

# Analyzing Public Sentiment and Polarization in U.S. 2024 Presidential Election Discussions on Reddit and 4chan

Krish Shetty  
kshetty@binghamton.edu  
SUNY Binghamton  
Binghamton, New York, USA

Rahulnarayan Iyer  
riyer3@binghamton.edu  
SUNY Binghamton  
Binghamton, New York, USA

Sarthak Zende  
szende@binghamton.edu  
SUNY Binghamton  
Binghamton, New York, USA

## Abstract

This report investigates the evolution of online political discourse during the US 2024 Presidential Election, focusing on Reddit and 4chan. Building on previous projects, we developed an interactive dashboard to analyze activity patterns, thematic discussions, and toxicity levels. The study reveals significant increases in engagement on Election Day, contrasting moderation effectiveness, and evolving discourse patterns across platforms. The findings indicate structured and civil interactions on Reddit compared to the volatile and high toxicity content of 4chan. This research underscores the critical role of platform dynamics in shaping political engagement, leveraging tools such as the ModerateHateSpeech API for real-time toxicity analysis [4, 5].

## Keywords

Social Media Analysis, Political Discourse, Sentiment Analysis, Toxicity Measurement, Reddit, 4chan, ModerateHateSpeech API, Real-time Data Processing, U.S. Presidential Election, Data Visualization

## 1 Introduction

The US 2024 presidential election witnessed unprecedented levels of political discourse across social media platforms, reflecting the growing role of digital spaces in shaping public opinion. Platforms like Reddit and 4chan, despite their contrasting structures and norms, serve as key arenas for political engagement, mobilization, and polarization. While Reddit encourages structured and moderated discussions through its voting-based content ranking system [2], 4chan's /pol/ board is characterized by anonymity, minimal moderation, and reactive posting, often leading to polarized interactions [3].

Understanding how these platforms influence and reflect political discourse is crucial, especially during critical events such as elections. Previous studies highlight the importance of analyzing temporal patterns, sentiment distributions, and toxicity levels in social media discussions to uncover platform-specific user behaviors [1, 4]. However, cross-platform comparative analyses during major political milestones remain underexplored.

Building on the data collection pipeline established in Project 1 and the exploratory analyses conducted in Project 2, this study addresses the research question: *How do activity patterns and discussion topics on Reddit and 4chan evolve in the lead-up, during and after major political events?* Using this data, we developed an interactive dashboard to visualize and analyze user engagement, thematic evolution, and toxicity dynamics during this period.

The dashboard incorporates advanced data visualization techniques, such as volume analysis, toxicity classification, and word

frequency trends, enabling dynamic exploration of discourse patterns. Tools like the ModerateHateSpeech API facilitate real-time toxicity scoring, enhancing the granularity of our analysis [5]. By examining synchronized activity spikes, evolving discussion themes, and moderation efficacy, this study provides actionable insights into the socio-political dynamics of online communities.

Through this comprehensive analysis, we aim to deepen the understanding of how social media platforms influence political engagement and community behavior during pivotal moments. These findings not only contribute to the broader field of social media research but also highlight the implications of platform design and moderation policies in fostering civility and reducing toxicity in online discourse.

## 2 Dataset Description

The dataset used in this study is based on the data collection pipeline established in Project 1 and refined in Project 2.

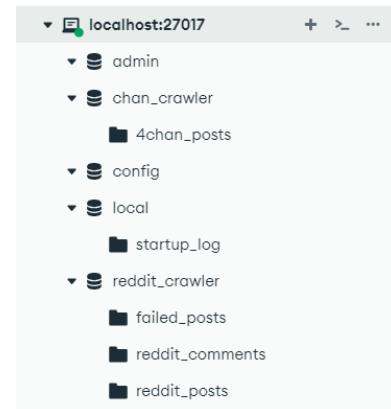


Figure 1: Dataset

It focuses on online political discourse during the US 2024 presidential election. This period captures user engagement around critical events such as Election Day on November 6.

