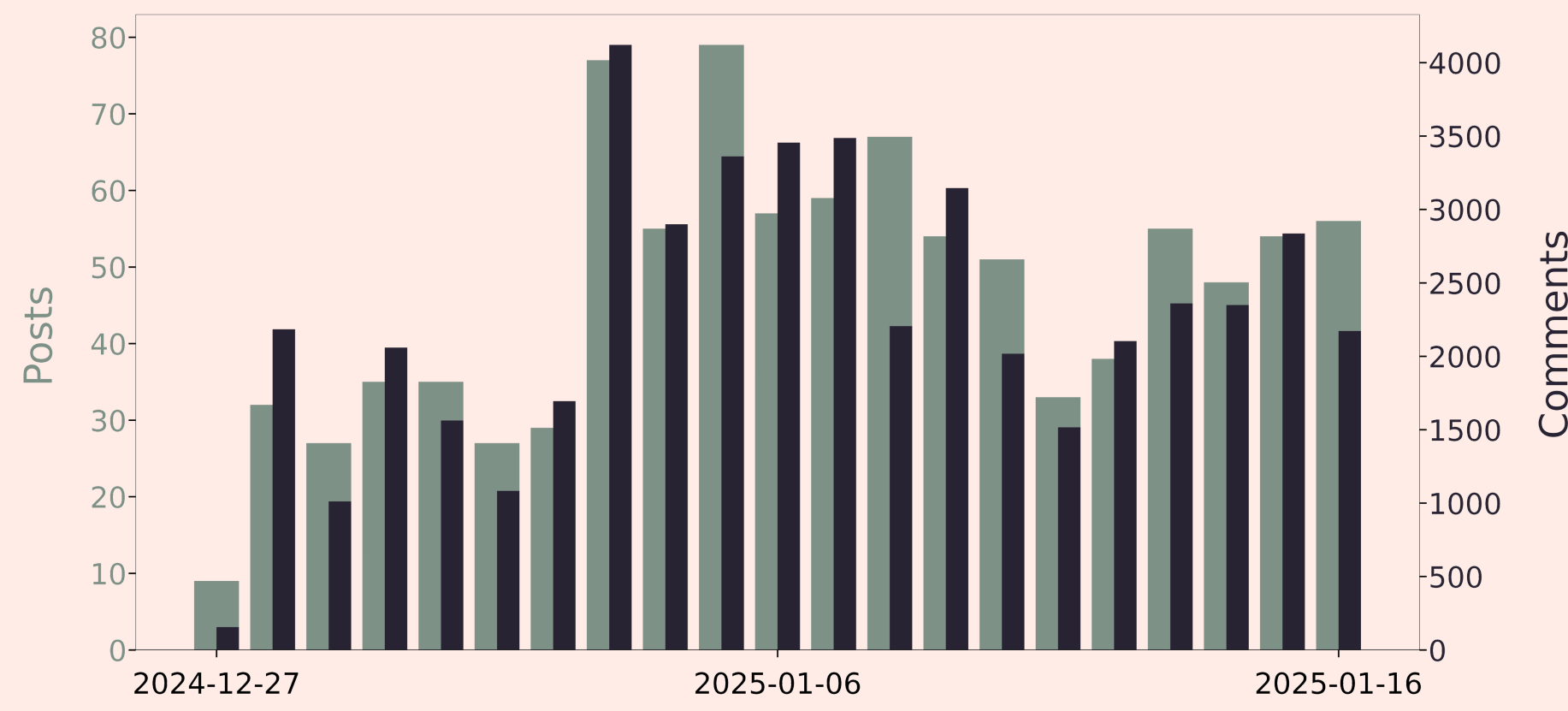




## 1 INTRODUCTION

Reddit's popularity in Austria is rising, with users projected to hit over **2 million** by 2028 (5.1% annual growth) ([1]). The subreddit r/Austria ([2]) hosts over 591,000 members, acting as a hub for Austrian discussions. This study analyzed 977 posts and 47,775 comments from 8,241 users over three weeks, uncovering engagement patterns in the community.

Daily posts and comments distribution in r/Austria subreddit



## 2 RESEARCH FOCUS

Politics dominates r/Austria, with a third of all comments in this category, despite only **18.7%** of posts having the politics flair. This highlights the central role of political discussions in engaging Austrians. Our analysis focused exclusively on posts and comments tagged with the "politics" flair for deeper insights.

