

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Climate Change Pulse: An Agentic RAG-Enhanced Framework for Interactive Sentiment Analysis and Global Disaster Data Visualization

ALAN ZHENG¹, CARLOS GONZALEZ²

¹West-Windsor Plainsboro High School North, Plainsboro Township, NJ 08536 USA (e-mail: alanzheng240@gmail.com)

²Computer Science Department, California State Polytechnic University, CA 91768 USA (e-mail: carlos01oscar@gmail.com)

Corresponding author: Alan Zheng (e-mail: alanzheng240@gmail.com).

This work was supported in part by the Climate Change Twitter Dataset research initiative by Dimitrios Effrosynidis et al. (2022).

ABSTRACT Climate change is an urgent global issue, with natural disasters becoming more severe and frequent due to human activities. Understanding public sentiment around these events can inform climate awareness and policy. We developed Climate Change Pulse, a web-based tool that visualizes natural disasters alongside Twitter data to analyze how proximity and time influence climate-related sentiments. Using the Climate Change Twitter Dataset, we examined over 15 million tweets, mapping them with disaster data through an interactive UI. Challenges included missing geospatial data and sentiment classification limitations, addressed by refining data filters and leveraging embedded tweets. Our experiments tested how distance and time around disasters affect sentiment, revealing that proximity intensifies negative emotions, and climate change deniers exhibit surprisingly strong negative sentiments. Compared to prior methodologies focused on data collection or basic sentiment analysis, our approach emphasizes user interactivity and behavioral analysis. Climate Change Pulse offers a dynamic way to understand climate discourse, bridging data insights with public engagement.

INDEX TERMS Natural Language Processing, Machine Learning, Sentiment Analysis, Climate Change, Disaster Data, Data Visualization, RAG Systems

I. INTRODUCTION

CLIMATE change is one of the most prominent and critical issues we are facing today. Extreme heat and climate change-induced natural disasters directly and indirectly impact people worldwide. Climate change from human activities makes disasters more intense and frequent. Hot seasons now keep breaking temperature records, and more severe heat waves, floods, and droughts are normal in many countries. Ice sheets are melting, ocean levels are rising, and warmer oceans generate stronger hurricanes. Importantly, with rising temperatures, tipping points are reached and changes in systems occur that are irreversible. From The Climate Book by Greta Thunberg, "IPCC estimates that global warming will reach 3.2°C by 2100." When Hurricane Sandy hit New York, 8 billion dollars of damage from the storm surge were attributed to climate change. In 2003, a heatwave in Europe caused more than 70,000 premature deaths, and climate change has doubled the chance of occurrence. If global warming is kept under 2°C over 50 years, it could prevent 4.5 million premature deaths in the U.S. alone.

Technology can provide us with information about people's ideas—tools like data analysis can assess individuals' perception of different issues. For climate change, how does being near a disaster impact a person's emotions, or their sentiment? And how can we use data from social media platforms such as X, formerly known as Twitter? Answering these questions is key to better inform the general public and help drive change in policy-making in our society.

Current examples in tweets show potential areas of interest that we aim to further explore. We aim to observe whether natural disasters amplified by climate change impact people's sentiment. Compared to surveys that evaluate individuals' stance and awareness on current issues, people tend to express themselves more intensely and personally on social media. This advantage exists because social media is a less formal environment. Additionally, social media has more younger users and can be more diverse than sample populations used in surveys. As seen from the Pew Research Center report in 2023, when observing user activity among various social platforms such as TikTok, Instagram, Snapchat, YouTube,

and Facebook, "a third of teens use at least one of these five sites almost constantly." This advantageous presentation of information allows us to extract more meaningful insights from the readily collected data.

Climate Change Pulse is a web tool that visualizes disasters onto a world map, overlaid with various tweets. There is a scroll feature for the user to visualize which natural disasters occurred per year. Red dots on the map indicate a recorded disaster, and upon selection, tweets within a 1000-mile radius to the disaster will be displayed. The world map combines tools such as JavaScript and D3 from the following GitHub repository.