### 2.1 Reddit Dataset

The Reddit dataset comprises around 20000 posts and 1.2 million comments collected from political subreddits, including *r/politics*, *r/Conservative*, and *r/liberal*. These subreddits provide a diverse range of perspectives, enabling cross-sectional analyses of political engagement. Posts and comments are hierarchically structured, allowing for the study of discussion depth and user interaction

dynamics. Key challenges included handling large volumes of comments and normalizing timestamps to ensure accurate temporal alignment [2].

## 2.2 4chan Dataset

The 4chan dataset includes over 4.6 million posts from the /pol/ board, a platform known for its minimal moderation and high-volume, reactive posting culture. Data was organized by threads, with additional processing to address issues such as missing timestamps and thread-level aggregation. Due to MongoDB document size limitations, optimization techniques were employed to efficiently store and analyze unstructured data [3].

## 2.3 Preprocessing

Data preprocessing involved several critical steps to ensure consistency and reliability:

- **Timestamp Normalization:** All timestamps were standardized to a unified time zone, enabling consistent temporal analysis [1].
- **Data Cleaning:** Duplicate and irrelevant entries were removed, and malformed records were corrected.
- **Content Parsing:** Posts and comments were tokenized and normalized, preparing the data for sentiment and toxicity analysis [5].

To assess content toxicity, the ModerateHateSpeech API was integrated into the pipeline, allowing for real-time scoring of toxic interactions. Despite challenges such as API rate limits and system crashes, multi-threaded processing ensured timely and robust data collection.

## 2.4 Summary of Dataset

Table 1 provides a summary of the datasets used in this study.

**Table 1: Dataset Summary (November 1–14, 2024)**

Metric	Reddit Posts	Reddit Comments	4chan Posts
Total Volume	18,244	1,110,597	3,631,695
Daily Average	1,303	79,328	259,406
Peak Day Volume	878	145,459	262,058
Peak Day Date	November 6, 2024 (Election Day)		

The dataset forms a robust foundation for analyzing temporal activity patterns, thematic evolution, and toxicity dynamics across platforms during this politically charged period.

## 3 Methodology

This study leverages a multi-phase approach to analyze the evolution of online political discourse during the U.S. 2024 Presidential Election. The methodology includes three core components: data analysis, interactive tool development, and technical implementation. These steps are designed to address the research question: *How do activity patterns and discussion topics on Reddit and 4chan evolve in the lead-up to, during, and after major political events?*

### 3.1 Data Analysis

The data analysis phase focused on understanding platform-specific dynamics through three main dimensions:

- **Temporal Activity Patterns:** User engagement trends were analyzed across platforms, identifying spikes and lulls in activity around Election Day. These patterns were used to compare the intensity and duration of discussions [2, 3].
- **Toxicity Analysis:** Using the ModerateHateSpeech API, the prevalence of toxic content was quantified, revealing variations in moderation efficacy and civility levels across platforms [5].
- **Thematic Analysis:** Word frequency trends were analyzed to track the evolution of discussion topics, identifying platform-specific discourse patterns and their adaptation to political events [1].

## 3.2 Interactive Tool Development

An interactive web-based dashboard was developed using Flask to facilitate real-time visualization and exploration of the data. The dashboard integrates advanced visualization libraries, such as Plotly and D3.js, and provides the following features:

- Temporal activity plots comparing user engagement on Reddit and 4chan.
- Stacked bar charts visualizing toxicity trends across platforms.
- Interactive word frequency analysis with search and filtering capabilities.

These tools allow users to dynamically query and explore the evolution of discourse patterns across platforms during the election period.

## 3.3 Technical Implementation

The dashboard was implemented using a scalable and efficient architecture:

- **Backend:** MongoDB was used for data storage and retrieval, ensuring efficient handling of large datasets. Multi-threaded processing enabled real-time preprocessing and data normalization.
- **Frontend:** The frontend prioritized responsiveness and accessibility, leveraging frameworks such as Flask for integration with visualization tools.
- **Toxicity Scoring:** The ModerateHateSpeech API was integrated to provide real-time toxicity analysis, enhancing the granularity of content evaluation.
- **Performance Optimization:** Caching mechanisms were implemented to improve query performance for frequent dashboard interactions.