### Analysis Steps

- General Data Overview:** Examination of the dataset, including the number of posts, comments, unique users, and activity distribution over time.
- Word Cloud Analysis:** Visualization of the most frequent terms (excluding stopwords) across all contributions, providing a snapshot of prominent topics.
- Topic Analysis:** Identification of overarching political themes
- Sentiment Analysis:** Evaluation of the overall sentiment to capture the tone of discussions in the political flair.

## 3 METHODS AND DATA

Data was collected using the **PRAW** (Python Reddit API Wrapper), with the maximum number of posts retrievable via **subreddit.new()**.

### Data Facts

Total Members of subreddit	591 000 (16.1.2025)
Total Unique Users	8 241
Total Posts / Comments	977 / 47 775
Average Posts/Day	46.52 (STD: 17.25, MIN: 18, MAX: 89)
Average Comments/Day	2 275 (STD: 926.47, MIN: 619, MAX: 4 120)

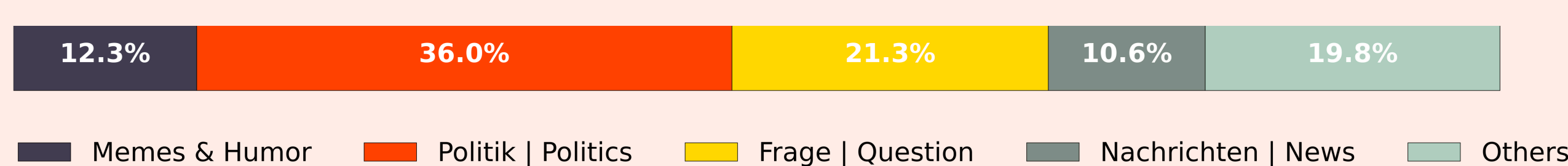
**Statistical analysis** included mean, standard deviation, correlation, and percentage calculations. **Word clouds** were generated with Python's **WordCloud** library, filtering stopwords via spaCy (German) and custom additions. **Topic modeling** used **LDA** via scikit-learn, with coherence evaluated in Gensim. **Sentiment analysis** employed a BERT-based model for language detection (English/German) and sentiment classification.

## 4 RESULTS

### Engagement Statistics

In this analysis, we examined engagement statistics based on the flair categories. Upvotes and comments are key engagement indicators on Reddit. Upvotes show community appreciation and boost visibility, while comments measure interaction and discussion.

Comments Distribution by Flair Category



- Frage | Question:** This flair accounts for 40.02% of posts, 22.41% of comments, and 5.01% of upvotes.
- Politik | Politics:** With 18.73% of posts, 34.94% of comments, and 21.11% of upvotes, it emerges as a dominant topic both in discussions and upvoting behavior.
- Memes & Humor:** While making up 14.84% of posts and 12.68% of comments, this flair receives 49.17% of all upvotes, highlighting its strong appeal to users.
- Nachrichten | News:** Representing 7.47% of posts, 10.65% of comments, and 5.63% of upvotes, this flair holds a moderate presence within the overall discussions.

