



ENCOURAGING VACCINATION ADOPTION THROUGH SOCIAL NETWORKS

MADT8103: Study of Specific Topics in Applied Data Science
(Advanced Data Technology Seminar)
Aj.Worapol Pongpech



AGENDA

- Introduction and Motivation
- Research Problems and Questions
- Objectives
- Theoretical Framework and Hypothesis
- Methodology and Tools
- Results
- Interventions
- Future work

Introduction and Motivation

INTRODUCTION

7.7 % increase
in vaccination refusal or delays

Social network significantly impacts vaccination decision.

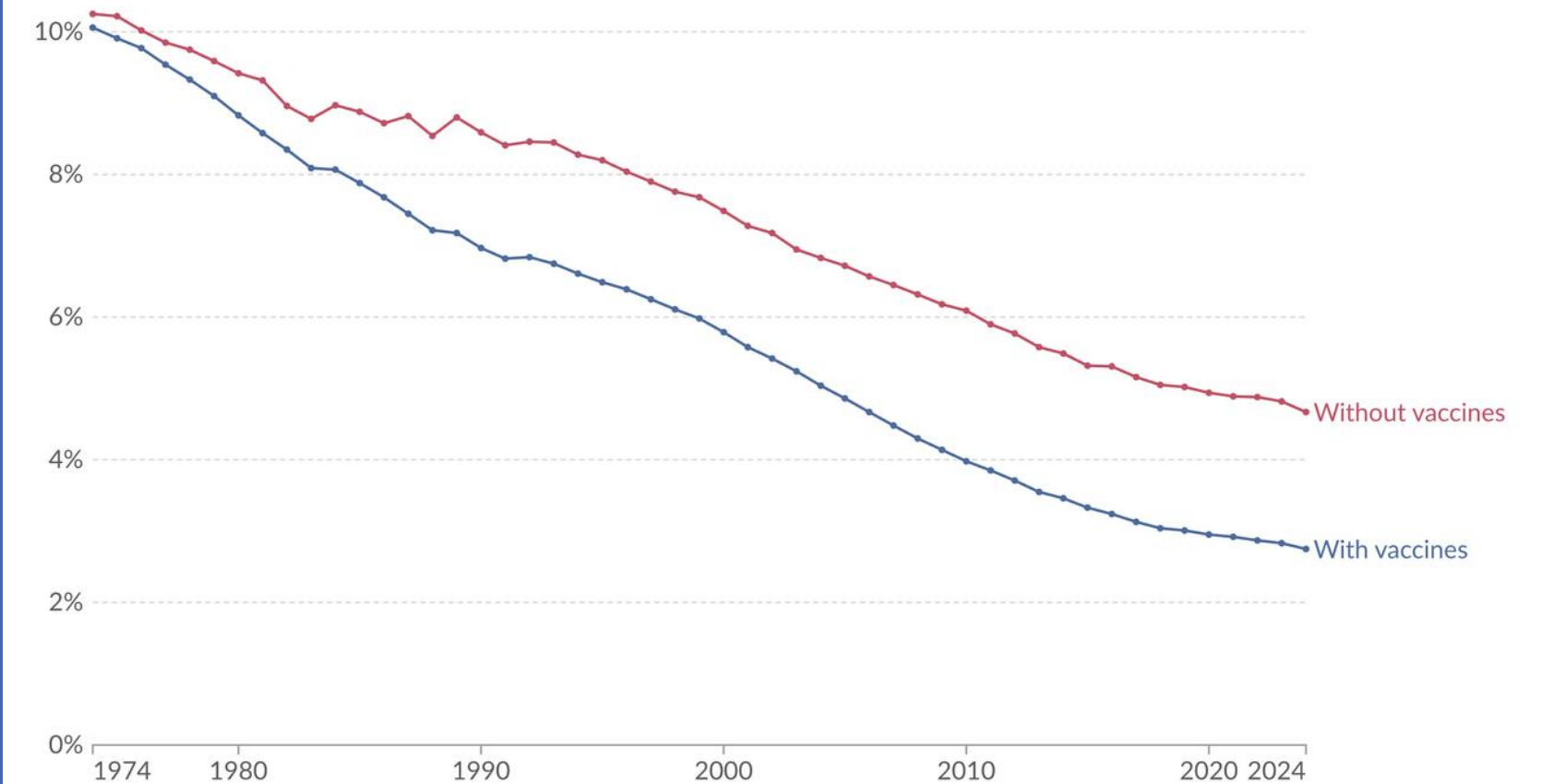
Motivation

- Social network shapes health opinions, influencing pro- and anti-vaccination views.
- Guide targeted interventions to combat misinformation and promote vaccination.

Global infant mortality rate with and without vaccines

Our World
in Data

Infant mortality rates measure the share of newborns who die before reaching their first birthday. They are shown here as the actual observed change in the global rate (in blue), and a hypothetical scenario if vaccinations hadn't been rolled out (in red).



Data source: Shattock et al. (2024). Contribution of vaccination to improved child survival: modelling 50 years of the Expanded Programme on Immunization.

CC BY

Research Problems and Questions

Research Problems



Mis-information
spreading



Social network
influencers impact
vaccination decisions

Research Questions

How does social network analysis help creating vaccination decision intervention?

Who are the key nodes (vertices) driving vaccine-related information?

Objectives

Objectives

- **Identify influencers**
- **Analyze information diffusion patterns**
- **Propose interventions**

Theoretical Framework and Hypothesis

Theoretical Framework

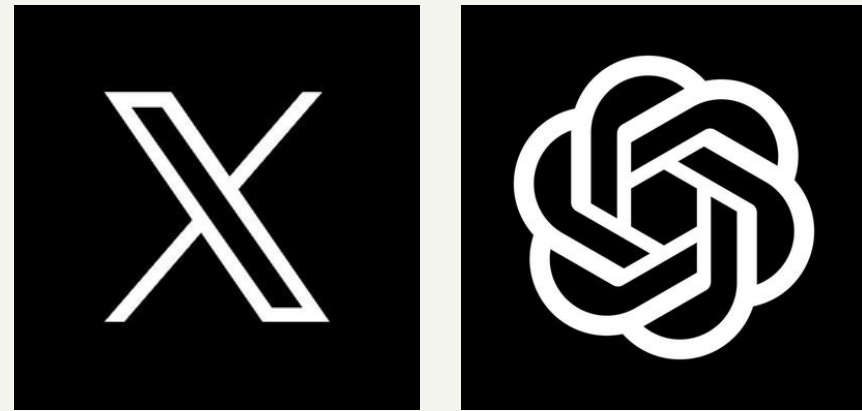
- Network theory
 - Nodes
 - Edges
 - Graph
 - Properties
- SNA
 - Centrality
 - Density
 - Clustering
 - Path Length

Hypothesis

- Influencers with a large follower base typically generate higher levels of content sharing.
- When an influencer is also the poster, the post will get higher number of shares.
- Posts from individuals with a higher modularity class are likely to have higher engagement levels.

Methodology and Tools

Methodology



Data Source

X (Prototype)
ChatGPT (Synthetic)



Data Processing

Filter relevant vaccine-
related posts



Analytical Scope

- Influencer and User Role Analysis
- Info-spread Analysis
- Sentiment and Topic Analysis

Data Source Limitation

X API V2 Products

Changes across our Free, Basic, and Pro packages are now live. View our [announcement](#) or check out our updated [product offerings](#) to learn more.

If you need higher levels of access, sign up for the Enterprise API today! Start [here](#).

