

March , 2024

# COLLUSION ANALYSIS

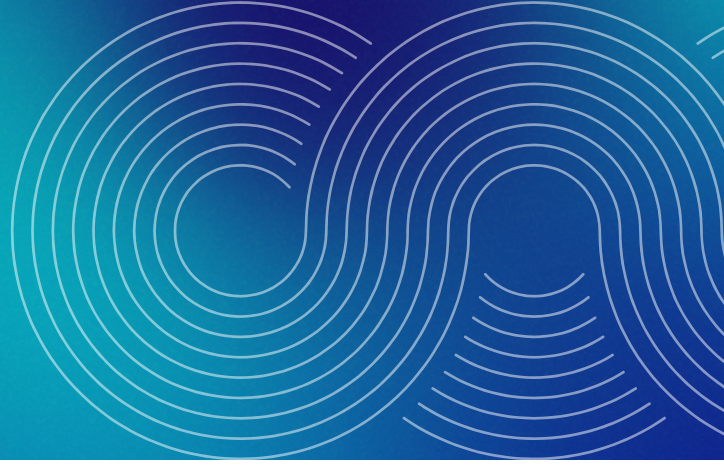
 ON THE ARBITRUM

# GOVERNANCE FORUM

(User Activity & Collaboration)



# TABLE OF CONTENTS



<b>1</b>	<b><u>EXECUTIVE SUMMARY</u></b>	<b>02</b>
<b>2</b>	<b><u>INTRODUCTION</u></b>	<b>02</b>
<b>3</b>	<b><u>METHODOLOGY</u></b>	<b>02</b>
<b>4</b>	<b><u>ANALYSIS FINDINGS</u></b>	<b>03</b>
	<u>Most Active Proposal Categories</u>	03
	<u>Frequency of Collaboration in Categories</u>	04
	<u>Cross-Category Participation</u>	05
	<u>Top 10 Most Active Users</u>	06
	<u>User Networks (Clusters)</u>	07
	<u>Collaboration Based on Trust Level</u>	08
	<u>Trends in User Collaboration Patterns</u>	09
	<u>Distribution of Topics Across Categories</u>	10
	<u>Distribution of Topics by Trust Levels</u>	11
	<u>Average Number of Comments per Topic Over Time</u>	12
	<u>Distribution of Users Across Trust Levels and Categories</u>	13
	<u>Flow of Interactions Between Users</u>	14
<b>5</b>	<b><u>CONCLUSION</u></b>	<b>15</b>
<b>6</b>	<b><u>RESOURCES</u></b>	<b>15</b>

# Executive Summary

The purpose of this report is to analyze user activity and collaboration patterns on the Arbitrum governance forum, focusing on detecting collusion among users. The analysis involves examining user behavior, interaction frequencies, and collaborative patterns across different proposal categories.

## Introduction

The Arbitrum governance forum serves as a platform for users to propose, discuss, and vote on various governance-related topics. Understanding user activity and collaboration dynamics is crucial for maintaining transparency and integrity within the community. This report aims to provide insights into user behavior and identify potential instances of collusion.

## Methodology



1

### Data Collection

Utilized the Discourse API and SQL queries to extract forum data, including user profiles, topics, and posts.

2

### Data Preprocessing

Cleaned and prepared the data by merging relevant datasets, converting data types, and handling missing values.

3

### Analysis Approach

Employed descriptive statistics, data visualization techniques, and network analysis to explore user activity and collaboration patterns.



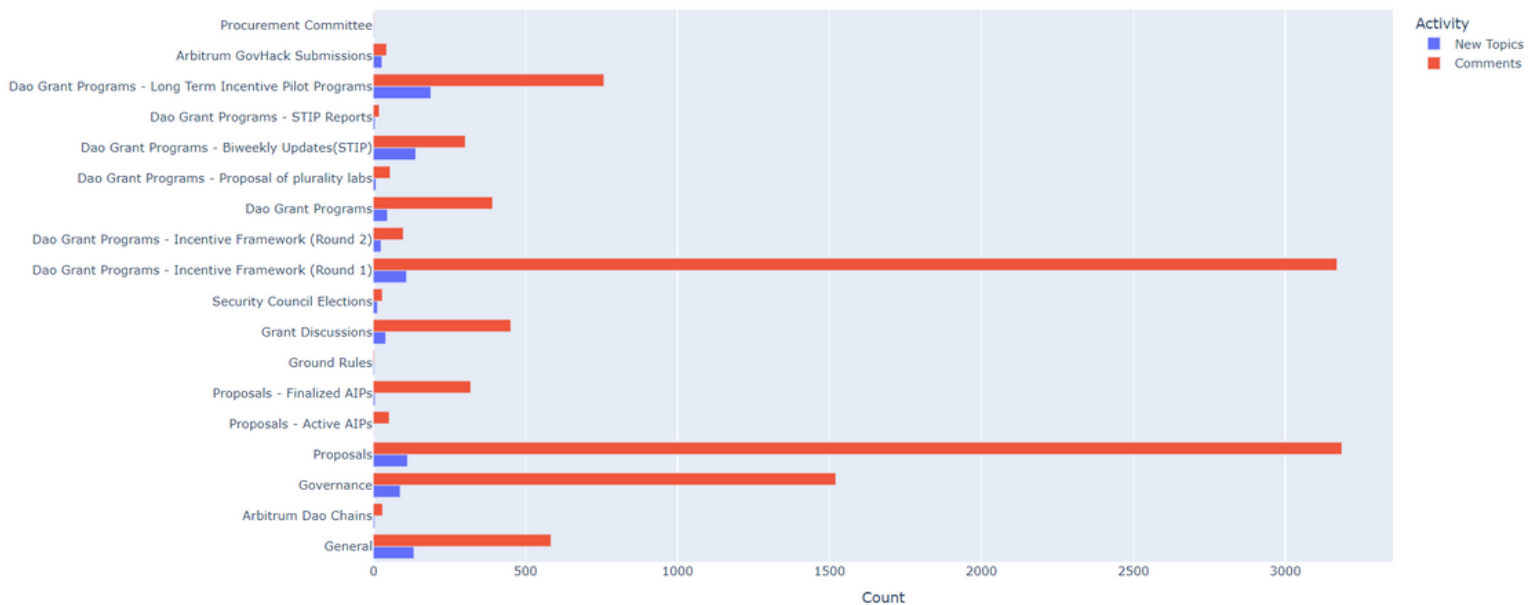
# ANALYSIS FINDINGS

## Most Active Proposal Categories

### Objective

Analyze user activity in terms of creating new topics and posting comments across different proposal categories to identify most active proposals.

