Haiti when I was younger and I saw unimaginable things such as toileting in the streets and eating animals! I saw it! It is true! (@laurabellew9040, 2024, Sep 27)	('fake news', 0.44), ('cnn stop', 0.07), ('hey cnn', 0.06), ('cnn still', 0.06), ('maga moron', 0.05), ('lol cnn', 0.05), ('usa usa', 0.05), ('see cnn', 0.05), ('word salad', 0.05), ('cnn propaganda', 0.04)
L9FiWz07t2o	Springfield neighbors describe immigrant situation; say they believe pet-snatching story	WSYX ABC News, 2024a, Sep 11	'Dont fuck with cats and dogs. People will find you just like " Dont fuck with cats" on netflix lol (@arvinstz, 2024, Oct 14)'	('get first', 0.08), ('mean right', 0.08), ('spoke english', 0.08), ('dear springfield', 0.08), ('still thing', 0.08), ('bring home', 0.08), ('dei hire', 0.08), ('bad move', 0.08), ('logansport indiana', 0.08), ('vote evil', 0.08)
LKAul2Gtj4g	JD Vance exposed for knowingly lying about 'pet-eating' rumors in Springfield, OH	MSNBC, 2024a, Sep 18	Vance is a puppet. (@samlott99, 2024, Sep 23)	('mlk fought', 0.1), ('mlk fight', 0.09), ('fake news', 0.08), ('go home', 0.07), ('random guy', 0.04), ('enough enough', 0.04), ('send back', 0.04), ('back haiti', 0.03), ('much easier', 0.03), ('got paid', 0.03)
M547oAXe3O4	'You think this is the last cup of B.S.': Steele slams Trump campaign for lying about migrants	MSNBC, 2024b, Sep 18	MAGA REPUBLICAN GRIFTER AND DONALD TRUMP, JD VANCE ARE BOTH ARE DICTATOR RULE UNITED STATES OF TRUMPIZE (@stevenseabrook8003, 2024, Sep 25)	('fake news', 0.47), ('ohio true', 0.08), ('told truth', 0.07), ('comment section', 0.06), ('facebook hoax', 0.05), ('trust bro', 0.05), ('abc news', 0.04), ('video true', 0.04), ('video fact', 0.04), ('covid came', 0.04)
mr5Bv-evY9A	Reporter Presses JD Vance About 'Debunked False Claims' About Haitians In Springfield, Ohio	Forbes Breaking News, 2024, Sep 23	Hes a joke! (@sonyahutchins6197, 2024, Oct 13)	('fake news', 0.44), ('abc news', 0.34), ('news abc', 0.13), ('freedom speech', 0.08), ('fake fake', 0.07), ('take note', 0.06), ('anti propaganda', 0.05), ('listen abc', 0.05), ('money money', 0.05), ('maga maga', 0.05)
oBCSCrhnfu0	JD Vance defends amplifying false claims about immigrants	Face the Nation, 2024, Sep 15	Looks like Face The Nation is a Democrat owned Network Defending KAREN ! VANCE tries to get his point across by the Democrat Narrator keeps interrupting Vance ! (@davidblanchard5252, 2024, Sep 30)	('non profit', 0.43), ('good luck', 0.3), ('go home', 0.2), ('free speech', 0.18), ('freedom speech', 0.12), ('deport deport', 0.12), ('back haiti', 0.12), ('great news', 0.11), ('sue biden', 0.1), ('fake news', 0.1)