Pro	Basic	Free
Unleash the full potential of X's v2 API	Begin your journey with X's v2 API	Get limited access to X's v2 API
Apps 3 environments	Apps 2 environments	Apps 1 environment
DMs Over 300K requests per month per user	DMs 75K requests per month per user	Posts Retrieve up to 100 Posts and 500 writes per month
Posts Retrieve up to 1M Posts per month	Posts Retrieve up to 15K Posts per month	
Realtime Posts stream Access to realtime posts stream	Realtime Posts stream No access	
Users Over 8M requests per month per user	Users 50k requests per month per user	
<div><div>SAVE 10%</div><div>\$4500/month</div><div>\$54000 billed annually</div><div>Upgrade in 1-Click</div></div> <div><div></div><div>\$5000/month</div><div>\$5000 billed monthly</div><div>Upgrade in 1-Click</div></div>	<div><div>SAVE 12.5% - LIMITED TIME</div><div>\$175/month</div><div>\$2100 billed annually</div><div>Upgrade in 1-Click</div></div> <div><div></div><div>\$200/month</div><div>\$200 billed monthly</div><div>Upgrade in 1-Click</div></div>	<div><div>✓</div><div>You have access to this plan</div><div>Downgrade to free access?</div><div>Downgrade</div></div>

X has changed their API policy, requiring at least \$175/month to access the necessary data.

Nodes and Edges in X

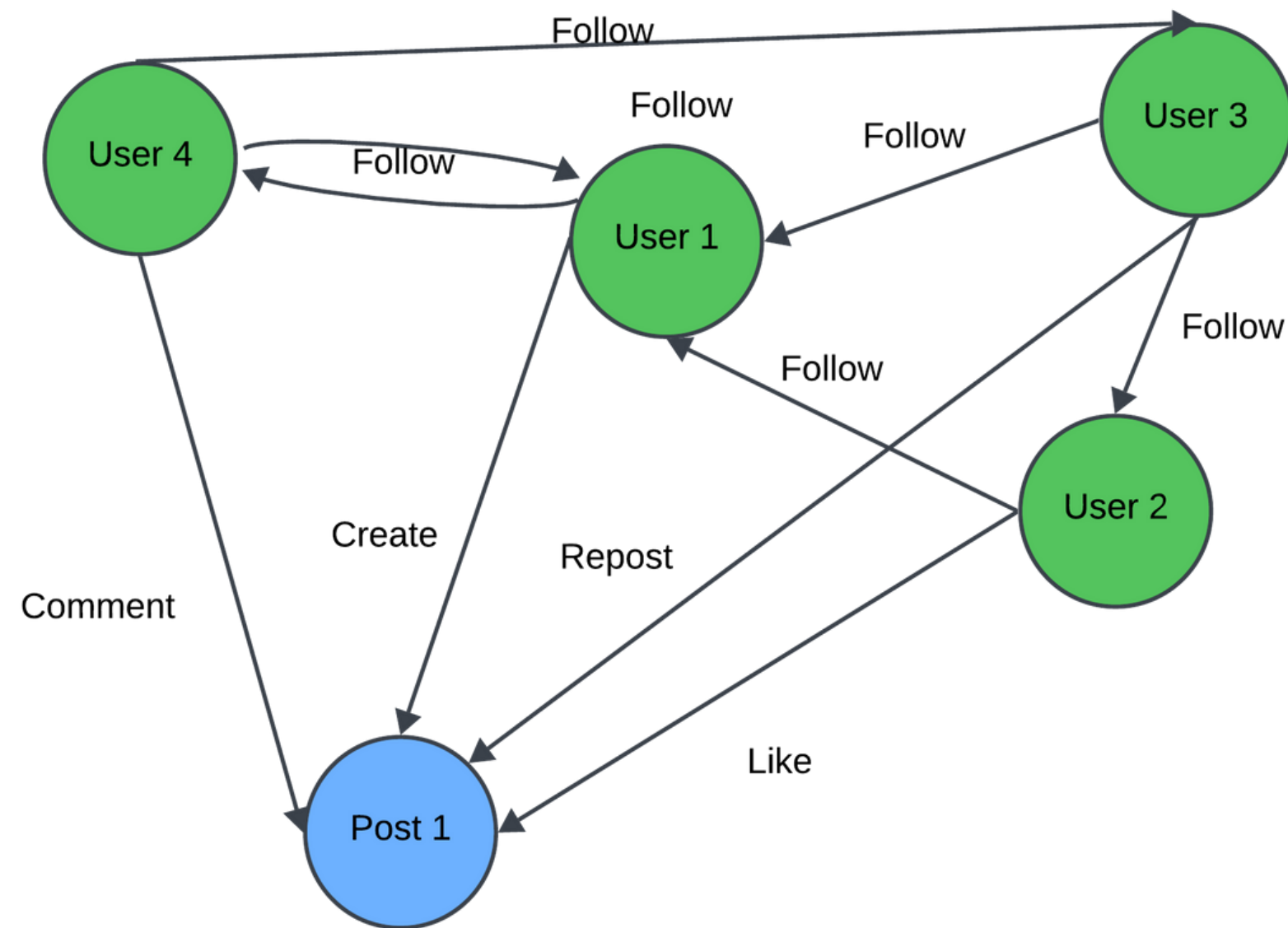
Nodes (Vertices)

- Users
 - Regular
 - Organization
 - Media
- Posts
- Communities
- Hashtags
- Topics
- Lists
- Spaces

Edges

- Connection
 - Follow
 - Join
- Activities
 - Post
 - Repost
 - Quote
 - Like
 - Reply
 - Mention
 - Hashtags

Interested Nodes and Edges in X



Nodes and Edges Matrix

	To Users	To Posts
From Users	Follow	Create, Like, Repost, Comment
From Posts	Has No Relationship on Users	No Relationship