Our research aims to use the datasets to see if there are any connections between sentiment and disaster. Both the tweets and disaster data were pooled from the "Climate Change Twitter Dataset" research initiative by Dimitrios Effrosynidis et al. (2022). Over fifteen million data points spanning over thirteen years related to climate change were sourced from the social media platform, where data such as gender, stance, sentiment, and disaster type are included. The tweets are from different people and organizations but are all discussing climate change.

II. RELATED WORK

The approach outlined by Barachi, May El, et al. used a Bi-directional LSTM model for sentiment classification on climate-related social media data. It effectively categorized sentiments but relied heavily on API data collection, which we lacked. Its focus was limited to model training, while our project advanced this by analyzing already labeled data for deeper insights.

The proposed method for Lu, Yafeng, et al. visualized social media sentiment during disasters, using geographic data from the Ebola Twitter dataset. While effective in mapping sentiment, it lacked user interactivity and broader analytical factors. Our project improved on this with a user-friendly UI and an emphasis on behavioral differences between climate change deniers and believers.

Mouronte-López and Subirán analyzed climate change sentiments on Twitter using sentiment models like VADER and topic modeling. Although comprehensive, it didn't explore user behavior or disaster proximity effects. Climate Change Pulse builds on this by integrating interactive visualizations and focusing on how disasters influence sentiment over time and space.

Drawing lessons from existing methods, our project moved beyond the limitations of traditional RAG and fixed data collection by adopting an agentic RAG framework for dynamically generating SQL queries on our static, tabular disaster and tweet data. This strategic shift makes use of the iterative nature and multi-domain capabilities of agentic systems, enabling a more interactive and behavior-focused conversational chatbot than approaches centered on basic sentiment classification or static visualizations.

III. CHALLENGES AND SYSTEM DESIGN

A. WORLD MAP INTERFACE INTEGRATION

One primary challenge is incorporating the dataset with the world map interface. For example, the world map has a dropdown menu showing the names of all the countries; however, some countries have different naming conventions in the datasets. This must be addressed because it can cause issues in the pairing of those countries and the tweets in the area. Other data is also formatted so that JavaScript code cannot work with it. The disaster data's date records needed to be converted into standardized date formats to work with the code. Additionally, creating the interface itself has its own challenges. Disasters are represented with icons, and the map needs to be constantly updated to remove old icons and add new icons. Also, since Elon Musk acquired Twitter, the data cannot be accessed directly. This means getting the contents of tweets by hydrating them is no longer viable.

B. TWITTER DATA ACCESSIBILITY

Another challenge was ensuring Twitter and disaster data were properly structured and accessible for analysis. A significant portion of data had incomplete properties, so they cannot be visualized. Missing latitude and longitude values means that tweets and disasters cannot be placed on the map. This means that in order to render the data, we had to preprocess it by parsing dates, grouping records, and filtering invalid entries. At the same time, if we filtered out invalid entries, it excludes a substantial portion of the data, which is misleading. So an approach to get usable data would also need methods to fill in errors in order to keep as much data as possible.

C. RAG ARCHITECTURE

When deciding on the architecture of our LLM, we explored different methods, such as large-context windows and fine-tuned LLMs. We opted for a RAG-based solution, instructing the model to act as a "data analyst" to interact with the dataset. Due to the complexity of the task and the limitations of accessible data, training and fine-tuning a model seemed less promising compared to the RAG-based solution, where all we need to do is instruct the LLM to generate queries to index our data. We also considered the advantages and drawbacks between an in-memory pandas dataframe versus an out-of-memory SQLite database. Due to the limitations of available cloud-computing resources and the performance differences for this specific task, designing an LLM to generate SQL queries made more sense for the task of retrieving data efficiently and accurately.

IV. SYSTEM OVERVIEW

The web application is a map visualization tool with a conversational chatbot. Users can change the year and observe differences in tweet sentiment and disaster severity across various countries. When the user clicks on a disaster, the system will collect from the tweets dataset and select those within a specific time and distance. For more clarification, the user can ask context-driven questions regarding the dataset.



