Analyzing the Dominant Themes and Sentiments in discourse around "Redpill" content on YouTube

Lexiang (Daniel) Hu

dhu578@akane.waseda.jp
University of Toronto
Student No. 97241065-2

Data And Social Media Analysis (Waseda University)
Prof. Emily Öhman
January, 2025

1 Background

The "manosphere" refers to a network of online communities that explore and promote diverse interpretations of masculinity, often in reaction to feminist ideologies and modern gender roles. This loosely connected ecosystem includes subgroups such as men's rights activists (MRAs), incels (involuntary celibates), pick-up artists (PUAs), and "alpha male" influencers, each with their own focus but united by shared grievances regarding gender dynamics and a belief in male disenfranchisement (Ging, 2019; Farrell et al., 2021).

At its core, the manosphere is characterized by a narrative of victimhood, framing masculinity as being under attack in contemporary society. Many of its participants believe that feminism has created a cultural environment that devalues men and undermines traditional gender roles. Common concepts within the manosphere include hypergamy (the belief that women are naturally predisposed to seek partners of higher status), sexual market value (SMV) (a metric for evaluating desirability in the dating market), and redpill philosophy (Ging, 2019; Ging & Siapera 2024).

A unifying concept across much of the manosphere is the "redpill," a metaphor derived from the 1999 sci-fi film The Matrix. In the film, the protagonist is given a choice between a red pill, representing awakening to a harsh reality, and a blue pill, symbolizing remaining in blissful ignorance. Within the manosphere, the redpill signifies a perceived awakening to "truths" about societal structures believed to disadvantage men. Advocates of redpill ideology often view this realization as a tool for empowerment and self-improvement, emphasizing hyper-masculinity, dominance in relationships, and critiques of feminism (Ging & Siapera 2024).

The manosphere's ideologies have been widely spread across digital platforms, with Reddit and TikTok receiving significant scholarly attention. Subreddits like r/TheRedPill and r/MGTOW served as hubs for text-based discussions that reinforce in-group identities and critique modern gender norms (Farrell et al., 2021; Labbaf, 2020). On TikTok, short-form videos target younger audiences, using engaging formats to propagate themes of self-improvement and critiques of feminism (Swallowing and Spitting Out the Red Pill, 2023). However, YouTube remains a relatively underexplored platform for analyzing redpill content, despite being a central space for manosphere influencers to reach broad audiences.

YouTube offers a distinct environment where long-form video content is combined with interactive features such as comment sections and live streams. Influencers like Andrew Tate exemplify how manosphere figures leverage the platform to craft compelling narratives and foster dynamic engagement. Research such as Hosseinmardi et al. (2021) has challenged the perception that YouTube's algorithm alone drives users toward radical content, arguing that individual preferences play a larger role. Nonetheless, YouTube's infrastructure amplifies manosphere ideologies, making it a critical medium for understanding redpill discourse. This study focuses on YouTube as a medium for redpill narratives, aiming to fill a gap in the existing literature by analyzing how its unique features shape the spread and reception of manosphere ideologies. Overall, if the discourse in "Red Pill" and related communities reflects specific ideological themes, then these themes will emerge as distinct topics through topic modeling and exhibit varying sentiment patterns. Specifically, themes related to self-improvement are expected to show predominantly positive sentiment, while critiques of feminism and gender dynamics will show more negative sentiment.