Data Synthesis Framework

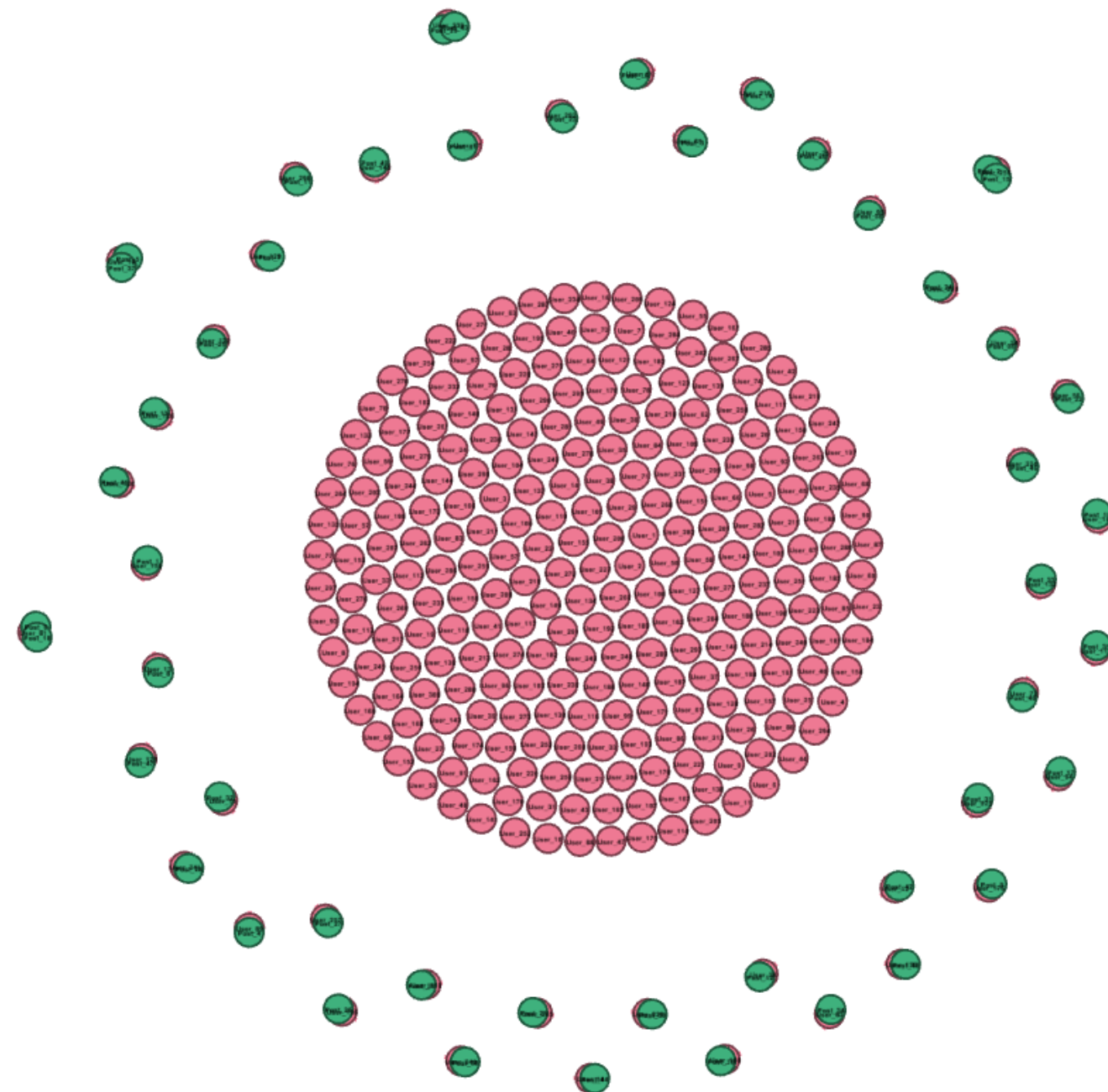
- User to Post

- 50 Creators
 - A user can create up to 2 posts
- 8,234 Likes
 - Free-for-All
 - Distributed randomly (7 Peaks)
- 2,251 Shares
 - Poster can't share
 - Distributed similarly to Like
- 1,791 Comments
 - Free-for-All
 - Distributed randomly

- User to User

- 30,180 Follows
 - Can't follow self
 - Distributed randomly
 - Can be bidirected

Visualization - Create Layer

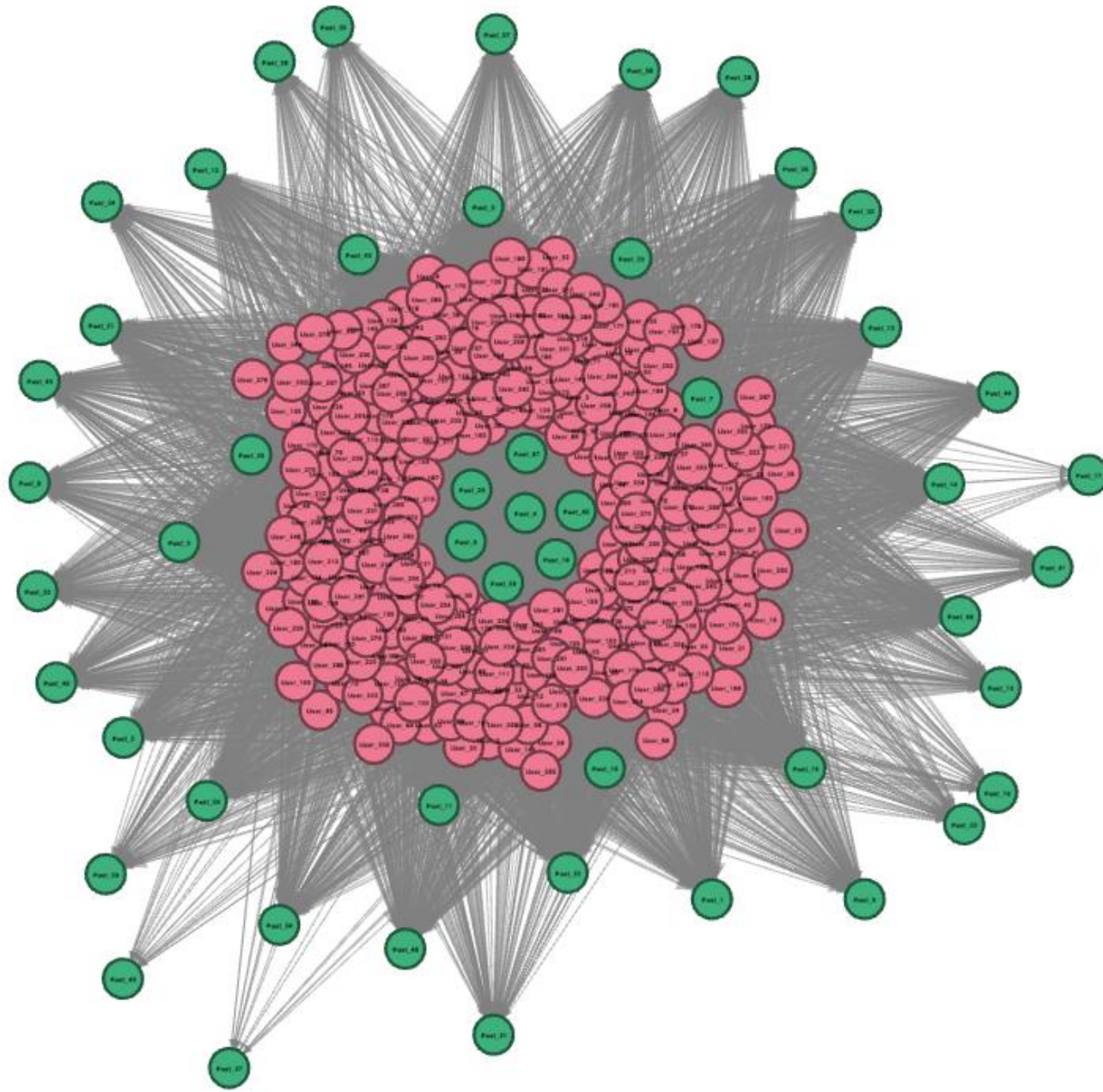


50 Edges

Quick Insights:

- 2 Posts: 4 Users
- 1 Post: 42 Users
- 0 Post: 250 Users

Visualization - Like

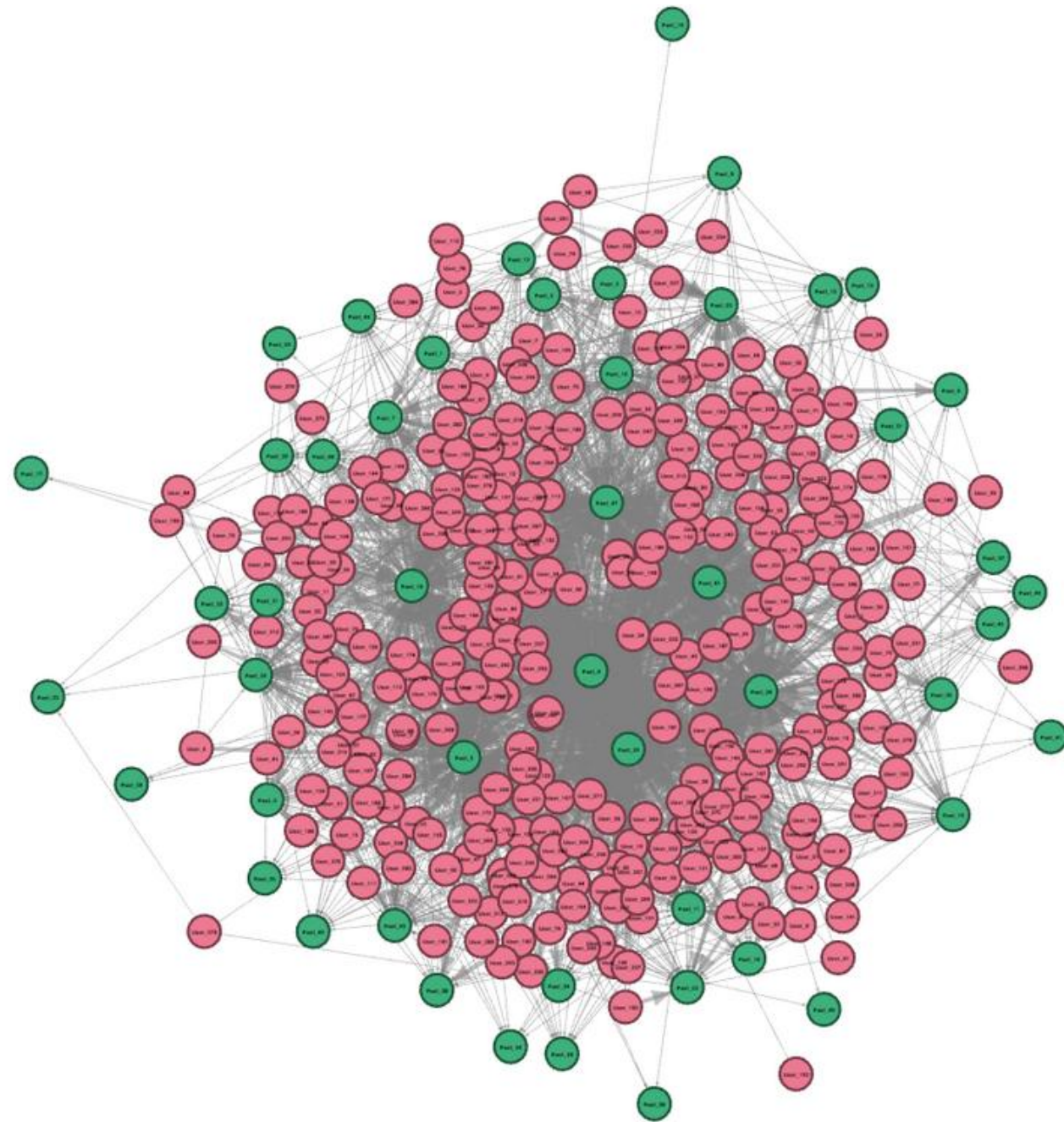


8,234 Edges

Quick Insights:

- 7 Most-liked Posts

Visualization - Share Layer

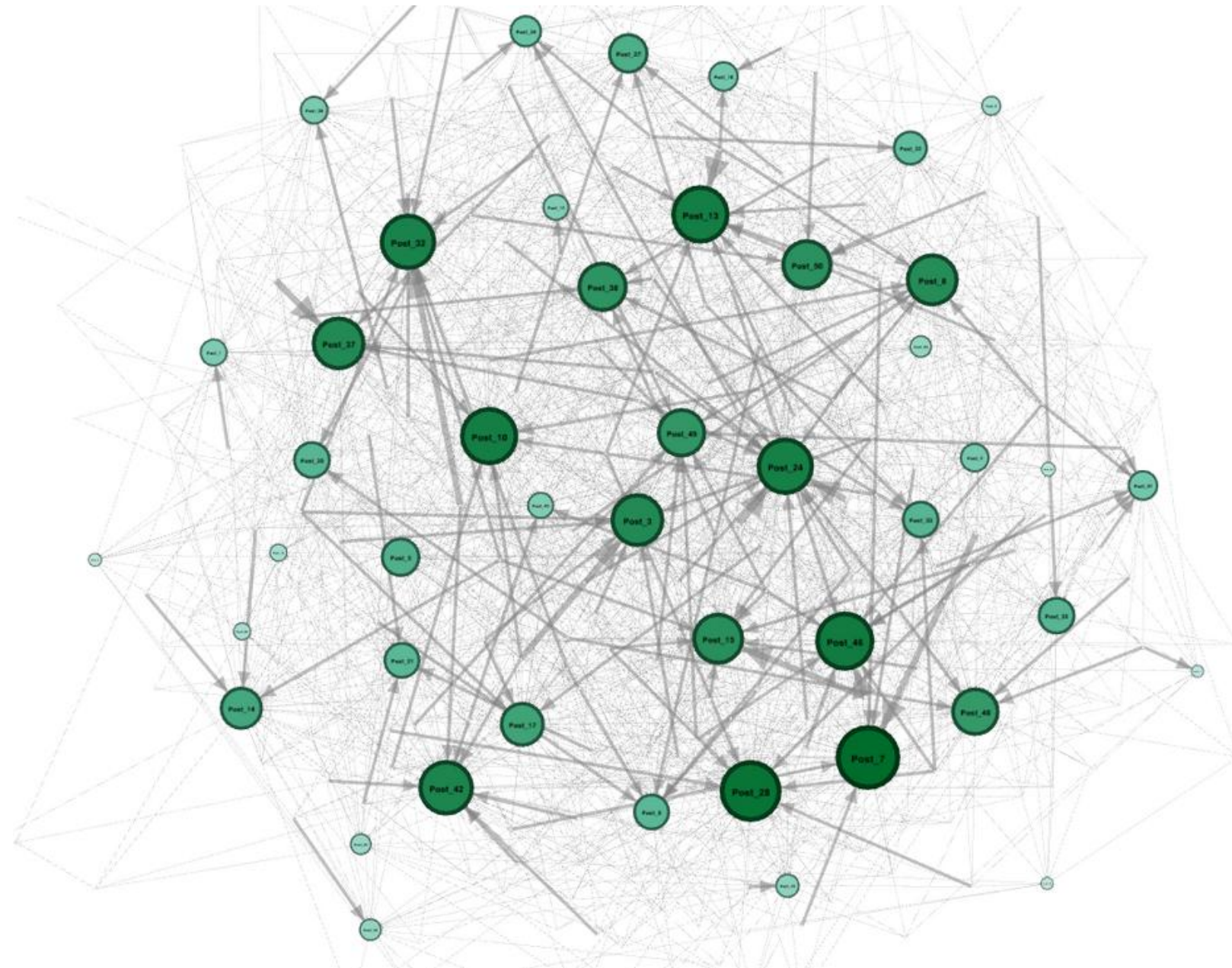


2,251 Edges

Quick Insights:

- 7 Most-Shared Posts

Visualization - Comment Layer

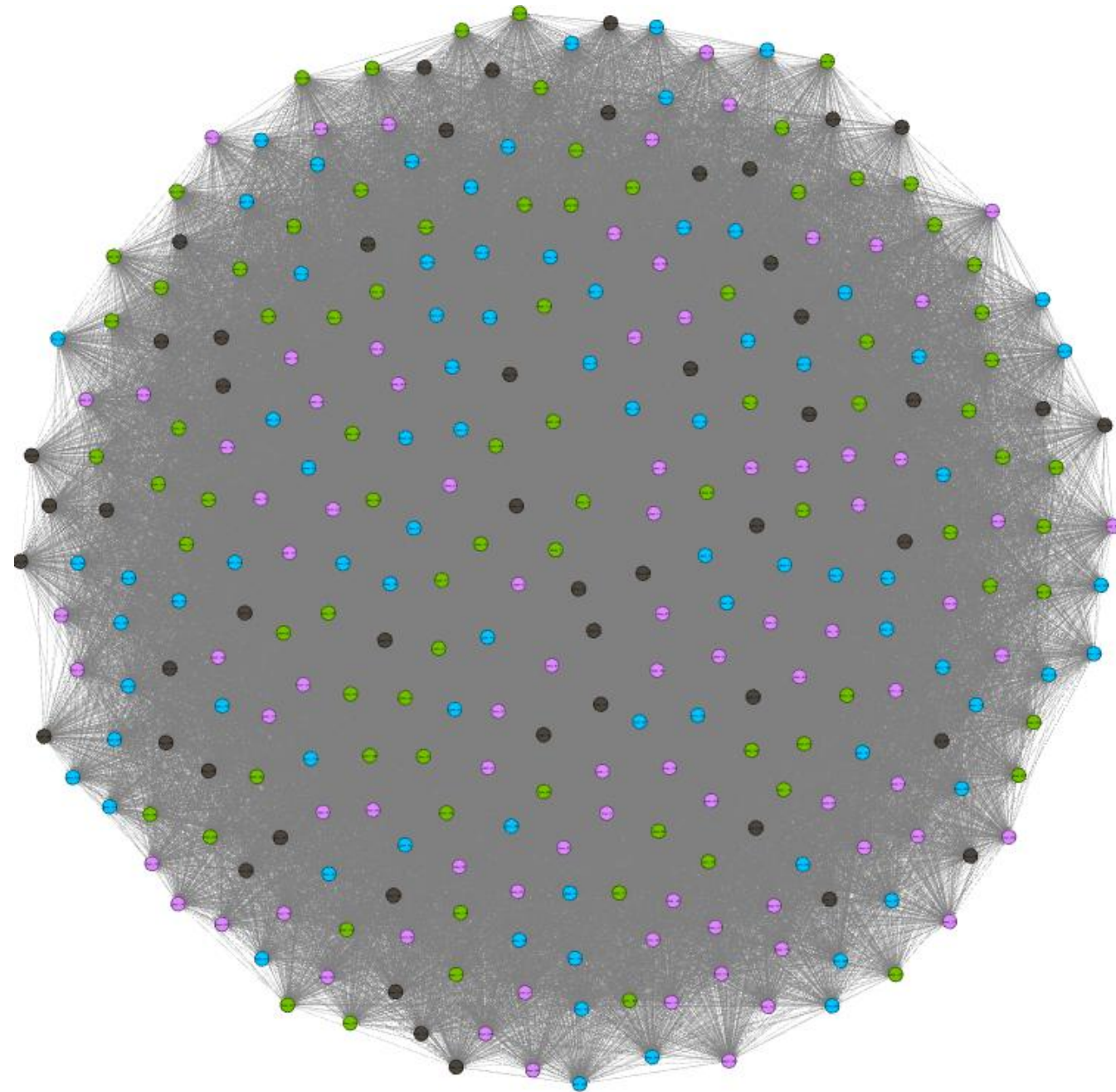


1,791 Edges

Quick Insights:

- Comments didn't follow Like and share

Visualization – Follow Layer



30,180 Edges

Quick Insights:

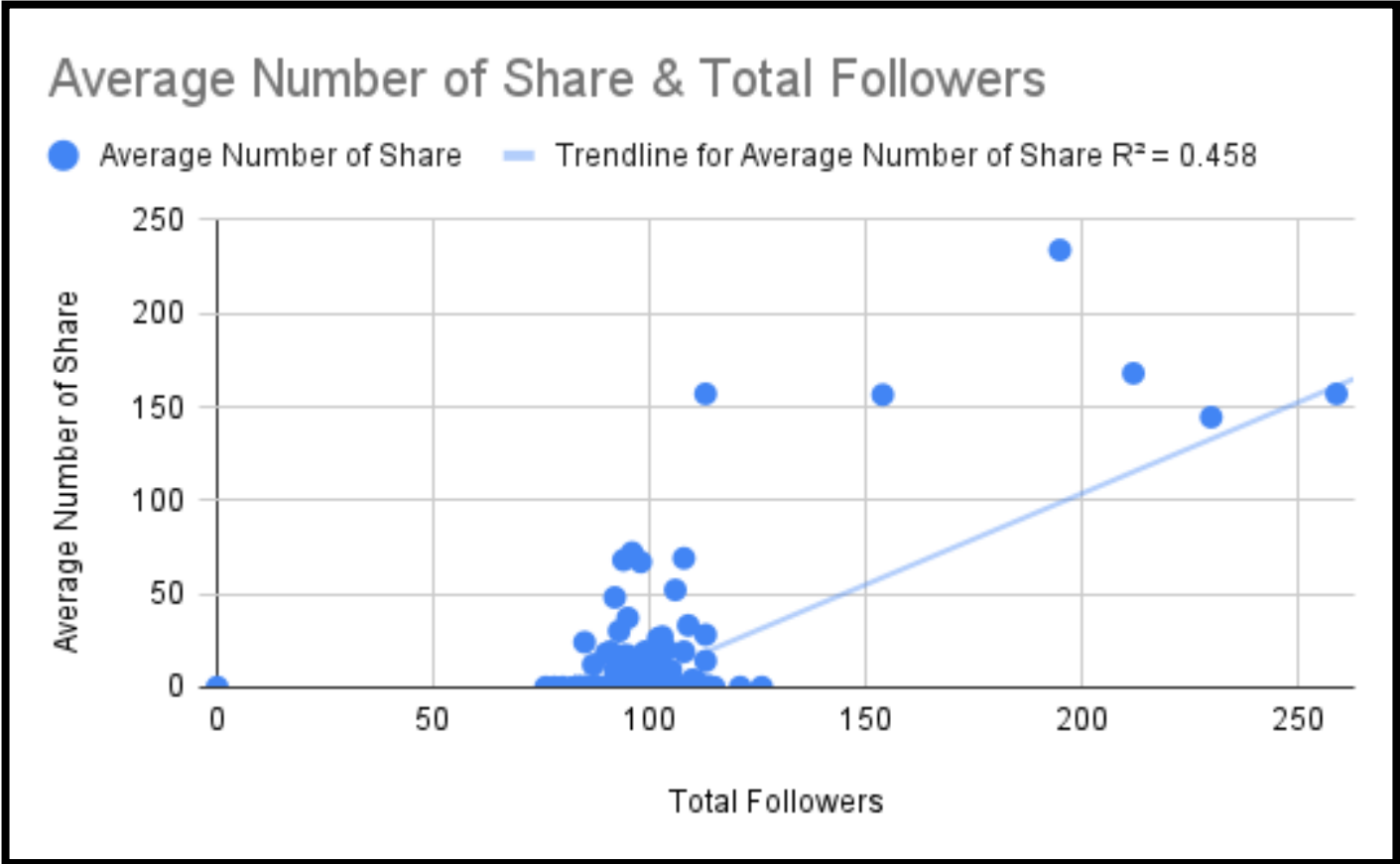
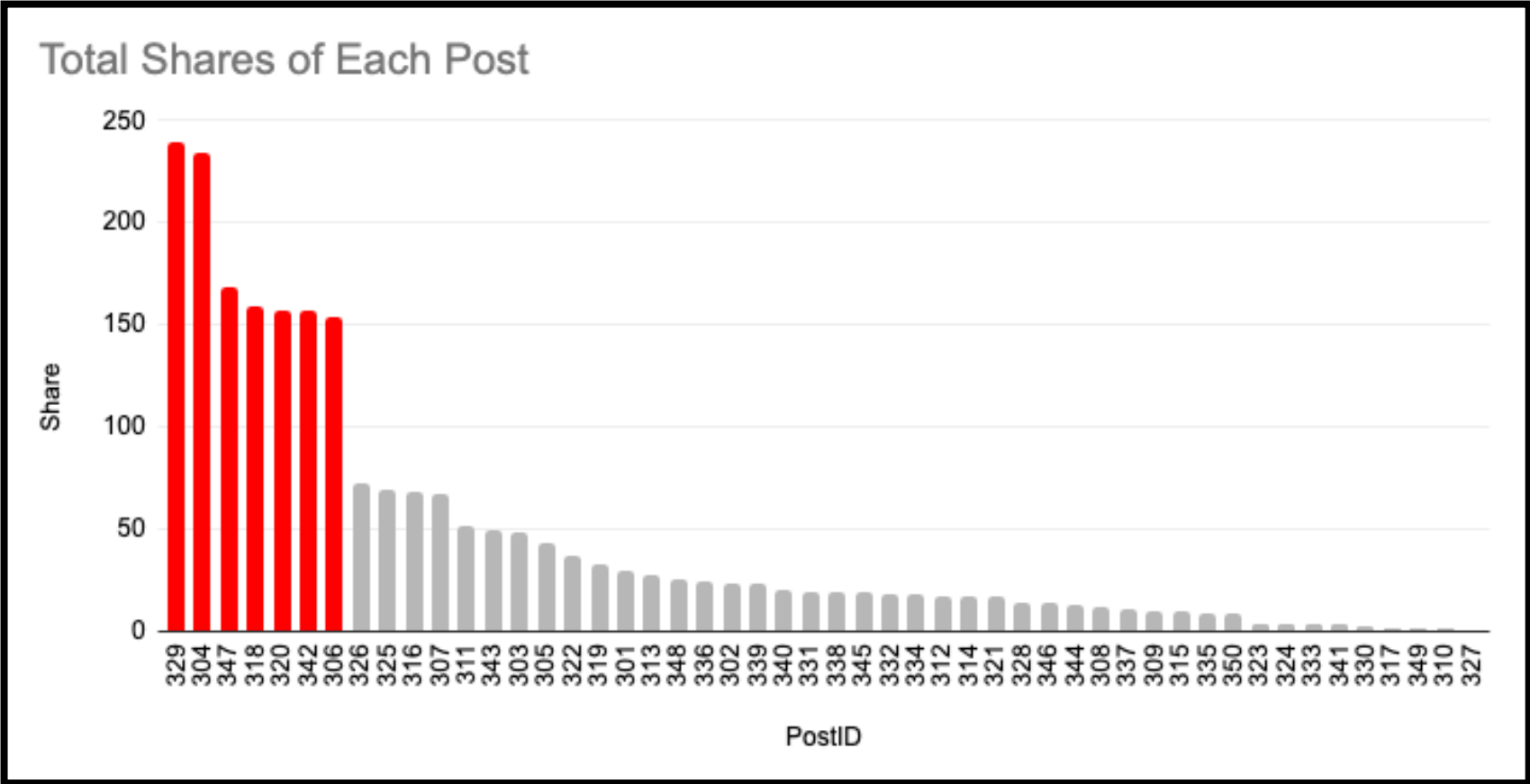
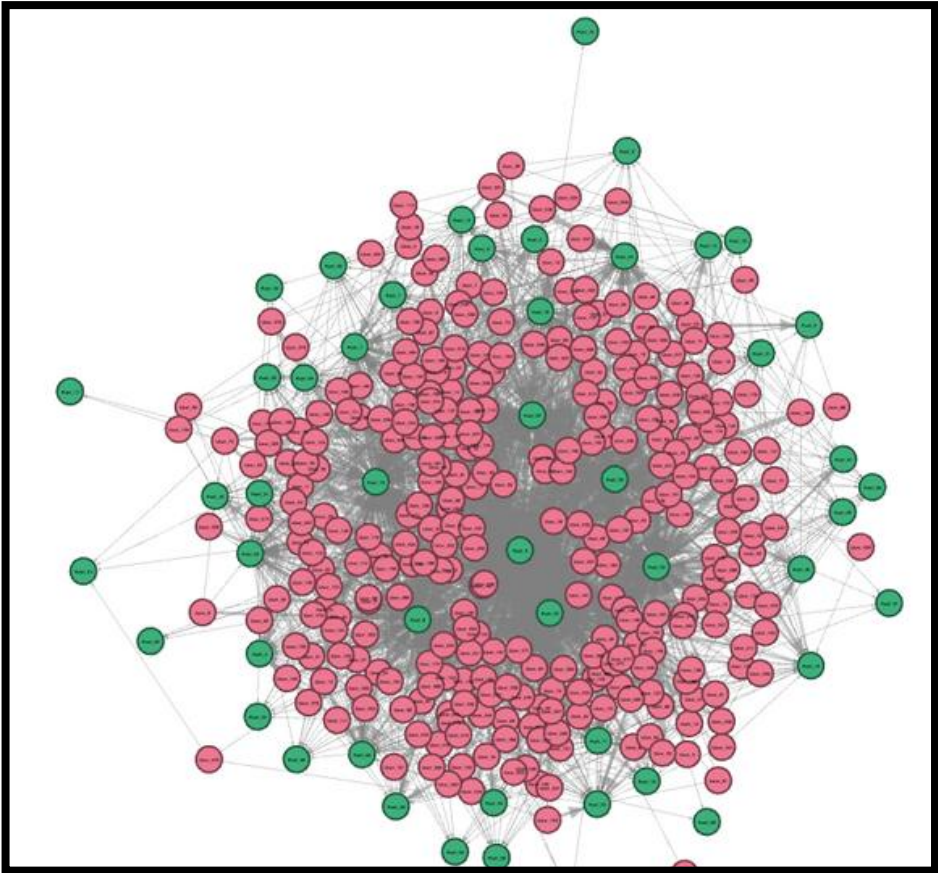
- Modularity Class
 - 0: 24%
 - 1: 23%
 - 2: 24%
 - 3: 15%
 - Others: 14%

Analytical Methods and Scopes

- **Social Network Analysis (SNA)**
 - Influencer and User Role Analysis
 - Info-spread Analysis
- **Topic Modeling**
 - Sentiment and Topic Analysis

Results

Hypothesis 1: Influencers with a large follower base typically generate higher levels of content sharing.