This methodology combines robust data analysis techniques with interactive tools to provide actionable insights into platform-specific behaviors and the dynamics of online political discourse.

## 4 Results and Analysis

This section presents the findings from the analyses conducted using the interactive dashboard, focusing on temporal activity patterns, toxicity trends, and thematic evolution across Reddit and 4chan during the U.S. 2024 Presidential Election period.

### 4.1 Volume Analysis Comparison

**Objective:** To compare activity patterns on Reddit and 4chan before, during, and after Election Day.

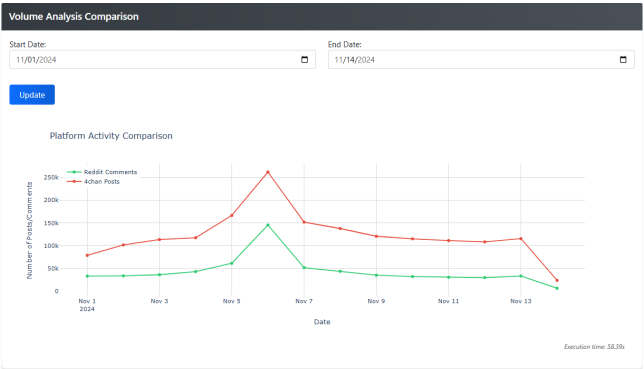


Figure 2: Volume Analysis

**Findings:**

- Both platforms exhibited synchronized spikes in activity on November 6 (Election Day), reflecting the universal impact of major political events on user engagement.
- Reddit:** Maintained a consistent baseline activity with a peak of 145,459 comments on Election Day, highlighting sustained user engagement.
- 4chan:** Displayed more volatile daily fluctuations, with posts peaking at approximately 262,058 on Election Day, indicating a reactive and high-velocity posting culture.

### 4.2 Toxicity Analysis

**Objective:** To evaluate the proportion of toxic content across platforms and assess moderation efficacy.

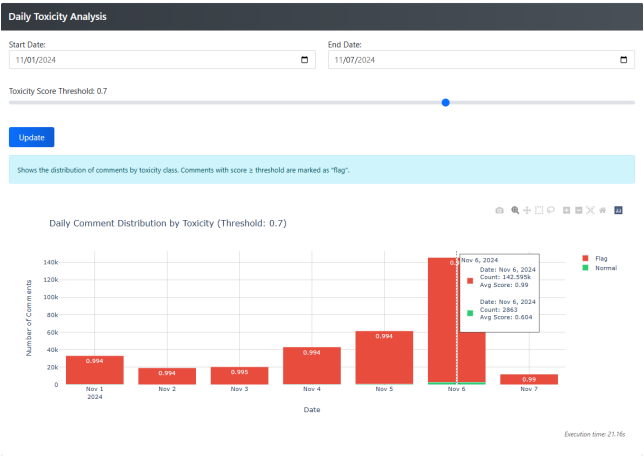


Figure 3: Toxicity Analysis

**Findings:**

- Only 3% of Reddit comments were flagged as toxic, indicating effective moderation compared to 4chan.

- Toxicity levels on Reddit remained consistent throughout the election period, suggesting that increased activity did not lead to higher toxicity ratios.
- 4chan exhibited significantly higher levels of toxicity, aligning with its minimal moderation policies and reflecting a less structured discourse environment.

