# Exploring Echo Chamber Effect in the r/depression Subreddit

Sifat Redwan Wahid
University of Minnesota Duluth
Duluth, MN, USA

## Abstract

Echo chamber refers to an environment where people are exposed to ideas or beliefs that reflect their own, filtering out diverse perspectives. Echo chambers can form on real-life or in online spaces. It can form based on user's preferences or the platform's recommendations as the algorithms prioritize to maximize user interactions resulting in recommending similar contents to the users which reinforces similar ideas over and over again. In the long run this recursion of recommendations can lead to radicalization, extremism or polarization on based on controversial issues. Previous studies on echo chambers mainly focused on controversial topics like COVID-19, Russia-Ukraine war, abortion etc. Hardly any study exists that focuses on mental health communities. In this project I have tried to explore the echo chamber effect in the r/depression subreddit, an online community for peer support for people with depressive disorder. I have tried to explore if people in this subreddit get diverse viewpoints from the community or are they merely echo chambers.

**ACM Reference Format:**
Sifat Redwan Wahid. 2025. Exploring Echo Chamber Effect in the r/depression Subreddit. In . ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 Introduction

Mental health communities on Reddit, such as r/depression, provide platforms where users can share personal experiences, express frustrations, and seek peer support. However, these communities may unintentionally foster echo chambers—environments where users reinforce similar narratives or perspectives, limiting exposure to diverse viewpoints that could otherwise help them cope with mental health challenges. Members of these communities commonly post to either seek support, request practical advice, or simply vent about their situations.

While each user's circumstances are unique, many posts can be grouped into common themes or associated with certain keywords. Community responses vary: some offer direct advice, while others share relatable personal experiences intended to comfort the original poster (often referred to as OP, or Original Poster). These different responses and post themes can shape the type of discussion that emerges. Some interactions lead to productive, meaningful conversations beneficial to the OP, whereas others become repetitive or homogeneous, adding little value. In less productive discussions, participants may primarily reinforce their existing opinions or ideas rather than providing genuine support. Such exchanges are unlikely to benefit—and may even harm—the OP, negatively affecting the community as a whole. These environments, where users repeatedly encounter similar beliefs while filtering out diverse perspectives, exemplify echo chambers.

Not every discussion results in an echo chamber. Previous studies indicate that echo chambers typically form around controversial topics, often driven by strong pre-existing beliefs, group identity reinforcement, or platform algorithmic biases. Political and religious subreddits are common examples. However, echo chambers can also emerge within mental health-related communities if negative sentiments or pessimistic viewpoints about sensitive topics are continuously reinforced. Because individuals facing mental health challenges are often particularly sensitive and vulnerable, becoming confined within such echo chambers can have detrimental effects. Hence, it is crucial to investigate community norms, cultures, and user interaction patterns. Research in this direction can provide insights into emotional dynamics within peer-support communities and help inform algorithmic moderation practices and community design.

No universally accepted method exists for detecting echo chambers, as definitions and indicators vary greatly depending on the domain and context of the study. Some studies employ natural language processing (NLP) techniques—such as semantic similarity or sentiment homogeneity—to assess the echo chamber effect, while others use network-based methods like analyzing user interactions, retweet cascades, or cross-community engagement. The choice of method largely depends on the specific platform and discourse type (e.g., political, social, health-related). In this project, semantic similarity analysis was chosen to assess echo chambers in r/depression.

## 2 Related Work

Echo chambers have been studied extensively in the context of political and controversial discourse for social media for quite a long time. However, different researchers have studied echo chamber differently and also proposed their methods. Ghafouri et al.[4] in their proposed two new metrics for quantification of echo chamber. They measured the effect of an echo chamber through the inverse effect of user diversity, and (2) polarization by means of user separability between two echo chambers in a topic. Their results indicate a data driven findings and they observed a direct relationship between echo chamber and polarization. Alatawi et al. [1] proposed another novel metric, Echo Chamber Score (ECS), a metric that assesses the cohesion and separation of user communities by measuring distances between users in the embedding space. ECS then measures (i) cohesion—the average embedding distance among users inside


ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

each detected community—and (ii) separation—the distance between users across communities; a high-cohesion/ high-separation graph gets a high ECS, signaling an echo chamber. Their proposed model tracks echo-chamber intensity at both topic and community level and correlates with supervised baselines without labeled data. Cinelli et al. [2] explored the key differences between social media platforms and how they influence the formation of echo chambers. They focused on two main dimensions: 1) homophily in the interaction networks and 2) bias in the information diffusion toward like-minded peers.

