MDPI

Article

Public Discourse Surrounding the 2025 California Wildfires: A Sentiment and Topic Analysis of High-Engagement YouTube Comments

Dmitry Erokhin (ORCID)

International Institute for Applied Systems Analysis, 2361 Laxenburg, Austria; erokhin@iiasa.ac.at

Abstract: This study explores public discourse surrounding the January 2025 California wildfires by analyzing high-engagement YouTube comments. Leveraging sentiment analysis, misinformation detection, and topic modeling, this research identifies dominant emotional tones, thematic patterns, and the prevalence of misinformation in discussions. The results show a predominantly neutral to positive sentiment, with notable emotional intensity in misinformation-related comments, which were rare but impactful. The thematic analysis highlights concerns about governance, environmental issues, and conspiracy theories, including water mismanagement and diversity-related critiques. These findings provide insights for crisis communication, policymaking, and misinformation management during disasters, emphasizing the importance of aligning strategies with public concerns.

Keywords: California wildfires; public discourse; YouTube comments; sentiment analysis; topic modeling; misinformation detection; crisis communication; conspiracy theories; governance; environmental concerns

1. Introduction

Wildfires have become increasingly frequent and severe in recent years [1], posing significant threats to ecosystems, human lives, and property. In January 2025, the California wildfires emerged as one of the most devastating events at the beginning of the year, attracting widespread public attention and discourse. As climate change continues to exacerbate the frequency and intensity of such disasters [2], understanding how people perceive and discuss these events is critical for effective crisis communication, policymaking, and misinformation management.

The rise of social media platforms has transformed how information about disasters is disseminated and discussed [3]. By fundamentally shifting crisis communication from traditional expert-to-receiver models to decentralized, real-time engagement [4], social media now serves as a hub for public discourse that offers valuable insights into societal responses, emotional reactions, and the spread of information. Social media enables rapid information dissemination, public participation, and self-organization during crises, which can facilitate community resilience and psychosocial healing [4,5]. However, alongside these constructive discussions, social media also provides fertile ground for misinformation [6] and unreliable user-generated content, potentially undermining public trust, hindering disaster response efforts, exacerbating societal divides, and even being used maliciously to harm organizations [7]. This shift from a unidirectional flow of information in traditional media to a multidirectional one in social media necessitates adaptations in crisis communication strategies, prompting researchers to propose modifications to existing models to better integrate social media's evolving role in disaster communication cycles.

The aim of this study is to provide a focused analysis of public engagement during the January 2025 California wildfires. Using a robust methodological approach that combines sentiment analysis, misinformation (this study focuses on misinformation—defined as false or inaccurate information created and disseminated, often by twisting or misinterpreting facts—rather than disinformation, which is intentionally crafted to deceive and harm [8], because the covert nature and complex intent behind disinformation make it especially challenging to detect reliably) detection, and topic modeling, this research seeks to uncover the emotional tone, thematic content, and prevalence of misinformation in online discussions. Specifically, this study addresses the following questions: (1) What are the dominant emotional tones in the public discourse? (2) What thematic patterns emerge in discussions about the wildfires? (3) To what extent is misinformation present in this discourse? The findings yield insights for refining crisis communication, enhancing misinformation management, and guiding policy interventions tailored to the unique challenges of wildfire events.

2. Literature Review

The increasing prevalence of wildfires globally has prompted a growing body of research on their social, environmental, and economic impacts [9]. In this context, social media platforms have emerged as vital tools for public engagement, information dissemination, and situational awareness during crises and natural disasters [10,11]. These platforms enable real-time updates, facilitate information exchange, and support public participation in disaster response efforts [12], playing a significant role in shaping societal responses, perceptions, and decision making [13,14]. Studies also highlight social media's ability to rapidly spread critical updates during emergencies, often serving as primary news sources [15]. However, most existing research on social media has predominantly focused on textual platforms like X (previously Twitter), leaving a gap in understanding how video-centric platforms such as YouTube are leveraged (a search on Web of Science (as of 27 January 2025) highlights this disparity: studies on Twitter yield 45,244 results, and Facebook 75,285, while YouTube lags with only 18,069 results and TikTok with a mere 3254, underscoring the academic preference for text-based platforms over video-centric or emerging social media). Though existing research on textual platforms has provided valuable insights into information diffusion, sentiment dynamics, and the rapid spread of emergency-related content, the inherent limitations of text such as the absence of non-verbal cues and contextual imagery highlight the necessity of exploring video-centric platforms for a more comprehensive understanding of disaster communication.

Recent research highlights the growing importance of multimodal data analysis in disaster response and sentiment analysis, particularly on platforms like YouTube. While traditional approaches have focused on textual content [16], studies now emphasize the value of integrating audio, visual, and textual data for more comprehensive insights [17,18]. Unlike textual platforms, video-centric media can capture nuanced emotional expressions, detailed situational contexts, and complex narratives that combine sound, imagery, and text. This differentiation underscores the importance of dedicated research into how video content contributes uniquely to public understanding and crisis management. This multimodal approach has shown improved accuracy in emotion prediction and sentiment analysis. Social media platforms, including YouTube, offer rich sources of information during disasters, providing situational updates and public perspectives.

At the same time, the proliferation of misinformation on social media platforms has been a growing concern, particularly during crises where the stakes are high [19]. Research has documented how misinformation spreads rapidly on social media platforms, often amplified by algorithmic recommendations and viral sharing behaviors [20,21]. Misin-

formation during wildfire events [22] can manifest in the form of conspiracy theories, misleading claims about the causes of fires, or inaccurate safety information.

Recent studies have explored YouTube's role in misinformation spread, focusing on both video content and user comments. Ref. [23] developed a model using video captions to classify misinformation with high accuracy. Ref. [24] examined cross-platform misinformation, finding that social media posts linking to YouTube videos are particularly effective in spreading false information. They emphasized the importance of using multi-platform data for more effective detection. Ref. [25] analyzed video metadata and user engagement to identify channels promoting conspiracy theories, demonstrating the potential of these metrics in detecting malicious content. Ref. [26] focused on YouTube comments, revealing that conspiracy theories flourish in this relatively unmoderated space.