User Activity by Proposal Categories



Source:- [Visualization Link](#)

### Visualization Explanation

To gain insights into the level of user engagement across various proposal categories, we utilized bar charts as a visualization tool. Each bar in the chart represents a specific proposal category, and the height of the bar indicates the level of user activity within that category, measured by the number of topics created and comments posted.

From the visualization, it is evident that certain proposal categories exhibit higher user activity compared to others. The categories "DAO Grant Programs - Incentive Framework (Round 1)" and "Proposals" stand out as the most active categories, with 3,170 comments and 109 topics, and 3,186 comments and 112 topics, respectively. On the other hand, the "Procurement Committee" category has the least user activity, with only 2 comments and topics.

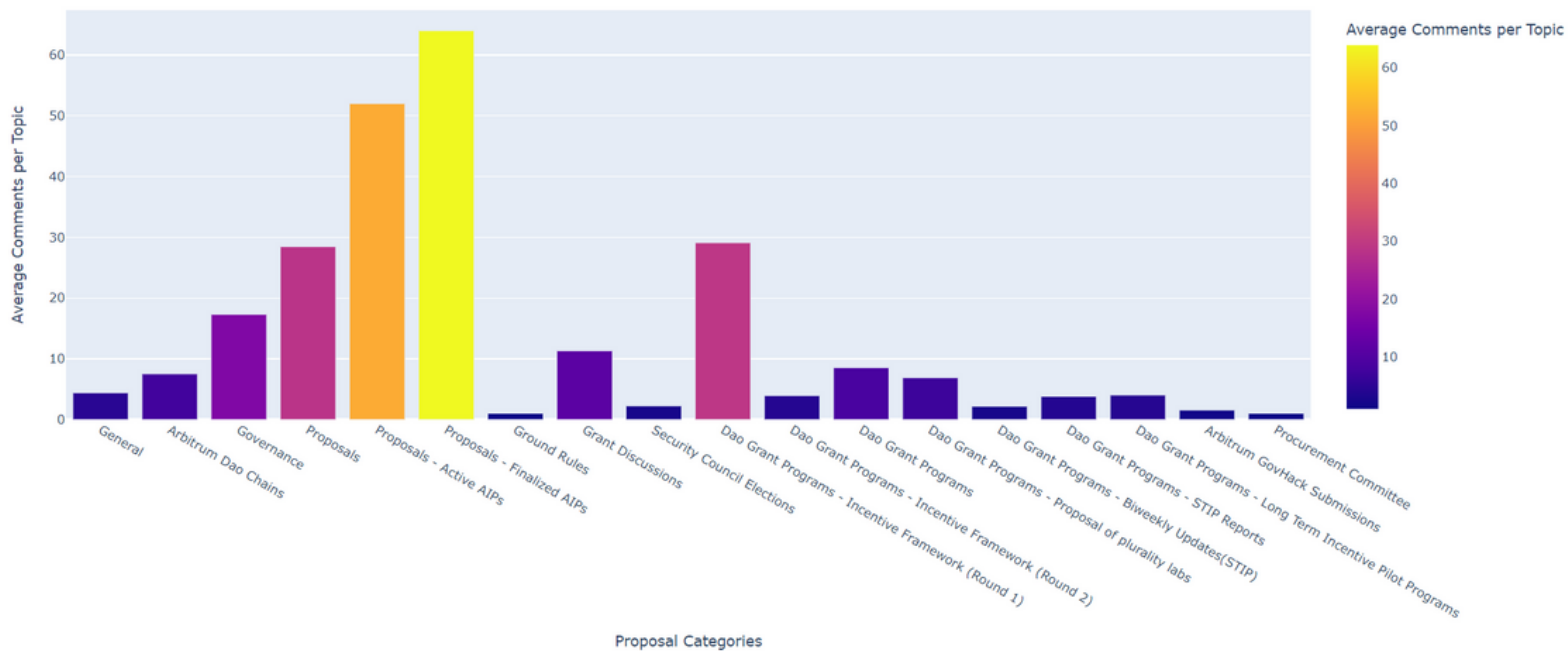


# Frequency of Collaboration in Categories

## Objective

Explore categories and subcategories where users tend to collaborate more frequently by consistently commenting on proposals.

Average Comments per Topic by Proposal Categories



Source:- [Visualization Link](#)

## Visualization Explanation

To delve deeper into understanding collaboration patterns among users within proposal categories and subcategories, we employed bar charts as an effective visualization tool.

From the visualization, we observed that the "Finalized AIPs" and "Active AIPs" subcategories within the "Proposals" category exhibited the highest average number of comments per topic. Additionally, the "Incentive Framework (Round 1)" subcategory of the "DAO Grant Programs" category ranked third in terms of collaboration frequency.