id ID	TITLE	SOURCE/DATE	EXAMPLE COMMENT	TOP-10 TFIDF TERMS
oITtLIAYln4	Ohio residents rage at officials over Haitian migrant surge: 'No vacancy'	Fox News, 2024b, Sep 12	Turns out, it was a 64 year old white guy, ha ha ha ah !!!!!!! (@ davidheaslip4413, 2024, Oct 13)	('fake news', 0.12), ('great nation', 0.11), ('dont want', 0.1), ('truth social', 0.08), ('proven true', 0.06), ('great job', 0.05), ('media media', 0.05), ('tampon tim', 0.05), ('blah blah', 0.04), ('good job', 0.04)
ra2PLZlP1o0	Dana Bash and JD Vance clash over baseless claims about Haitian immigrants (Part 1/2)	CNN, 2024c, Sep 15	This is a garbage interview. You are a garbage reporter. (@ amcurella, 2024, Oct 16)	('face nation', 0.34), ('fake news', 0.21), ('made china', 0.13), ('good job', 0.1), ('great job', 0.09), ('minimum wage', 0.07), ('bon appetit', 0.06), ('fake media', 0.06), ('tim walz', 0.06), ('silicon valley', 0.05)
Ru34 gTyHKlw	Trump spreads more lies about Springfield, Ohio	CNN, 2024a, Sep 14	Not legal and not wanted. They are a burden to American society. We want them deported! (danielkim7841, 2024, Sep 15)	('deport deport', 0.29), ('send back', 0.27), ('vote red', 0.08), ('god bless', 0.08), ('smirk face', 0.07), ('back haiti', 0.07), ('save act', 0.06), ('close border', 0.06), ('fake news', 0.06), ('back came', 0.06)
si1Tbz1kbczw	JD Vance reacts to Springfield father saying he politicized his son's death (Part 2/2)	CNN, 2024d, Sep 15	This has brought to you by Pfizer (@brettschultz5150, 2024, Oct 9)	('dana bash', 0.56), ('sperm count', 0.07), ('dana got', 0.07), ('fake news', 0.07), ('jim crow', 0.05), ('white supremacist', 0.05), ('put place', 0.05), ('great job', 0.05), ('past eight', 0.05), ('good job', 0.04)
TQm_EAmLhJw	Springfield, Ohio resident: This is a dystopian nightmare	Fox News, 2024c, Sep 12	TRUMP 2024 🗳️🇺🇸 (@ WiseUpRiseUp42, 2024, Oct 15)	('dana bash', 0.29), ('ever thought', 0.11), ('half time', 0.11), ('job journalist', 0.11), ('love fact', 0.11), ('stir hate', 0.11), ('total bullshit', 0.11), ('asylum claim', 0.11), ('dana got', 0.1), ('mop floor', 0.1)
xc0lkVmiN8o	Haitian community in Springfield, Ohio, under threat after debunked comments	ABC News, 2024a, Sep 16	'ABC is ILYING!! I get real true news from the local people-NOT the dishonest lying bought and paid for fake news media. (@ RonaldMacDonald-hd8 gh, 2024, Sep 23)	('fake news', 0.34), ('abc news', 0.3), ('shame abc', 0.1), ('back haiti', 0.09), ('still job', 0.09), ('abc joke', 0.07), ('governor said', 0.06), ('abc lie', 0.06), ('far left', 0.06), ('news channel', 0.06)

(5) IMPLICATIONS/APPLICATIONS

The YouTube video and comment data provide several opportunities for reuse in the study of (1) natural language processing, (2) network text analysis, (3) political disinformation on social media, (4) disinformation/misinformation on YouTube, (5) news framing effects, and (6) public discussions of immigration.

(5.1) NATURAL LANGUAGE PROCESSING

As suggested by the TFIDF analysis above, the corpus is suitable for natural language processing. The comments can inform topic modelling research or demonstrations. SpaCy named entity information can inform fine-grained linguistic analysis of named agents in the corpus. Emoji tags can be used to assess what and how often YouTube commenters employ emojis in their arguments (Gray & Holmes, 2020). VADER scores can be used to assess the sentiment of comments. Engagement data provided by like and comments counts, and date of publication information can also scaffold time series analyses and novel explorations into social media metrics.

(5.2) NETWORK TEXT ANALYSIS

Reply data captured in the *in_reply_to* and *authorship* fields of *youtube_haitian_disinformation_comment_reply_metadata.csv* can be used for network modelling, uncovering relationships between comments and between commenters. Such modelling can reveal what types of comments facilitate the most conversation by measuring the in-degree and out-degree of replies.

(5.3) POLITICAL DISINFORMATION ON SOCIAL MEDIA

For researchers interested in quantitative and qualitative approaches to disinformation on social media, transcript, video description, video title, and comment data are amenable to manual, qualitative coding and automated approaches such as sentiment analysis (Han, 2022) or topic modelling (Latorre & Amores, 2021; Shekhar, et al., 2021).

Table 1 Top-10 TFIDF Bigrams for each YouTube Video in the *YouTube Disinformation Corpus*.

