COSC 2671 – Social Media and Network Analysis
Assignment 2

A Social Media & Network Analysis of the Russia and Ukraine War

Done by:
Avrio Lopes (s4102909)
Mayank Kantharia (s4083242)
Tanishq Bhatia (s4039994)

1. Introduction

1.1 Background

On February 24, 2022, the Russian Federation launched a full-scale military invasion of Ukraine, marking the most significant armed conflict in Europe since World War II *(Harding, 2022; Mearsheimer, 2022)*. As the physical war unfolded, a parallel digital battle emerged across online platforms, where competing narratives, eyewitness accounts, and strategic misinformation campaigns proliferated. This conflict has generated a vast reservoir of user-generated content, offering researchers a unique opportunity to analyze public sentiment, narrative framing, and information flow in real time*(Serrato et al., 2022; Cinelli et al., 2020)*.

Two platforms, Reddit and YouTube, stand out for their scale and complementary affordances in digital discourse:

Reddit is a community-oriented discussion forum that supports in-depth, threaded conversations and decentralized moderation. For this project, we collected data from two prominent war-focused subreddits:

- **r/UkraineRussiaReport** – A subreddit emphasizing curated news links, verified updates, and aggregated situation reports.
- **r/RussiaUkraineWar2022** – A more dynamic forum for real-time updates, user-submitted eyewitness content, and breaking developments.

YouTube, the world's most widely used video-sharing platform, functions as both a media broadcast channel and a participatory comment forum*(Bärtl, 2018)*. We collected video metadata and full comment threads by querying the platform using the keyword phrase **"Russia Ukraine war."** This search returned a range of content, from professional news broadcasts and independent analysis to emotionally charged personal testimonies.

Reddit's structure fosters reflective, interactive discourse, allowing users to vet information collaboratively. In contrast, YouTube comments often capture raw, emotive responses to visual stimuli. By combining these two modalities, this study seeks to uncover not only what narratives dominate but also how emotion and engagement dynamics differ across platforms.

1.2 Significance

This project is situated at the intersection of conflict communication, digital sociology, and computational social science. The insights generated from this dual-platform analysis have implications across several domains:

1. **Humanitarian and Journalistic Monitoring**

    a. Identifying early warnings of humanitarian crises (e.g., civilian casualties, refugee flows) that may precede mainstream media coverage.
    b. Leveraging crowd-sourced reporting to validate or challenge official narratives.
2. **Misinformation and Narrative Warfare**
    a. Detecting disinformation campaigns through anomalous sentiment spikes or unnatural topic distributions.
    b. Understanding the architecture of digital echo chambers and narrative virality.
3. **Policy and Governance Relevance**
    a. Measuring public sentiment in response to international policy decisions such as sanctions, military aid, or peace negotiations.
    b. Capturing public discourse trends that may influence or reflect geopolitical alignments.
4. **Academic and Methodological Contribution**
    a. Enhancing the robustness of cross-platform analysis using combined NLP (VADER, LDA) and SNA methodologies.
    b. Demonstrating scalable frameworks for analyzing digital reactions to real-world crises.

1.3 Research Questions

To guide our investigation, we propose the following research questions:

RQ1: Cross-Platform Sentiment Dynamics

1. How do sentiments (positive, neutral, negative) vary between Reddit discussions and YouTube video comments related to the Russia–Ukraine war?
    a. Are submission-level sentiments on Reddit more neutral than comment-level sentiments?
    b. Is YouTube more emotionally polarized than Reddit?
    c. How does sentiment relate to user interaction metrics (likes, upvotes, replies)?

RQ2: Topic Diversity and Thematic Robustness

2. What are the dominant latent topics discussed on Reddit and YouTube, and how robust are these themes to model configuration changes?
    a. Do recurrent themes emerge across platforms (e.g., frontline updates, foreign aid, sanctions, propaganda)?
    b. How do topics shift with different vectorizers (CountVectorizer vs TF-IDF) or several topics?
    c. Are there noticeable differences in polarity among key topics?

RQ3: Community Structures and Content Polarization

3. What kinds of community clusters emerge on Reddit and YouTube, and how do they differ in sentiment orientation, topical focus, and engagement?
    a. Do Louvain communities on Reddit align with sentiment extremes or topic clusters?
    b. Can YouTube comment threads reveal similar communities via comment trees or user co-commenting behavior?
    c. Are some communities more susceptible to echo-chamber effects?

1.4 Report Structure

To address these questions comprehensively, our report is structured as follows:

1. **Introduction**: Provides background on the conflict and articulates the motivation for analyzing online discourse, and outlines our key research questions.
2. **Data Collection & Preprocessing**: Details on the Reddit and YouTube data pipelines, API usage, comment cleaning, and dataframe preparation.
3. **Sentiment Analysis**: Application of the VADER sentiment model across both platforms, including sentiment classification, distribution, and comparison.
4. **Topic Modeling**: Implementation of Latent Dirichlet Allocation (LDA) on both datasets; includes parameter sensitivity analysis and thematic labeling.
5. **Graph-Based Analysis**: Construction of Reddit's user interaction graph (reply and mention edges), centrality metrics, and structural interpretations.
6. **Community Detection**: Reddit's Louvain modularity clustering and YouTube's inferred community heuristics via user interaction mapping.
7. **Information Diffusion Simulation** (Reddit only): Modeling cascade dynamics under different seed node strategies using the Independent Cascade model.
8. **Findings & Discussion**: Synthesized analysis addressing RQ1–RQ3 with visual evidence and interpretive insight.
9. **Limitations & Future Work**: Discussion on data scope limitations, platform API constraints, and opportunities for future expansion.
10. **Conclusion**: Summary of insights with reflection on real-world relevance and potential policy applications.
11. **References**: Comprehensive list of academic sources, platform documentation, and dataset provenance.

Each section interleaves narrative, code excerpts, and visual artifacts like word clouds, histograms, heatmaps, bar charts, and network diagrams to build a rich, actionable portrait of how the Russia–Ukraine war is discussed, framed, and propagated across Reddit and YouTube.

2. Data Collection & Preprocessing

To ensure a robust and comparable analysis across Reddit and YouTube, we assembled two parallel datasets covering war-related discourse. Below, we describe each source, the sampling strategy, and the initial cleaning steps that produced our working DataFrames.

2.1 Reddit Data Collection

1. **API and Credentials**: We used the Python Reddit API Wrapper (PRAW) with a dedicated user agent.
2. **Subreddit Selection**: We used the following four subreddits:
   a) r/UkraineRussiaReport
   b) r/RussiaUkraineWar2022
3. **Sampling Strategy**: We retrieved the top 1,000 "hot" submissions. For each submission, we recorded its metadata(ID, author, creation timestamp, title, body text, score, upvote ratio, comment count, URL, and permalink).
4. **Comment Retrieval**: From each submission, we fetched up to 50 comments (excluding "load more" placeholders), capturing each comment's ID, author, parent ID, creation time, score, and text body.
5. **Final Data:** After merging data from both Reddit sources, we obtained a total of 1,945 posts and 37,981 comments.

2.2 YouTube Data Collection

1. **API Configuration**: We used the Google Data API (YouTube v3) with a valid developer key.
2. **Search Queries**: The phrase, "Russia Ukraine war," was submitted to retrieve the most recent 300 videos matching each query.
3. **Video Metadata**: For each video ID, we collected the title, description, channel name, publish timestamp, tags, view count, like count, and comment count.
4. **Comment Threads**: We then extracted all top-level comments with replies (up to 1000 per video), including comment ID, author display name, text content, like count, and publish time. Where comments were disabled, we logged and skipped those videos.
5. **Final Data:** After extracting the final dataset, we obtained metadata for 175 videos along with 51,079 comments and replies.

2.3 Cleaning & Filtering

To ensure clean and structured data for analysis, we applied the following preprocessing steps:

1. **Null and Removal Filters:** Eliminated rows where the text field was null, exactly "[removed]" or "[deleted]," or matched common bot-generated messages (e.g., "I am a bot").

2. **Language and Noise Reduction:**
    a. Stripped out URLs and usernames to focus on textual content.
    b. Removed HTML tags and non-ASCII characters, including emojis and special symbols.
3. **Tokenization and Stopword Removal:**
    a. Converted text to lowercase and expanded common chat abbreviations for consistency.
    b. Tokenized on word-like patterns and removed English stopwords to retain meaningful terms.
    c. Applied stemming to reduce words to their root forms and unify variations.
4. **Frequency Filtering:**
    a. Discarded tokens appear only once globally to reduce noise.
    b. Excluded the top 20 most frequent tokens to enhance focus on substantive terms.

Following these steps, each DataFrame included a new column of preprocessed text optimized for sentiment analysis, topic modeling, and network construction.

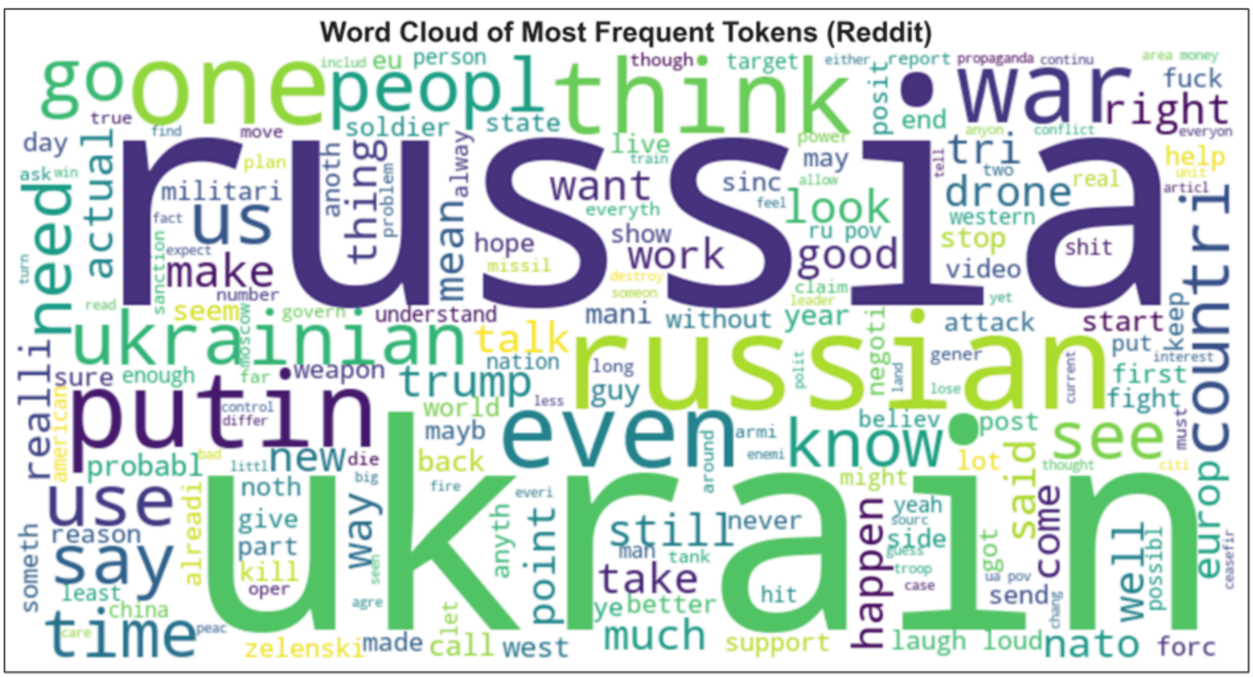

Figure 2.3: Word Cloud of Most Frequent Tokens for Reddit

The preprocessing of word clouds for each platform reveals both commonalities and contrasts in the dominant language used. In *Figure 2.3*, the YouTube cloud highlights how viewers repeatedly mention "ukrain," "russian," and "putin," reflecting direct reactions to video content and key actors, along with more colloquial terms like "think" and "look." By comparison, *Figure 2.4* shows the Reddit cloud emphasizing similar core terms ("ukrain," "russian," "war," "putin") but also a wider spread of peripheral discussion words such as "take," "claim," and "stop," which points to Reddit's broader, forum-style debate where users elaborate on multiple aspects of the conflict rather than simply reacting to visual media.

Figure 2.4: Word Cloud of Processed Text for YouTube

With the data fully assembled and preprocessed, we proceed in Section 3 to apply sentiment analysis and visualize the emotional contours of the online discourse.

3. Sentiment Analysis

In this section, we quantify the emotional tone of online discourse around the Russia–Ukraine war on both Reddit and YouTube. We describe our chosen methodology, present distributional results, and interpret what these reveal about public sentiment.

3.1 Methodology

VADER Sentiment Lexicon: We applied the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analyzer, which is optimized for social media text and balances rule-based heuristics with a sentiment lexicon. VADER returns four scores per document with "positive," "negative," "neutral," and a single "compound" score ranging from -1 (most negative) to +1 (most positive).

Preprocessing Recap: Before scoring, every text was converted to lowercase, with punctuation and URLs removed to eliminate unnecessary elements. It was then tokenized and stemmed to standardize word forms. Finally, rare and overly common tokens were filtered out to reduce noise and focus on meaningful content.

This ensures that VADER sees the core sentiment-bearing words without noise.

Classification Thresholds: We mapped VADER's compound score to three categories:

1. Positive: score ≥ +0.01
2. Negative: score ≤ –0.01
3. Neutral: –0.01 < score < +0.01

These thresholds balance sensitivity (capturing slight sentiment) with specificity (avoiding mislabeling neutral statements).

3.2 Reddit Sentiment Distribution

Figures 3.1 and 3.2 offer two interconnected viewpoints on the sentiment distribution within YouTube comments.

In *Figure 3.1*, the histogram of VADER compound scores reveals a pronounced clustering around neutral values, evidenced by the towering bar at zero, with more gradual tails extending into both positive and negative regions. This central peak suggests that many posts and comments convey balanced or informational language, while a non-trivial number of highly positive or negative outliers reflect moments of strong emotional expression.