Perera et al. [8] in their work proposed a new measure, Signed Echo (SEcho) that quantifies opinion reinforcement and rejection properties of echo chambers and an echo chamber detection algorithm, Signed Echo Detection Algorithm (SEDA) based on this measure, which preserves the connectivity among echo chamber members. Markgraf et al. [6] in their research paper first proposed an echo chamber metric that handles many parties instead of binary parties (democrat vs republic). They first used the metric cosine similarity to measure homophily in multi party systems. They found that the members of the political right experience the least amount of crosscutting communication and the highest degrees of homophily. Impicciche et al. [5] in their work provided a comparative analysis of the echo chamber detection metrics, with a focus on their application in a cross-platform scenario constituted by the two major social media platforms, i.e., Twitter (now renamed X) and Reddit. Some of the primary echo chamber detection metrics considered encompass network analysis, content analysis, and hybrid solutions.

Villa et al. [10] proposed a two-phase framework for detecting echo chambers on Twitter's early-COVID-19 discourse. Morini et al. [7] proposed a four-step pipeline for detecting echo chambers: (i) the identification of a controversial issue; (ii) the inference of users' ideology on the controversy; (iii) the construction of users' debate network; and (iv) the detection of homogeneous meso-scale communities. In summary, they detected small group of communities based on controversial issues.

# 3 Methodology

In this project, I have chosen semantic similarity as the primary metric for detecting potential echo-chamber effects. Having actively participated in the community for some time, I noticed several recurring themes expressed through specific keywords. Posts and comments were grouped based on these identified keywords for analysis. The selected keywords and themes include: **alone**, **depression**, **anxiety**, **ADHD**, **PTSD**, **stress**, **debt**, **poverty**, **suicide**, **self-harm**, and **death**. While additional themes could have been explored, I limited this study to these specific topics due to the scale and time constraints of the project. The key methodologies employed are briefly described below:

a) Semantic similarity among a thread's comments is calculated after the thread is grouped by specific keywords. High average scores are interpreted as evidence of unusually homogeneous discussion, suggesting that an echo chamber may be present.

b) Semantic similarity across posts is evaluated to determine whether keyword-grouped posts convey similar meanings, suggesting that users who share them may have similar experiences.

c) Sentiment analysis across time get an idea od the overall mood or expression of the subreddit.

## 3.1 Dataset

For this study, I collected a year-long dataset from the Reddit platform via the ArcticShift archive, a publicly accessible repository for Reddit data. The dataset comprises posts and comments from the r/depression subreddit, covering the period from April 2024 to April 2025. In total, it includes approximately 154027 posts and 428174 comments, capturing a wide range of user interactions, emotional expressions, and topic trends over time. Both the posts and comments dataset had a lot of irrelevant attributes. Only the relevant fields were kept for a cleaner analysis.

| Field Name | Description |
|---|---|
| `id` | Unique identifier for the comment |
| `parent_id` | ID of the parent comment or post |
| `link_id` | ID of the original post associated with the comment |
| `body` | The textual content of the comment |
| `author` | Username of the comment's author |
| `score` | Net upvotes received by the comment |
| `created_utc` | Time of creation in UTC timestamp |
| `subreddit` | Name of the subreddit (for validation or filtering) |

**Table 1: Selected fields from the comment dataset of r/depression subreddit used in the analysis**

| Field | Purpose |
|---|---|
| `id` | Unique post identifier |
| `title` | Post headline; |
| `selftext` | Main body of the post |
| `author` | Username of the poster. |
| `score` | Net up-votes retained as a simple engagement signal. |
| `num_comments` | Comment count |
| `created_utc` | UTC timestamp of submission |
| `subreddit` | Subreddit name |

**Table 2: Selected fields from the post dataset of r/depres- sion subreddit used in the analysis**

## 3.2 Preprocessing

As social media comments are always noisy, informal and unique it is hard to model them. I performed several preprocessing steps to clean and standardize the datasets. The texts were normalized by converting to lowercase, removing URLs, punctuation, special characters, and stopwords. Tokenization and lemmatization were applied using the spaCy [3] library to standardize word forms and reduce lexical variation.