Identify Key Influencers

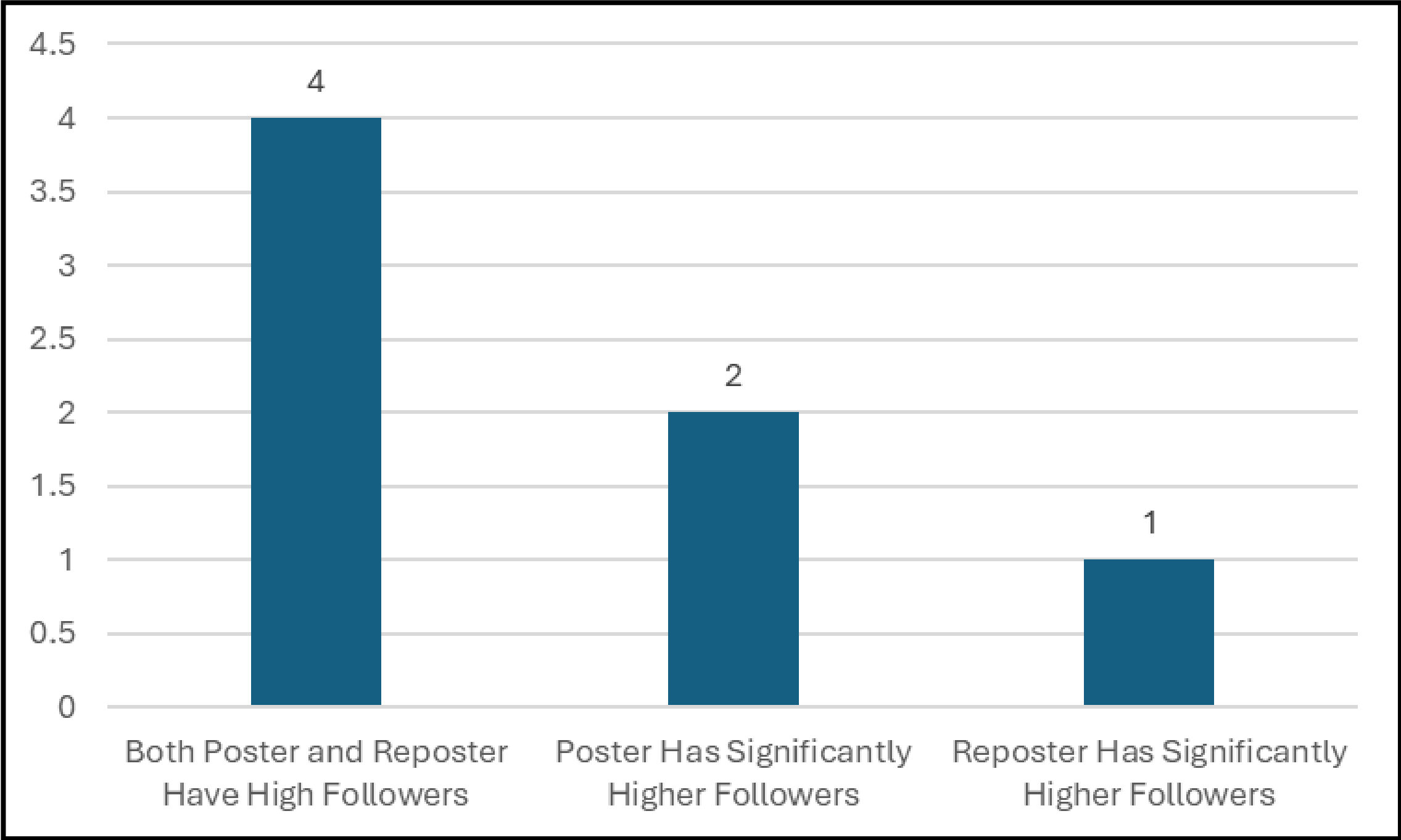
Criteria

- A high number of followers
 - Follower equal or more than 195 (Percentile 99)
- A high average number of shares per post
 - Average of share equal or more than 157 (Percentile 99)

User ID	Number of Follower	Average No. of Share	Follower Segmentation	Share Segmentation
12	259	157	High	High
125	212	168	High	High
89	195	234	High	High

Hypothesis 2: When an influencer is the poster, the post will get higher number of shares.

Rank	User ID	Post ID	No. of Share
1	239	329	239
2	89	304	234
3	125	347	168
4	81	318	159
5	75	320	157
6	12	342	157
7	81	306	154

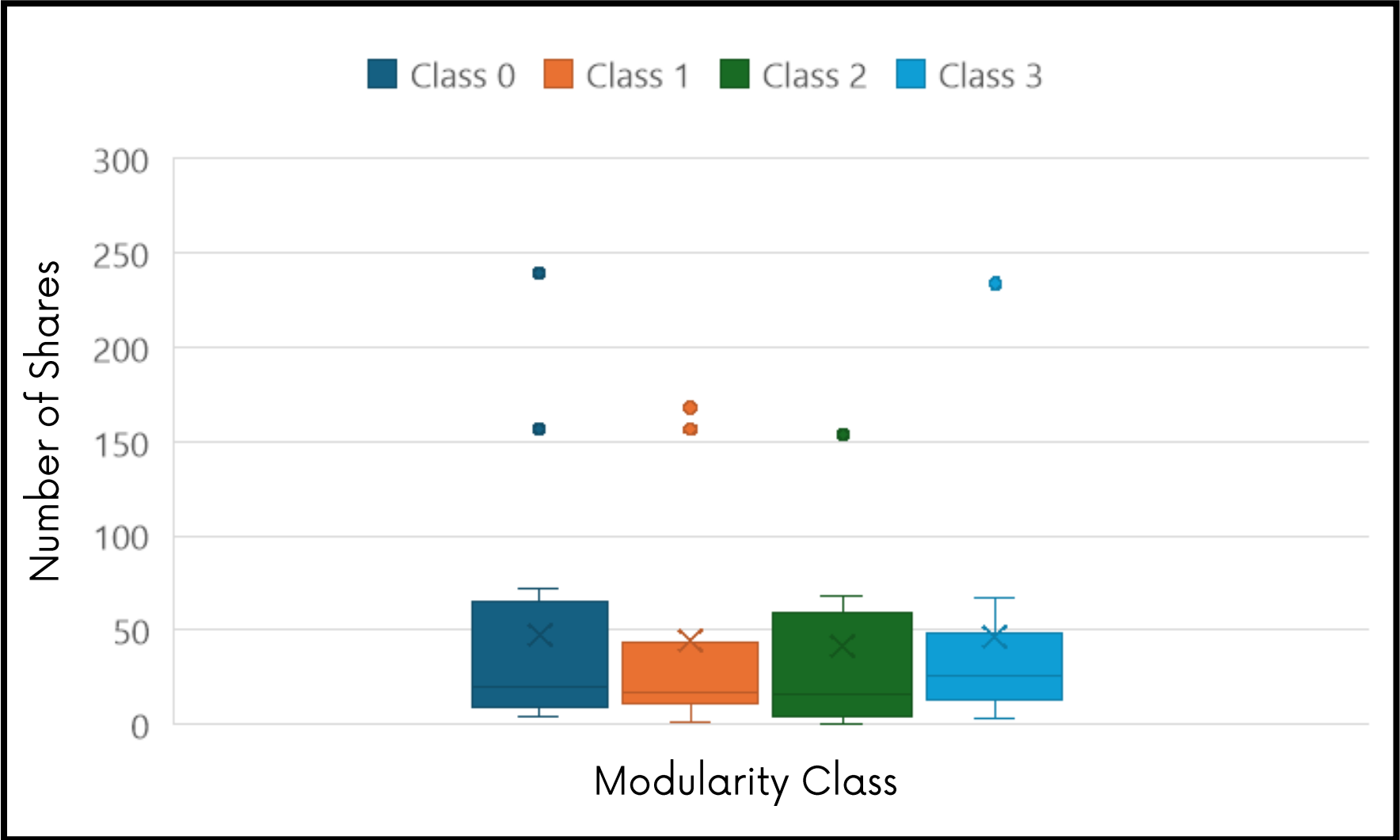


Hypothesis 3: Posts from individuals with a higher modularity class are likely to have higher engagement score.