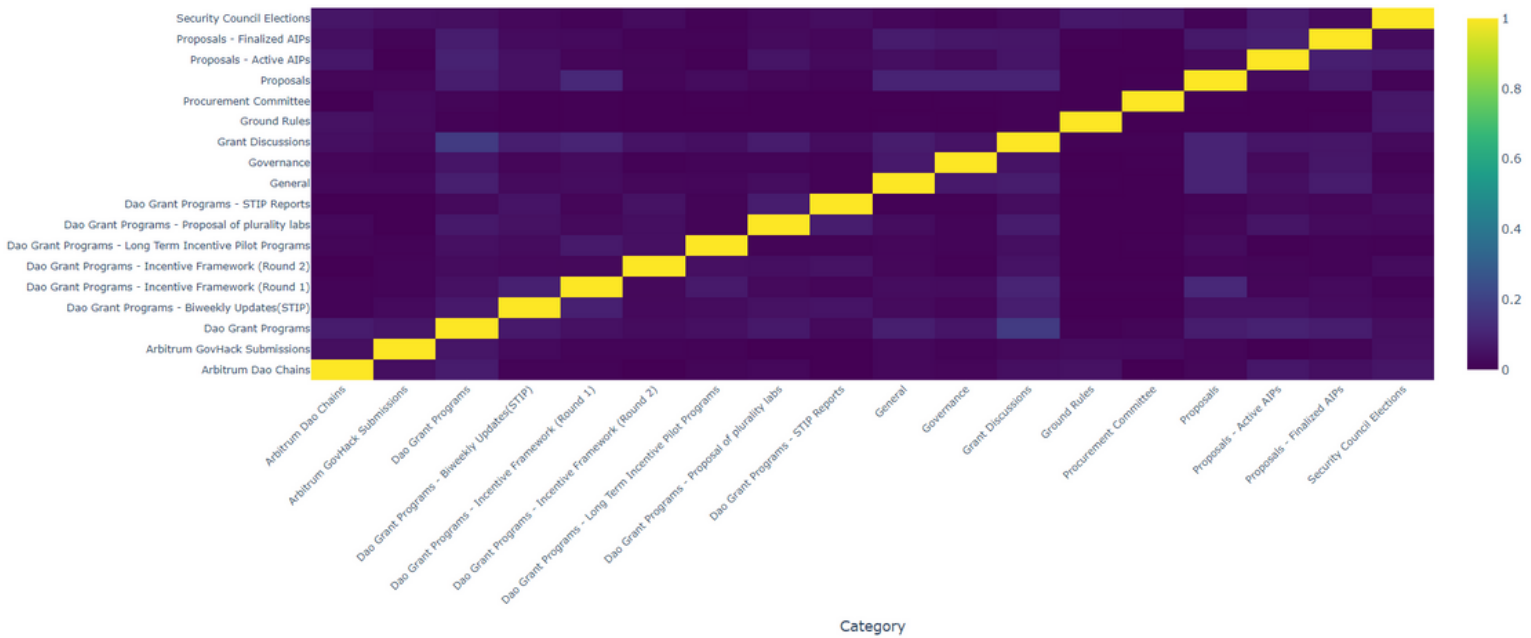


# Cross-Category Participation

## Objective

Investigate if users actively participating in one category also engage in other categories.

Jaccard Index Between Categories



## Percentage of users participating in multiple categories



## Source

[Visualization Link 1](#)

[Visualization Link 2](#)

## Visualization Explanation

With this analysis our objective was to delve into user engagement patterns across different proposal categories, specifically examining whether users who actively participate in one category also engage in other categories.

From the visualization, we noted that approximately 0.43% of users actively participated in multiple categories. This finding suggests that while there is some cross-category participation among users, the majority of users tend to focus their engagement within specific categories.

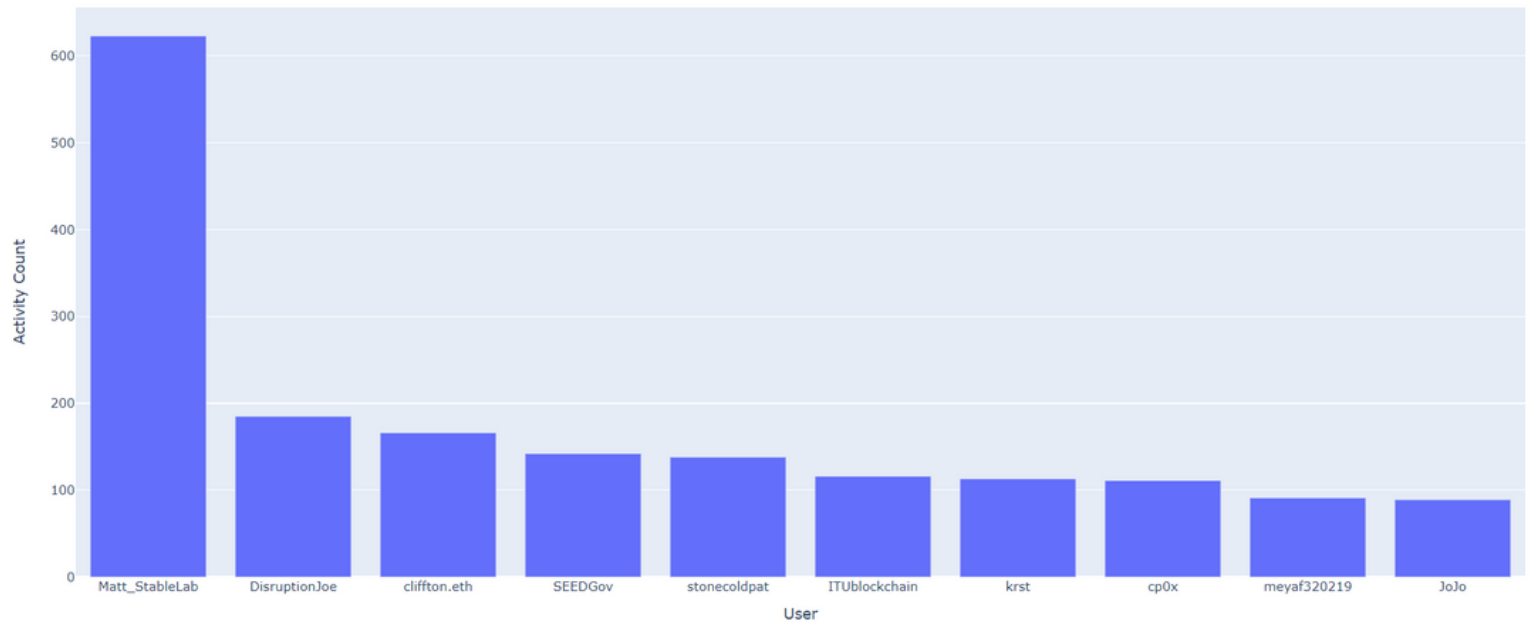


# Top 10 Most Active Users

## Objective

Identify the top 10 most active users based on their contributions to creating topics and posting comments.

Top 10 Most Active Users Across All Proposal Categories



Source:- [Visualization Link](#)

## Visualization Explanation

Visualized user activity distribution using bar charts, highlighting the top 10 most active contributors. The bars represent users, and their heights indicate the level of activity.

The visualization unequivocally reveals Matt\_StableLab as the most active participant in the forum, showcasing their significant engagement through contributions to topic creation and comment posting.