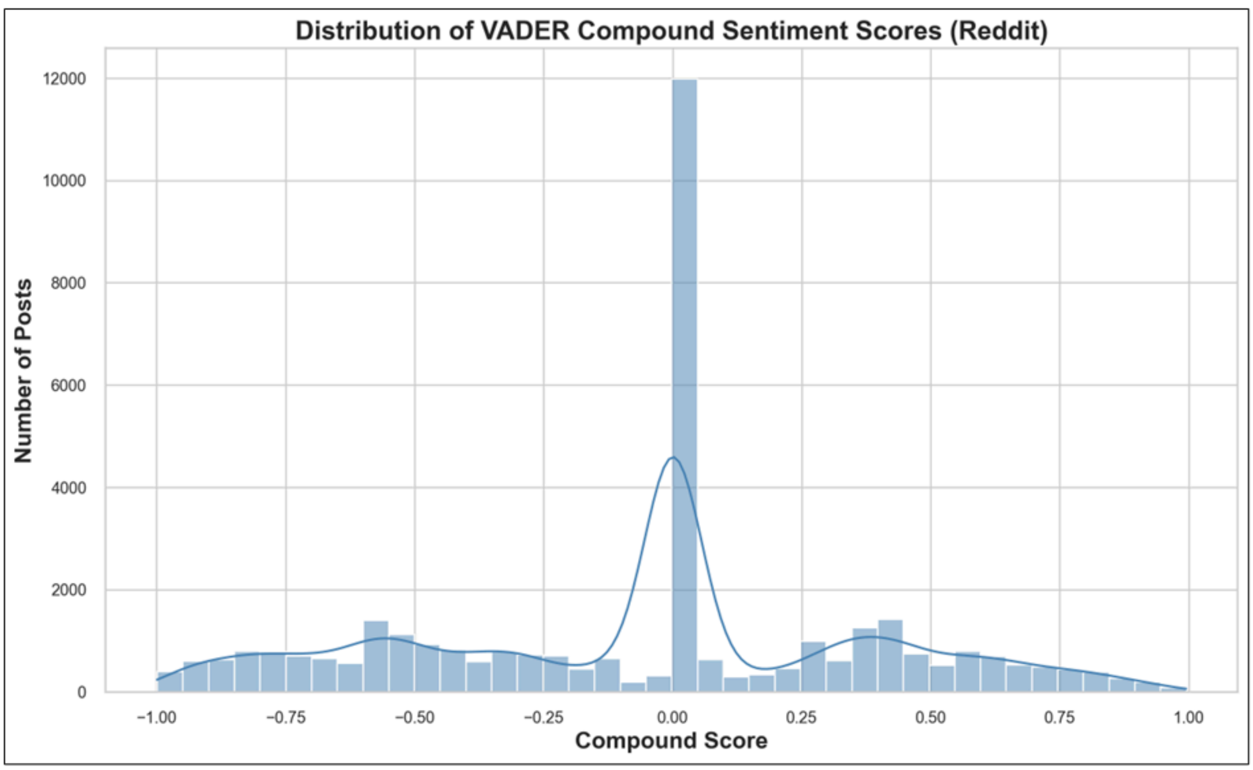

Figure 3.1: Histogram of Reddit compound scores

Figure 3.2 illustrates the proportion of posts classified as positive, neutral, and negative. Among the total dataset, **negative sentiment** posts are the most prevalent, accounting for **37.49%**. The **neutral sentiment** posts represent **31.49%**, while **positive sentiment** posts comprise **31.02%**. This distribution suggests that discussions in the dataset skew toward a more negative tone, potentially indicating concerns, criticisms, or dissatisfaction expressed within the community. The relatively balanced proportion between positive and neutral sentiment posts implies a mix of constructive engagement and impartial discourse. Understanding this trend provides deeper insights into the general mood of the discussion, highlighting areas that may require further analysis, such as factors driving negative sentiment or themes underlying neutral and positive interactions.

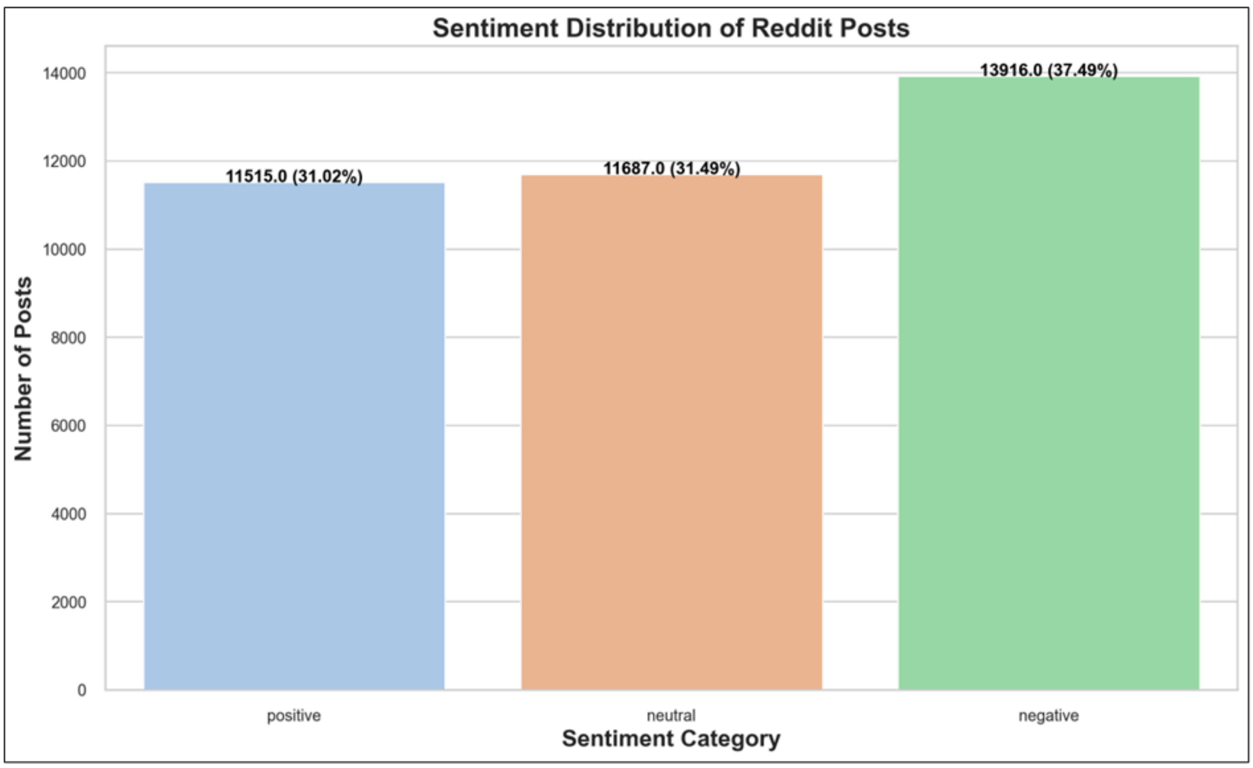

Figure 3.2: Bar chart of Reddit sentiment counts.

3.3 YouTube Sentiment Distribution

Between *Figure 3.3* and *Figure 3.4*, we can observe two complementary perspectives on YouTube comment sentiment.

Figure 3.3 presents the full distribution of VADER compound scores for all comments, plotted as a histogram with an overlaid kernel density estimate. The distribution is sharply peaked at zero, highlighting a strong neutral bias, while the tails extend nearly symmetrically into both positive and negative territory. The summary box reports a slightly negative **mean (≈–0.075)**, a median exactly at 0.0, a **standard deviation** around **0.442**, and a modest **negative skew (≈–0.051)**, indicating that although neutrality dominates, there is a small lean toward negative expressions.

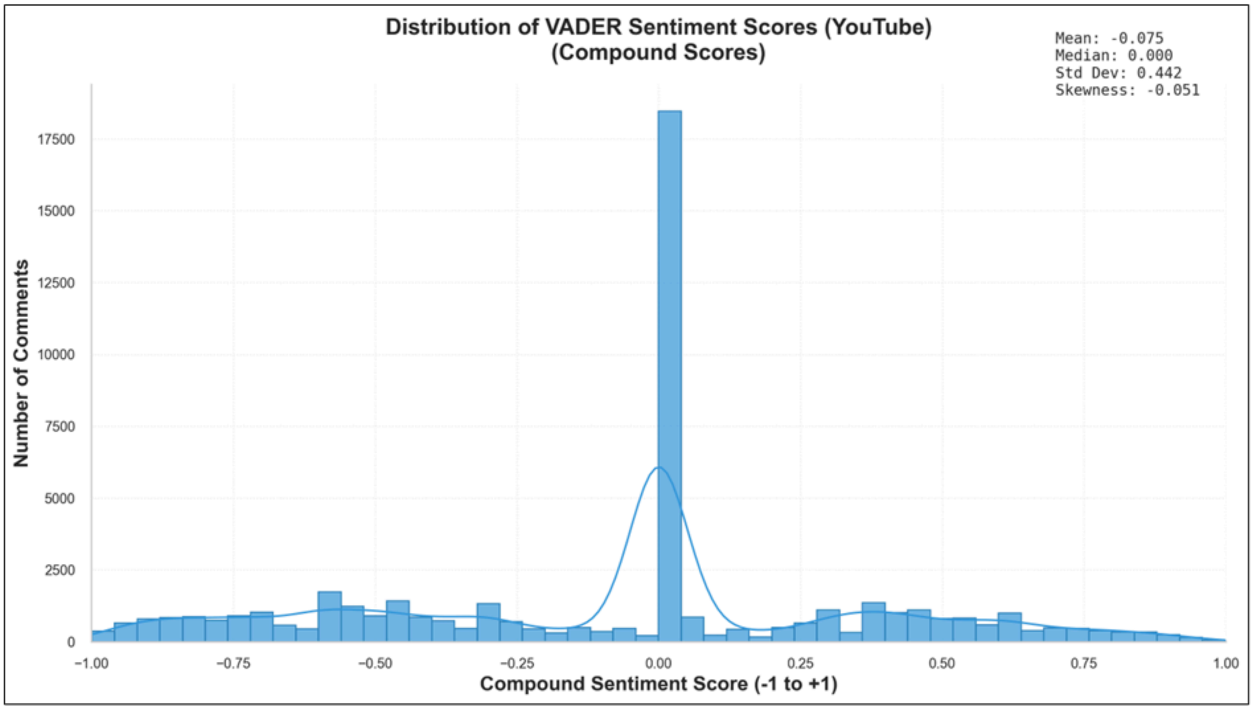

Figure 3.3: Histogram of YouTube compound scores

Moving to *Figure 3.4,* the bar chart illustrates the sentiment distribution of comments related to the Russia–Ukraine conflict, measured using dynamic thresholds. **Neutral comments** make up a substantial portion of the discussion, accounting for **35.54%**, while **negative sentiment** slightly surpasses it at **36.92%**. **Positive comments**, though fewer, still represent a notable **27.55%** of the total dataset. This distribution highlights the prominence of neutral engagement, indicating a measured approach to discourse, while the prevalence of negative sentiment suggests an ongoing intensity in public reactions. Despite being the least frequent, positive comments remain a significant factor, reflecting support and optimism in certain discussions. Together, these insights underscore the complex and polarized nature of sentiment surrounding the conflict.

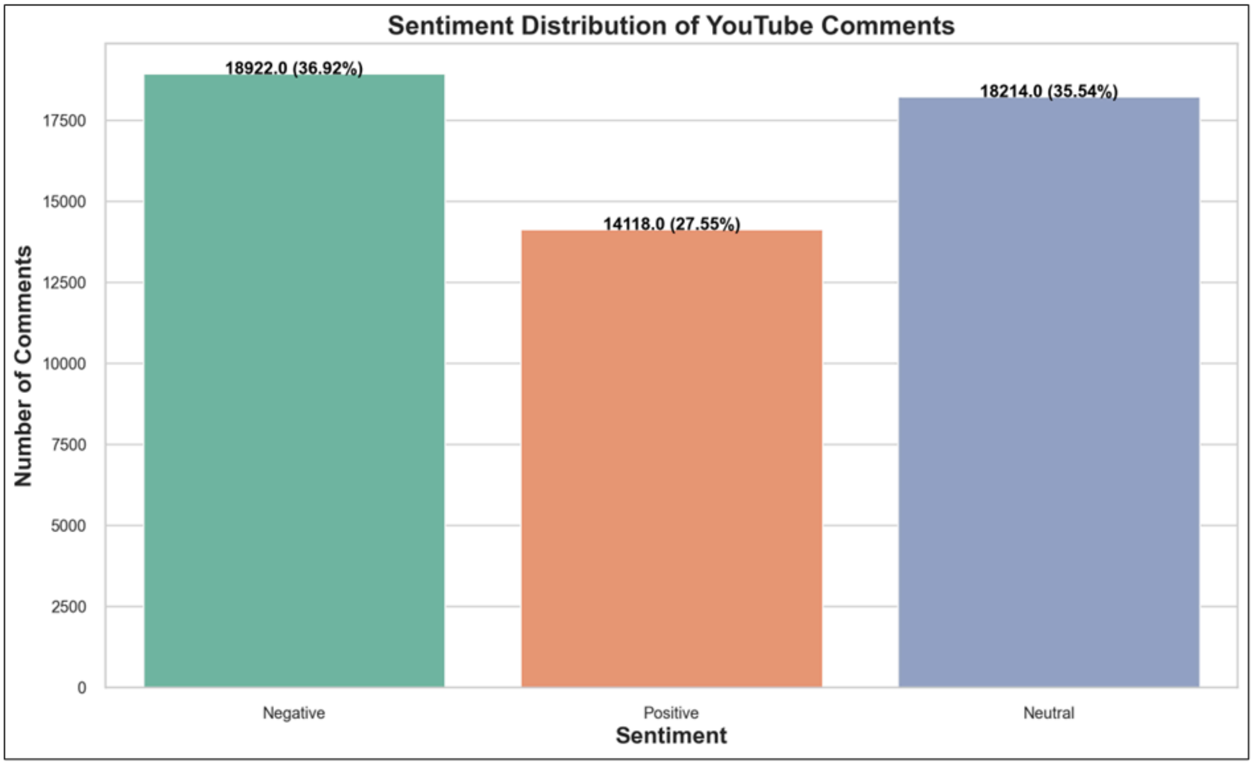

Figure 3.4: Bar chart of YouTube sentiment counts.