FIGURE 1. Flowchart showing the code for the application

The data were grouped based on attributes such as year and country, and a map object was created to efficiently store tweets based on ID. The visualization update function incorporates the data to be displayed. We filled in missing data points from our dataset using the Google Maps API. Notably, we filled in the missing latitude and longitudinal coordinates by using the region or address where a disaster occurred—that way, we ensure that we can overlay every disaster into the visualization. We used a Google Maps API key to find the locations based on the data and calculated the coordinates. We also worked more on the data; we removed tweets from the Twitter data that did not have coordinates. Due to the large sample size and the lack of other geospatial information from the Twitter dataset, finding the location for these tweets was not possible.

Pandas was used for extensive data visualization and analysis to better understand the underlying distributions and patterns in the data. For example, one visualization compared the mean sentiment of tweets and their aggressiveness with the worst disasters. The purpose of showing tweets when you click on a disaster is to see if there are changes in sentiment. Specifically, to see if being near a disaster around the time of it is related to sentiment on climate change.

A. COMPONENT A: INTERACTIVE MAP INTERFACE

On the website, cooler countries are in shades of blue while hotter countries are in shades of red. The symbols used for disasters are red dots; when the map is running, they "flicker" on and off. Hovering over a disaster also gives information about the number of deaths from the disaster. In order to show tweets for disasters, windows are used; the windows have a scroll feature to see all the embedded tweets. To exit a tweet window, click outside the window or on another disaster.

The embedTweets function expects a numpy array of tweet IDs as its parameter. This function is immediately called upon execution of the program when the website is first launched. It uses the tweetsElement array which is in code dealing with the Twitter data. After that, it has some attributes that affect the appearance of the embedded tweets. Setting conversation to none means it does not show replies to the tweet. Cards hide images and polls that the tweet would have. Align centers the tweet and theme makes the tweet in light mode.

B. COMPONENT B: GEOSPATIAL DATA PROCESSING

Many data points in the disaster data were missing coordinates, which meant the data was not complete. However, many disasters have a geographic location. Using the Google Maps API, we checked the location of the data to find possible

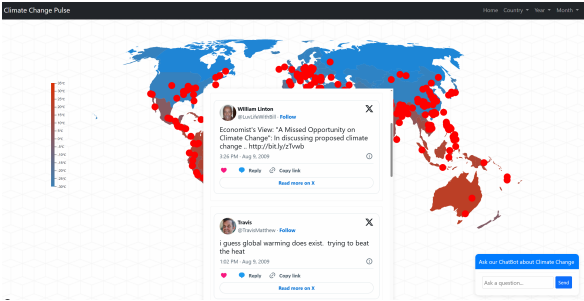


FIGURE 2. Climate Change Pulse web application interface

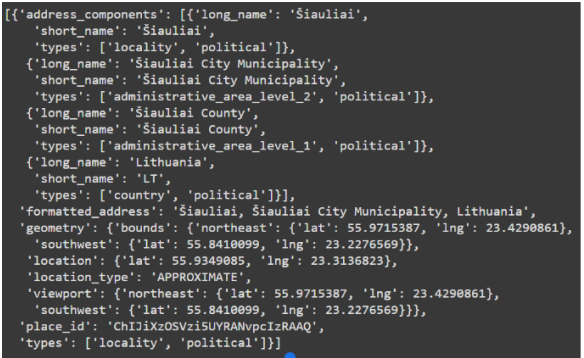


FIGURE 3. Example of possible locations after calling geocode function

locations. Then, taking all of those location coordinates into account, we calculated the average.

In get_coords, we are calling the Google Maps geocode function to get possible locations, then adding them to latitude and longitude sets. Afterwards, we get the average latitude and longitudes. Also, some locations have special characters, so we had to encode and decode them using two character encoding standards. Finally, the modified data is added back into the data file.

C. COMPONENT C: DATA ANALYSIS AND VISUALIZATION

We were particularly interested in answering the following question: When a disaster occurs, does this increase the discussion of climate change in the region? Combining the tools of programming with the large corpus of data, we aim to further explore the data itself by performing various exploratory data analysis techniques, such as aggregating, filtering, and augmenting the data to paint a clearer picture of the granularity of the datasets.