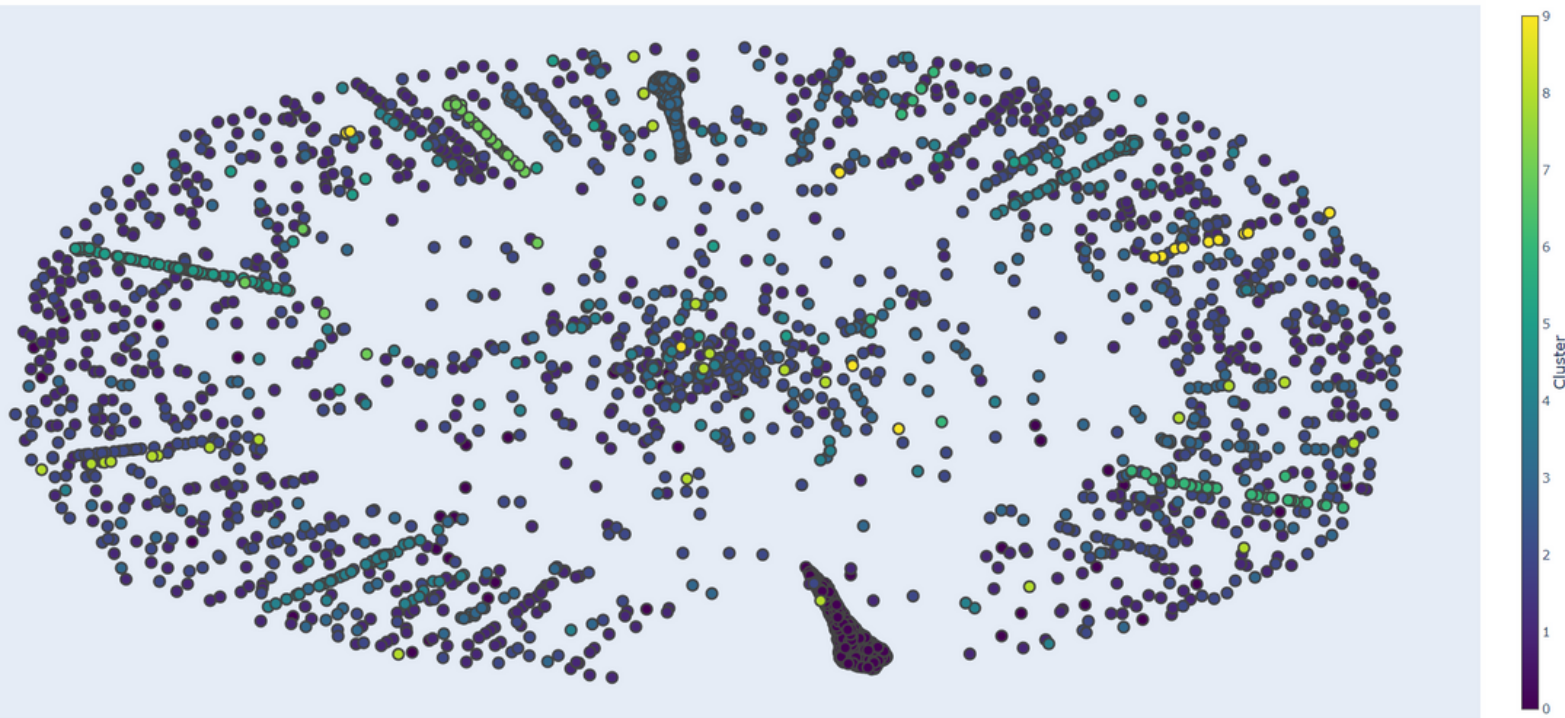


# User Networks (Clusters)

## Objective

Explore user networks or clusters where members frequently interact with each other's posts across different or same categories.

Top Clusters Using Louvain Algorithm



Source:- [Visualization Link](#)

## Visualization Explanation

For this analysis, we utilized the Louvain algorithm, which is specifically designed to detect communities within large networks. By applying this algorithm, we were able to identify clusters of users who frequently interact with each other's posts across various proposal categories.

The visualization reveals the presence of approximately 80 clusters within the user network, showcasing instances where members frequently interact with each other's posts across various proposal categories. Each cluster is distinctly visible, allowing for a clear understanding of the interconnectedness and collaborative patterns among users on the forum.



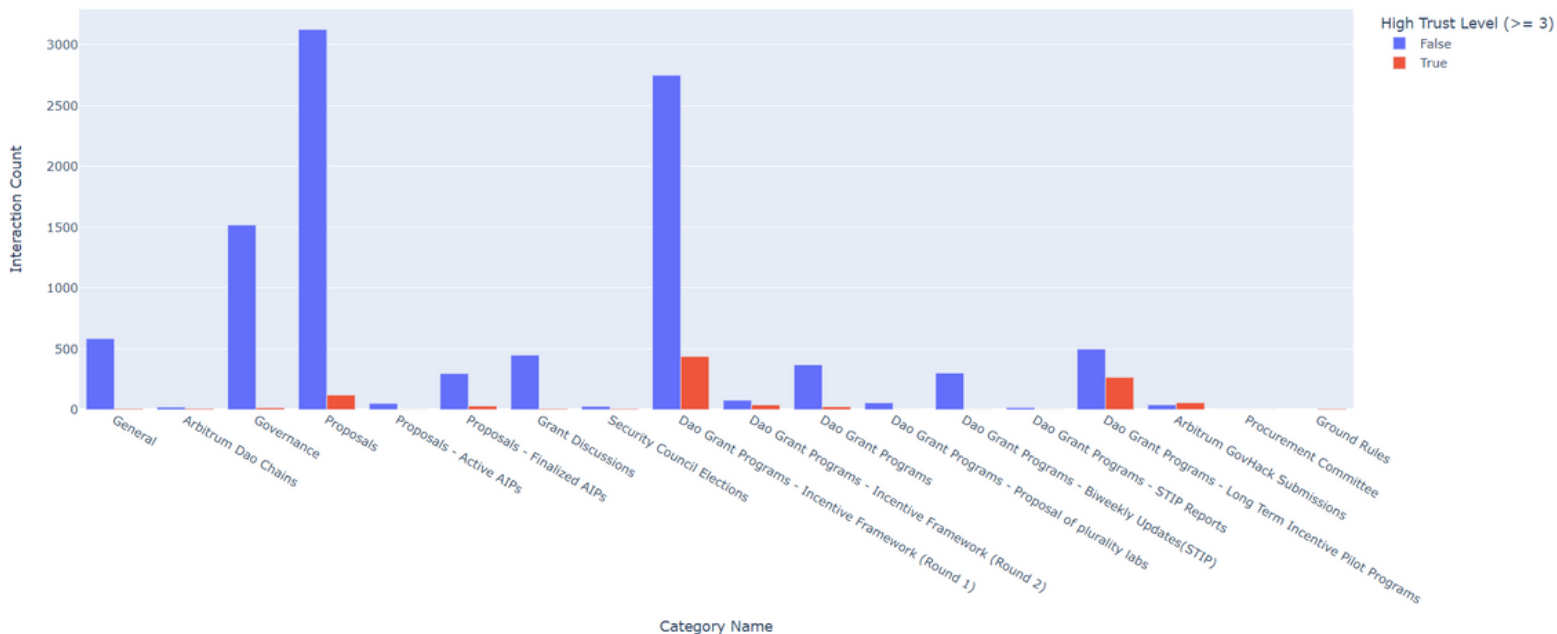


# Collaboration Based on Trust Level

## Objective

Analyze if users with high "Trust Level" collaborate more frequently compared to regular users across all proposal categories.