3.4 Comparative Insights

Having already examined the sentiment distribution in *Figures 3.1 to 3.4*, we can now step back and compare the overall breakdown across both platforms. *Table 3.1* consolidates these findings, showing the proportion of positive, neutral, and negative comments on Reddit versus YouTube. This comparison highlights how sentiment trends differ between the two communities, even when analyzed using the same dynamic thresholds.

Platform	**Positive**	**Neutral**	**Negative**
Reddit	31.02%	31.49%	37.49%
YouTube	27.55%	35.54%	36.92%

Table 3.1: High-level sentiment breakdown for both platforms.

Key Takeaways:

1. Reddit's sentiment distribution remains relatively balanced across positive, neutral, and negative categories, reinforcing its role as a space for discussion and debate rather than purely emotional reactions.

2. YouTube's comment sections show a stronger polarization, with a higher proportion of both negative and positive sentiment. This reflects how video content and real-time engagement often provoke more intense expressions of support or criticism.
3. Neutral sentiment is more prevalent on YouTube, suggesting that while discussions can be polarized, a significant portion of users engage in measured discourse rather than extreme reactions.
4. Reddit's comments skew slightly more negative than its posts, indicating that peer-to-peer interactions may amplify critical viewpoints as users challenge each other's perspectives.

3.5 Implications

1. **Crisis Detection:** Negative sentiment spikes can flag urgent developments (e.g., major offensives, civilian atrocities) faster than manual monitoring.
2. **Content Moderation:** Platforms might prioritize fact-checking or human review for threads exhibiting sustained negative sentiment, as these may be fertile ground for rumors or sensationalism.
3. **Aid and Policy Messaging:** Humanitarian organizations can tailor outreach, positive reinforcement, or empathetic appeals based on the emotional climate of target communities.

With sentiment patterns established, we next turn in Section 4 to uncover the underlying thematic structure through topic modeling.

4. Topic Modeling

To uncover the latent thematic structure in our Reddit and YouTube corpora, we employed **Latent Dirichlet Allocation (LDA)**. LDA treats each document as a mixture of topics and each topic as a probability distribution over words, allowing us to surface the primary themes driving online discourse about the Russia–Ukraine war.

4.1 Methodology

1. **Vectorization:** We transformed each preprocessed text string into a document–term matrix using Scikit-learn's **CountVectorizer**, carefully tuning key parameters such as max_df, min_df, and max_features to optimize LDA performance based on the characteristics of the data.
2. **LDA Training**: An **LDA** model was fitted with six topics using the "online" learning algorithm for scalability and a fixed random seed for reproducibility.
3. **Topic Extraction**: For each topic, we identified the top five words by weight and distilled concise labels summarizing their themes.

4.2 Model Selection & Parameter Tuning

To validate our choice on the number of topics, we varied the number of topics k from 2 to 10 and recorded two standard metrics:

1) **Perplexity** (lower values indicate better model fit to the data)
2) **Coherence (c_v)** (higher values indicate more semantically coherent topics)

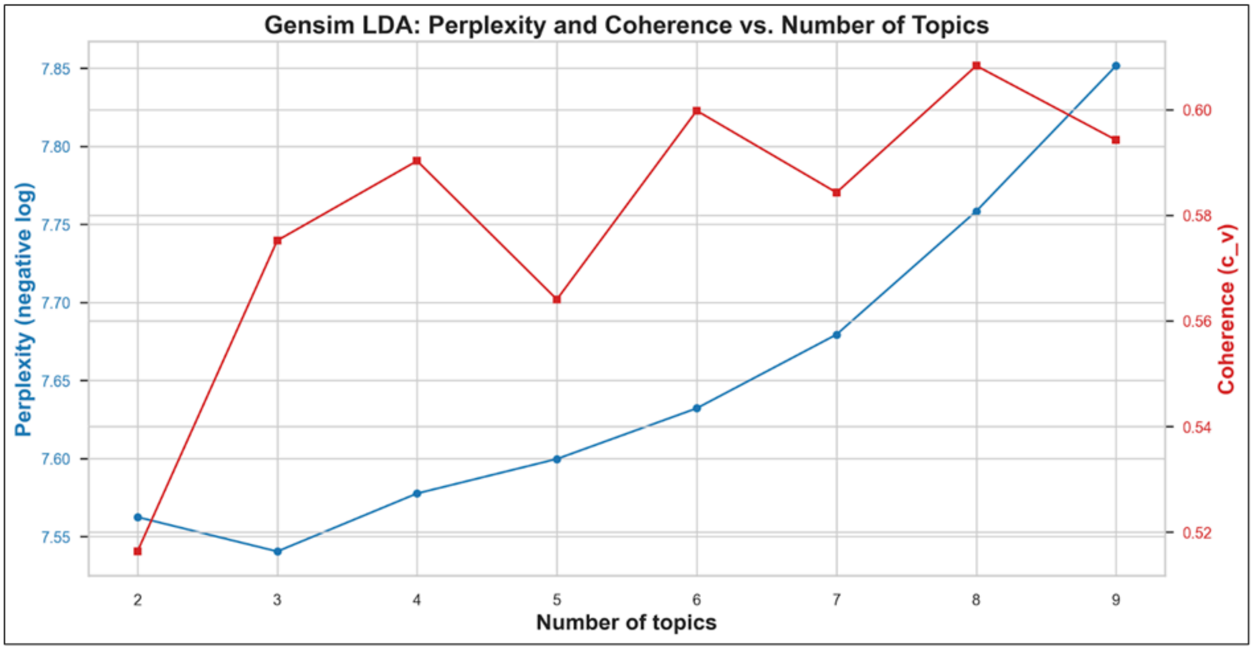

Figure 4.1: Perplexity (blue) and Coherence (red) plotted against the number of topics.

The best choice here is **6 topics**. At six, the model achieves a strong balance: its perplexity is still low (≈ 7.63) while coherence (≈ 0.60) is very close to its peak. Picking fewer topics (e.g., 3) gives slightly lower perplexity but at the cost of noticeably lower coherence, meaning the topics are less semantically tight. Going out to eight or nine topics only marginally improves coherence but drives perplexity much higher, indicating overfitting and less reliable topic assignment. In short, six topics deliver the sweet spot between a well-fitting model and interpretable, meaningful themes.

4.3 Topic Extraction & Visualization

4.3.1 Reddit Topics

We applied LDA to the **r/RussiaUkraineWar2022** corpus and obtained six distinct topics. The top five words for each topic, along with our interpretive labels, are presented in *Table 4.1.*

Topic	Top-10 Words	Summary Label
1	russia, ukrain, war, want, countri, nato, laugh, europ, stop, loud	General war overview & calls to action, Discussions framing the conflict and urging halts or sanctions.

2	putin, fuck, peopl, ukrain, guy, know, trump, die, man, fight	Anger toward leadership & casualties, Strong emotional reactions to leaders and loss of life.
3	like, peopl, look, russian, know, say, think, thing, make, someth	Personal opinions & commentary, Users expressing individual viewpoints and reflections.
4	ukrain, putin, said, russia, trump, russian, talk, state, presid, war	Political leadership narratives, Quotes and discussion of speeches by heads of state.
5	ukrainian, russian, drone, pov, soldier, forc, ru, ua, kill, civilian	On-the-ground conflict & civilian impact, Accounts of drone strikes, troop actions, and civilians.
6	russia, use, ukrain, missil, russian, weapon, year, drone, time, like	Military hardware & timeline of attacks, Focus on missiles, drones, and chronology of eve

Table 4.1: Reddit LDA Topics.

Figure 4.2 is the combined PyLDAvis view showing the inter-topic distance map on the left (with Topic 1 highlighted) and, on the right, a bar chart of the top 30 most relevant terms for Topic 1, illustrating both overall frequency and topic-specific saliency.

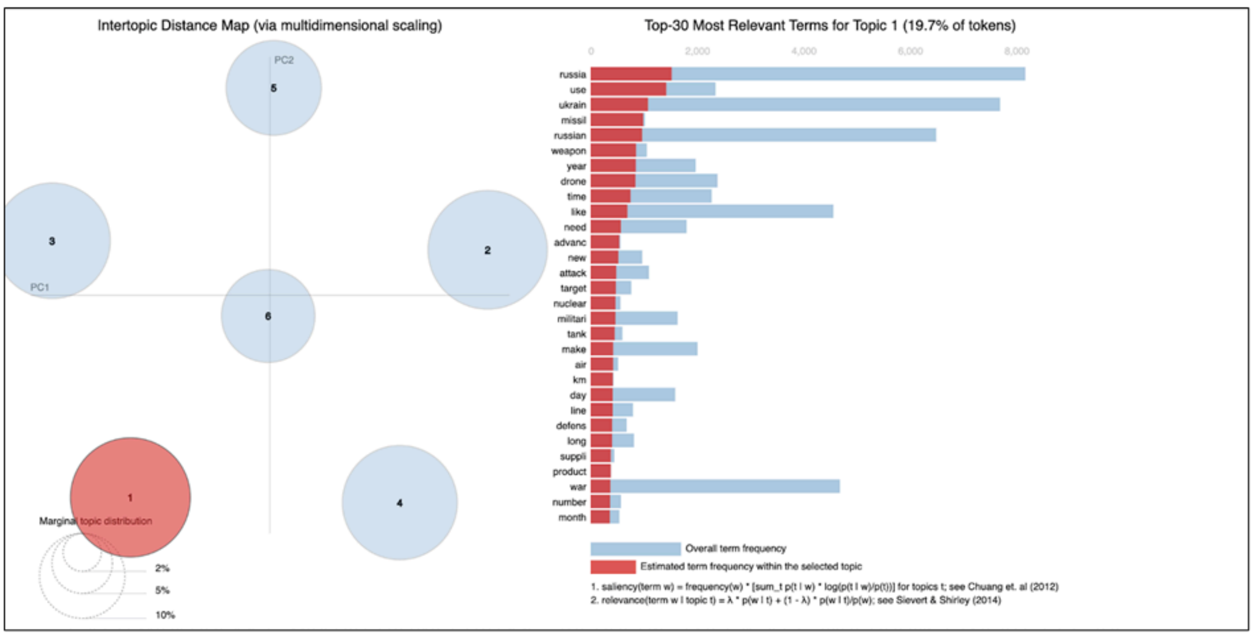

Figure 4.2: Horizontal bar charts of the top words for the top Reddit topic.

In *Figure 4.3* are the six Reddit LDA topic word clouds, with each panel (Topic 1 through Topic 6) visualizing the most salient terms scaled by frequency associated with that topic.

Figure 4.3: Word clouds for Reddit Topics 1–6.

4.3.2 YouTube Topics

Separately, we ran LDA on the YouTube comment corpus, extracting six topics. *Table 4.2* lists the top ten words for each topic along with concise thematic labels.

Topic	Top-10 Words	Summary Label
1	putin, trump, like, man, play, thank, better, god, love, truth	Tributes and morale (support and solidarity), Expressions of admiration, and positive reinforcement.
2	believ, deal, continu, media, power, care, ye, propaganda, person, watch	Media critique and propaganda discussions, Debates over truth, bias, and information control.
3	russia, ukrain, russian, ukrainian, nato, finland, year, militari, war, attack	Broader geopolitical context and NATO expansion, Discussions situating the war in regional security.
4	putin, trump, war, know, talk, zelenski, presid, like, say, stop	Leaders & speeches, sentiments, Reactions to public addresses by Putin, Zelenskiy, Trump, etc.
5	russia, ukrain, want, countri, war, eu, europ, nato, fight, start	Calls for action & alliances, Arguments for military aid, sanctions, or alliance support.
6	news, china, laugh, new, india, loud, channel, good, world, live	Global news & channel commentary, Meta-discussion on news sources, channels, and live updates.

Table 4.2: YouTube LDA Topics.

In *Figure 4.4*, similar to our Reddit topic bar chart, we have highlighted the top YouTube topic.

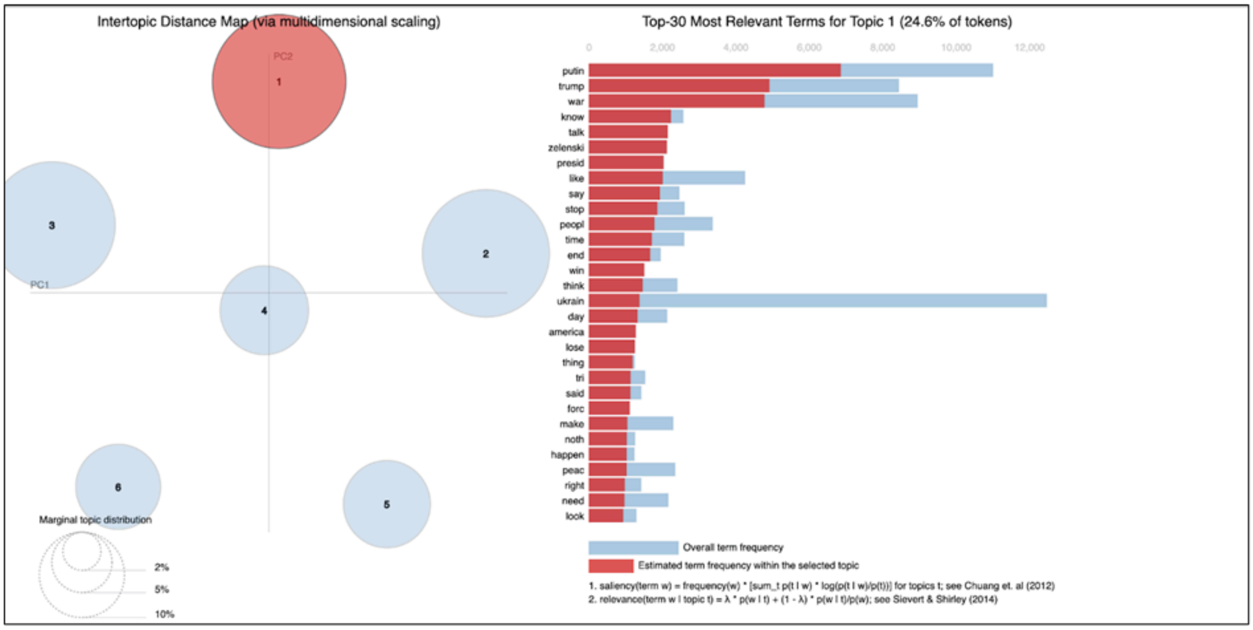

Figure 4.4: Horizontal bar charts of the top words for the top YouTube Topic.