### 4.3 Word Frequency Analysis

**Objective:** To track the evolution of discussion topics and identify platform-specific discourse patterns.

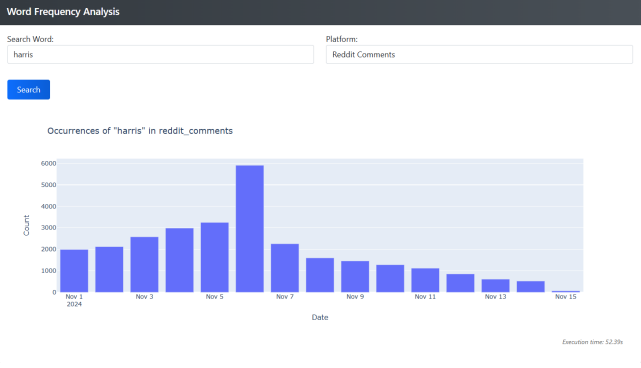


Figure 4: Word Frequency Analysis

**Findings:**

- Terms like *Trump*, *vote*, and *election* dominated Reddit comments, reflecting a strong focus on election-related themes.
- Platform-specific patterns were evident, with Reddit favoring structured discussions on political figures and policies, while 4chan leaned towards reactive and often incendiary topics.
- Thematic shifts aligned closely with major political milestones, demonstrating the adaptability of online discourse to real-world events.

### 4.4 Cross-Platform Comparisons

The findings highlight distinct user behaviors and moderation dynamics:

- Engagement Patterns:** Reddit facilitated collaborative discussions with higher comment-to-post ratios, while 4chan exhibited reactive and high-volume posting.
- Toxicity Trends:** Reddit's effective moderation policies resulted in lower toxicity, whereas 4chan's minimal oversight contributed to higher levels of harmful content.
- Discussion Themes:** Both platforms showed adaptability to political milestones, though thematic focus varied significantly.

### 4.5 Key Insights

- Election Day (November 6) emerged as a focal point for user engagement, emphasizing the universal influence of significant political events.
- Effective moderation plays a crucial role in fostering civility and structured discussions, as demonstrated by Reddit's sustained engagement and lower toxicity levels.

- The interactive dashboard proved invaluable for analyzing cross-platform discourse dynamics and uncovering patterns in user behavior during politically charged events.

## 5 Discussion

The findings of this study provide significant insights into the dynamics of online political discourse during the U.S. 2024 Presidential Election. By examining user engagement patterns, toxicity levels, and thematic shifts across Reddit and 4chan, this study highlights the influence of platform-specific characteristics on political discussions.

### 5.1 Platform-Specific Dynamics

Reddit and 4chan demonstrated distinct user behaviors and engagement patterns:

- **Reddit:** Sustained and structured engagement with higher comment-to-post ratios reflects the platform's moderation policies and community-driven discussions. This aligns with its design to foster deliberative and collaborative discourse [2].
- **4chan:** Volatile and high-velocity posting, coupled with significantly higher toxicity levels, underscores the impact of minimal moderation and anonymity on user behavior [3]. These characteristics create a reactive and often polarizing environment, particularly during politically charged events.

### 5.2 Moderation and Toxicity Trends

The study highlights the critical role of moderation in shaping the civility of online discourse:

- Platforms with robust moderation, such as Reddit, demonstrated lower toxicity levels and maintained consistent civility, even during periods of heightened engagement.
- Conversely, 4chan's lack of moderation resulted in higher prevalence of toxic content, reflecting the need for tailored moderation approaches to mitigate harmful interactions [5].

### 5.3 Thematic Evolution and Political Events

The analysis of discussion themes revealed that both platforms adapted their discourse to align with major political milestones, such as Election Day. Key terms like *Trump*, *vote*, and *election* dominated the conversation, reflecting a shared focus on the electoral process. However, the nature of discussions varied:

- Reddit facilitated nuanced and policy-focused debates, indicating a preference for structured dialogue.
- 4chan exhibited a more reactive discourse, often marked by incendiary and polarizing topics, highlighting the divergent roles of these platforms in shaping political narratives.

### 5.4 Broader Implications

The findings underscore the broader socio-political implications of platform design and moderation policies:

- **Civility in Political Discourse:** Platforms with effective moderation foster more civil and meaningful discussions, which are crucial for constructive political engagement [1].

- **Adaptability to Real-World Events:** Thematic shifts observed in response to political milestones highlight the ability of online communities to reflect and amplify real-world discourse.
- **Interactive Dashboards:** The use of dynamic visualization tools proved effective in uncovering temporal, thematic, and toxicity trends, providing actionable insights into user behavior and platform dynamics.