Interaction Counts by Trust Level and Category



Source:- [Visualization Link](#)

## Visualization Explanation

The analysis utilized a grouped bar chart to compare the frequency of collaboration between users with different trust levels across all proposal categories. The analysis uncovered a notable trend that users with trust levels lower than 3 exhibit a higher propensity for collaboration compared to those with trust levels of 3 and above.

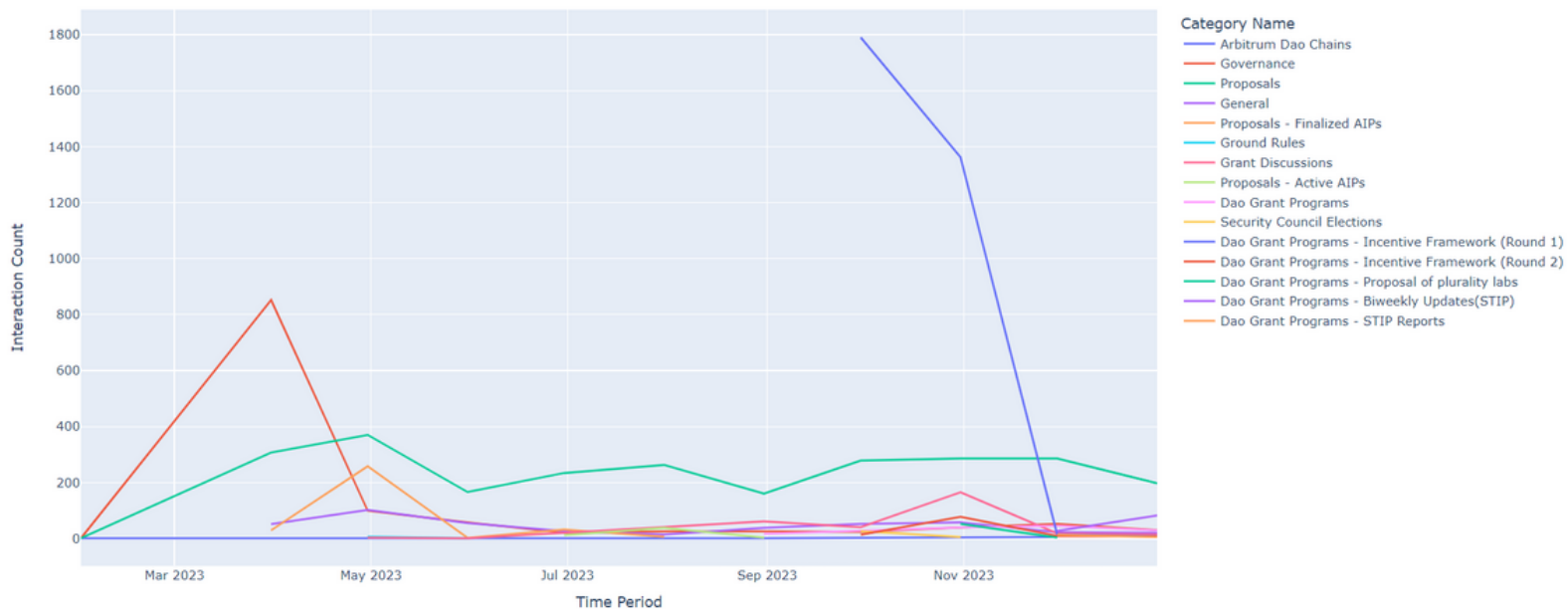


# Trends in User Collaboration Patterns

## Objective

Investigate trends in user collaboration patterns over time across different proposal categories.

User Collaboration Patterns Over Time Across Proposal Categories



Source:- [Visualization Link](#)

## Visualization Explanation

Presented temporal trends using line charts, where the x-axis represents time and the y-axis represents the interaction count per topic. Each line represents a proposal category, showing how collaboration patterns evolve over time.

The visualization indicates that the "Proposal" and "General" categories exhibit greater fluctuations compared to other categories. These fluctuations may indicate varying levels of user activity or engagement within these specific proposal categories over time.