We utilize the matplotlib module in Python to create this comprehensive visualization, which is an example of several graphs we generated to better understand the underlying contents of the data. Several tasks needed to be completed, such as aggregating the sentiment and aggressiveness scores into months per year; labeling the extreme points in the disasters dataset, such as the most costly disaster (the amount, in USD, incurred in damages as a result of the disaster); most deaths occurred; and most people affected. We then overlaid a line

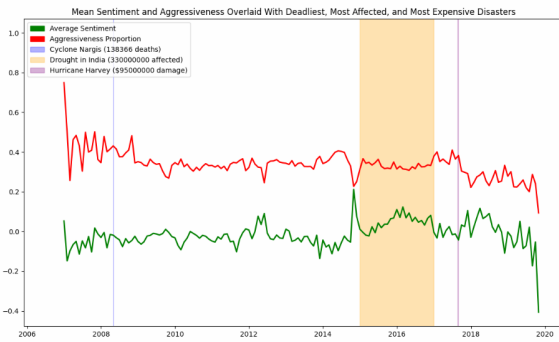


FIGURE 4. Line plot of Twitter sentiment and aggressiveness data and disasters of interest



FIGURE 5. Example of a conversation with AI feature

plot of the mean sentiment and aggressiveness scores onto a graph, where the shaded regions indicate our disaster points of interest. Once all these data points are collected, we simply combine all the data into this one graph.

D. COMPONENT D: AGENTIC RAG SYSTEM

The AI feature is a contextually-aware RAG-based system, where the user can ask questions about the datasets and receive accurate up-to-date information about the data. We first load the databases in-memory using SQLite, retrieving the column names and rows from the original source files, and loading them into two separate databases. The program checks the database to determine which one is appropriate to use, given some user query, to then generate a SQL query. Next, the system validates the SQL expression, retrying if it is invalid or cannot be parsed. We provide error handling for instances such as these, ensuring the LLM can revise its mistakes upon future iteration. The LLM is presented with metadata from the databases, such as the column names, data types, and granularity to ensure accurate retrieval of information. We ultimately are looking at the use case of LLMs to query and interact with large databases, exploring the possibilities of LLMs assisting distributed systems to answer users' questions via an agentic framework.

We ask the LLM to generate a SQL query based on the user's query and dataset of choice. We call the an-

swer_with_table function which takes in the user query and returns a SQL expression that retrieves the information required to answer the user query. The LLM could fail to retrieve information because of errors, due to a nonsensical query received from the user (e.g., 'Please tell me about my finances.'), or a fault from the LLM (e.g., we index from a column which does not exist in the database). We consider these scenarios by feeding the LLM error feedback and modify the prompt to attempt to generate a new query. The retry_count variable keeps track of the number of attempts so they do not exceed max_retries. This process continues until we either successfully retrieve information from the database or until we fall back to some error. The LLM prompts the user with updated information about their query, with the client receiving a textual response.

V. EXPERIMENTAL RESULTS

In the experiment, we aimed to test how distance and time windows around natural disasters influence public sentiment on climate change as expressed on Twitter. We systematically varied distance thresholds (500 km, 1000 km, 2000 km) and time windows (1, 3, and 7 days before and after a disaster) to identify patterns. Using the Climate Change Twitter Dataset, we filtered tweets based on proximity and time relative to disasters and calculated average sentiment scores.

Our most significant findings showed that tweets closer to disasters tend to have stronger negative sentiments, highlighting the emotional impact of proximity. Sentiment also became more negative as time progressed after a disaster, reflecting prolonged concern. Interestingly, climate change deniers exhibited the most negative sentiment, even more than believers, which was unexpected. This suggests emotional responses are influenced by both proximity to disasters and pre-existing beliefs. The experiment underscored the importance of parameter selection and revealed complex interactions between distance, time, and user stance.