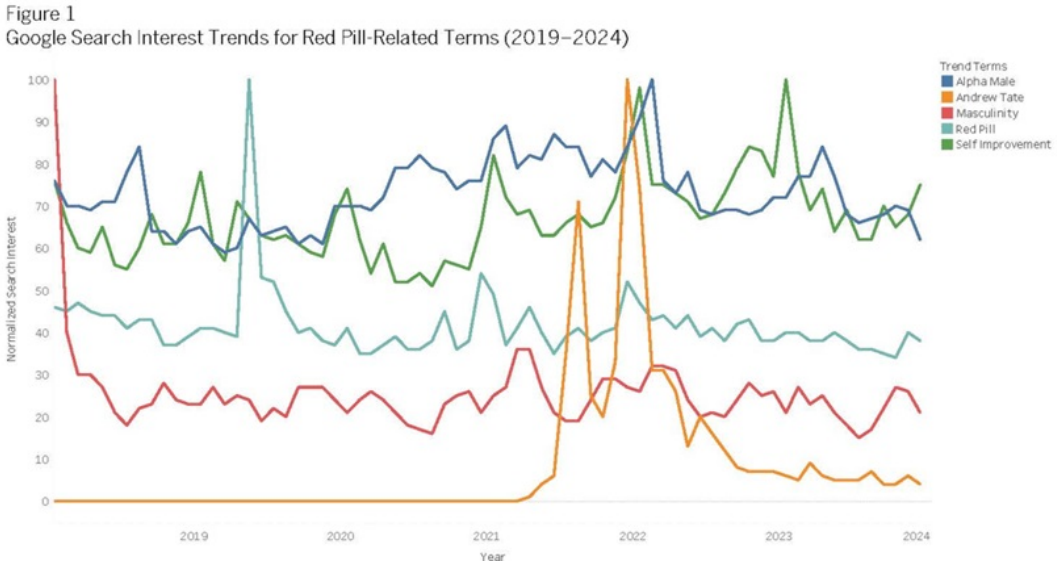


Figure 1: Google search interest trends for terms associated with redpill discourse (Alpha Male, Andrew Tate, Masculinity, Red Pill, and Self Improvement) from 2019 to 2024. Data scraped from Google Trends using the Pytrends library, illustrates relative popularity over time. The plot was created in Tableau Desktop 2024.1.

2 Data

Figure 1 illustrates the normalized Google search interest for terms that have been associated with manosphere discourse in previous studies—"Alpha Male", "Andrew Tate", "Masculinity", "Red Pill", and "Self Improvement"—in the five-year period between January 1, 2019, and December 31, 2024. Google Trends provides a measure of search interest by normalizing data on a scale of 0 to 100. A value of 100 represents the peak popularity of a term in the specified time frame and location, while 0 indicates minimal or no relative interest. It is important to note that these values are relative and do not reflect absolute search counts. The normalization process ensures comparability across terms and time periods, making it an effective tool for tracking cultural and social trends (Google, n.d.). Collectively, the terms exhibit consistently high relative search interest, underscoring their prominence in contemporary discussions about masculinity and manosphere ideologies. Notable patterns emerge from the data. "Alpha Male" and "Self Improvement" exhibit consistent search interest, reflecting their enduring relevance in discussions surrounding masculinity and personal development. In contrast, "Andrew Tate" shows a dramatic spike in 2022, coinciding with his rise to prominence and subsequent controversies, before sharply declining. Meanwhile, "Red Pill" maintains moderate but steady interest over time, suggesting its establishment in manosphere ideology. Interestingly, "Masculinity" shows a gradual decline, indicating either reduced interest in general discourse about gender roles or a shift toward more specific terminology. From previous studies, Redpill and other manosphere content primarily clustered around masculinity, women and self-improvement, thus there were the three search results that were utilized. A new youtube account was created, and a video was randomly selected from inputting one of the three queries: "Redpill #Masculinity, Redpill #Dating, Redpill #Self-improvement".