In *Figure 4.5* are the six YouTube LDA topic word clouds, with each panel (Topic 1 through Topic 6) visualizing the most salient terms scaled by frequency associated with that topic.

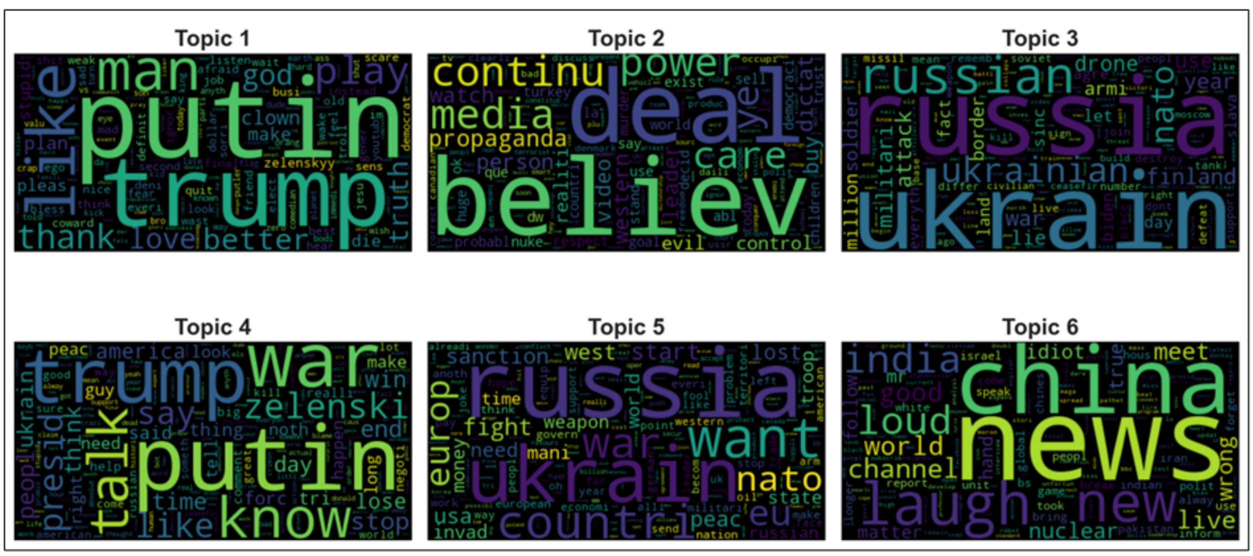

Figure 4.5: Word clouds for YouTube Topics 1–6.

4.4 Interpretation & Insights

1. **Comparative Emphases:**
    a. Reddit's topics range from broad conflict discourse (Topic 1: *russia, ukrain, war, nato, europe*) and raw leadership reactions (Topic 2: *putin, trump, die, fight*) to

reflective commentary (Topic 3: *like, think, say*) and formal political debate (Topic 4: *ukrain, putin, presid, talk*), with human-cost narratives (Topic 5: *drone, soldier, civilian*) and tactical hardware discussion (Topic 6: *missil, weapon, drone*).
   b. YouTube's clusters skew toward moral leadership narratives (Topic 1: *putin, truth, love*), media and propaganda critique (Topic 2: *media, propaganda, power*), hard-news conflict reporting (Topic 3: *russia, militari, attack, finland*), political dialogue (Topic 4: *zelenski, talk, say*), broader war framing (Topic 5: *eu, fight, start*), and platform-driven engagement themes (Topic 6: *news, channel, live*).
2. **Shared Themes:**
   Both platforms foreground leadership storylines (Reddit Topic 2 and YouTube Topic 1), wider geopolitical alignments (Reddit Topic 1 & YouTube Topic 3), and human-impact or equipment debates (Reddit Topic 5 and YouTube Topic 5).
3. **Actionable Findings:**
   a. Peace-and-diplomacy messaging could tap into the political debate clusters (Reddit Topic 4, YouTube Topic 4) to reach audiences already focused on negotiation updates.
   b. Misinformation monitoring efforts should zero in on propaganda streams (YouTube Topic 2) and detailed hardware discussions (Reddit Topic 6) to intercept and correct false narratives.

5. Graph-Based Analysis

To understand how users interact with Russian and Ukrainian content on social platforms, we constructed directed reply networks for both Reddit and YouTube *(Stieglitz et al., 2018)*. In these graphs, each node corresponds to a unique user, and a directed edge from A to B indicates that A replied to B's comment (with edge weights tracking the total number of replies). Below, we describe key structural properties, identify central "hubs," and explore how different centrality measures relate to one another.

5.1 Reddit Reply Network

After filtering out deleted or automated accounts and retaining only genuine user-to-user replies, the Reddit network comprises **5,726 nodes** and **14,476 edges**, with an average degree of **5.06** and **0 isolated** users who neither reply nor receive replies. The graph is almost entirely weakly connected (99.1 % of nodes), while strong connectivity (mutual reply pairs) remains low, reflecting predominantly one-way reply patterns (*Figure 5.1*).

	In-Degree	Out-Degree	Degree Centrality	Eigenvector Centrality	Closeness Centrality
Count	5726.000000	5726.000000	5726.000000	5.726000e+03	5726.000000
Mean	2.528117	2.528117	0.000883	3.125711e-03	0.071725
Std. Dev	7.050324	6.650803	0.002304	1.284137e-02	0.066798
Min	0.000000	0.000000	0.000175	2.499803e-23	0.000000
25th Percentile	0.000000	1.000000	0.000175	2.499803e-23	0.000000
50th Percentile	1.000000	1.000000	0.000349	9.772254e-07	0.099828
75th Percentile	2.000000	2.000000	0.000699	5.565556e-04	0.133848
Max	251.000000	257.000000	0.088734	2.416588e-01	0.219046

Figure 5.1: Summary statistics for the full Reddit reply network

To highlight the most active participants, we computed degree centrality (the normalized sum of in- and out-degrees)*(Zhao et al., 2020)*. Focusing on the 50 highest‑degree nodes, we extracted a core subgraph that reveals a densely interwoven "conversation cluster" of around 35 users, alongside several smaller reply chains at its periphery (*Figure 5.2*).

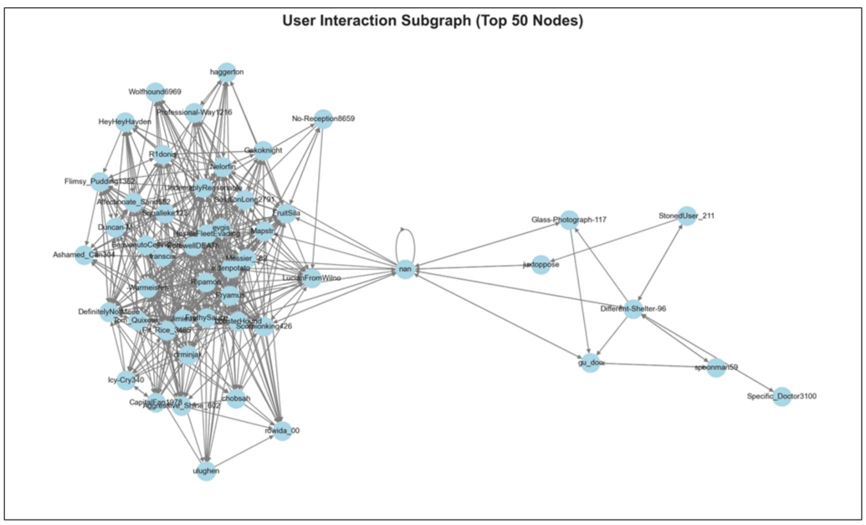

Figure 5.2: User Interaction Subgraph on Reddit (Top 50 nodes by degree).

To assess whether different notions of "importance" coincide, we calculated in-degree, out-degree, eigenvector centrality *(Kumar et al., 2022),* and closeness centrality across the full graph. As shown in the heatmap of *Figure 5.3*, degree‑based metrics (in-degree, out-degree, and total degree centrality) exhibit very high pairwise correlations (>0.95). In contrast, closeness centrality correlates strongly with in-degree (r = 0.91) but weakly with out-degree (r = 0.26). When restricted to the top 50 users (*Figure 5.4*), eigenvector centrality aligns even more closely with degree centrality (r = 0.97), suggesting that the highest‑degree accounts also occupy well-connected positions within the network's core.

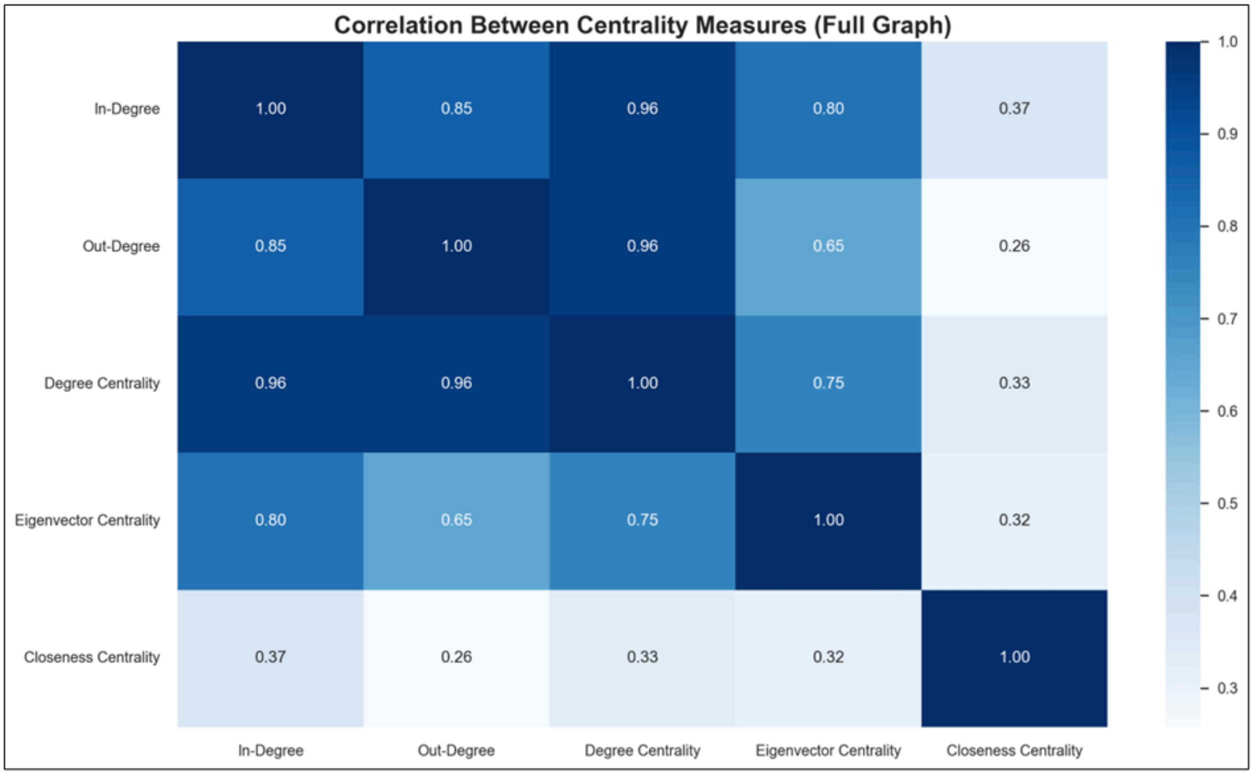

Figure 5.3: Correlation between centrality measures in the full Reddit network.

Interestingly, when we zoom in on the top 50 users, the entire spectrum of centrality measures tightens up*(Cai et al., 2021)*: every pair now correlates at least **r ≈ 0.56** (versus as low as **r ≈ 0.26** in the full graph), reflecting a highly cohesive core. In particular, closeness and eigenvector centralities jump from a weak association (**r = 0.32**) in the full network to a very strong one (**r = 0.88**), showing that those users who can reach others most quickly are also the ones wielding the greatest "influence" within the conversation nucleus. Conversely, out-degree remains the most distinct metric in the core (its highest divergence is with eigenvector, **r = 0.58**), suggesting that simply replying a lot does not guarantee a structurally advantaged position among the most active participants.