(5.4) DISINFORMATION/MISINFORMATION ON YOUTUBE

The corpus can underpin research into YouTube-specific forms of political disinformation, political polarization (Munger & Phillips, 2022), and racial/xenophobic agitation akin to the work of Inwood and Zappavigna (2023a; 2023b), who view YouTube as a platform conducive to the propagation of conspiracy theories and describe YouTube comments as critical for examining multiple discursive activities (see also Dynel, 2014; Heydari et al., 2019; Cinelli et al., 2021).

(5.5) NEWS FRAMING EFFECTS

Relationships between textual data categories can be mined. For example, textual similarity between YouTube transcript data and attendant comments could be calculated to determine how video content is being recontextualized by YouTube audiences (Boyd, 2014). From another perspective, the YouTube video metadata included with the comments can be studied for their “news framing effects” (Gamson & Modigliani, 1989; Entman, 1993). Theories of news framing propose that news outlets guide audience reception of journalistic content by emphasizing, deemphasizing, and/or contracting elements of the reporting (de Vreese, 2005, p. 52; Lecheler et al., 2015, p. 814; Lecheler & de Vreese, 2012, p. 186). This orchestration of priorities is achieved through textual, visual, or audio “framing devices” (Gamson & Modigliani, 1989; see also de Vreese, 2005, p. 54). Such devices include news headlines, charts, pull quotes, and corporate branding (de Vreese, 2005, p. 54). Within the corpus, tags, transcripts, video descriptions, and title information function as comparable devices and can inform traditional news framing analysis. The effects of these news framing devices can then be assessed with the comment data through qualitative or quantitative methods. For research into the recognized news frame of “immigration” (Eberl & Meltzer, 2021; Quinsaat, 2014; Eberl et al., 2018; Kim & Want, 2018; Yang, 2020; Guo et al., 2023; Reedy et al., 2023), the corpus provides opportunities to trace the use of immigration framing in media in the context of a historic national election.

(5.6) PUBLIC DISCUSSIONS OF IMMIGRATION

The corpus can inform research into public discussions of immigration in the era of social media and Trumpism by building upon existing studies of video content as Lybecker et al. (2015) and Laaksonen et al. (2020) do. Following Lybecker et al. (2015), an analysis on the construction of Haitian immigrants through video content could be undertaken using video transcript data. Incorporating comment data would represent an extension of Lybecker et al.’s (2015) analysis of depictions of the U.S.-Mexican Border, which uses video content as its data. Similarly, the qualitative discursive analysis of video content by Laaksonen et al. (2020) to reveal major patterns of movement-building argumentation in Finnish anti-immigration discussions could be applied to the corpus. Qualitative and/or quantitative analysis of comments could then be used to inspect direct reactions to the anti- or pro-immigrant stances.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR CONTRIBUTIONS

- Conceptualization
- Data curation
- Formal Analysis
- Investigation
- Methodology
- Project administration

- Resources
- Software
- Visualization
- Writing – original draft
- Writing – review & editing

AUTHOR AFFILIATIONS

Ryan M. Omizo  orcid.org/0000-0002-2796-6281
Department of English, Temple University, Philadelphia, United States