## 3.3 Evaluation Metrics

There is no officially extablished metric to measure echo chamber. Existing methods for echo chamber detection are either one-dimensional, only considering the network behavior of users while ignoring their semantic behavior, or require massive supervised labeling, which is both expensive and less generalizable. In this project I chose to use semantic similarity as the primary quantifier. I have used the cosine similarity library available from PyTorch [9] to explore homophily in a single thread. The cosine similarity metric can be expressed by the following equation:

$$\text{cosine similarity} = \frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\max\left(\|\mathbf{x}_1\|_2 \, \|\mathbf{x}_2\|_2,\ \varepsilon\right)}.$$

This equation returns the cosine similarity between $\mathbf{x}_1$ and $\mathbf{x}_2$, embeddings for two pieces of texts.

## 4 Results

Figure 1 displays the mean pairwise cosine similarity of comments within posts that contain each keyword. The scores range from approximately 0.20 to 0.33, indicating that comment sections are seldom identical in wording yet can still be semantically similar. *Hopelessness* yields the highest average similarity (≈ 0.33), followed by *medication* and *living with parents* (≈ 0.30). Threads on these themes therefore host the most linguistically uniform discussions, suggesting that contributors tend to echo the same concerns or coping narratives. The lowest similarity scores occur for *ADHD*, *PTSD*, and *debt* (≈ 0.20–0.22). Comment exchanges under these topics are more diverse, implying a broader range of personal experiences or viewpoints. In summary the representation of the Figure 1 supports the hypothesis of echo chambers that: topics tied to immediate emotional states or widely shared treatment routines exhibit tighter semantic clustering, whereas discussions of chronic diseases remain more unique.

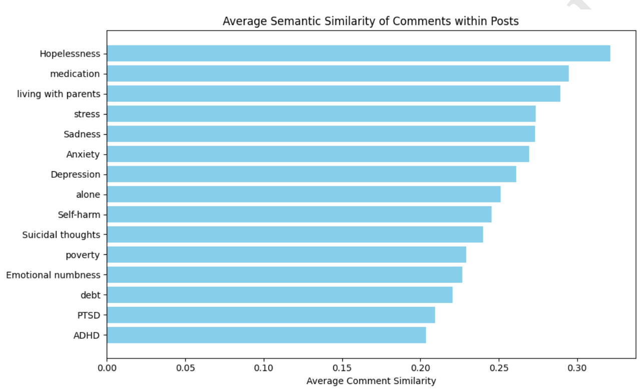


**Figure 1: Semantic similarity within threads grouped by certain keywords**

Figure 2 ranks keywords by the average cosine similarity between distinct posts that mention them. Similarity values span a fairly narrow band ( 0.39–0.46). *Loneliness* produces the most homogeneous post set ( 0.46), closely followed by *poverty* and *debt* ( 0.44–0.45). Posts that invoke this stressors exhibit kind of similar linguistic patterns.

The lowest cross-post similarity is observed for broad affective states such as *anxiety* and *depression* ( 0.39–0.40), indicating greater thematic diversity in how these conditions are framed.

An additional linguistic pattern emerged in the mental-health subreddit: posts generally fall into two functional categories. Users either (i) *vent*—expressing emotions or desires—or (ii) *seek advice* by posing explicit questions. To test whether these formats foster different degrees of echoing replies, I split the dataset accordingly and compared the semantic similarity of the resulting comment threads. Figure 3 shows box-plots of the similarity scores for the two groups. Interestingly there is not much difference for such posts. The two distributions almost overlap: both medians sit near 0.22, and their inter-quartile ranges are almost similar.