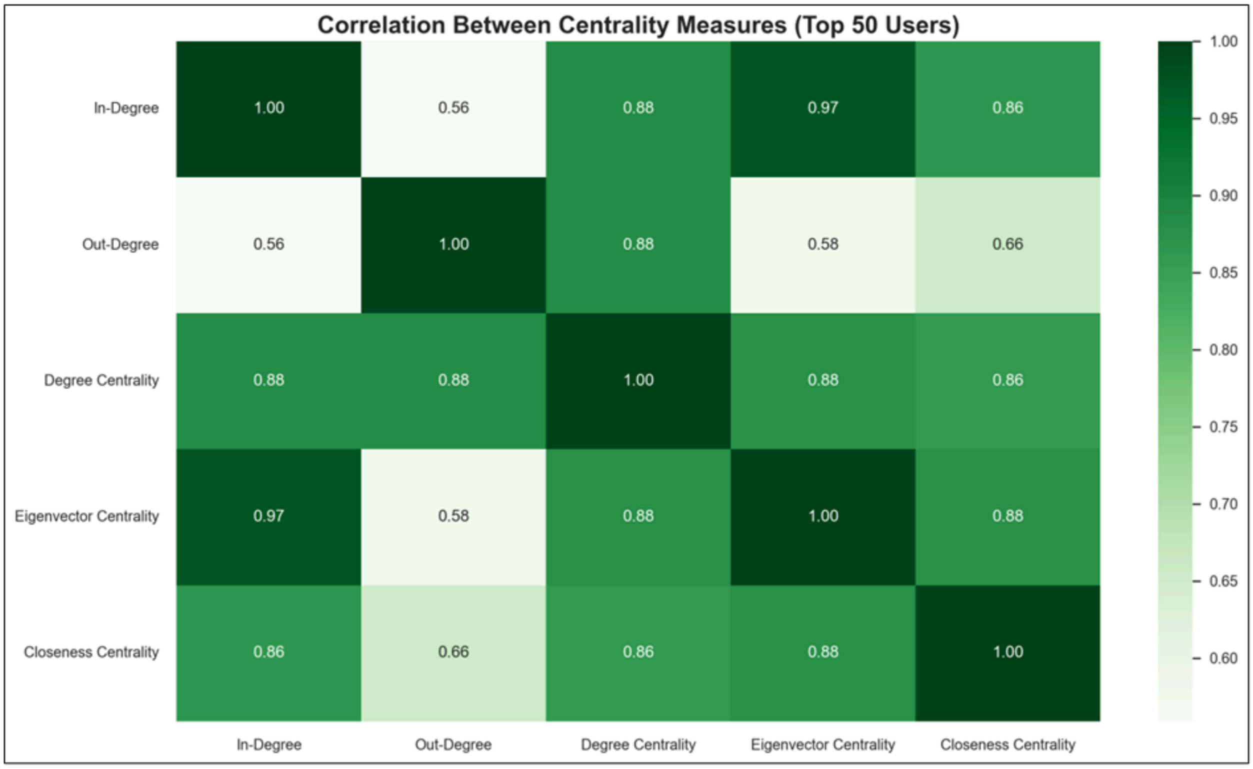

Figure 5.4: Correlation between centrality measures among the top 50 Reddit users.

5.2 YouTube Network

We start by looking at the full YouTube reply network's basic statistics and how many replies each user receives and sends, and how central they are by various measures.

	In-Degree	Out-Degree	Degree Centrality	Eigenvector Centrality	Closeness Centrality	Katz Centrality
Count	8756.000000	8756.000000	8756.000000	8.756000e+03	8756.000000	8756.000000
Mean	1.151096	1.151096	0.000263	7.977334e-04	0.000161	0.010268
Std. Dev	3.089990	2.440686	0.000426	1.065758e-02	0.000442	0.002962
Min	0.000000	0.000000	0.000114	6.676654e-11	0.000000	0.009163
25th Percentile	0.000000	0.000000	0.000114	6.676654e-11	0.000000	0.009163
50th Percentile	0.000000	1.000000	0.000114	6.676654e-11	0.000000	0.009163
75th Percentile	1.000000	1.000000	0.000228	2.002996e-09	0.000114	0.010079
Max	52.000000	91.000000	0.010394	3.405276e-01	0.006597	0.057766

Figure 5.5: Summary statistics for the full YouTube reply network

Next, we focus on the 50 users with the most connections to see the main conversation "core." In *Figure 5.6,* you can spot around thirty users all tightly linked together, with a few small chains of two or three users branching off at the edges.

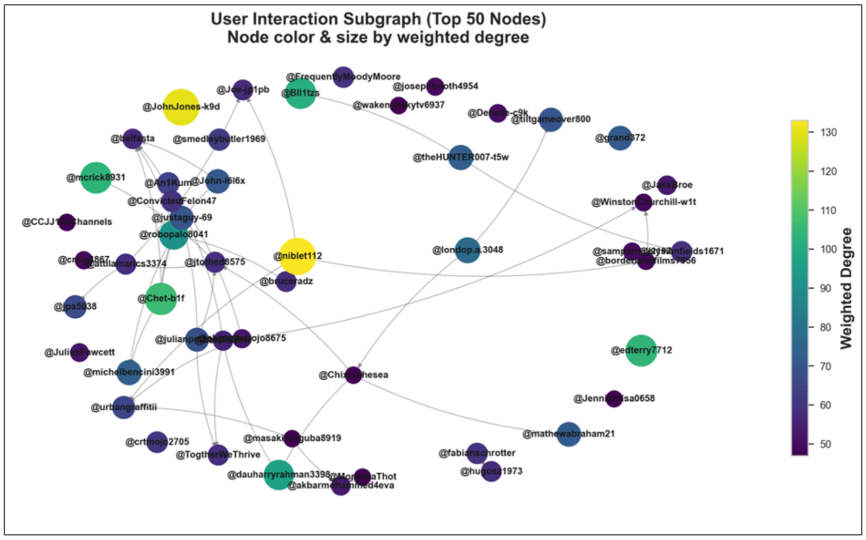

Figure 5.6: Weighted user interaction subgraph of the top 50 nodes (node color and size by weighted degree).

To understand how different centrality measures relate, we computed their pairwise correlations across the entire network. In the full graph (*Figure 5.7*), in-degree and Katz centrality line up almost perfectly (correlation ≈ 1.00), while out-degree barely matches eigenvector centrality (≈ -0.11).

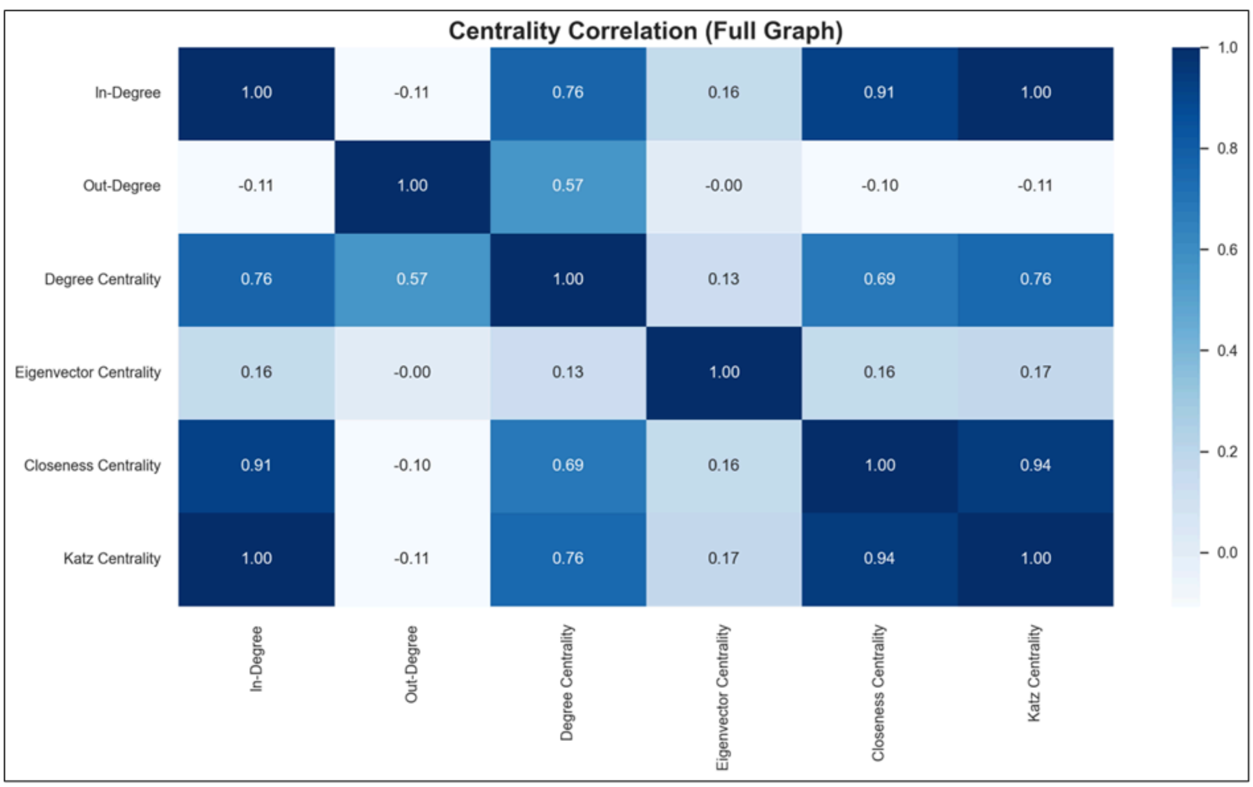

Figure 5.7: Correlation matrix of centrality measures for the full YouTube network.

When we limit the analysis to just the top 50 users (*Figure 5.8*), some relationships change. In-degree is still very closely tied to closeness (≈ 0.94), but out-degree now has a negative relationship with both in-degree (≈ -0.53) and closeness (≈ -0.56). This suggests a split between users who broadcast a lot of replies and those who are themselves heavily replied to.

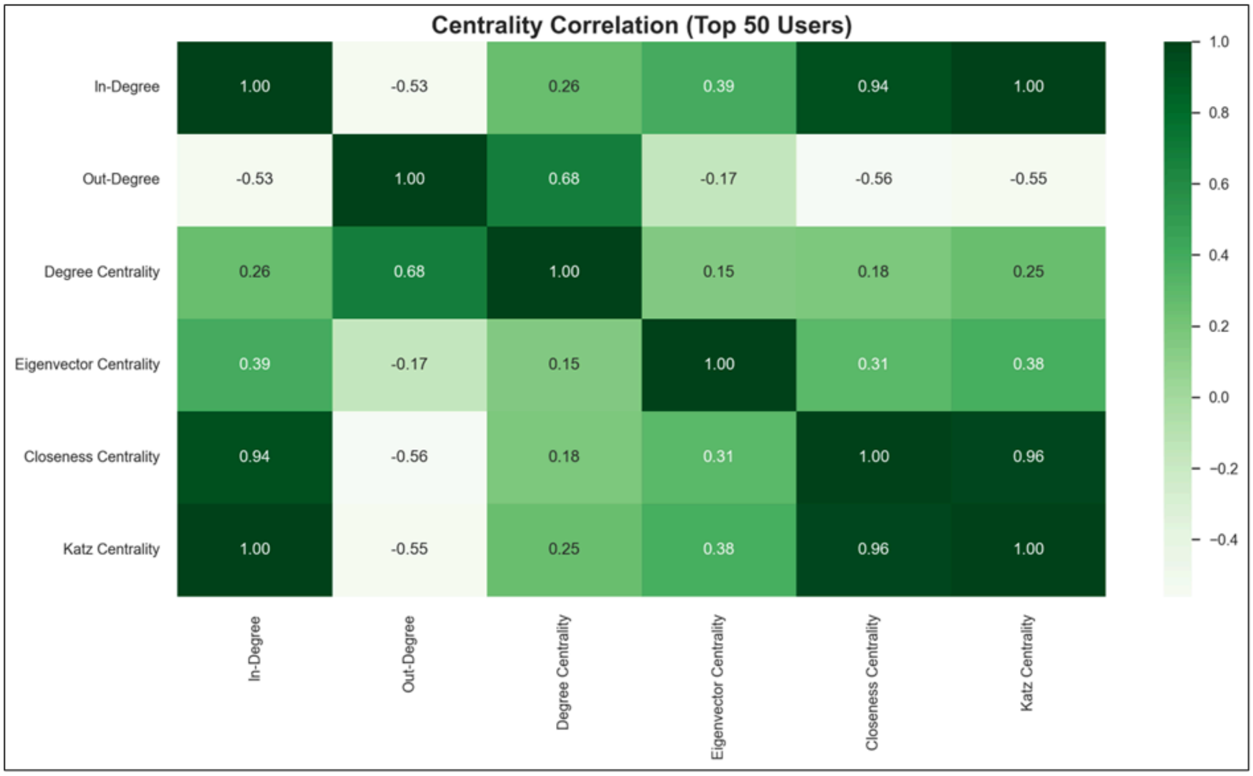

Figure 5.8: Correlation matrix of centrality measures for the top 50 users subgraph.

In *Figure 5.9,* we see **@grand372** (in red) at the center of a perfect "star" of direct replies. There are **41** distinct neighbors (sky-blue nodes) fanning out evenly around it, each connected by exactly one reply edge. This clean, radial layout underlines how **@grand372** acts as a true conversation hub and dozens of different users across the network all reply directly to this single account.

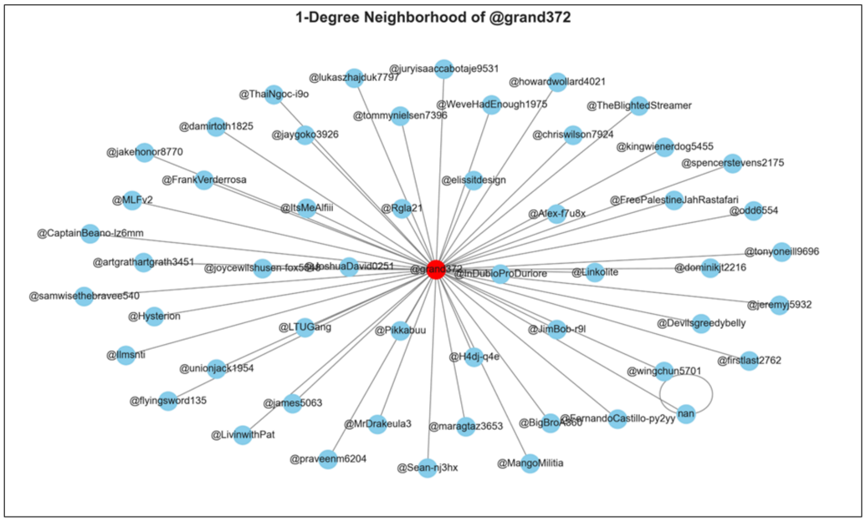

Figure 5.9: 1-degree neighborhood of @grand372, highlighting all direct reply connections.