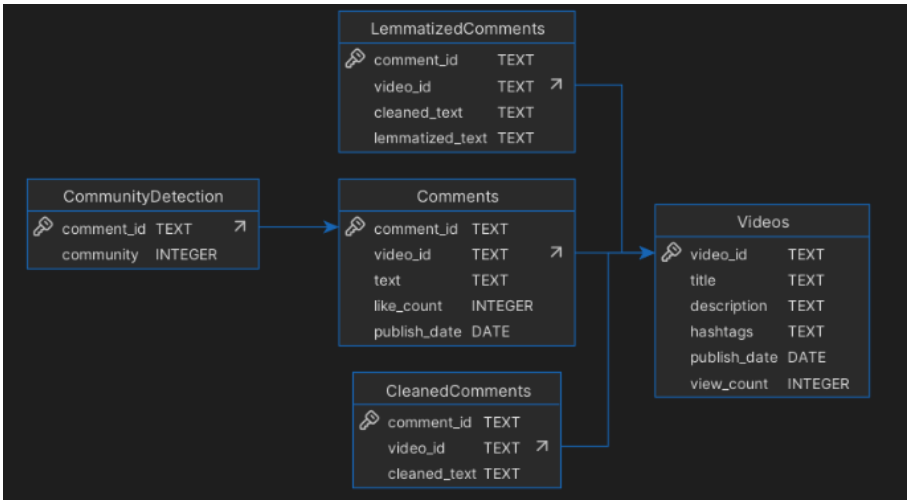


Figure 2: Hierarchical grouping of the database schema. Visualized using DbVisualizer 24.3.3

Video Title	Description
Reject Modernity, Embrace True Masculinity	Published on June 1, 2023. Duration: 3m 20s. The video begins with short-form clips (TikTok, reels) critiquing masculinity before transitioning to visuals of bodybuilders, soldiers, and athletes.
Women STILL ALWAYS Tell On Themselves in 2025—Just Let Them Talk Long Enough...	Uploaded on January 14, 2025, by *Manosphere Highlights*, with 45,000+ views. Features commentary over a *Dr. Daf Show* video discussing whether men should pay for first dates and why women don't initiate contact.
When You Learn The REAL Red Pill — Women, Dating, Marriage And Self Improvement Become EASIER	Uploaded on April 7, 2023, by *Legion of Men*, accumulating 59,000+ views. Commentary on *Red Pill: What to Expect on the Path to Understanding*, discussing redpill ideology, self-improvement, and traditional gender roles.

Table 1: Summary of YouTube Videos Selected for Analysis

After selecting these videos, the YouTube API was employed to extract further details such as the video descriptions, view and like counts, and the top 100 comments for each video. The comments were then subjected to cleaning and lemmatization before being saved into the SQLite database for subsequent analysis.

3 Methods

Following data collection, the data underwent tokenization, and lemmatization. A Bidirectional Encoder Representations from Transformers (BERT)-based model was employed for these tasks due to its ability to process text in context, distinguishing it from traditional lemmatization approaches such as spaCy's rule-based models. Unlike conventional methods that tokenize and lemmatize words in isolation, BERT considers the surrounding linguistic structure, thereby capturing the entire context of a comment (Devlin et al., 2019). This distinction is particularly relevant in analyzing manosphere discourse, where terms may have context-dependent meanings. Additionally, BERT is optimized for handling informal and non-standard language, which is prevalent in YouTube comments. Internet slang, abbreviations, and emerging terminology often challenge rule-based models, whereas BERT's transformer-based embeddings enable a more accurate understanding of meaning and intent. BERT enhances feature extraction, improving the performance of subsequent analyses, including sentiment classification and topic modeling.

To identify clusters of semantically similar comments, we applied the Louvain method, a widely used community detection algorithm based on modularity optimization (Blondel et al., 2008). First, the lemmatized comments were vectorized using Term Frequency-Inverse Document Frequency (TF-IDF) to represent textual content numerically. A cosine similarity threshold of 0.2 was selected to balance semantic coherence and network connectivity, ensuring that only moderately sim-

ilar comments were linked while filtering out weak associations. Prior research suggests that similarity thresholds in the range of 0.15–0.25 effectively preserve meaningful discussion clusters while preventing excessive fragmentation or overly large, incoherent communities (Coscia et al., 2011; Manning et al., 2008). Lower thresholds led to broad, loosely defined communities, whereas higher thresholds isolated many nodes, reducing interpretability. Thus, 0.2 provided an optimal balance, allowing for the formation of distinct yet meaningful communities.