A potential blind spot in the program is the selection of distance and time windows around disasters. If the chosen thresholds (e.g., 500 km, 1000 km, or 2000 km) are too narrow or too broad, they might miss key sentiment patterns or dilute meaningful insights. It is critical to test these parameters to ensure the program accurately captures the relationship between disasters and public sentiment.

To test the impact of distance and time windows, the experiment systematically varies these parameters. Distance thresholds are set at 500 km, 1000 km, and 2000 km, while time windows are defined as 1, 3, and 7 days before and after a disaster. These ranges are chosen to reflect realistic geographical and temporal scopes of disaster impact.

The experiment uses the "Climate Change Twitter Dataset" by Dimitrios Effrosynidis et al. (2022), which includes over 15 million tweets related to climate change. Control data is sourced from the same dataset, ensuring consistency in tweet sentiment, stance, and disaster context. By filtering tweets based on distance and time, the program calculates average sentiment scores for each combination of parameters.

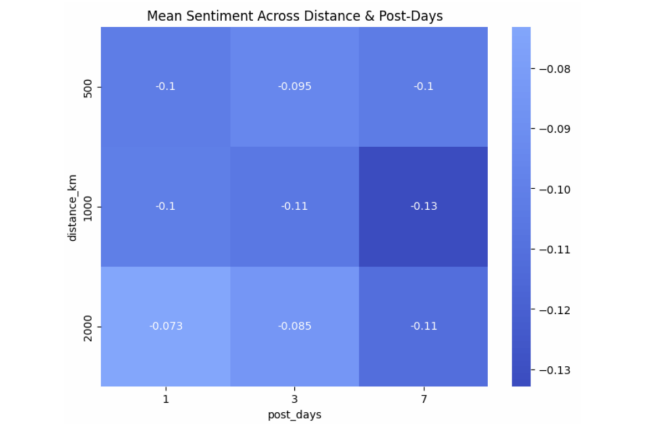


FIGURE 6. Correlation matrix of the mean sentiment across distance (km) and days post-disaster

This setup allows us to observe how sentiment varies with proximity and time relative to disasters. For example, do tweets closer to disasters show stronger negative sentiment? Does sentiment become more negative as time progresses after a disaster? The experiment is designed to isolate these variables and identify optimal thresholds for capturing meaningful sentiment patterns.

The analysis reveals several key insights about the relationship between disasters and public sentiment on Twitter. First, sentiment becomes more negative as the distance threshold increases, suggesting that tweets closer to disaster locations reflect stronger emotional responses. This aligns with the expectation that proximity to a disaster intensifies public concern and emotional expression.

Second, the time window significantly influences sentiment. Pre-event sentiment becomes more negative as the number of days before a disaster increases, potentially reflecting growing anxiety or anticipation. Post-event sentiment also trends more negatively over time, possibly due to prolonged discussions or the accumulation of negative news.

A surprising finding is the stark contrast in sentiment between different stances. Deniers exhibit the most negative sentiment, which may reflect frustration or skepticism toward climate change discourse. Believers, while also negative, show less extreme sentiment, possibly indicating a more measured or concerned tone. Neutral users, as expected, exhibit the least negative sentiment, suggesting a lack of strong emotional investment.

The biggest effect on results appears to be the combination of distance and stance. Tweets closer to disasters from deniers show the most negative sentiment, while neutral users remain relatively unaffected. This suggests that emotional responses to disasters are not only influenced by proximity but also by individuals’ pre-existing beliefs about climate change.

Overall, the experiment highlights the importance of carefully selecting distance and time parameters to capture meaningful sentiment patterns. It also underscores the role of stance in shaping public discourse around disasters, offering