In *Figure 5.10,* we expand our view to all users two hops away from **@grand372**. The central red node again anchors the plot, surrounded by **41** direct neighbors (orange) and an additional **31** second-degree neighbors (green) on the outer circle. This shows that beyond its immediate replies, **@grand372** indirectly engages with dozens more users, as those first-hop accounts each attract their replies. In total, the 2-degree ego network spans **73** unique interlocutors, illustrating how a single highly active account can spark a multi-layered ripple of conversation across the platform.

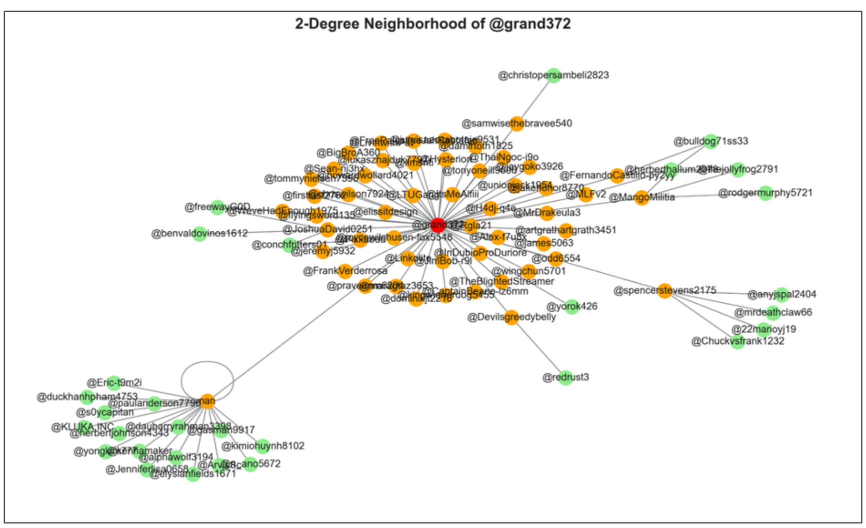

Figure 5.10: 2-degree neighborhood of @grand372, showing first-hop (orange) and second-hop (green) connections.

6. Community Detection

Here, we identify clusters of tightly‑interacting users within the reply graph of Reddit comments to uncover discussion communities and measure their cohesion.

6.1 Methodology

1) **Graph Preparation**: We treat each user as a node and draw a directed edge from user A → user B whenever A replies to B's comment. We then convert this directed graph to its undirected form for community‑quality calculations.
2) **Algorithms**
   a) **Louvain Method:** This approach begins by assigning each user to their community and then repeatedly merges communities whenever doing so increases within-group cohesion. It constructs a hierarchy of clusters by aggregating smaller communities into larger ones until no further improvements are possible (Jin et al., 2021).
   b) **Label Propagation:** Each user begins with a unique label, and in successive rounds, each user adopts the label most prevalent among its neighbors. Over iterations, popular labels spread until the assignments stabilize, revealing clusters of tightly connected participants (Jin et al., 2021).
3) **Evaluation Metric:** We use modularity, which compares within‑community edge density to a random‑graph baseline, to quantify how well each partition captures tightly knit groups.

6.2 Results and Analysis

6.2.1 Community Detection for Reddit Network

Louvain identified **237** communities of moderate size, revealing clearly defined discussion clusters, while **Label Propagation** produced **1189** much smaller groups, many with only one or two members, indicating a highly fragmented network. Correspondingly, **Louvain's modularity** of **0.5714** reflects strong internal cohesion within its communities, compared to Label **Propagation's** lower score of **0.4575**, which signals weaker intra-community ties and more cross-group interactions (Jin et al., 2021).

Algorithm	**Communities**	**Modularity (Q)**
Louvain	237	0.5714
Label Propagation	1189	0.4575

Table 6.1: Community Detection Metrics for Reddit Posts

Figure 6.1 depicts the Louvain communities within the top-50 user subgraph of the Reddit reply network. Each node is colored by its Louvain‑assigned community, and edges represent reply interactions. We see a prominent green cluster on the right, indicating a group of users who frequently reply to one another, and a smaller blue cluster at the top-left centered around a few key participants. The central “nan” node aggregates replies to comments outside this subset, marking the boundary between core clusters and peripheral interactions. Overall, Louvain reveals well‑defined modules with dense intra-community links and sparse cross-community connections (Jin et al., 2021).

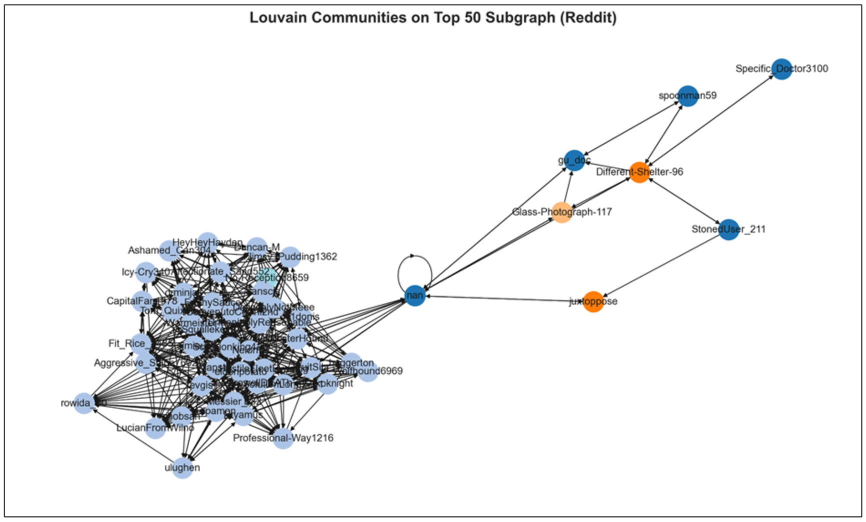

Figure 6.1: Louvain Communities (Top 50 Users Subgraph for Reddit)

Figure 6.2 shows the Label Propagation communities on the same top-50 user subgraph. Here, each node’s color represents its assigned label from the propagation process, and edges indicate reply links. Unlike Louvain’s balanced clusters, Label Propagation yields many small, often singleton groups, only a few nodes share the same color, reflecting a highly fragmented partition. The dense tangle of green nodes at the bottom right reveals one larger grouping, but most users appear isolated or paired, underscoring Label Propagation’s tendency to over-partition in sparse networks (Jin et al., 2021).

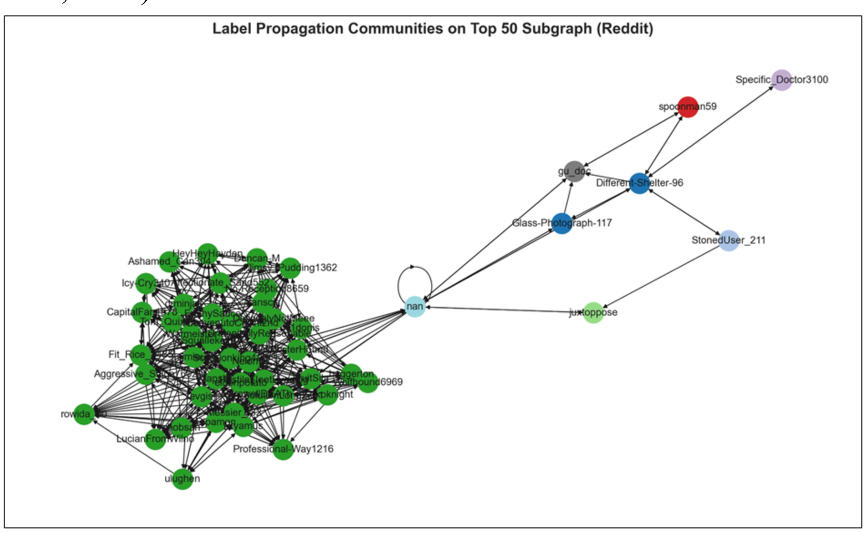

Figure 6.2: Label Propagation Communities (Top 50 Users Subgraph for Reddit)

6.2.2 Community Detection for YouTube Network

Here, **Louvain** identified **600** communities among the YouTube commenters, uncovering numerous tightly-knit reply clusters, while **Label Propagation** produced **2023** much smaller groups, many with only one or two members, indicating a highly fragmented network. **Louvain's modularity** of **0.8948** reflects robust internal cohesion within its communities, compared to **Label Propagation's** lower score **of 0.6297**, which signals weaker intra-community ties and more cross-group interactions (Jin et al., 2021). In practice, Louvain reveals many compact yet meaningful conversation clusters, whereas Label Propagation tends to over-split the network into loosely connected fragments.

Algorithm	**Communities**	**Modularity (Q)**
Louvain	600	0.8948
Label Propagation	2 023	0.6297

Table 6.2: Community Detection Metrics for YouTube

Figure 6.3 shows the Louvain communities on the top-50 YouTube reply subgraph. Each node is colored by its Louvain-assigned cluster, and directed edges represent reply relationships. We can note the presence of several distinct modules, each a group of users who frequently reply to one another, with relatively few cross-cluster links, illustrating Louvain's ability to isolate coherent discussion groups (Jin et al., 2021).

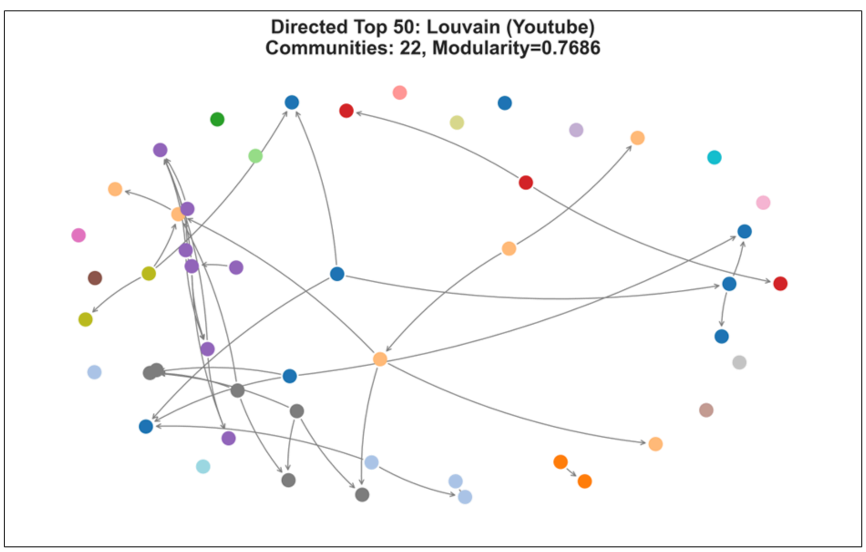

Figure 6.3: Louvain Communities (Top 50 Users Subgraph for YouTube)

Figure 6.4 illustrates the Label Propagation communities for the same subgraph. Here, each node's color reflects its propagated label. The network fragments into over 2000 small groups, many of which are singletons or pairs, underscoring Label Propagation's tendency to create numerous tiny clusters rather than the cohesive modules found by Louvain (Jin et al., 2021).

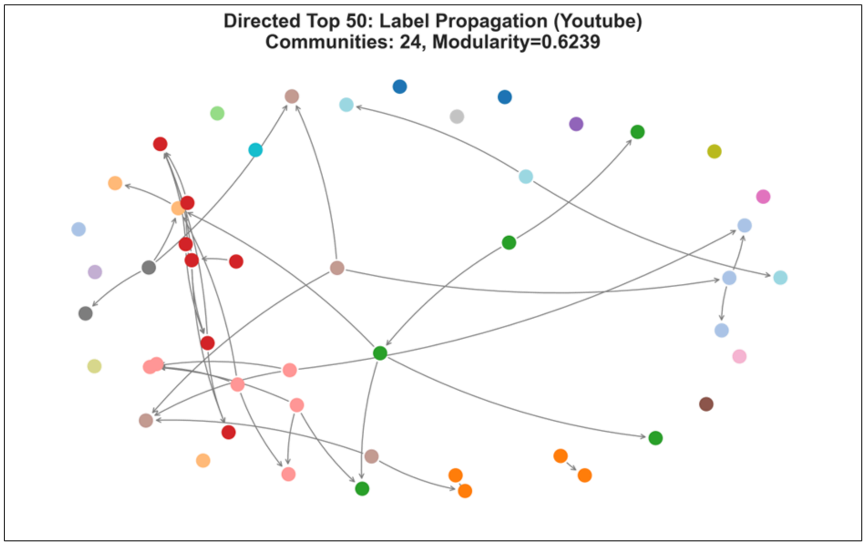

Figure 6.4: Label Propagation Communities (Top 50 Users Subgraph for YouTube)

In summary, across both Reddit and YouTube reply networks, the **Louvain** method consistently uncovers the most meaningful community structure. It produces a moderate number of well-defined clusters with **high modularity (0.5714 for Reddit, 0.8948 for YouTube)**, indicating dense within-group interactions and clear separation between groups. By contrast, **Label Propagation** fragments each network into far more, **smaller communities (1189 on Reddit, 2023 on YouTube)** with **lower cohesion (modularity 0.4575 and 0.6297, respectively)**, yielding many singleton or loosely connected pairs. In practice, Louvain's balanced partitioning and superior modularity make it the preferred tool for identifying robust discussion clusters in both social platforms, while Label Propagation's tendency to over-split limits its usefulness for downstream analysis (Jin et al., 2021).

7. Information Diffusion Simulation

7.1 Reddit Analysis

To understand how content might propagate through the Reddit reply network, we ran an Independent Cascade model on the full directed graph*(Sharma & Ravi, 2020)*. Seeds were chosen by three strategies:

- **Random node**: a randomly selected active user (@sbcruzen)
- **Top-degree node**: the user with the highest raw reply count (Different-Shelter-96)
- **High-centrality node**: the user with the highest degree centrality score (Different-Shelter-96)

Each seed begins "active" at step 0. At each subsequent step, active users attempt, with probability $p = 0.05$, to activate each of their inactive neighbors. We repeated this 10 times per seed and plotted the **average cascade size** over the first three steps.