The Louvain method grouped these nodes into communities, maximizing modularity to reveal coherent discussion clusters. The network visualization of detected communities is presented in Figure 3, where nodes represent comments, and colors indicate different communities. The relative sizes of these communities are shown in Figure 4b, which presents the distribution of comment counts among the detected groups. To further explore the linguistic themes within the largest communities, Figure 5 visualizes the most frequently used words within the five largest communities.

To identify overarching themes within the discussions, we employed Latent Dirichlet Allocation (LDA), a probabilistic generative model that assigns topics based on word co-occurrence probabilities (Blei et al., 2003). The preprocessed text was transformed into a document-term matrix, and LDA assigned probability distributions of words to topics. The optimal number of topics was determined using the coherence score which evaluates topic quality by measuring semantic similarity within topics. Sentiment analysis was performed using the Valence Aware Dictionary and sEntiment Reasoner (VADER), a lexicon-based sentiment model optimized for social media text (Hutto & Gilbert, 2014). Each comment received a compound sentiment score, which was categorized as positive, neutral, or negative. Figure 6a displays the sentiment distribution across topics, highlighting whether manosphere discussions tend to be optimistic, neutral, or pessimistic. Additionally, sentiment scores were aggregated at both the community level (Figure 4b) and the video level (Figure 4c) to identify differences in emotional tone between user groups and content.

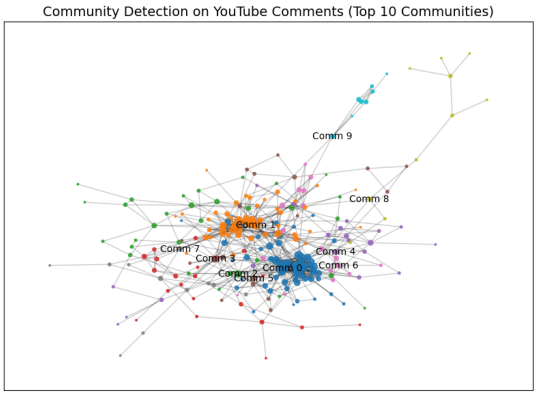
To explore the relationship between topics and detected communities, topic assignments were mapped onto community structures. Figure 4b presents the topic distribution across the top 10 communities, highlighting which thematic discussions dominate each group. Furthermore, Figure 6a shows topic alignment with specific videos, allowing an examination of how content influences discourse structures within the manosphere.

4 Results

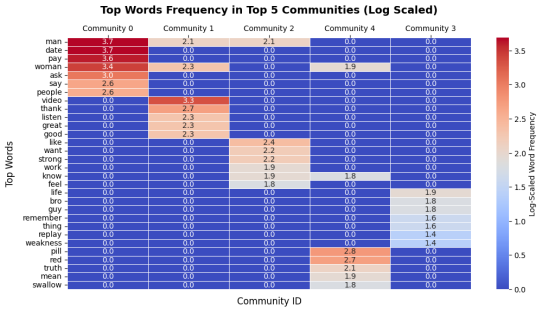
The community detection algorithm identified a total of 61 distinct communities. The modularity score confirmed strong clustering, indicating well-defined groups of users engaging in specific discussions. Each community exhibited specific clusters of words(Figure 3b). As expected, words related to the discourse of the manosphere (Man, want, Woman, date, etc.) are displayed in Figure 3c.

LDA modeling identified eight distinct topics, which were manually labeled based on the most prominent words. The topics included themes such as self-improvement, gender dynamics, and relationship advice. The topic-word distribution is presented in Figure 5, where each column represents a topic, and the rows display the main words associated with each theme. This visualization provides insight into the recurring discussions within manosphere content.