### Topic Analysis

Coherence Score: **0.47** | Analyzed: 16665 cleaned texts over 3 weeks

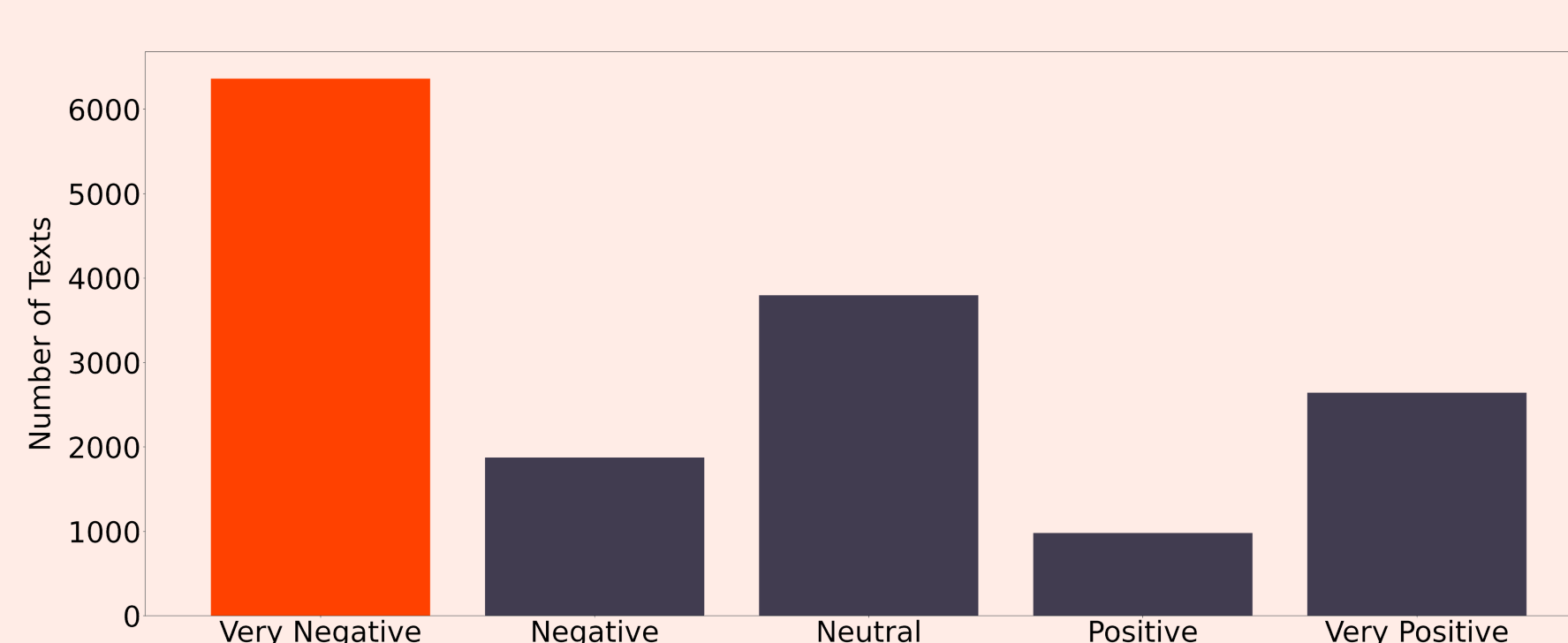
### Identified Key Themes

- Austria & Economy:** Discussions about money, living costs, geopolitical issues.
- Parties & Politicians:** Focus on SPÖ, ÖVP, FPÖ, Neos, Babler, Kickl.
- Social Concerns:** Taxes, financial burdens, societal challenges.
- Government & Power:** Strategies, parties, and power dynamics.

### Sentiment Analysis

The sentiment analysis successfully processed a total of 15657 texts from political posts and comments. The results indicate a predominance of negative sentiment, with the majority of texts rated as "**Very Negative**" (6362). In contrast, positive sentiment is less common, with only 2642 texts receiving a "Very Positive" rating. The distribution of sentiments is as follows: Very Negative: 40.6%, Neutral: 24.2%, Very Positive: 16.9%, Positive: 6.3%, and Negative: 12%. This suggests that while some posts and comments express positivity, the overall discourse in the political context tends toward critical sentiment.