Sentiment analysis and topic modeling are valuable tools for understanding public responses to natural disasters on social media. Studies have employed VADER for sentiment analysis and Latent Dirichlet Allocation (LDA) for topic modeling to examine emotional reactions and identify key themes during crises [27–30]. Research has shown that public sentiment often fluctuates during disasters, with negative impacts observed in affected regions. Common topics identified include crisis-related information, disaster impacts, and public feedback to authorities. These analyses can reveal evolving public concerns and emotions throughout disaster periods, helping crisis managers design effective response strategies.

By focusing on YouTube, this study contributes to the growing field of social media research in disaster contexts and underscores the importance of integrating diverse platforms into the analytical framework. By acknowledging the unique strengths and challenges of both textual and video-centric platforms, researchers, policymakers, and communicators can develop robust crisis communication models that tailor strategies to each medium's specific capabilities, ultimately deepening our understanding of public engagement during wildfires and offering valuable insights to effectively address public concerns and combat misinformation.

3. Methodology

This study analyzed public discourse surrounding the January 2025 California wildfires by focusing on YouTube (this study focused on YouTube because its robust API allows efficient collection and analysis of large datasets, and its diverse content and broad user base make it ideal for examining public discourse; however, this focus may limit the generalizability of the findings due to platform-specific biases, suggesting that future research should include additional social media platforms) videos with at least one million views as of 24 January 2025. Videos were selected based on their view count to ensure they represent widely consumed content, reflecting key narratives, themes, and public engagement with the topic. The YouTube Data API was utilized to extract user comments from these videos. Using a Python 3.13.1 script, all top-level comments were retrieved, ensuring comprehensive coverage through pagination. Each comment's metadata, including author name, text, publication date, and number of likes, were collected and stored in a structured format for further analysis.

Once extracted, the data underwent preprocessing to ensure consistency and usability. Missing or non-string values in the comment text were replaced with empty strings, and all text data were converted to string format. Stop words were removed using the Natural Language Toolkit (NLTK) to facilitate keyword analysis and topic modeling [31].

Sentiment analysis (in the context of text analysis, sentiment refers to the general polarity of an expressed opinion, typically classified as positive, negative, or neutral, and it captures the overall tone of a statement rather than the specific emotional states it conveys;

emotion analysis, on the other hand, goes beyond sentiment to identify distinct affective states such as happiness, anger, sadness, or fear, providing a more nuanced understanding of how people react to events) was conducted using the VADER Sentiment Intensity Analyzer [32], which provided a compound sentiment score for each comment. This score, a composite of positive, neutral, and negative sentiment measures, was used to classify comments into five emotional categories: happiness, excitement, neutral, sadness, and anger. The classification thresholds were defined based on the compound score.

A compound score of 0.6 or greater indicated strong positive sentiment, categorized as "Happiness", while scores between 0.3 and 0.6 reflected moderately positive sentiment, labeled as "Excitement". Scores ranging from −0.3 to 0.3 represented emotional neutrality, categorized as "Neutral", capturing the absence of strong positive or negative emotions. Moderately negative sentiment, with scores between −0.6 and −0.3, was classified as "Sadness", and strong negative sentiment, with scores of −0.6 or lower, was labeled as "Anger". For scores falling outside these predefined ranges, the emotion was classified as "Unknown", accounting for ambiguity or the inability to classify the emotion.

To detect misinformation, a keyword-based approach was employed. Comments containing terms such as "conspiracy", "hoax", "fake", "fraud", "misleading", or "scam" were flagged as potentially containing misinformation. This method provided a straightforward mechanism for identifying comments related to misinformation themes (it is important to note that while the identified comments are related to misinformation, they do not necessarily support misinformation narratives; these comments may also oppose misinformation or take a neutral stance, engaging in discussion rather than endorsing or amplifying false claims). The selection of keywords for detecting misinformation was based on previous literature on misinformation detection and included the most commonly used terms and their closest synonyms to ensure comprehensive coverage of misleading narratives.

In addition, this study examined specific topics that users and news media linked to potential reasons behind the January 2025 California wildfires, including some that were conspiratorial in nature: water mismanagement, directed energy weapons (DEWs), arson or deliberate ignition, diversity and inclusion initiatives, misuse of palm tree imagery, and theories involving Sean "Diddy" Combs. (In late 2024, Combs faced multiple allegations of sexual misconduct, leading to significant media coverage and public discourse [33]. Amidst this, unfounded claims surfaced on social media suggesting that the Los Angeles wildfires were intentionally set to destroy evidence related to Combs' alleged crimes. These theories posited that the fires served as a cover-up for his legal troubles.) For each topic, an extensive list of keywords was developed to capture variations and synonyms. The text of each comment was searched for the presence of these keywords, ensuring case-insensitive, full-word matching to avoid partial matches. The frequency of comments linked to each topic was then quantified to assess their prominence in wildfire-related discussions. This approach provided insights into how individuals framed the wildfires' causes, highlighting the prevalence of potentially conspiratorial narratives within the discourse.

Topic modeling was performed using Latent Dirichlet Allocation (LDA) to uncover thematic patterns within the comment corpus [34]. Text data were vectorized using CountVectorizer, with stop words excluded and document frequency thresholds applied to filter noise. LDA identified five distinct topics, with each comment assigned to the topic it most closely aligned with. The most representative words for each topic were extracted to aid interpretation and provide insight into the underlying themes.

The processed dataset included a range of attributes for each comment, including sentiment scores, emotional classifications, misinformation labels, and assigned topics with their corresponding representative words.

This methodological approach, leveraging a combination of sentiment analysis, misinformation detection, and topic modeling, provided a comprehensive framework for examining public sentiment and thematic content in the context of online discussions about the 2025 California wildfires. The use of Python libraries and tools such as the YouTube Data API, NLTK, VADER, and scikit-learn enabled efficient data extraction and analysis, ensuring a robust exploration of the dataset.