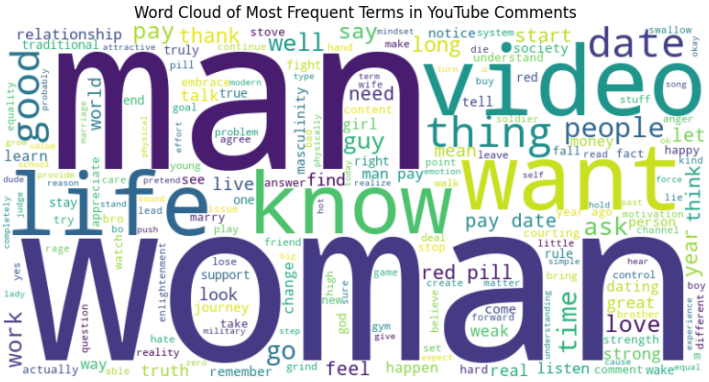
The sentiment analysis results suggest that manosphere discussions surprisingly exhibit a predominantly positive sentiment(Figure 6a). The radar plots (Figures



(a) Community Grouping of YouTube Comments, showing top 10 of 61 Communities.



(b) Heatmap of the Word Frequency in the Top 5 Communities.



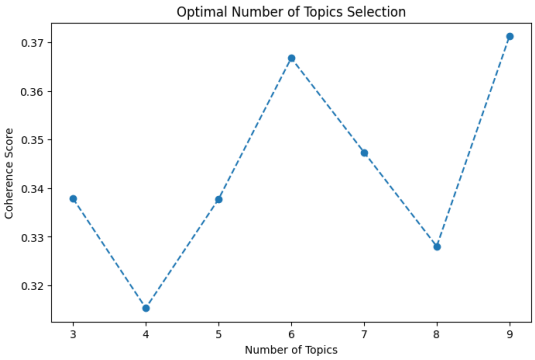
(c) Word Cloud of Most Frequent Words in YouTube Comments.

Figure 3: Community detection and lexical analysis of YouTube comments. The first subfigure (a) represents the detected communities, while (b) and (c) illustrate the most frequent words in discussions.

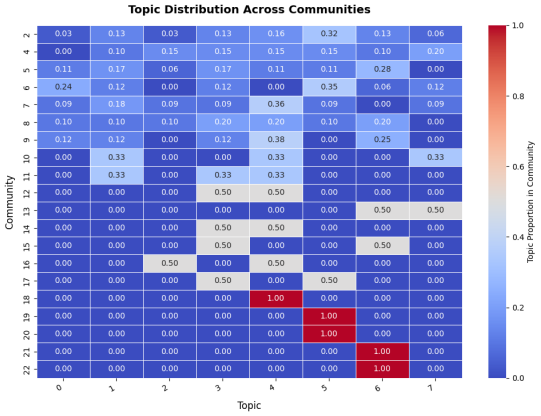
6b, 6c) further emphasize that most topics contain a higher proportion of positive comments, with a minority showing strong negative polarity. When sentiment was examined at both the community level and the video level, it appears there is more varied sentiment in different communities, while the comments of all three videos appear to be more similar to one another. Overall, there is a stronger neutral/positive sentiment among the comments.