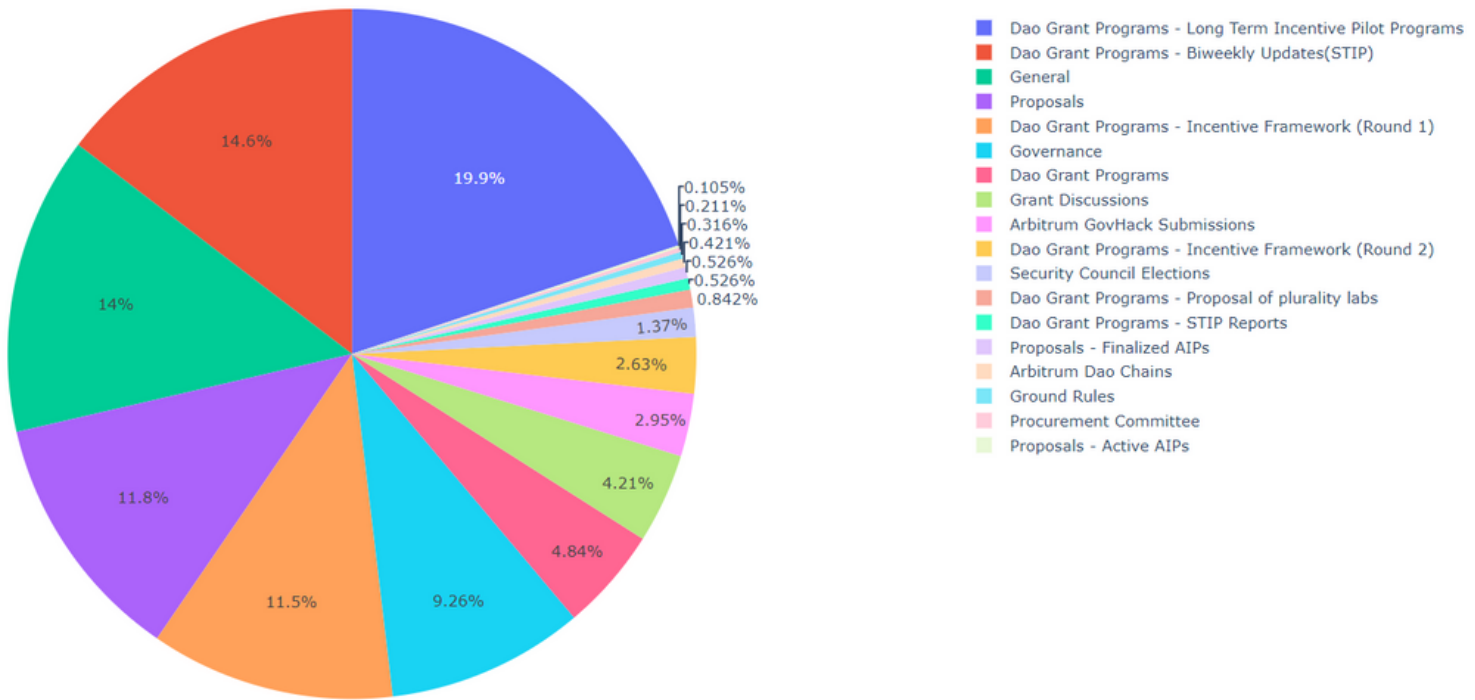


# Distribution of Topics Across Categories

## Objective

Examine the distribution of topics across different proposal categories.

Distribution of Topics Across Proposal Categories



Source:- [Visualization Link](#)

## Visualization Explanation

Visualized topic distribution using pie charts to understand category-wise topic distribution. Each slice represents a category, and its size indicates the proportion of topics in that category.

The pie chart highlights the distribution of topics across various proposal categories. It reveals that DAO Grant Program - Long Term Incentive Pilot Program boasts the highest number of topics, accounting for 19.9% of the total, followed by DAO Grant Program - Biweekly Updates with 14.6%, and General category with 14%.

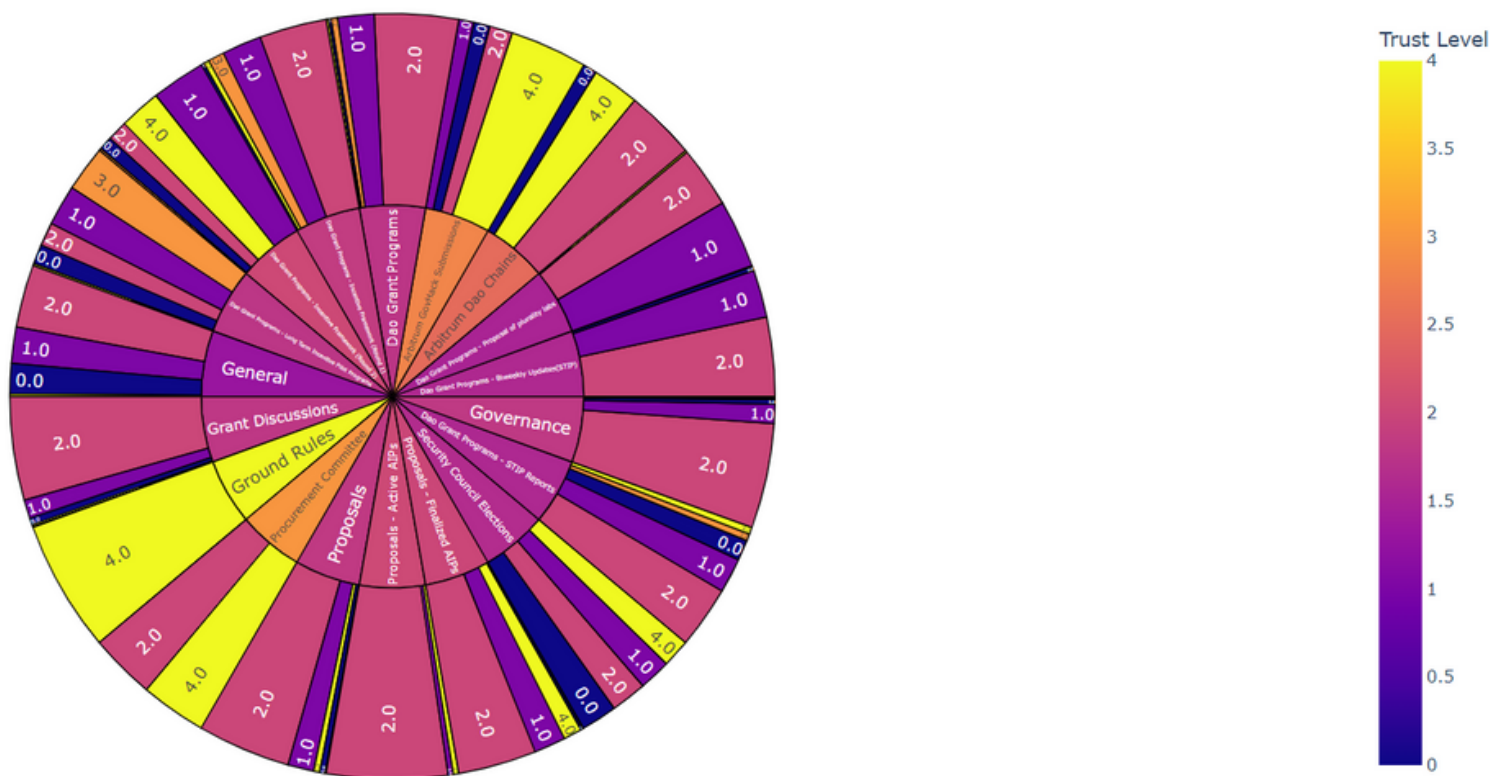


# Distribution of Topics by Trust Levels

## Objective

Explore the distribution of topics created by users with different trust levels across proposal categories.