4. Results

A total of 76 YouTube videos related to the 2025 California wildfires, each with over one million views, were analyzed. After preprocessing, the dataset comprised 263,638 comments, with the number of comments per video ranging from 79 to 19,552 (see Figure 1). On average, each video contained approximately 3469 comments, reflecting significant variation in audience engagement across the selected videos. This extensive dataset provided a robust foundation for sentiment analysis, misinformation detection, and thematic exploration.
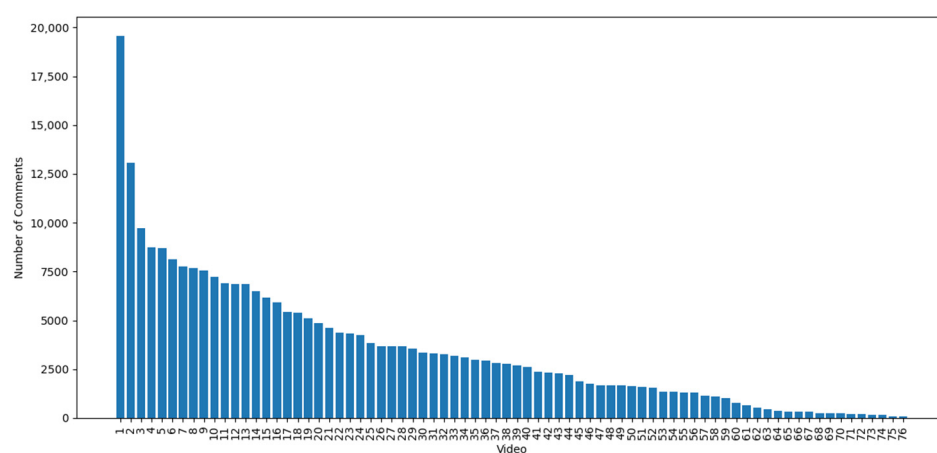

Figure 1. Number of comments per video.

The number of likes on comments ranged from 0 to 30,626, with an average of 7.52 likes per comment, indicating a wide disparity in user engagement. Sentiment scores, calculated using VADER, ranged from −1 to 1, with an average score of 0.06, suggesting a generally neutral to slightly positive sentiment overall (see Figure 2). A positive and statistically significant correlation was observed between the number of likes and sentiment scores (correlation coefficient = 0.0062, *p*-value = 0.0016), indicating that comments with a more positive sentiment tended to receive more likes.

The emotional analysis of the comments revealed a diverse distribution of emotions across the dataset (see Table 1). Neutral comments dominated the discussion, accounting for 36.32% (95,760 comments) of the total, followed by expressions of happiness at 21.15% (55,748 comments) and excitement at 14.94% (39,380 comments). Anger and sadness were less prevalent, representing 14.14% (37,271 comments) and 13.46% (35,476 comments), respectively. A negligible number of comments (0.00%, three comments) fell into the "unknown" category, highlighting the effectiveness of the sentiment classification process. This distribution reflects a predominantly neutral tone in the discourse, with notable pockets of positive and negative emotions.

The analysis of misinformation-related content revealed that only a small fraction of comments, 0.98% (2577 comments), contained keywords associated with misinformation, while the vast majority, 99.02% (261,061 comments), did not. This finding indicates that explicit misinformation was relatively rare in the dataset. However, the presence of nearly 2600 potentially misleading comments highlights the need for continued vigilance and

targeted strategies to address misinformation in discussions about critical topics like the 2025 California wildfires (also see the word cloud for misinformation-related comments in Figure 3).

Figure 2. Sentiment score distribution.

Table 1. Emotion distribution for all comments.

Emotion	% of Comments
Neutral	36.32
Happiness	21.15
Excitement	14.94
Anger	14.14
Sadness	13.46

Figure 3. Word cloud for misinformation-related comments.

The sentiment analysis of comments flagged as related to the topic of misinformation revealed a markedly negative emotional tone (see Table 2). Sadness was the most prevalent emotion, accounting for 36.36% of these comments, closely followed by anger at 35.27%. A closer look at these comments revealed that anger largely stemmed from distrust (media, politicians, and insurance), perceived injustice, and polarization. Sadness arose from disillusionment, helplessness, and grief over systemic failures or personal losses. Both emotions reflect a broader societal crisis of trust and a desire for accountability in the face

of disasters and misinformation. In contrast, positive emotions like happiness (6.60%) and excitement (5.67%) were much less common, suggesting that such comments rarely elicited positive reactions. Neutral comments made up 16.10% of the total, reflecting a smaller portion of emotionally disengaged responses. This distribution underscores the heightened emotional intensity and predominantly negative sentiment associated with misinformation in the discourse.

Table 2. Emotion distribution for misinformation-related comments.

Emotion	% of Comments
Sadness	36.36
Anger	35.27
Neutral	16.10
Happiness	6.60
Excitement	5.67

In addition, the analysis revealed varying levels of engagement with potentially conspiratorial topics linked to reasons behind the January 2025 California wildfires (see Table 3). Among the keyword groups, references to water were the most prevalent, appearing in 9772 comments, indicating significant public discourse around water-related issues. This was followed by mentions of diversity, equity, and inclusion (DEI) initiatives, which were found in 4058 comments, reflecting a notable association made by users between these initiatives and the wildfire response. Arson, intentional, or deliberate actions were referenced in 1315 comments, highlighting the continued relevance of arson-related narratives. Mentions of directed energy weapons (DEWs) appeared in 526 comments, while 417 comments referenced theories involving Sean "Diddy" Combs, suggesting limited but notable traction for these conspiracy theories. Finally, palm trees or palms were mentioned in 134 comments, indicating relatively low public engagement with narratives linking palm trees to the wildfires. These findings underscore the diverse range of topics discussed and the prominence of both factual and conspiratorial narratives in public discourse surrounding the wildfires.

Table 3. Distribution of potentially conspiratorial comments.