5 Discussion

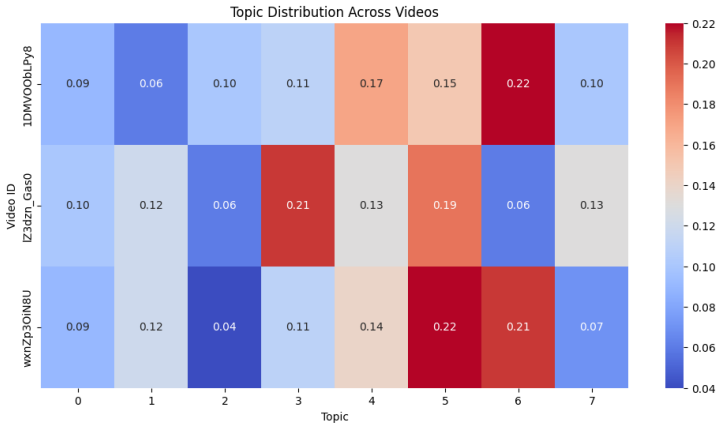
The findings indicate that manosphere discussions on YouTube could be divided into distinct communities with varying thematic focus and sentiment(Figure 3). Sentiment analysis suggest that the majority of comments skew neutral or positive, regardless of the type of video(women, masculinity, self-improvement). This finding contrasts with prior research, such as Farrell et al. (2021), who found that manosphere discourse on Reddit contained a high prevalence of hostility, resentment, and gendered grievance narratives. Similarly, Hosseinmardi et al. (2021) observed that radical content consumption on YouTube was associated with an increase in hostile and reactionary comments, particularly against women and feminism. However, in the present study, many comments appear supportive of the videos themselves, which may have influenced the



(a) Optimal Number of Topics Selected via Coherence Score.



(b) Topic Distribution Across Communities.



(c) Topic Distribution Across Videos.

Figure 4: Topic modeling results from LDA analysis. Subfigure (a) shows topic selection, while (b) and (c) display topic distributions across communities and videos, respectively.

sentiment analysis. For example, the comment "You got my support... You got God's support... Keep spreading this mindset" reflects positive reinforcement of manosphere narratives rather than overt hostility. Another comment, "I wish I'd heard of this type of positivity years ago", further underscores this pattern of affirmation rather than adversarial discourse.

These results may reflect platform-specific differences in manosphere discourse. On Reddit and other text-based forums, discussions are structured as threaded debates, allowing for deeper ideological engagement (Farrell et al., 2021). In contrast, YouTube comments are shorter, more reactive, and less conducive to long-form arguments, potentially explaining the prevalence of neutral-to-positive sentiment in this dataset. Furthermore, as Ging (2019) noted, some manosphere factions, particularly those focusing on self-improvement and dating strategies, present their ideology under the guise of rationality and positivity, even while promoting problematic views on gender roles.

Another consideration is that not all manosphere content is overtly extreme. Labbaf (2019) found that incel discourse contained the highest levels of negative sentiment, whereas communities focused on self-improvement or men's rights advocacy often framed discussions in neutral or positive terms. As this study did not collect data from a more extreme platform and material(e.g incel-related discussions), this

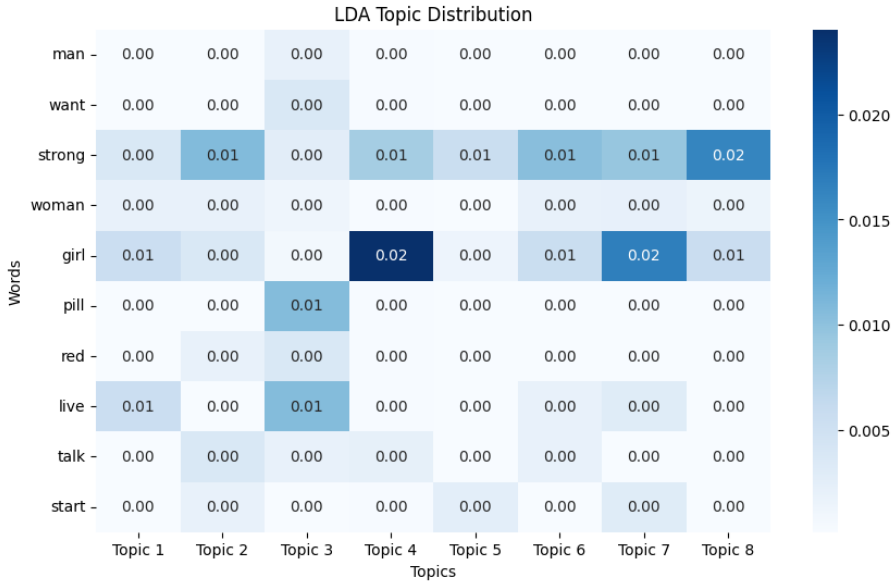


Figure 5: Topic Distribution as Determined by LDA

could explain why extreme negativity was less prevalent.