## 5.5 Limitations and Future Work

While the study provides valuable insights, certain limitations must be acknowledged:

- The analysis focuses on data from a limited time frame (November 1–14, 2024). Extending the study to cover longer periods could provide a more comprehensive understanding of discourse dynamics.
- The study primarily examines Reddit and 4chan, leaving other platforms like Twitter and Facebook unexplored. Future work could expand the scope to include additional platforms for a more holistic analysis.
- Technical constraints, such as system crashes and API rate limits, introduced minor gaps in data collection. Enhancing the robustness of data pipelines could address these challenges.

By addressing these limitations, future research can build upon this study to explore the evolving dynamics of online political discourse across a broader range of platforms and events.

## 6 Conclusion and Future Work

This study provides a comprehensive analysis of online political discourse during the U.S. 2024 Presidential Election, focusing on Reddit and 4chan. Through the development of an interactive dashboard, we explored temporal activity patterns, toxicity trends, and thematic shifts across these platforms. The findings reveal platform-specific dynamics, with Reddit fostering sustained and civil discussions and 4chan exhibiting reactive and high-toxicity interactions. Election Day (November 6) emerged as a focal point for user engagement, demonstrating the universal impact of significant political events on online discourse.

The study highlights the critical role of platform design and moderation policies in shaping user behavior and discourse quality. Reddit's structured moderation facilitated meaningful discussions, while 4chan's minimal oversight contributed to polarized and incendiary content. The thematic analysis demonstrated the adaptability of both platforms to real-world political milestones, emphasizing their role as mirrors and amplifiers of socio-political events.

While the study offers valuable insights, it also identifies limitations that open avenues for future research:

- **Extended Timeframes:** Expanding the analysis to cover a broader timeline would provide a more comprehensive understanding of discourse evolution across different phases of the election cycle.
- **Inclusion of Additional Platforms:** Future work could incorporate platforms such as Twitter, Facebook, and YouTube to explore cross-platform interactions and their collective influence on political narratives.

- **Enhanced Technical Infrastructure:** Addressing challenges such as system crashes and API rate limits by implementing more robust data pipelines would improve the reliability and scalability of the analysis.
- **Sentiment and Emotion Analysis:** Incorporating advanced sentiment analysis techniques could deepen the understanding of user emotions and their impact on discourse dynamics.

The interactive dashboard developed in this study serves as a scalable framework for examining digital discourse during critical socio-political events. By integrating real-time visualization and dynamic querying capabilities, this tool provides actionable insights into user behavior and platform dynamics. Future enhancements could include more granular user engagement metrics, cross-linguistic analysis, and broader sentiment classifications to extend its applicability to diverse contexts.

In conclusion, this study underscores the transformative potential of interactive tools in analyzing online discourse, offering a foundation for understanding the evolving dynamics of social media in politically charged environments. By addressing the identified

limitations and expanding the scope, future research can further illuminate the complex interplay between digital platforms, political events, and societal behavior.

## References

- [1] Kartik Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2018. Quantifying and Understanding Polarization in Social Media. In *Proceedings of the 2018 World Wide Web Conference (WWW)*. ACM, 913–922. <https://doi.org/10.1145/3178876.3186139>
- [2] Eric Gilbert and Karrie Karahalios. 2013. Widespread Emotion in Online Social Networks: Reddit's Emotional Arc. *Proceedings of the 2013 ACM Conference on Computer Supported Cooperative Work (2013)*, 235–244.
- [3] Gabriel Hine, Jeremiah Onaolapo, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Roxana Samaras, Gianluca Stringhini, and Jeremy Blackburn. 2017. Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Role in the 2016 US Election. In *Proceedings of the 26th International Conference on World Wide Web Companion (WWW)*. ACM, 505–518. <https://doi.org/10.1145/3041021.3055134>
- [4] Anna Schmidt and Michael Wiegand. 2017. A survey on hate speech detection using natural language processing. *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media (2017)*, 1–10.
- [5] Zeerak Waseem and Dirk Hovy. 2016. Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. *Proceedings of the NAACL Student Research Workshop (2016)*, 88–93.