Distribution of User Trust Levels Across Proposal Categories



Source:- [Visualization Link](#)

## Visualization Explanation

The analysis utilized a sunburst chart generated using Plotly. This visualization effectively displays the distribution of topics by trust levels across different proposal categories. Each level of the sunburst represents a hierarchy, with the outer ring representing proposal categories, and the inner rings representing trust levels within each category.

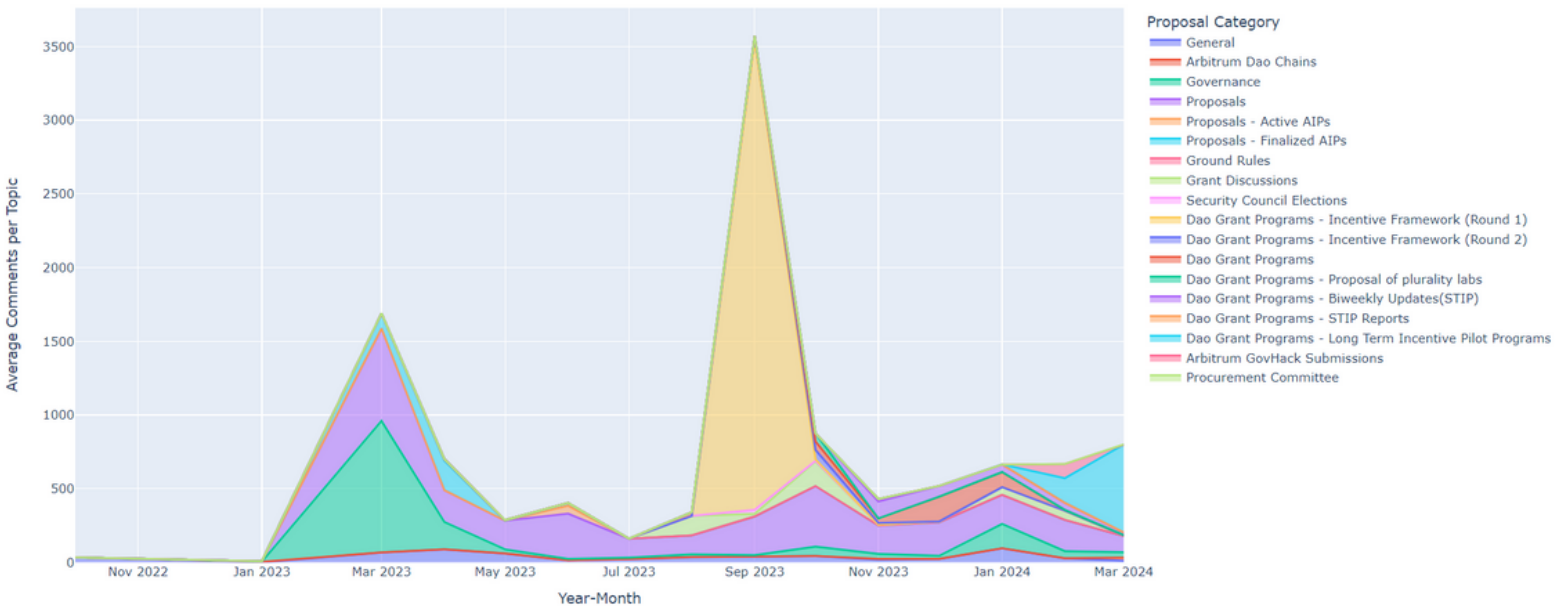


# Average Number of Comments per Topic Over Time

## Objective

Analyze how the average number of comments per topic varies over time across different proposal categories.

Average Number of Comments per Topic Over Time by Category



Source:- [Visualization Link](#)

## Visualization Explanation

The analysis employed an area chart to visualize the average number of comments per topic over time across different proposal categories. In the area chart, each category is represented by a distinct colored area, allowing for easy comparison of comment activity trends over time.

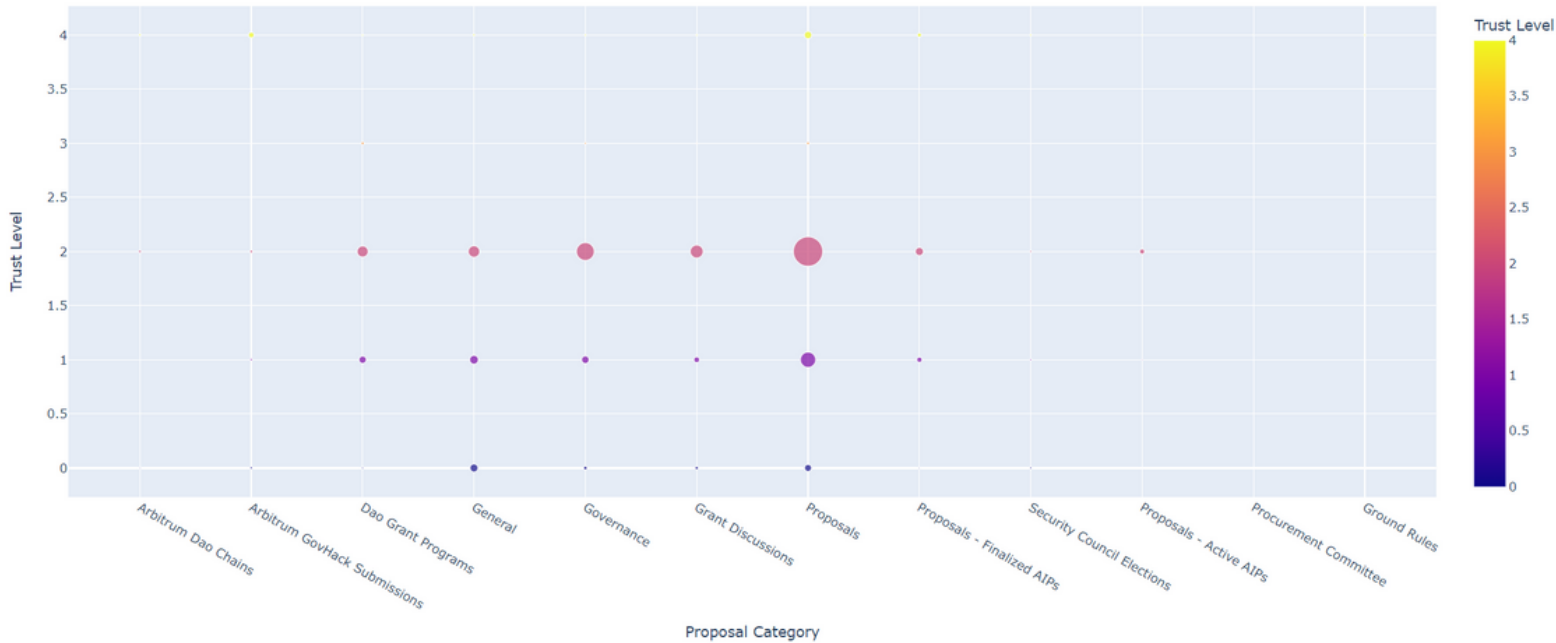


# Distribution of Users Across Trust Levels and Categories

## Objective

Investigate the distribution of users across different trust levels and proposal categories based on their total activity.