The connection between topics, communities, and video content suggests that content creators play a role in shaping audience sentiment and engagement. The distribution of topics in videos (Figure 4c) shows that certain themes attract different sentiment profiles, indicating that the discourse of the manosphere is not uniform, but rather dependent on the type of content being produced. This aligns with previous research indicating that algorithmic recommendations can reinforce ideological echo chambers (Hosseinmardi et al., 2021).

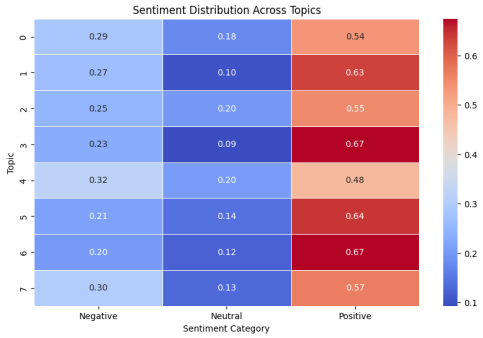
Overall, this study underscores the diversity of manosphere discourse on YouTube, illustrating the interplay between community structures, thematic discussions, and emotional tone. Future research could explore the sentiments and the role of YouTube's recommendation algorithms in amplifying certain types of discourse. Additionally, applying BERT-based sentiment classification rather than lexicon-based approaches (e.g., VADER) may provide more nuanced insights into the subtleties of manosphere rhetoric.

6 Limitations

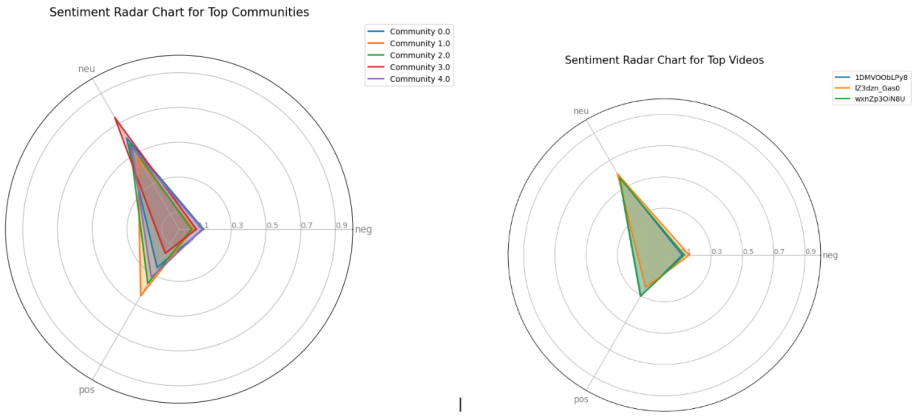
This study relied on VADER for sentiment analysis; while effective for short-form social media text, it does not capture contextual sentiment nuances (Hutto & Gilbert, 2014). Since VADER is lexicon-based, it struggles with sarcasm, irony, and manosphere-specific terminology that may shift in meaning depending on context. Future studies could integrate BERT-based sentiment classification, which would allow for deep contextual analysis rather than simple lexicon matching.

Additionally, BERT was used for lemmatization and tokenization, allowing for context-aware text processing. However, this pre-processing step does not account for semantic shifts in sentiment classification, as sentiment scores were still computed using VADER's predefined dictionary rather than a transformer-based sentiment model. Implementing fine-tuned BERT sentiment classification could improve sentiment detection, particularly for manosphere discourse where polarizing language and coded terminology are common.