Keyword Group	Count
Water	9772
Diversity/Equity/DEI/Inclusion	4058
Arson/Intentional/Deliberate	1315
DEW/Weapon	526
Diddy/Combs/Sean	417
Palm/Palms	134

The topic modeling analysis identified five distinct themes within the dataset, with varying levels of prevalence (see Table 4). Topic 2 was the most dominant, encompassing 28.50% (75,136 comments) of the total, followed by Topic 1, which accounted for 24.92% (65,704 comments). Topics 0 and 3 were nearly equally represented, making up 17.53% (46,226 comments) and 17.73% (46,731 comments), respectively. Topic 4 was the least prevalent, covering 11.32% (29,841 comments). This distribution highlights the thematic diversity in the comments, with significant engagement across all identified topics, suggesting that discussions about the 2025 California wildfires spanned a wide range of focal points.

Table 4. Comment distribution by topic.

Topic	% of Comments
Topic 2 (Governance and Financial Aspects)	28.50
Topic 1 (Personal Connections and Emotions)	24.92
Topic 3 (Global and Emotional Reactions)	17.73
Topic 0 (Environmental and Situational Discussions)	17.53
Topic 4 (Religious and Spiritual Narratives)	11.32

The thematic analysis of the dataset provided insights into the key discussions in the comments (see Table 5). The most prevalent theme, Governance and Financial Aspects (Topic 2), included keywords such as "California", "government", "insurance", and "money", reflecting concerns about policy and economic impacts of the wildfires. The Personal Connections and Emotions theme (Topic 1) was characterized by expressions of support and empathy, featuring words like "love", "hope", "family", and "safe". The Environmental and Situational Discussions theme (Topic 0) focused on the physical and ecological dimensions of wildfires, with frequent references to "water", "fires", "trees", and "winds". The Global and Emotional Reactions theme (Topic 3) encompassed broader concerns beyond the immediate disaster, with keywords like "climate change", "sad", and "world". Finally, the Religious and Spiritual Narratives theme (Topic 4) illustrated faith-based responses to the disaster, with terms such as "God", "Jesus", "pray", and "repent". These themes underline the multifaceted nature of public discourse, blending practical concerns, emotional expressions, and broader ideological perspectives (also see word clouds by theme in Figure 4).

Table 5. Most frequent words by topic.

| Topic 1	Topic 2	Topic 0	Topic 3	Topic 4
God (7157)	People (14,273)	Fire (10,643)	People (9565)	God (10,722)
Love (5866)	California (10,645)	Water (7972)	Gaza (7226)	Jesus (3772)
Safe (5493)	Fire (10,071)	Fires (5427)	Fire (4860)	Lord (2368)
Like (5480)	Trump (6073)	Like (3294)	Like (4626)	Fire (1818)
People (5153)	Get (5815)	House (2928)	God (4568)	Pray (1684)
Hope (4979)	Like (5685)	Houses (2782)	Allah (3256)	Free (1565)
Guys (4880)	Money (5123)	California (2282)	Fires (2715)	People (1501)
Thank (4767)	State (4907)	Would (2202)	Climate (2569)	Christ (1495)
Bless (4539)	Insurance (4813)	People (2089)	Sad (2406)	Amen (1468)
Fire (3700)	Newsom (4622)	Trees (1836)	Feel (2389)	Mary (1199)
Stay (3307)	Fires (4475)	Homes (1807)	Change (2377)	Hollywood (1123)
Good (3048)	Water (4352)	Get (1742)	World (2354)	Repent (1119)
Family (3015)	Need (4174)	Build (1646)	Palestine (2215)	Thank (987)
Prayers (2877)	Dei (3809)	Winds (1636)	Sorry (2148)	Palestine (945)
Everyone (2744)	Government (3710)	Could (1595)	America (2055)	Please (938)
Man (2697)	Would (3686)	Years (1531)	Lost (1987)	Name (908)
Dog (2579)	Help (3652)	Burn (1510)	Homes (1811)	Mother (875)
Sorry (2498)	Time (3249)	Right (1371)	Think (1809)	🙏 (856)
Get (2492)	Going (3062)	Wood (1330)	Know (1692)	World (807)
Heart (2445)	Know (3037)	Area (1278)	California (1646)	Gods (765)
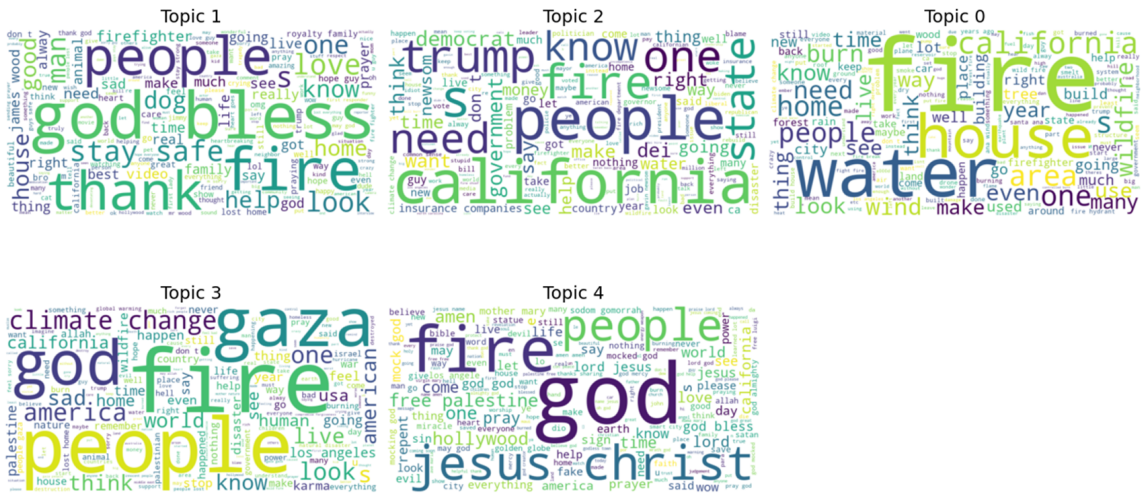
Figure 4. Word clouds by topic.