Distribution of Users Across Trust Levels and Proposal Categories Based on Total Activity



Source:- [Visualization Link](#)

## Visualization Explanation

The analysis utilized a bubble chart to illustrate the distribution of users across different trust levels and proposal categories based on their total activity. In the bubble chart, each bubble represents a combination of a trust level and a proposal category. The size of each bubble corresponds to the total activity level, while the color indicates the trust level of users.

The visualization clearly demonstrates that users with a trust level of 2 exhibit the highest activity across all categories, as evidenced by the larger size and greater frequency of bubbles associated with trust level 2. Conversely, users with a trust level of 4 display the least activity, with smaller and fewer bubbles representing this group across various categories.

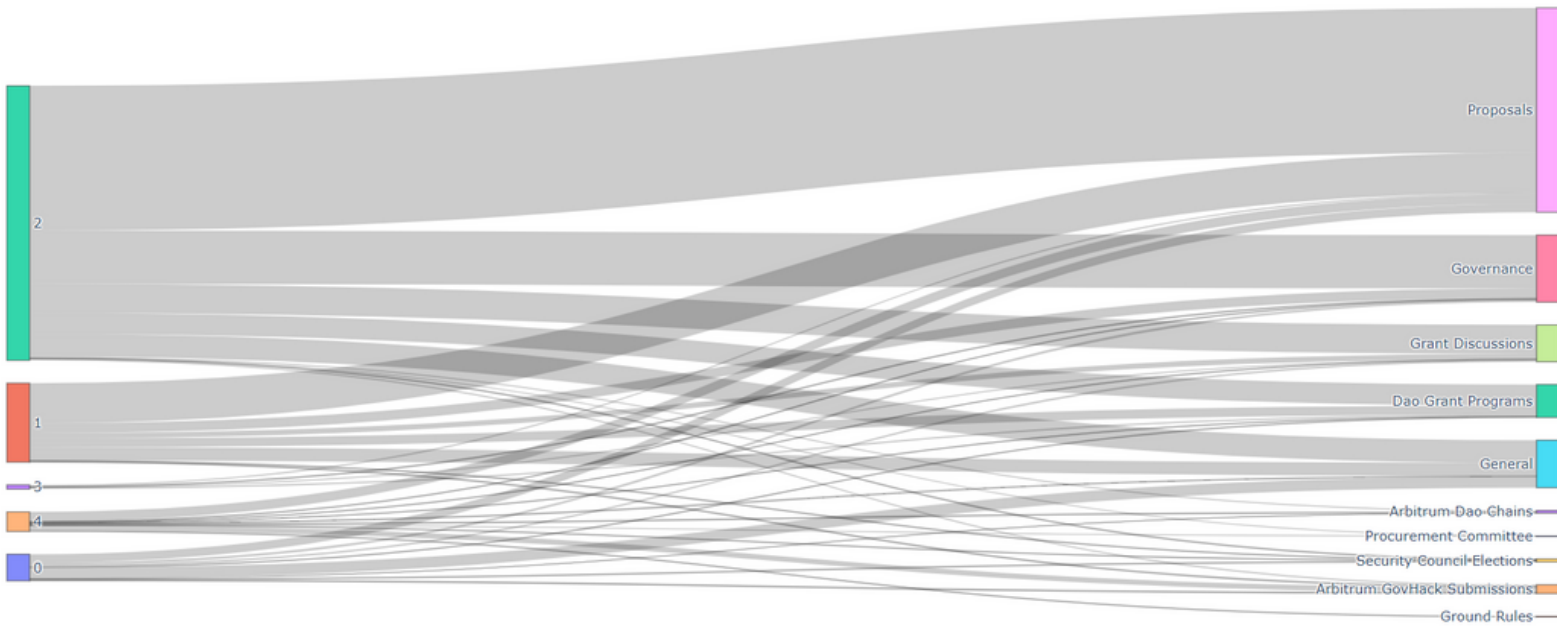


# Flow of Interactions Between Users

## Objective

Explore the flow of interactions between users with different trust levels across various proposal categories.

Flow of Interactions Between Users with Different Trust Levels Across Proposal Categories



[Source:- Visualization Link](#)

## Visualization Explanation

Utilized Sankey diagrams to visualize interaction flows between users. Nodes represent trust levels or proposal categories, and the flow between nodes represents interactions. The width of the flow indicates the volume of interactions between user groups or categories.



# CONCLUSION

In conclusion, the analysis of user activity and collaboration patterns within the Arbitrum governance forum reveals valuable insights into the dynamics of community engagement and interaction. By examining user behavior across different proposal categories, it is evident that certain categories attract higher levels of activity and collaboration, indicating areas of particular interest or significance within the community. The identification of top contributors and collaborative clusters underscores the importance of community-driven initiatives in fostering dialogue and driving the governance process forward. Additionally, the observed fluctuations in collaboration patterns over time highlight the dynamic nature of user engagement within the forum. Understanding these patterns can inform forum moderators and community managers in devising strategies to encourage participation, detect potential collusion, and ensure the integrity of the governance process. Furthermore, the distribution of topics across categories and the analysis of user trust levels provide valuable insights into the composition and behavior of the forum's user base. Overall, this analysis serves as a foundation for ongoing efforts to promote transparency, inclusivity, and accountability within the Arbitrum governance ecosystem.

## Resources

**01** **Discourse API Docs (For getting arbitrum forum data)**  
<https://docs.discourse.org>

**02** **Visualization Understanding**  
<https://visme.co/blog/data-visualization-types/>

**03** **Lighthouse (For hosting visualization file)**  
<https://docs.lighthouse.storage/lighthouse-1>

**04** **Python Plotly Docs**  
<https://plotly.com/python/>