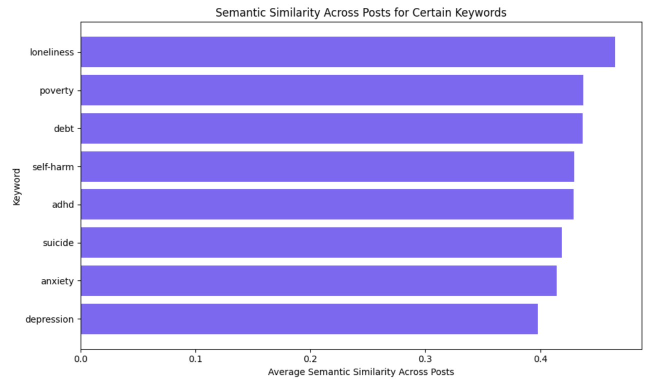


**Figure 2: Semantic similarity across posts grouped by certain keywords**

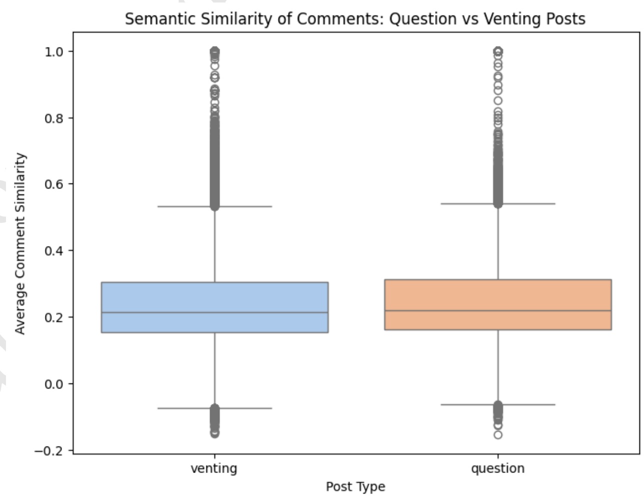


**Figure 3: Semantic similarity for posts asking questions vs. venting out**

Additionally, I performed sentiment analysis across the subreddit for the full year dataset, to get insights about any temporal patterns of the subreddit. Figure 4 tracks the subreddit's mean post sentiment over the 12-month window. Sentiment remains overall negative across the year with some slight positive tone for a few months. In short, the community's overall mood stays consistently negative, with brief intervals of slightly more positive language rather than a sustained trend.

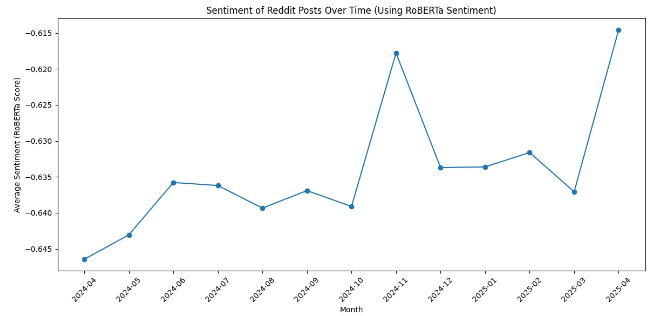


**Figure 4: Monthly average sentiment of r/*depression* posts, scored with a RoBERTa sentiment model**

## 5 Discussion and Limitations

The findings from this project indicate potential echo chamber effects for certain keywords in the r/depression subreddit. From Figure 1, we observe that certain topics such as hopelessness, medication, and living with parents, exhibited higher semantic similarity within comments, suggesting that discussions around these themes tend to reinforce homogeneous narratives. This aligns with the echo chamber hypothesis, where users may repeatedly encounter and amplify similar perspectives, limiting exposure to diverse viewpoints. Conversely, topics like ADHD, PTSD, and debt displayed lower semantic similarity, indicating greater diversity in responses. This could be attributed to the varied personal experiences associated with chronic conditions or situational stressors, which may not lend themselves to uniform discussions.