The analysis of sentiment scores across topics revealed significant variations, underscoring the distinct emotional tone of discussions within each thematic category (see Figure 5 for sentiment and Figure 6 for emotion intensity by topic). Topic 1, which focused on supportive and empathetic expressions, had the highest average sentiment score (0.3086), indicating a predominantly positive tone. Similarly, Topic 4, characterized by religious and spiritual themes, also displayed a positive sentiment (0.2197), suggesting that comments in this category were generally uplifting or encouraging. In contrast, Topic 0, related to Environmental and Situational Discussions, had an almost neutral average sentiment score (-0.0019), reflecting a descriptive or fact-based tone. Topics 2 and 3, centered on governance and global perspectives, had negative average sentiment scores (-0.0748 and -0.0912, respectively), highlighting a more critical or somber tone in comments associated with these themes.
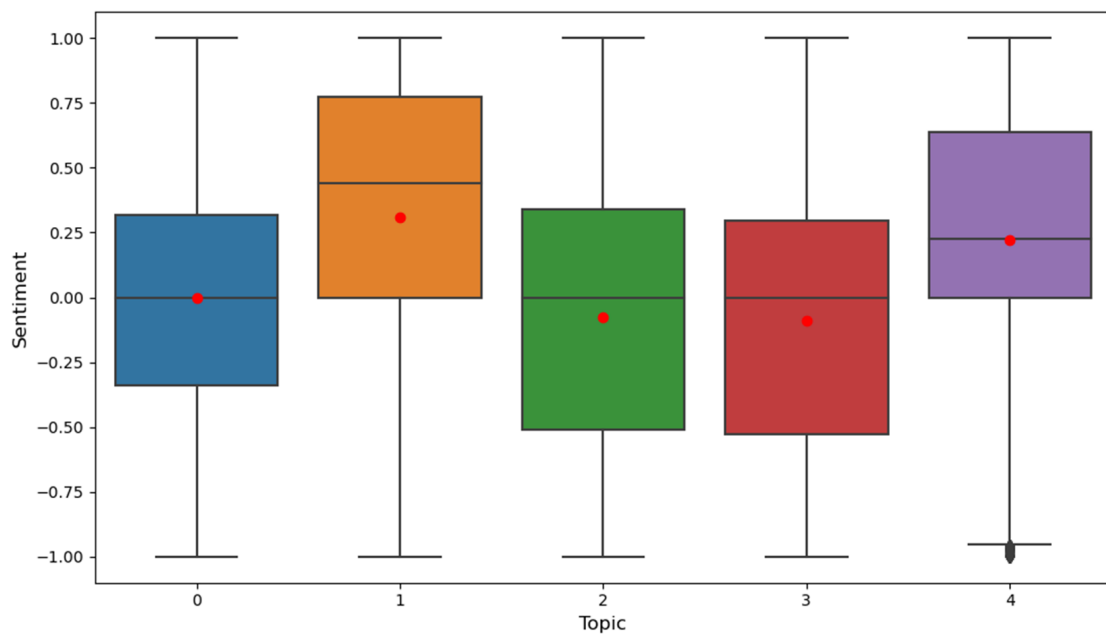

Figure 5. Sentiment by topic (the outliers in Topic 4 have a mean of -0.9775, with a minimum value of -0.9999 and a maximum value of -0.9546).
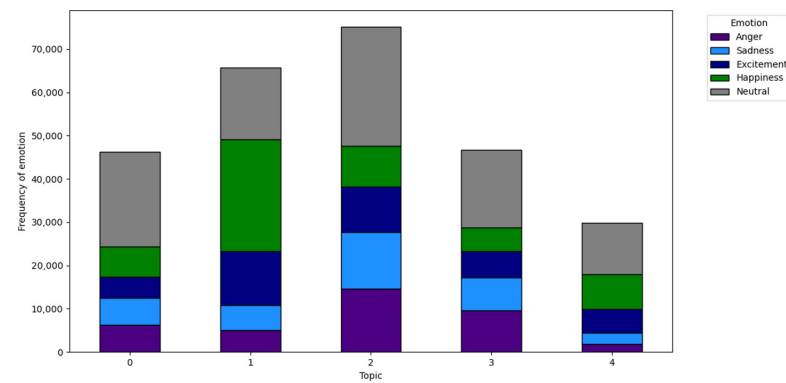
Figure 6. Emotion intensity by topic.

5. Discussion

The results of this study provide valuable insights into the public discourse surrounding the January 2025 California wildfires, as reflected in YouTube comments on videos with high engagement. The emotional analysis revealed that neutral comments dominated the discourse, with notable pockets of happiness and excitement, suggesting a generally measured or positive tone in public sentiment. Anger and sadness were present but less prevalent, indicating that negative emotions, while significant, did not overshadow the discussions. This contributes to the growing research on social media sentiment during disasters, which reveals complex patterns. Studies show that negative sentiments often dominate, with off-site users expressing more negativity than those directly affected [35]. However, positive sentiments also emerge, reflecting hope and encouragement [36]. Sentiment analysis can be categorized into various types, including anxiety, disappointment, and relief, providing a nuanced understanding of public reactions [37]. Spatiotemporal analysis demonstrates that sentiment varies across locations and disaster stages, with higher concentrations of posts in populous areas and near disaster sites [38]. While negative tweets spread faster, positive and neutral messages tend to have broader reach [35].

Additionally, the low prevalence of misinformation-related comments highlights the relatively minimal presence of explicit misinformation within the dataset. However, the identified instances underscore the importance of monitoring and addressing misinformation in critical discussions, especially given its non-negligible incidence and causal impacts on beliefs and behaviors [39].

Thematic analysis through topic modeling revealed diverse focal points in the discourse. The most dominant themes included governance and financial concerns, personal connections, environmental and situational issues, global perspectives, and religious or spiritual responses. These findings illustrate the multifaceted nature of public engagement, encompassing both practical concerns and emotional or ideological dimensions.

The findings carry several implications for policymakers, communication professionals, and researchers. First, the predominance of neutral and positive sentiments suggests that public discussions can be leveraged to foster constructive engagement during crises. For instance, the empathetic and supportive tone identified in Topic 1 highlights opportunities to amplify messages of solidarity and community resilience. Conversely, the negative sentiments associated with governance-related discussions (Topic 2) indicate areas where public trust may need to be strengthened through transparent communication and effective policies.