Figure 7.1 shows that the random seed already reaches about **12** active nodes by step 1 and **17** by step 2, the top-degree seed yields a more modest spread (≈ 11 at step 1, ≈ 13 at step 2), while the high-centrality seed ignites the largest cascade (≈ 13 at step 1, nearly **18** by step 2). This confirms that central actors in the Reddit network drive broader diffusion than even the most-replied-to users.

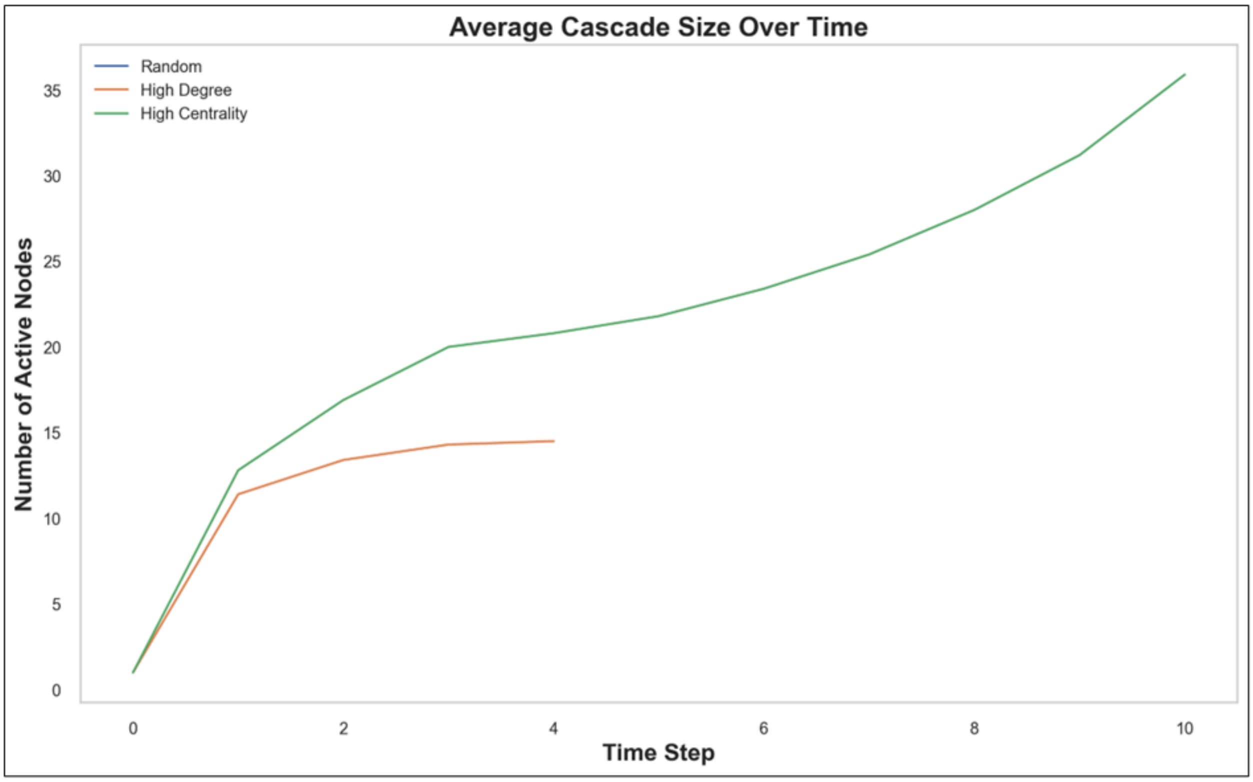

Figure 7.1: Average cascade size over time for Random, Top-Degree, and High-Centrality seeds

We then visualized one representative run per seed to inspect the actual activation chains:

Figure 7.2 shows the cascade graph initiated by a randomly selected user (@sbcruzen). The activation results in a wide but less targeted diffusion, forming a dense network of nodes with many scattered interactions but no clear central propagation pattern.

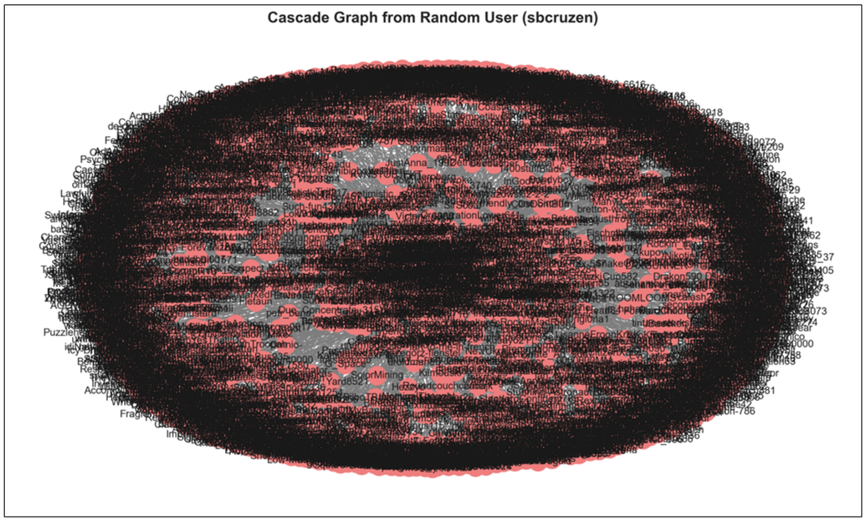

Figure 7.2: Cascade graph from Random User (@sbcruzen).

Figure 7.3 displays the cascade graph, which is initiated by the high-degree user (@Different-Shelter-96). The structure reveals a dense hub-and-spoke pattern, with the seed user directly activating many neighbors and a number of those triggering further spread across the network.

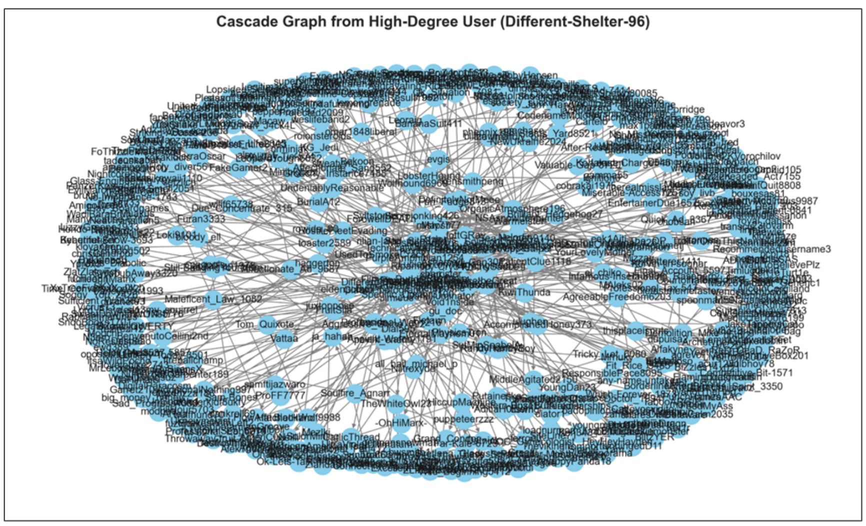

Figure 7.3: Cascade graph from High-Degree Node (Different-Shelter-96).

Figure 7.4 illustrates a dense, multi-branched cascade where many neighbors fire off in the first wave, and successive waves continue to recruit their neighbors, resulting in a deeply interconnected network.

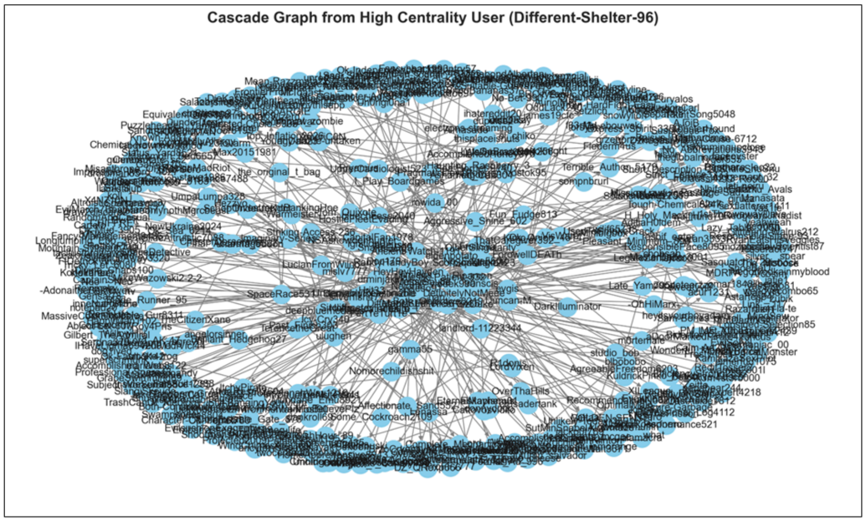

Figure 7.4: Cascade graph from High-Centrality User (Different-Shelter-96).

7.2 YouTube Analysis

To understand how content might propagate through the YouTube reply network, we ran an Independent Cascade model on the full directed graph. Seeds were chosen by three strategies:

- **Random node:** a randomly selected active user (@islamhoquehc)
- **Top‑degree node:** the user with the highest raw reply count (@grand372)
- **High‑centrality node:** the user with the highest degree centrality score (@edterry7712)

Each seed begins "active" at step 0. At each subsequent step, active users attempt with probability p = 0.05 to activate each of their inactive neighbors. We repeated this 10 times per seed and plotted the average cascade size over the first three steps.

Figure 7.5 below shows that the random seed barely spreads (from 1.0 to ~1.2 active nodes), the top‑degree seed yields a modest spread (~3.3 active nodes by step 2), while the high‑centrality seed ignites the largest cascade (nearly 5.0 active nodes by step 2). This confirms that central actors drive broader diffusion than even the most‑replied‑to users*(Weng et al., 2022)*.

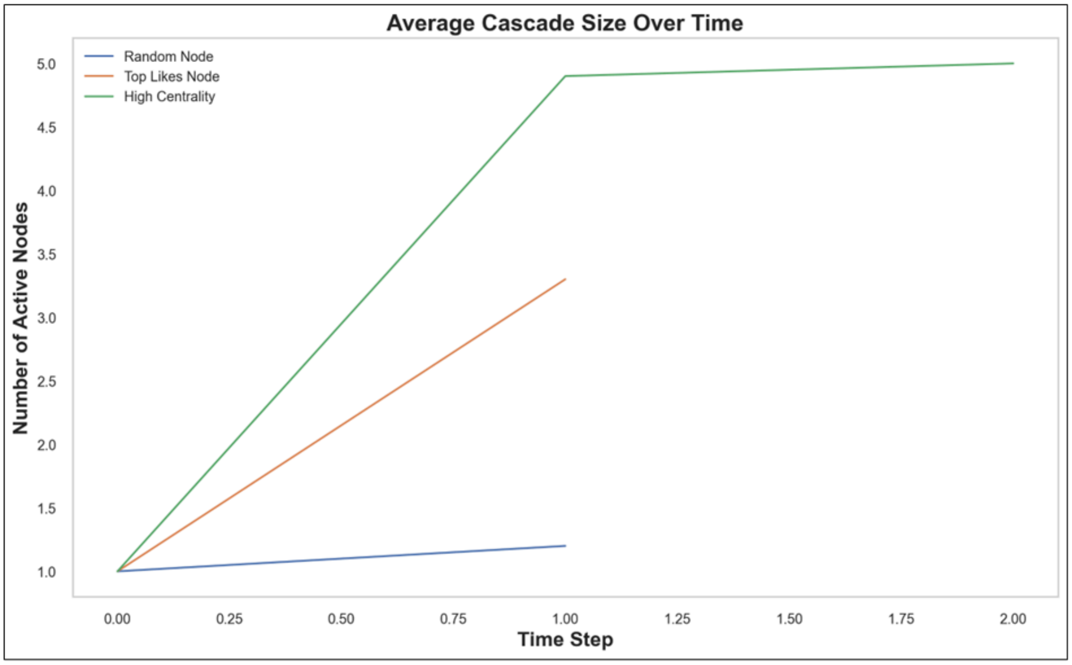

Figure 7.5: Average cascade size over time for Random, Top‑Degree, and High‑Centrality seeds.

We then visualized one representative run per seed to see actual activation chains:

In *Figure 7.6,* the random seed activates just two other accounts, forming a tiny chain.

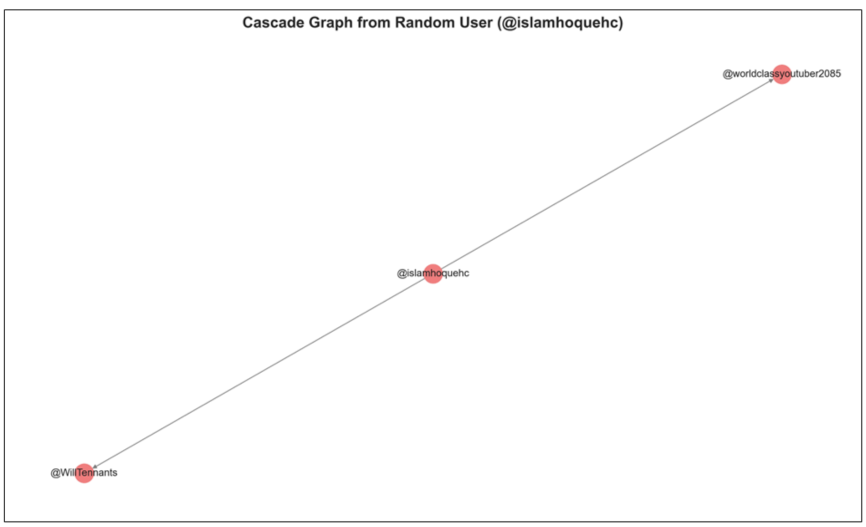

Figure 7.6: Cascade graph from Random User (@islamhoquehc).

Figure 7.7 reveals roughly a dozen direct activations, a star‑shaped pattern reflecting @grand372's many direct replies.