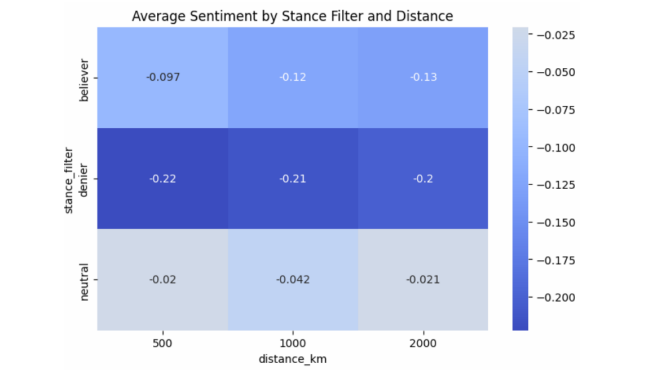


FIGURE 7. Correlation matrix of the average sentiment by stance and distance

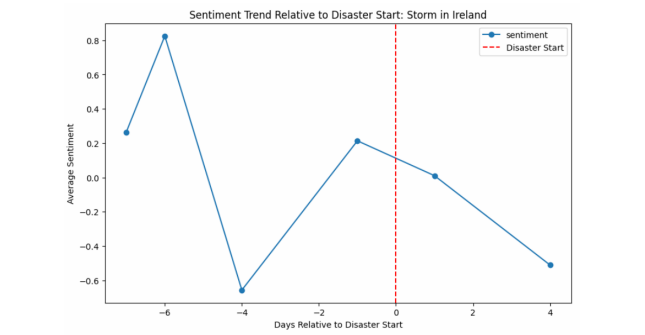


FIGURE 8. Line plot of sentiment across time. The example shows the changes in tweet sentiment across days before and after the disaster

valuable insights for targeted communication strategies.

VI. METHODOLOGY COMPARISON

A. COMPARISON WITH BI-DIRECTIONAL LSTM APPROACH

A similar project was "A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change" by Barachi, May El, et al. The authors had access to data similar to the Climate Change Twitter dataset and some API for data collection (which we did not readily have). It employs a Bi-directional LSTM that assigns and categorizes sentiments into different categories. Today, state-of-the-art models would improve the quality of the performance. Their task was training a model specifically to assign labels, which we build upon.

We used already labeled data from a similar model in order to drive analysis and understanding of the data and draw conclusions from it. We have different scopes, as they were trying to collect data while we focused on data analysis. They reported the accuracy for different emotions and positions. Our application has the potential for internet users and experts alike to offer critique on the issue.

B. COMPARISON WITH GEOGRAPHIC SENTIMENT VISUALIZATION

Lu, Yafeng, et al. addresses a similar problem of using social media data to see how disasters impact sentiment. Using the Ebola Twitter dataset, they are primarily concerned with investigating sentiment trends and behaviors in geo-located Twitter data. Their approach follows by using pre-trained classifiers to label the raw data with a sentiment score, assigning labels based on a majority vote, and then taking the Twitter data classes and passing them into the entropy function. They aim to primarily investigate the quantifiable levels of disagreement among the classes.

Our project builds upon this because we have a UI and more factors in our analysis. Since we have a UI, it is more accessible. Also, we were concerned with investigating the behaviors of deniers and believers.

C. COMPARISON WITH VADER AND TOPIC MODELING

A relevant study which aims to address the analysis of climate change discourse on social media examines sentiments expressed in Twitter interactions related to climate change, analyzing 92,474 tweets to assess sentiment polarity and underlying topics (Mouronte-Lopez and Subiran, 2022). The study employs algorithms such as VADER and TextBlob, alongside unsupervised machine learning techniques, to determine sentiment polarity and utilizes Latent Dirichlet Allocation (LDA) for topic modeling. The findings reveal that discussions on climate change are predominantly negative across various topics, including activism, biodiversity, and sustainability. Not only this, but they also explore differences in sentiment by geography, gender, and account type. Our project builds upon this foundation by providing an interactive web tool that visualizes natural disasters on a world map, overlaid with relevant tweets. Users can explore disasters by year and location, with red dots indicating recorded events.

D. COMPARISON WITH AGENTIC RAG SYSTEMS

The survey by Singh, et al., "Agentic Retrieval-Augmented Generation: A Survey on Agentic RAG," offers a comprehensive overview of various RAG system implementations and their respective strengths and weaknesses. The key goal of their paper is to discuss the advantages of designing one framework over another, which resonates directly with our development journey.