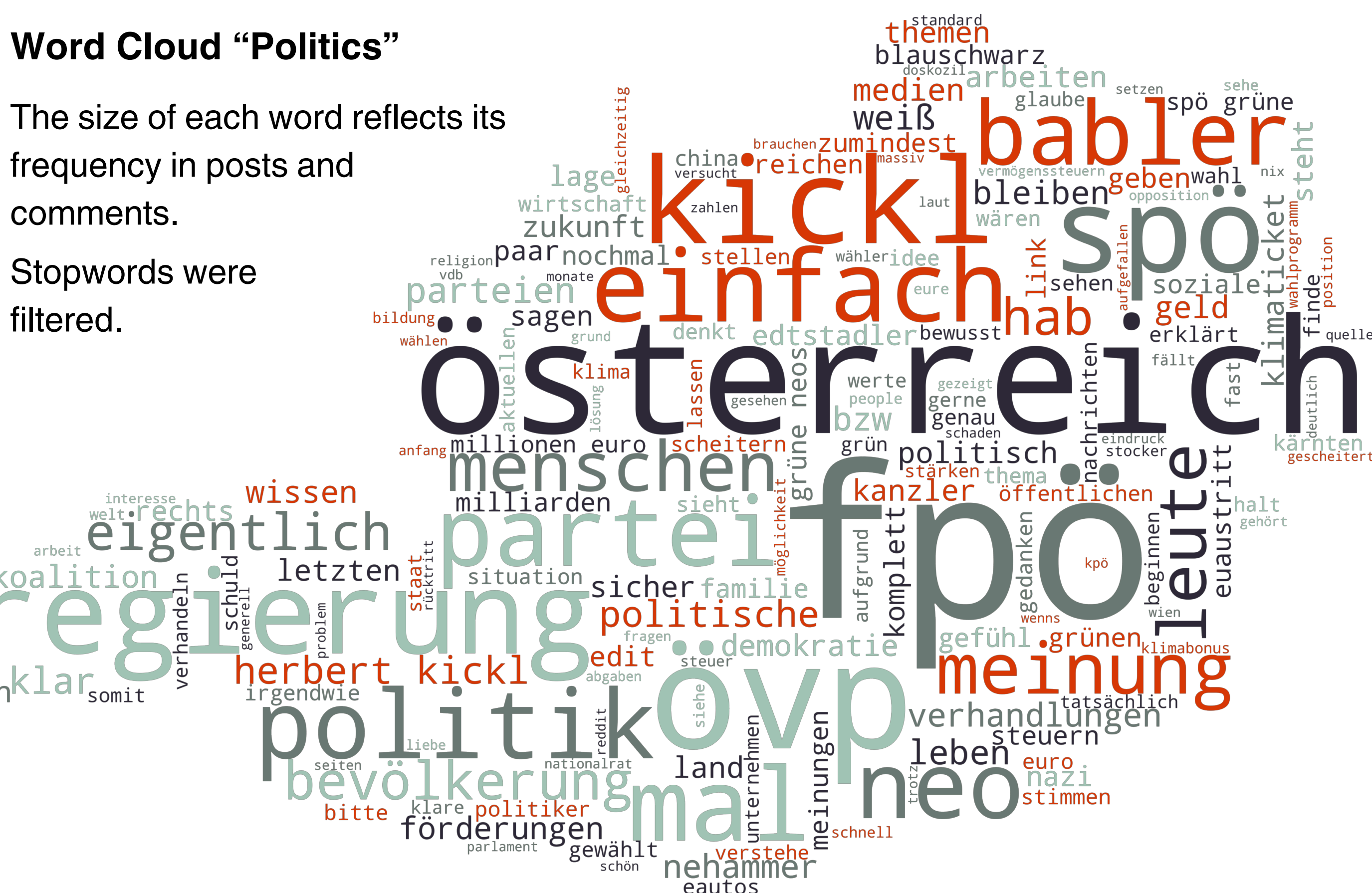
Distribution of sentiments in political posts and comments



### Word Cloud "Politics"

The size of each word reflects its frequency in posts and comments.

Stopwords were filtered.



## 5 CONCLUSION

Findings show that while flairs like Memes & Humor attract the most upvotes (49.17%), they generate less discussion. In contrast, posts tagged as Politik | Politics drive both significant upvotes and comments (34.94%), highlighting political content as a key driver of engagement and debate on the platform.

Topic modeling identified four main themes in political discussions: Austria's economy, political parties and politicians, social concerns, and government power. Despite a moderate coherence score (0.47) suggesting topic overlap or fragmentation, these themes reflect core community concerns. Sentiment analysis revealed a **predominantly negative tone**, with over 40% of texts classified as Very Negative, underscoring a critical view of Austria's political landscape.

Reddit data provides insights into public discourse but is limited by its user demographics. Longer studies and analysis and longer data periods could improve trend insights, while using posts to train language models may uncover nuanced patterns in political dialogue.

## 6 REFERENCES

- [1] Statista GmbH. "Forecast of the number of Reddit users in Austria from 2020 to 2028 (in millions)." - statista.com Website. URL: <https://www.statista.com/> (visited on 2015-01-08)
- [2] Reddit Netherlands B.V. reddit.com Website. URL: <https://www.reddit.com/r/Austria/> (visited on 2025-01-15)

Resources and further details regarding our project are available on **GitHub**.  
<https://github.com/BettyInnovates/redditAustriaAnalysis>