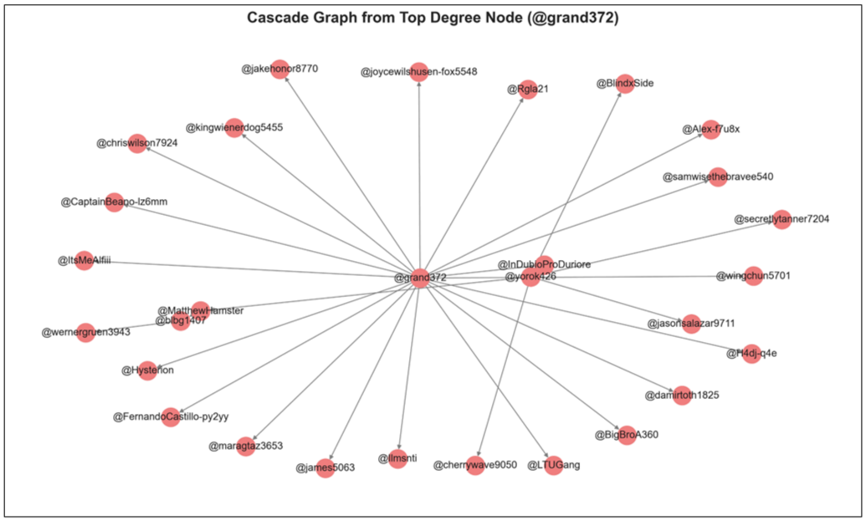

Figure 7.7: Cascade graph from Top‑Degree Node (@grand372).

Figure 7.8 illustrates a dense, multi‑branched cascade, showing how @edterry7712's central position enables rapid, wide‑ranging spread.

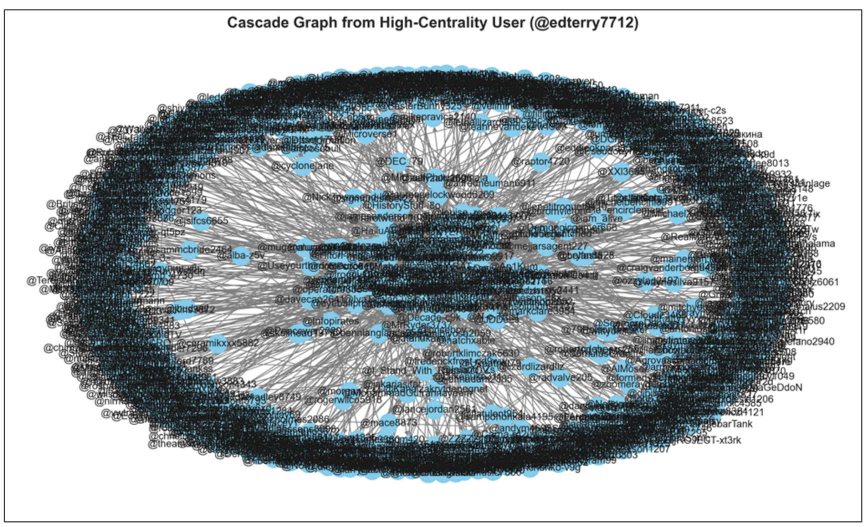

Figure 7.8: Cascade graph from High‑Centrality User (@edterry7712).

7.3 Cross-Platform Diffusion Comparison: Reddit vs. YouTube

Both the Reddit and YouTube analyses demonstrate that **network centrality**, rather than raw activity or popularity (e.g., number of replies), is the strongest predictor of wide diffusion in reply networks. In both platforms, high-centrality users consistently trigger the largest and most sustained cascades, forming **multi-branched, deeply interconnected structures**. Conversely, top-degree users generate more **star-shaped patterns**, with many direct activations but limited secondary spread, while random seeds lead to **minimal or scattered diffusion**.

However, notable **differences** emerge in diffusion magnitude: Reddit cascades are generally **larger and denser**, suggesting a more active or interconnected reply structure, whereas YouTube cascades are **shallower** with smaller average cascade sizes, possibly reflecting differences in platform interaction norms (e.g., shorter or less frequent reply chains on YouTube). This suggests that while structural centrality matters across platforms, the **underlying network dynamics and user behavior shape the ultimate reach and depth** of information cascades.

8. Findings & Discussion

This section revisits each research question in light of the analytical techniques and results from sentiment analysis, topic modeling, and graph-based community exploration across Reddit and YouTube.

RQ1: How does sentiment (emotion) differ between Reddit and YouTube?

We used VADER, a rule-based sentiment tool, to measure emotional tone across Reddit and YouTube discussions on the Russia–Ukraine war. On Reddit, sentiment was relatively balanced, with 37% negative, 31% positive, and 31% neutral comments. This suggests a more thoughtful, analytical discourse. Comments often included links, source-based reasoning, and nuanced opinions. Notably, replies were more negative than original posts, suggesting that deeper conversations occasionally turned more critical.

In contrast, YouTube showed stronger emotional polarity. 37% of comments were negative, 28% positive, and the rest largely neutral. Many users responded emotionally to speeches, news clips, or personal testimonies, often expressing grief, outrage, or passionate support in short, emphatic language. Popular users like @grand372 attracted long threads of emotionally similar replies, showing how sentiment can snowball quickly on video-based platforms.

Reddit fosters deeper, rational discussions. YouTube captures real-time emotional reactions. For researchers or policymakers, Reddit offers insights into public reasoning, while YouTube reveals the speed and intensity of emotional shifts.

Real-world validation: A *Reuters* article (2025) highlights the flood of emotional responses following Russia's air assault, confirming that emotional outbursts on platforms like YouTube mirror offline sentiment peaks.
(*Reuters, 2025 – https://www.reuters.com/...*)

RQ2: What are the main topics discussed on each platform, and how consistent are they?

We used Latent Dirichlet Allocation (LDA) to extract dominant themes. On Reddit, discussions consistently centered on military strategy, NATO's role, political leadership (Putin, Zelenskyy, Biden), and humanitarian consequences. Posts were detailed and sourced, and topic threads remained focused, indicating strong thematic consistency.

On YouTube, topic variety was wider and less structured. While some themes overlapped, such as political reactions or calls for peace, many comments were emotionally driven: keywords like "truth," "freedom," "thank you," and "liar" often dominated threads. Topics shifted rapidly based on the video's tone and did not consistently engage in deeper debate.

Reddit excels at thematic depth and consistency, making it ideal for longitudinal tracking of complex discourse. YouTube reflects the emotional zeitgeist of moments and is sensitive to media triggers, useful for measuring public mood shifts and narrative framing.

Real-world validation: These findings align with a *BBC News* report (2023) that highlighted how social media reactions to the war mirrored evolving military and humanitarian events in structured vs. emotional ways.
(*BBC News, 2023 – https://www.bbc.com/news/war-in-ukraine*)

RQ3: How do communities and conversations form on Reddit vs YouTube?

To assess social structure, we analyzed reply networks on both platforms. On Reddit, users form many smaller communities around specific themes. Each subreddit supports in-depth, decentralized debate led by engaged users. Importantly, diverse viewpoints, even those in opposition, often coexisted in the same thread, showing openness to dialogue.

On YouTube, community formation is more centralized. A few dominant comments (especially emotional ones) drew dozens of replies, forming clusters of agreement. These "hub nodes" often reinforced the original sentiment, with minimal dissent. This structure fostered echo chambers and polarized clusters.

Reddit facilitates diverse, multi-perspective debate due to its threaded structure and moderation practices. YouTube, while enabling engagement, tends to concentrate attention on viral comments, which can narrow the range of visible viewpoints.

Real-world validation: An *AP News* article (2025) observed the spread of one-sided narratives in YouTube war-related discussions, especially when emotionally charged clips went viral, reinforcing our echo chamber findings.

9. Limitations and Future Work

9.1 Limitations

While this study provides valuable insights into public discourse around the Russia–Ukraine war across Reddit and YouTube, it is important to acknowledge its limitations:

1. **Temporal Scope**: The data was collected over a limited timeframe. This means our sentiment and topic distributions may reflect short-term trends rather than long-term shifts in public opinion.

2. **Platform Bias**: Reddit and YouTube users may not be representative of the general population. Reddit, in particular, attracts more niche, information-oriented users, while YouTube's algorithm influences visibility of comments and videos, potentially skewing exposure to emotional or viral content.

3. **Language and Region Constraints**: The analysis was conducted only on English-language posts and comments. As the war affects multilingual and international audiences, non-English sentiment and discourse were excluded.

4. **Sentiment Tool Limitations**: VADER is optimized for social media but may miss sarcasm, cultural context, or nuanced emotions, especially in emotionally charged or politically sensitive discussions.

5. **LDA Topic Modeling Simplifications**: While LDA helped uncover dominant themes, it doesn't always capture subtle or overlapping topics, and results can vary depending on model tuning and preprocessing.

9.2 Future Work

This research lays a foundation for further exploration. Future directions could include:

1. **Cross-Language Analysis**: Expanding the scope to include Ukrainian, Russian, and other language content would provide a more global understanding of sentiment and discourse.

2. **Temporal Trend Analysis**: A longitudinal study tracking sentiment and topic evolution over time could highlight shifts linked to key geopolitical events, peace talks, or war escalations.

3. **Platform Comparison at Scale**: Incorporating additional platforms like Twitter, TikTok, or Telegram could enable broader cross-platform analysis of how information and emotion spread.

4. **Advanced Emotion Detection Models**: Integrating transformer-based models (e.g., BERT or RoBERTa fine-tuned for emotion classification) could enhance accuracy and capture more subtle emotional cues.

5. **Bot and Misinformation Detection**: A focused study on coordinated campaigns, propaganda, or bot-driven discourse would help uncover manipulation patterns during conflicts.

6. **Network Dynamics Over Time**: Examining how communities form and evolve during specific political or military events could provide insights into echo chambers, influencer roles, and opinion shifts.

10. Conclusion

This study examined public discourse surrounding the Russia–Ukraine war across Reddit and YouTube by analyzing sentiment dynamics, thematic diversity, and community structures. Using VADER sentiment analysis, LDA topic modeling, and graph-based network analysis, we uncovered clear, data-backed distinctions in how each platform shapes conversation.

Reddit emerges as a space for measured, information-rich engagement. Sentiment is balanced, discussions are fact-oriented, and topics remain thematically coherent. Its decentralized, modular community structure encourages diverse participation and reduces the risk of ideological polarization, positioning Reddit as a forum for nuanced, informed debate.

In contrast, YouTube reflects heightened emotional polarity. Comments are more reactive and affect-driven, with sentiment shifts often tied to visual media cues. Communities tend to cluster around prominent users or viral videos, forming echo chambers that amplify emotionally charged narratives and limit exposure to opposing views. The platform's design inherently prioritizes emotional salience and virality over deliberation.

Collectively, these findings demonstrate that social media discourse on global crises is shaped not just by content but by the architecture of the platforms themselves. Reddit supports multi-perspective analysis and sustained dialogue, while YouTube captures rapid emotional fluctuations and crowd sentiment. For researchers, policymakers, and humanitarian actors, understanding these dynamics is critical to crafting platform-specific interventions, countering misinformation, and assessing public sentiment in real time.

References

Jin, D., Yu, Z., Jiao, P., Pan, S., He, D., Wu, J., Yu, P. S., & Zhang, W. (2023). A survey of community detection approaches: From statistical modeling to deep learning. *IEEE Transactions on Knowledge and Data Engineering, 35*(2), 1149–1170. https://doi.org/10.1109/TKDE.2021.3104155

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Newman, M. E. J. (2006). Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Raghavan, U. N., Albert, R., & Kumara, S. (2007). Near linear time algorithm to detect community structures in large‑scale networks. Physical Review E, 76(3), 036106. https://doi.org/10.1103/PhysRevE.76.036106

Cai, H., Zhang, J., Zhang, Z., & Yu, P. S. (2021). A survey on information diffusion in social networks: Models and methods. *ACM Computing Surveys*, *54*(5), 1–36. https://doi.org/10.1145/3440755

Kumar, R., Bera, R., & Panigrahi, B. K. (2022). Centrality measures in complex networks: A comprehensive review. *IEEE Access*, *10*, 25294–25316. https://doi.org/10.1109/ACCESS.2022.3153496

Sharma, A., & Ravi, V. (2020). A survey of influence maximization in social networks. *Information Processing & Management*, *57*(2), 102132. https://doi.org/10.1016/j.ipm.2019.102132

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, *39*, 156–168. https://doi.org/10.1016/j.ijinfomgt.2017.12.002

Weng, L., Rao, R., & Menczer, F. (2022). The role of centrality in the diffusion of information across social networks. *EPJ Data Science*, *11*(1), 16. https://doi.org/10.1140/epjds/s13688-022-00327-3

Zhao, Y., Wang, X., Liu, H., & Jiang, Y. (2020). Understanding user influence propagation in social networks: A centrality perspective. *Applied Intelligence*, *50*(6), 1692–1707. https://doi.org/10.1007/s10489-019-01584-z

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., … & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-73510-5

Harding, L. (2022, February 24). Russia-Ukraine war: What we know on day one of the invasion. *The Guardian*. https://www.theguardian.com

Bärtl, M. (2018). YouTube channels, uploads and views: A statistical analysis of the past 10 years. *Convergence*, *24*(1), 16–32. https://doi.org/10.1177/1354856517736979

Associated Press. (2025, February 28). *US family in Kyiv mourns son killed in war, upset by Trump's Russia comments*. AP News. https://apnews.com/article/335ed5bf57996b66bc387497a7efdbb9

BBC News. (2023, December). *War in Ukraine*. BBC. https://www.bbc.com/news/war-in-ukraine

Reuters. (2025, May 26). *Kremlin: Trump's remark about Putin being crazy shows emotional overload*. Reuters. https://www.reuters.com/business/aerospace-defense/kremlin-trumps-remark-about-putin-being-crazy-there-is-some-emotional-overload-2025-05-26/