Another limitation arises from the aggregation of sentiment at the community and video levels. While individual com-



(a) Sentiment Distribution Across Topics.



(b) Radar Chart of Sentiment Analysis Across Communities.

(c) Radar Chart of Sentiment Analysis Across Videos.

Figure 6: Sentiment analysis results across different dimensions. Subfigure (a) shows sentiment by topic, while (b) and (c) display sentiment patterns at the community and video levels, respectively.

ments were classified for sentiment, the process of averaging scores at a macro level may obscure ideological polarization. Communities with highly polarized discussions (mix of extremely positive and extremely negative comments) may incorrectly appear neutral when averaged, masking significant ideological divides. Future work should explore more throughout sentient clustering within communities to identify these polarized subgroups.

This study also focuses exclusively on long-form YouTube content, excluding short-form platforms such as TikTok and YouTube Shorts, which are increasingly influential in shaping manosphere discourse. A broader cross-platform analysis, incorporating other social media sites (such as Reddit or Twitter), would provide a more comprehensive understanding of manosphere engagement across digital spaces. Finally, the selected YouTube videos represent only a small sample of manosphere content, limiting generalizability. While the methodology allows for structured analysis, a larger dataset spanning multiple manosphere creators would improve robustness and reduce selection bias in identifying dominant manosphere narratives.

7 Conclusion

I will now need to clear my google and youtube history.

References

[1] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, 3*, 993–1022.

[2] 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.

[3] Coscia, M., Giannotti, F., & Pedreschi, D. (2011). A classification for community discovery methods in complex networks. *Social Networks, 34*(1), 1–11. https://doi.org/10.1016/j.socnet.2010.09.002

[4] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1*, 4171–4186.

[5] Farrell, T., Fernandez, M., & Novotny, J. (2021). Exploring the manosphere on Reddit: Understanding misogyny in the digital age. *Journal of Community & Applied Social Psychology, 31*(2), 156–170.

[6] Ging, D. (2019). Alphas, betas, and incels: Theorizing the masculinities of the manosphere. *Men and Masculinities, 22*(4), 638–657.

[7] Ging, D., & Siapera, E. (2024). Red-pilled: Tracing the political economy, affective infrastructures and ideological trajectories of the digital manosphere. *Men and Masculinities*. https://doi.org/10.1177/1097184X241286800

[8] Google. (n.d.). How Trends Data Is Adjusted. *Retrieved from Google Trends Documentation*.

[9] Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Representation learning on graphs: Methods and applications. *IEEE Data Engineering Bulletin, 40*(3), 52–74.

[10] Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D. M., & Watts, D. J. (2021). Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences, 118*(32), e2101967118. https://doi.org/10.1073/pnas.2101967118

[11] Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*.

[12] Labbaf, F. (2019). United by rage, self-loathing, and male supremacy: The rise of the incel community. *Invoke: Journal of Undergraduate Research, 5*. Retrieved from https://journals.library.ualberta.ca/invoke/index.php/invoke/article/view/48979

[13] Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. *Cambridge University Press*.

[14] Swallowing and Spitting Out the Red Pill: Young Men, Vulnerability, and the Manosphere (2023). *Feminist Media Studies*.

[15] YouTube. (2023). Reject Modernity, Embrace True Masculinity [Video]. Retrieved from https://www.youtube.com/watch?v=wxnZp3OiN8U.

[16] YouTube. (2025). Women STILL ALWAYS Tell On Themselves in 2025—Just

Let Them Talk Long Enough... [Video]. Retrieved from https://www.youtube.com/watch?v=lZ3dzn_Gas0&t=522s.

[17] YouTube. (2023). When You Learn The REAL Red Pill | Women, Dating, Marriage And Self Improvement Become EASIER [Video]. Retrieved from https://www.youtube.com/watch?v=1DMVOObLPy8.