The low but noteworthy presence of misinformation-related content underscores the need for proactive misinformation management strategies. On the one hand, the natural origin of wildfires likely contributes to the overall low volume of misinformation. On the other hand, various conspiracies still emerge, attempting to link the 2025 California

wildfires to human activities or deliberate plots. Conspiracy theories tend to emerge during crises and emergencies as people seek to make sense of uncertain situations and regain a sense of control [40,41]. These theories often arise from fear, uncertainty, and distrust in official information sources [42]. Social media platforms play a crucial role in the spread of conspiracy theories, with online communities forming around shared beliefs. The propagation of these theories can follow a power-law distribution, potentially leading to a critical transition from sporadic mentions to widespread dissemination. Conspiracy theories can have real-world consequences, as seen in the case of 5G–COVID-19 misinformation leading to arson attacks. Among the identified conspiracies in the case of the 2025 California wildfires, water mismanagement claims dominated the discourse, with nearly 10,000 comments alleging deliberate diversion or insufficient water resources for firefighting efforts. Similarly, over 4000 comments blamed funds allocated to diversity, equity, and inclusion (DEI) initiatives for alleged inefficiencies in wildfire response, reflecting a polarizing narrative that systemic policies detracted from disaster preparedness. Other conspiratorial theories, such as deliberate arson and the use of directed energy weapons (DEWs), gained traction, showcasing public speculation fueled by unfounded claims during crises. Additionally, less prominent but notable narratives included the theory involving Sean "Diddy" Combs, which alleged the fires were set to cover up evidence and claims linking burning palm trees to urban redevelopment agendas. These findings underscore the widespread dissemination of conspiratorial reasoning, emphasizing the need for targeted interventions to address misinformation in the wake of disasters.

This study demonstrates the potential of public sentiment and thematic analysis as tools for supporting crisis communication and policymaking. Understanding the emotional and thematic dimensions of public discourse enables stakeholders to tailor their responses to the needs and concerns of affected communities. Additionally, by monitoring misinformation, authorities can enhance their capacity to safeguard public trust and mitigate the adverse impacts of false narratives. These insights can also guide future communication strategies to ensure alignment with public sentiment and enhance the effectiveness of outreach efforts.

Despite its contributions, this study has limitations. The reliance on YouTube comments limits the scope to a specific subset of public discourse, excluding other platforms and offline interactions [43]. The use of keyword-based methods for misinformation detection, while straightforward, may not capture all forms of misinformation or provide nuanced insights into the context of flagged comments [44]. Similarly, sentiment analysis using the VADER tool, while effective, may not fully account for linguistic nuances [45]. The thematic categorization through LDA, while informative, relies on interpretive assumptions, which may influence the identification of topics [46]. Another potential limitation is that YouTube videos themselves may carry emotional bias, which could influence the sentiment of user comments. The analysis showed a moderate positive correlation between video sentiment and comment sentiment, with a statistically significant Pearson correlation coefficient of approximately 0.35 ($p \approx 0.002$). (For each analyzed video, transcript text was extracted, which was then processed using NLTK's VADER sentiment analyzer to compute compound sentiment scores. If a transcript was not available, the video's title and description were extracted and analyzed for sentiment.)

Future research could expand the scope of analysis by incorporating data from additional social media and digital platforms to capture a broader spectrum of public discourse. Advanced machine learning models, such as BERT-based classifiers [47], could enhance the accuracy of sentiment and misinformation detection by accounting for contextual nuances. Additionally, longitudinal studies could explore how public sentiment and misinformation evolve over time during and after crises, offering insights into long-term trends and recov-

ery processes [19]. Finally, integrating qualitative analyses, such as in-depth interviews or focus groups [48], could complement quantitative findings and provide richer insights into the motivations and perceptions of individuals engaging in public discourse.

6. Conclusions

This study provides a comprehensive analysis of public discourse during the 2025 California wildfires, emphasizing the emotional and thematic diversity in high-engagement YouTube comments. The findings reveal that, while neutral and positive sentiments predominated overall, discussions surrounding misinformation exhibited a significantly heightened negative emotional tone. In addition, the thematic analysis highlighted a mix of practical concerns and conspiratorial narratives, with water mismanagement claims notably resonating with the public.

These insights underscore the potential of sentiment and topic analyses as essential tools to guide crisis communication and policymaking. In particular, understanding public sentiment and the dynamics of misinformation can help inform more effective risk and crisis communication strategies.

Based on the analysis, crisis communication strategies should incorporate enhanced real-time monitoring of social media platforms. This involves employing advanced analytics and integrating data from multiple channels to promptly detect and address emerging misinformation and public concerns. Such a comprehensive monitoring approach ensures that any potential issues are identified quickly, allowing for timely responses.

Moreover, proactive communication strategies are essential. Clear and transparent messaging protocols need to be established to counteract misinformation, foster public trust, and address both the practical and emotional aspects of crisis events. By prioritizing clarity and openness, authorities can more effectively manage the flow of information during emergencies.

In addition, it is crucial for authorities to engage actively with the community. This means creating and maintaining two-way communication channels that allow them to listen to public concerns and tailor responses based on real-time feedback, thereby promoting community resilience. Active engagement ensures that the voices of those affected are heard and considered in the decision-making process.

Finally, fostering cross-sector collaboration is key. Building strong partnerships between government agencies, media organizations, and digital platforms can help coordinate unified risk-communication efforts, ensuring that authoritative and consistent messages reach the public during crises.

While this study is limited by its focus on YouTube data and reliance on keyword-based methods, the findings offer valuable perspectives on public engagement during disasters. Future research should broaden the analysis to include multiple social media platforms, utilize more sophisticated analytical models, and adopt longitudinal designs to further explore the evolution of digital discourse during and after crisis events.

Data Availability Statement: Data are publicly available.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Cunningham, C.X.; Williamson, G.J.; Bowman, D.M. Increasing frequency and intensity of the most extreme wildfires on Earth. *Nat. Ecol. Evol.* **2024**, *8*, 1420–1425. [CrossRef] [PubMed]
2. Jones, M.W.; Smith, A.J.; Betts, R.; Canadell, J.G.; Prentice, I.C.; Le Quéré, C. Climate Change Increases the Risk of Wildfires: January 2020. ScienceBrief. 2020. Available online: https://ueaeprints.uea.ac.uk/id/eprint/77982 (accessed on 1 February 2025).