Our solution has undergone several iterations of RAG-based systems, notably exploring the naive RAG approach, to then pivoting towards an agentic framework. Given the static and tabular nature of our external data, we discovered that traditional document chunking and storage were not the most effective solution. Instead, we investigated using RAG techniques to generate SQL queries to assist in filtering our data at a large scale. The inherent iterative nature and multi-domain task strengths of agentic systems show promising results for building a conversational chatbot in our use case, and thus, we have opted to integrate this system into our web application.

VII. CONCLUSION

A. LIMITATIONS AND FUTURE WORK

One limitation of our project is the reliance on pre-labeled data from the Climate Change Twitter Dataset, which may contain inherent biases from the original labeling process. Additionally, many tweets lack geospatial information, limiting the scope of our disaster-to-sentiment correlation analysis. The embedded tweet feature, while useful, restricts access to older or deleted tweets, potentially omitting valuable data. Furthermore, our current sentiment analysis does not fully capture nuanced emotions or the context behind tweets, which could affect the depth of our insights.

If given more time, we would improve the project by incorporating more robust natural language processing techniques, such as fine-tuned transformer models, to enhance sentiment classification. We would also explore geolocation inference methods to fill in missing location data. Expanding the dataset to include real-time data collection via APIs would provide more dynamic insights. Finally, adding advanced data visualization tools would improve user interaction and data interpretation.

B. CONCLUDING REMARKS

Climate Change Pulse bridges the gap between climate change discourse and real-world disaster impacts through data visualization and sentiment analysis. By mapping tweets alongside natural disasters, we provide unique insights into public perception. Our tool fosters awareness and can inform policy discussions, highlighting how climate events influence global sentiment.

ACKNOWLEDGMENT

The authors would like to thank Dimitrios Effrosynidis et al. for providing the Climate Change Twitter Dataset that made this research possible. We also acknowledge the open-source community for the various tools and libraries used in this project.

REFERENCES

- [1] D. Effrosynidis, A. Karasakalidis, A. Symeonidis, and P. Arampatzis, "The climate change twitter dataset," *arXiv preprint arXiv:2201.06024*, 2022.
- [2] G. Thunberg, *The Climate Book*. New York, NY, USA: Penguin Press, 2023.
- [3] M. E. Barachi, A. Al-Rubaie, and A. Al-Khalil, "A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change," *Journal of Cleaner Production*, vol. 312, p. 127708, 2021.
- [4] Y. Lu, X. Hu, X. Wang, and H. Liu, "Visualizing social media sentiment in disaster scenarios," in *Proc. 24th Int. Conf. World Wide Web*, Florence, Italy, 2015, pp. 1211-1216.
- [5] M. Mouronte-López and M. Subirán, "What do Twitter users think about climate change? Characterization of Twitter interactions considering geographical, gender, and account typologies perspectives," *Weather, Climate, and Society*, vol. 14, no. 4, pp. 865-877, 2022.
- [6] A. Singh, A. Ehtesham, and A. Kumar, "Agentic retrieval-augmented generation: A survey on agentic RAG," *arXiv preprint arXiv:2501.09136*, 2025.
- [7] M. Anderson, M. Faverio, and J. Gottfried, "Teens, social media and technology 2023," *Pew Research Center*, Dec. 2023. [Online]. Available: <https://www.pewresearch.org/internet/2023/12/11/teens-social-media-and-technology-2023/>

-
- [8] E. Torres, "Global temperature D3.js data visualization," *GitHub*, 2023.
[Online]. Available: <https://github.com/ema2159/global-temperature>

ALAN ZHENG is a student at West-Windsor Plainsboro High School North, Plainsboro Township, NJ, USA. His research interests include climate change data analysis, natural language processing, and interactive data visualization. He has contributed to the development of Climate Change Pulse, focusing on sentiment analysis and disaster data correlation.

CARLOS GONZALEZ is a student in the Computer Science Department at California State Polytechnic University, CA, USA. His research interests include machine learning, data science, and web application development. He has worked on the system architecture and RAG implementation for Climate Change Pulse.

...