For Figure 2 posts related to loneliness, poverty, and debt showed higher linguistic homogeneity, implying that users framing their struggles around these issues may adopt similar language or narratives. This could reflect shared cultural or socioeconomic stressors that shape how individuals articulate their experiences. Broader themes like anxiety and depression had lower cross-post similarity, possibly due to the wide-ranging manifestations of these conditions, leading to more varied expressions.

Surprisingly in Figure 4 the study found no significant difference in semantic similarity between posts that vented and those that sought advice. This suggests that the echo chamber effect is not strongly influenced by the post's intent but may instead stem from the community's collective tendency to respond in predictable ways

Although this project analyzed a large, year-long dataset using careful methods, it still has some limitations. Mental health discussions are sensitive, and people often hide or avoid sharing details publicly due to privacy. Language models often miss subtle clues or hidden meanings in social media texts —for instance, if someone says they're "just tired," the real underlying issues may not be captured by the model. Additionally, responses to a post also varies by the time of posting, length of the post or sometimes certain keywords like 'suicide', 'death' instantiates serious discussions. These factors can sometimes influence the *reliability* and *interpretability* of the metric. That's why measuring echo chambers using an one dimensional metric can have some limitations; In future, a hybrid approach of analyzing graph interaction of users and combining it with content analysis can provide a more robust metric for quantifying echo chamber.

## 6 Conclusion

This project focuses on exploring the echo chamber effect within the r/depression subreddit. Through the analysis of a year-long dataset, several potential echo chambers were identified around specific themes. Quantifying echo chambers is challenging, as detection methods can vary significantly based on the platform and topic context. While extensive research has examined echo chambers in political and controversial contexts, studies investigating mental health communities remain limited. Online mental health communities are essential spaces that offer valuable support and empathy to vulnerable individuals. Therefore, careful moderation and thoughtful community design are crucial to ensure the well-being of users. This project addresses this gap by providing insights and setting a foundation for future research on echo chambers in mental health forums.

## Links

- Dataset: https://arctic-shift.photon-reddit.com/download-tool
- Codebase: https://github.umn.edu/Wahid016/Echo-Chamber.git

## References

[1] Faisal Alatawi, Paras Sheth, and Huan Liu. 2023. Quantifying the echo chamber effect: an embedding distance-based approach. In *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*. 38–45.

[2] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the national academy of sciences* 118, 9 (2021), e2023301118.

[3] Explosion AI. 2025. spaCy—Industrial-Strength Natural Language Processing in Python. https://spacy.io/. Accessed 3 May 2025.

[4] Vahid Ghafouri, Faisal Alatawi, Mansooreh Karami, Jose Such, and Guillermo Suarez-Tangil. 2024. Transformer-Based Quantification of the Echo Chamber Effect in Online Communities. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–27.

[5] Paola Impicciché and Marco Viviani. 2024. Comparing Echo Chamber Detection Metrics: A Cross-modeling and Cross-platform Analysis of Twitter and Reddit. *ACM Transactions on the Web* (2024).

[6] Moritz Markgraf and Manfred Schoch. 2019. Quantification of echo chambers: a methodological framework considering multi-party systems. (2019).

[7] Virginia Morini, Laura Pollacci, and Giulio Rossetti. 2021. Toward a standard approach for echo chamber detection: Reddit case study. *Applied Sciences* 11, 12 (2021), 5390.

[8] Kushani Perera and Shanika Karunasekera. 2024. Quantifying Opinion Rejection: A Method to Detect Social Media Echo Chambers. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 57–69.

[9] PyTorch Project Contributors. 2025. torch.nn.CosineSimilarity. https://pytorch.org/docs/stable/generated/torch.nn.CosineSimilarity.html. Accessed 3 May 2025.

[10] Giacomo Villa, Gabriella Pasi, and Marco Viviani. 2021. Echo chamber detection and analysis: A topology-and content-based approach in the COVID-19 scenario. *Social Network Analysis and Mining* 11, 1 (2021), 78.