3. Erokhin, D.; Komendantova, N. Social media data for disaster risk management and research. *Int. J. Disaster Risk Reduct.* **2024**, *114*, 104980. [CrossRef]

4. Silver, A. The use of social media in crisis communication. In *Risk Communication and Community Resilience*; Routledge: London, UK, 2019; pp. 267–282. [CrossRef]

5. Hart, T.; Brewster, C.; Shaw, D. Crisis communication and social media: The changing environment for natural disaster response. In Proceedings of the Tenth Annual International Conference on Communication and Mass Media, Athens, Greece, 14–17 May 2012. Available online: https://publications.aston.ac.uk/id/eprint/22351/ (accessed on 1 February 2025).

6. Erokhin, D.; Komendantova, N. Earthquake conspiracy discussion on Twitter. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 1–10. [CrossRef]

7. Holmes, W.S. Crisis Communications and Social Media: Advantages, Disadvantages and Best Practices. 2011. Available online: https://trace.tennessee.edu/cgi/viewcontent.cgi?article=1003&context=ccisymposium (accessed on 1 February 2025).

8. Komendantova, N.; Erokhin, D.; Albano, T. Misinformation and its impact on contested policy issues: The example of migration discourses. *Societies* **2023**, *13*, 168. [CrossRef]

9. Santos, S.M.B.D.; Bento-Gonçalves, A.; Vieira, A. Research on wildfires and remote sensing in the last three decades: A bibliometric analysis. *Forests* **2021**, *12*, 604. [CrossRef]

10. Simon, T.; Goldberg, A.; Adini, B. Socializing in emergencies—A review of the use of social media in emergency situations. *Int. J. Inf. Manag.* **2015**, *35*, 609–619. [CrossRef]

11. Vongkusolkit, J.; Huang, Q. Situational awareness extraction: A comprehensive review of social media data classification during natural hazards. *Ann. GIS* **2021**, *27*, 5–28. [CrossRef]

12. Yin, J.; Lampert, A.; Cameron, M.; Robinson, B.; Power, R. Using social media to enhance emergency situation awareness. *IEEE Intell. Syst.* **2012**, *27*, 52–59. [CrossRef]

13. Boulton, C.; Shotton, H.; Williams, H. Using social media to detect and locate wildfires. In Proceedings of the International AAAI Conference on Web and Social Media, Cologne, Germany, 17–20 May 2016; Volume 10, pp. 178–186. [CrossRef]

14. Slavkovikj, V.; Verstockt, S.; Van Hoecke, S.; Van de Walle, R. Review of wildfire detection using social media. *Fire Saf. J.* **2014**, *68*, 109–118. [CrossRef]

15. Abedin, B.; Babar, A.; Abbasi, A. Characterization of the use of social media in natural disasters: A systematic review. In Proceedings of the 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, Sydney, NSW, Australia, 3–5 December 2014; pp. 449–454. [CrossRef]

16. Alam, F.; Ofli, F.; Imran, M. Crisismmd: Multimodal Twitter datasets from natural disasters. In Proceedings of the International AAAI Conference on Web and Social Media, Palo Alto, CA, USA, 25–28 June 2018; Volume 12. [CrossRef]

17. Chu, J.S. Integrative sentiment analysis: Leveraging audio, visual, and textual data. In Proceedings of the CS & IT Conference Proceedings, Xiamen, China, 18–20 October 2024; Volume 14. Available online: https://csitcp.com/abstract/14/142csit11 (accessed on 1 February 2025).

18. Imran, M.; Ofli, F.; Caragea, D.; Torralba, A. Using AI and social media multimodal content for disaster response and management: Opportunities, challenges, and future directions. *Inf. Process. Manag.* **2020**, *57*, 102261. [CrossRef]

19. Erokhin, D.; Yosipof, A.; Komendantova, N. COVID-19 conspiracy theories discussion on Twitter. *Soc. Media + Soc.* **2022**, *8*, 20563051221126051. [CrossRef]

20. Fernandez, M.; Bellogín, A.; Cantador, I. Analysing the effect of recommendation algorithms on the spread of misinformation. In Proceedings of the 16th ACM Web Science Conference, Stuttgart, Germany, 21–24 May 2024; pp. 159–169. [CrossRef]

21. Pathak, R.; Spezzano, F.; Pera, M.S. Understanding the contribution of recommendation algorithms on misinformation recommendation and misinformation dissemination on social networks. *ACM Trans. Web* **2023**, *17*, 1–26. [CrossRef]

22. Gimello-Mesplomb, F. *Decoding January 2025 Los Angeles Wildfires: How (and Why) the Emotional Power of Iconic Fires Revives Ancestral Fears and Fuels Misinformation*; Avignon University: Avignon, France, 2025. [CrossRef]

23. Jagtap, R.; Kumar, A.; Goel, R.; Sharma, S.; Sharma, R.; George, C.P. Misinformation detection on YouTube using video captions. *arXiv* **2021**, arXiv:2107.00941. [CrossRef]

24. Micallef, N.; Sandoval-Castañeda, M.; Cohen, A.; Ahamad, M.; Kumar, S.; Memon, N. Cross-platform multimodal misinformation: Taxonomy, characteristics and detection for textual posts and videos. In Proceedings of the International AAAI Conference on Web and Social Media, Atlanta, GA, USA, 6–9 June 2022; Volume 16, pp. 651–662. [CrossRef]

25. Hussain, M.N.; Tokdemir, S.; Agarwal, N.; Al-Khateeb, S. Analyzing disinformation and crowd manipulation tactics on YouTube. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 28–31 August 2018; pp. 1092–1095. [CrossRef]

26. Ha, L.; Graham, T.; Grey, J.W. Where conspiracy theories flourish: A study of YouTube comments and Bill Gates conspiracy theories. *Harv. Kennedy Sch. Misinformation Rev.* **2022**, *3*, 1–12. [CrossRef]