REFERENCES

- Boyd, M. S. (2014). (New) participatory framework on YouTube? Commenter interaction in US political speeches. *Journal of Pragmatics*, 72, 46–58. <https://doi.org/10.1016/j.pragma.2014.03.002>
- Catalini, M., Smyth, J., & Shopkowski, B. (2024, September 11). Trump campaign falsely accuses immigrants in Ohio of eating pets | AP News [News]. *AP NEWS*. <https://apnews.com/article/haitian-immigrants-vance-trump-ohio-6e4a47c52b23ae2c802d216369512ca5>
- Cinelli, M., Pelicon, A., Mozetič, I., Quattrocioni, W., Novak, P. K., & Zollo, F. (2021). Dynamics of online hate and misinformation. *Scientific Reports*, 11(1), 22083. <https://doi.org/10.1038/s41598-021-01487-w>
- Dynel, M. (2014). Participation framework underlying YouTube interaction. *Journal of Pragmatics*, 73, 37–52. <https://doi.org/10.1016/j.pragma.2014.04.001>
- Eberl, J.-M., & Meltzer, C. E. (2021). Media effects on attitudes toward immigration: Political sophistication as a shield. In *Media and Public Attitudes Toward Migration in Europe*. Routledge.
- Eberl, J.-M., Meltzer, C. E., Heidenreich, T., Herrero, B., Theorin, N., Lind, F., Berganza, R., Boomgaarden, H. G., Schemer, C., & Strömbäck, J. (2018). The European media discourse on immigration and its effects: A literature review. *Annals of the International Communication Association*, 42(3), 207–223. <https://doi.org/10.1080/23808985.2018.1497452>
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Gallagher, J. R., Wang, H., Modaff, M., Liu, J., & Xu, Y. (2023a). Analyses of seven writing studies journals, 2000–2019, Part II: Data-driven identification of keywords. *Computers and Composition*, 67, 102756. <https://doi.org/10.1016/j.compcom.2023.102756>
- Gallagher, J. R., Wang, H., Modaff, M., Liu, J., Xu, Y., & Beveridge, A. (2023b). Analyses of seven writing studies journals, 2000–2019, Part I: Statistical trends in references cited and lexical diversity. *Computers and Composition*, 67, 102755. <https://doi.org/10.1016/j.compcom.2023.102755>
- Gamson, W. A., & Modigliani, A. (1989). Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach. *American Journal of Sociology*, 95(1), 1–37. <https://doi.org/10.1086/229213>
- Gray, K., & Holmes, S. (2020). Tracing ecologies of code literacy and constraint in emojis as multimodal public pedagogy. *Computers and Composition*, 55, 102552. <https://doi.org/10.1016/j.compcom.2020.102552>
- Guo, L., Su, C. C., & Chen, H. T. (2023). Do News Frames Really Have Some Influence in the Real World? A Computational Analysis of Cumulative Framing Effects on Emotions and Opinions About Immigration. *The International Journal of Press/Politics*, 19401612231204535. <https://doi.org/10.1177/19401612231204535>
- Han, J. J. (2022). Reddit Dataset on Meme Stock: GameStop. *Journal of Open Humanities Data*, 8. <https://doi.org/10.5334/johd.85>
- Heydari, A., Zhang, J., Appel, S., Wu, X., & Ranade, P. G. (2019). YouTube Chatter: Understanding Online Comments Discourse on Misinformative and Political YouTube Videos.
- Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). SpaCy: Industrial-strengthened Natural Language Processing in Python. <https://doi.org/10.5281/zenodo.1212303>
- Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media* (Vol. 8, No. 1, pp. 216–225). <https://doi.org/10.1609/icwsm.v8i1.14550>
- Inwood, O., & Zappavigna, M. (2023a). Attitudes about Propaganda and Disinformation: Identifying discursive personae in YouTube comment sections. In *The Routledge Handbook of Discourse and Disinformation*. Routledge. <https://doi.org/10.4324/9781003224495-18>
- Inwood, O., & Zappavigna, M. (2023b). Conspiracy Theories and White Supremacy on YouTube: Exploring Affiliation and Legitimation Strategies in YouTube Comments. *Social Media + Society*, 9(1), 20563051221150410. <https://doi.org/10.1177/20563051221150410>