27. Vayansky, I.; Kumar, S.A.; Li, Z. An evaluation of geotagged twitter data during hurricane Irma using sentiment analysis and topic modeling for disaster resilience. In Proceedings of the 2019 IEEE International Symposium on Technology and Society (ISTAS), Medford, MA, USA, 15–16 November 2019; pp. 1–6. [CrossRef]

28. Yuan, F.; Li, M.; Liu, R. Understanding the evolutions of public responses using social media: Hurricane Matthew case study. *Int. J. Disaster Risk Reduct.* **2020**, *51*, 101798. [CrossRef]

29. Rahmadan, M.C.; Hidayanto, A.N.; Ekasari, D.S.; Purwandari, B. Sentiment analysis and topic modelling using the LDA method related to the flood disaster in Jakarta on Twitter. In Proceedings of the 2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta, Indonesia, 19–20 November 2020; pp. 126–130. [CrossRef]

30. Mendon, S.; Dutta, P.; Behl, A.; Lessmann, S. A hybrid approach of machine learning and lexicons to sentiment analysis: Enhanced insights from Twitter data of natural disasters. *Inf. Syst. Front.* **2021**, *23*, 1145–1168. [CrossRef]

31. Loper, E.; Bird, S. NLTK: The natural language toolkit. *arXiv* **2002**, arXiv:cs/0205028. [CrossRef]

32. Elbagir, S.; Yang, J. Sentiment analysis on Twitter with Python's natural language toolkit and VADER sentiment analyzer. In Proceedings of the IAENG Transactions on Engineering Sciences: Special Issue for the International Association of Engineers Conferences, Hong Kong, 13–15 March 2019; pp. 63–80. [CrossRef]

33. Swann, S. Conspiracy Theory Falsely Links Sean 'Diddy' Combs to Los Angeles Wildfires. PolitiFact. 2025. Available online: https://www.politifact.com/factchecks/2025/jan/16/instagram-posts/conspiracy-theory-falsely-links-sean-diddy-combs-t/ (accessed on 1 February 2025).

34. Jelodar, H.; Wang, Y.; Yuan, C.; Feng, X.; Jiang, X.; Li, Y.; Zhao, L. Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimed. Tools Appl.* **2019**, *78*, 15169–15211. [CrossRef]

35. Chen, S.; Mao, J.; Li, G.; Ma, C.; Cao, Y. Uncovering sentiment and retweet patterns of disaster-related tweets from a spatiotemporal perspective–A case study of Hurricane Harvey. *Telemat. Inform.* **2020**, *47*, 101326. [CrossRef]

36. Lu, Y.; Hu, X.; Wang, F.; Kumar, S.; Liu, H.; Maciejewski, R. Visualizing social media sentiment in disaster scenarios. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18–22 May 2015; pp. 1211–1215. [CrossRef]

37. Hwa Choi, S. Development of a disaster safety sentiment index via social media mining. *J. Public Policy Adm.* **2019**, *3*, 29. [CrossRef]

38. Han, X.; Wang, J. Using social media to mine and analyze public sentiment during a disaster: A case study of the 2018 Shouguang city flood in China. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 185. [CrossRef]

39. Ecker, U.K.; Tay, L.Q.; Roozenbeek, J.; Van Der Linden, S.; Cook, J.; Oreskes, N.; Lewandowsky, S. Why misinformation must not be ignored. *Am. Psychol.* **2024**. *advance online publication*. [CrossRef] [PubMed]

40. Soar, M.; Smith, V.L.; Dentith, M.R.; Barnett, D.; Hannah, K.; Dalla Dalla Riva, G.V.; Sporle, A. *Evaluating the Infodemic: Assessing the Prevalence and Nature of COVID-19 Unreliable and Untrustworthy Information in Aotearoa New Zealand's Social Media, January–August 2020, 6 September 2020*; Te Pūnaha Matatini: Auckland, New Zealand, 2020. Available online: https://www.waikato.ac.nz/assets/Uploads/News-and-events/News/2021/tdp-2020-paper.pdf (accessed on 1 February 2025).

41. Van Prooijen, J.W.; Douglas, K.M. Conspiracy theories as part of history: The role of societal crisis situations. *Mem. Stud.* **2017**, *10*, 323–333. [CrossRef] [PubMed]

42. Kou, Y.; Gui, X.; Chen, Y.; Pine, K. Conspiracy talk on social media: Collective sensemaking during a public health crisis. *Proc. ACM Hum.-Comput. Interact.* **2017**, *1*, 1–21. [CrossRef]

43. Hodson, J.; Veletsianos, G.; Houlden, S. Public responses to COVID-19 information from the public health office on Twitter and YouTube: Iimplications for research practice. *J. Inf. Technol. Politics* **2022**, *19*, 156–164. [CrossRef]

44. Kim, H.; Walker, D. Leveraging volunteer fact checking to identify misinformation about COVID-19 in social media. *Harv. Kennedy Sch. Misinformation Rev.* **2020**, *1*, 10. [CrossRef]

45. Dsouza, J. Sentiment analysis on IMDB movie reviews using VADER with lexical affinity and semantic sentiment expansion. In Proceedings of the 2023 4th IEEE Global Conference for Advancement in Technology (GCAT), Bangalore, India, 6–8 October 2023; pp. 1–7. [CrossRef]

46. Fan, A.; Doshi-Velez, F.; Miratrix, L. Assessing topic model relevance: Evaluation and informative priors. *Stat. Anal. Data Min. ASA Data Sci. J.* **2019**, *12*, 210–222. [CrossRef]

47. Elroy, O.; Erokhin, D.; Komendantova, N.; Yosipof, A. Mining the discussion of Monkeypox misinformation on Twitter using RoBERTa. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*; Springer Nature: Cham, Switzerland, 2023; pp. 429–438. [CrossRef]

48. Chilman, N.; Morant, N.; Lloyd-Evans, B.; Wackett, J.; Johnson, S. Twitter users' views on mental health crisis resolution team care compared with stakeholder interviews and focus groups: Qualitative analysis. *JMIR Ment. Health* **2021**, *8*, e25742. [CrossRef]