- Kim, J., & Wanta, W. (2018). News framing of the U.S. immigration debate during election years: Focus on generic frames. *The Communication Review*, 21(2), 89–115. <https://doi.org/10.1080/10714421.2018.1479616>
- Laaksonen, S. M., Pantti, M., & Titley, G. (2020). Broadcasting the movement and branding political microcelebrities: Finnish anti-immigration video practices on YouTube. *Journal of Communication*, 70(2), 171–194. <https://doi.org/10.1093/joc/jqz051>
- Latorre, J. P., & Amores, J. J. (2021). Topic modelling of racist and xenophobic YouTube comments. Analyzing hate speech against migrants and refugees spread through YouTube in Spanish. *Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21)*, 456–460. <https://doi.org/10.1145/3486011.3486494>
- Lecheler, S., Bos, L., & Vliegenthart, R. (2015). The Mediating Role of Emotions: News Framing Effects on Opinions About Immigration. *Journalism & Mass Communication Quarterly*, 92(4), 812–838. <https://doi.org/10.1177/1077699015596338>
- Lecheler, S., & de Vreese, C. H. (2012). News Framing and Public Opinion: A Mediation Analysis of Framing Effects on Political Attitudes. *Journalism & Mass Communication Quarterly*, 89(2), 185–204. <https://doi.org/10.1177/1077699011430064>
- Lozano, A. (2024, September 13). “It just exploded”: Springfield woman claims she never meant to spark false rumors about Haitians [News]. *NBC News*. <https://www.nbcnews.com/news/us-news/-just-exploded-springfield-woman-says-never-meant-spark-rumors-haitian-rcna171099>
- Lybecker, D. L., McBeth, M. K., Husmann, M. A., & Pelikan, N. (2015). Do New Media Support New Policy Narratives? The Social Construction of the U.S.–Mexico Border on YouTube. *Policy & Internet*, 7(4), 497–525. <https://doi.org/10.1002/poi3.94>
- Maher, K., & Boyette, C. (2024, September 15). JD Vance defends baseless rumor about Haitian immigrants eating pets | CNN Politics [News]. *CNN*. <https://www.cnn.com/2024/09/15/politics/vance-immigrants-pets-springfield-ohio-cnntv/index.html>
- Manning, C., & Schutze, H. (1999). Foundations of statistical natural language processing. MIT press.
- Munger, K., & Phillips, J. (2022). Right-Wing YouTube: A Supply and Demand Perspective. *Press/Politics*, 27(1), 186–219. <https://doi-org.libproxy.temple.edu/10.1177/1940161220964767>
- Olmstead, E. (2024 Sep. 16). J.D. Vance Admits He’s Telling Racist Lies for Attention. *The New Republic*. Retrieved September 25, 2024, from <https://newrepublic.com/post/185986/jd-vance-admits-migrants-conspiracy-racist-lies>
- Orazco, J. (2024, September 9). Springfield police say no reports of pets stolen, eaten, after viral social media post [News]. *Springfield-News-Sun*. <https://www.springfieldnewssun.com/news/springfield-police-say-no-reports-of-pets-stolen-after-viral-social-media-post/3WSIZQNHQVE4NP4TS5BVHBB2PY/>
- Quinsaas, S. (2014). Competing News Frames and Hegemonic Discourses in the Construction of Contemporary Immigration and Immigrants in the United States. *Mass Communication and Society*, 17(4), 573–596. <https://doi.org/10.1080/15205436.2013.816742>
- Reedy, J., O’Brien, B. G., & Hurst, E. H. (2023). Pandemic Politics: Immigration, Framing, and Covid-19. *Journal of Race, Ethnicity, and Politics*, 8(2), 246–266. <https://doi.org/10.1017/rep.2023.14>
- Shekhar, S., Akanksha, & Saini, A. (2021). Utilizing Topic Modelling To Identify Abusive Comments On YouTube. *2021 International Conference on Intelligent Technologies (CONIT)*, 1–5. <https://doi.org/10.1109/CONIT51480.2021.9498368>
- Timotija, F. (2024, September 14). Springfield woman regrets post that sparked false Haitian claims [Text]. *The Hill*. <https://thehill.com/homenews/state-watch/4879994-springfield-rumor-haitian-immigrants-facebook-post/>
- Vreese, C. H. (2005). News framing: Theory and typology. *Information Design Journal*, 13(1), 51–62. <https://doi.org/10.1075/idjdd.13.1.06vre>
- Yang, J. (2020). Framing Immigration and Illegal Immigration in the 2016 Presidential Campaigns: Comparing Donald Trump and Bernie Sander’s Position. *Athens Journal of Mass Media and Communications*, 7(1), 9–22. <https://doi.org/10.30958/ajmmc.7-1-1>

TO CITE THIS ARTICLE:

Omizo, R. M. (2025). Disinformation on Youtube: A Dataset of Youtube Comments on Videos Related to Claims Made by Trump and Vance on Haitian Immigrants. *Journal of Open Humanities Data*, 11: 12, pp. 9–9. DOI: <https://doi.org/10.5334/johd.283>

Submitted: 15 November 2024

Accepted: 02 January 2025

Published: 12 February 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Journal of Open Humanities Data is a peer-reviewed open access journal published by Ubiquity Press.