User ID	Modularity Class	Engagement Score
239	0	1,328
89	3	1,293
12	0	1,105
81	2	1,030
81	2	1,029
75	1	1,027
125	1	1,011

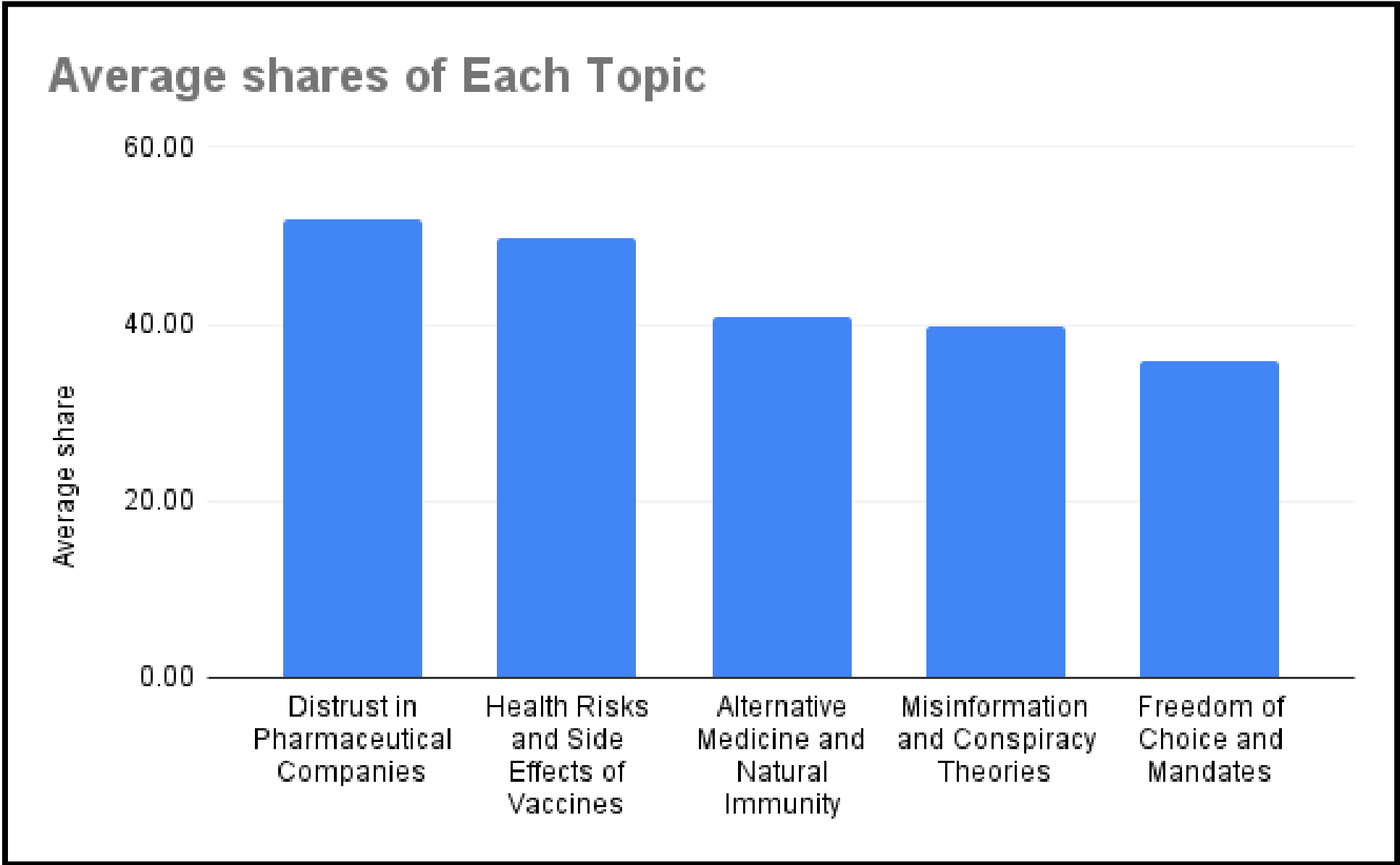
Engagement Weight Criteria

- Like = 1
- Comment = 3
- Share = 4



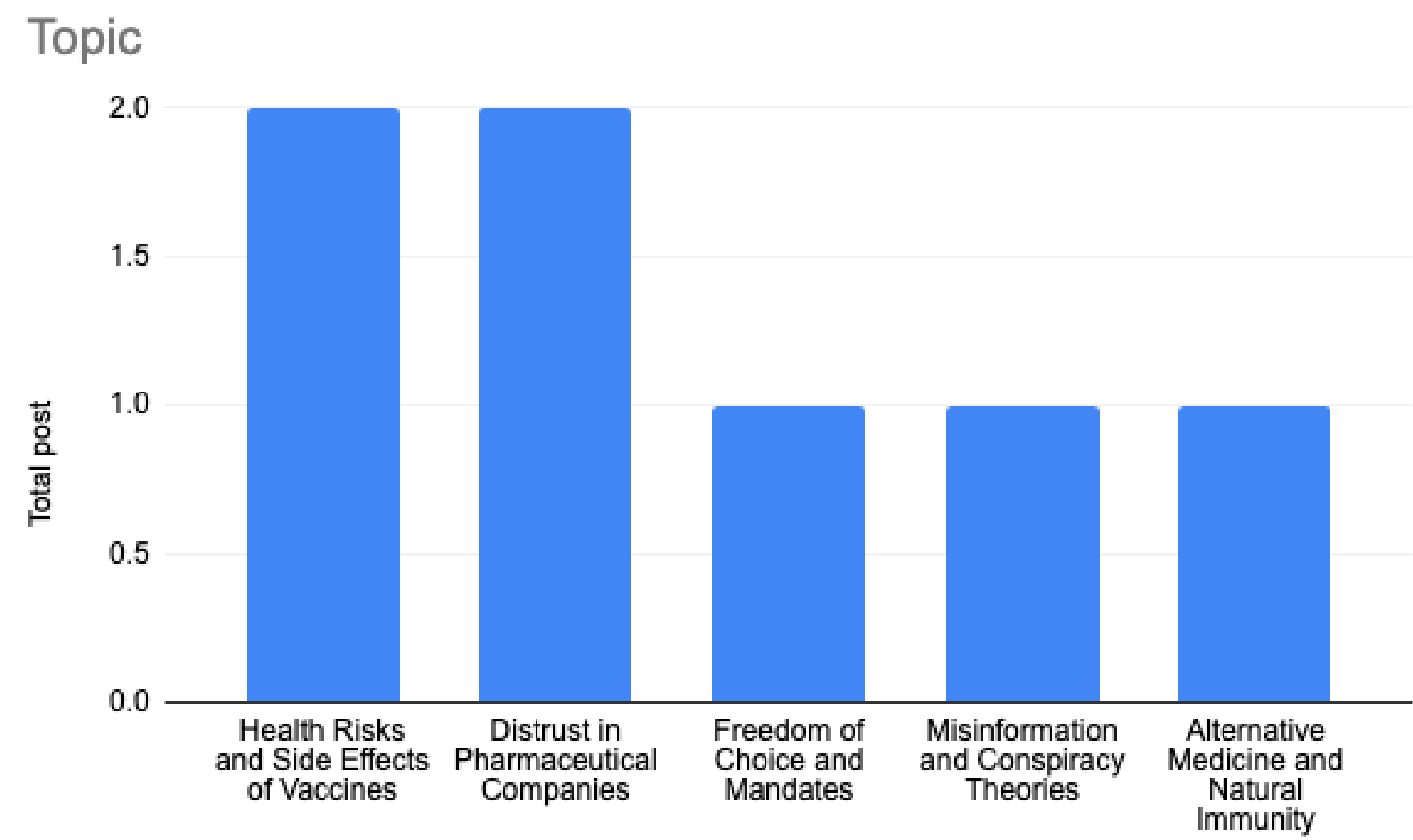
Topic Modeling : Average Number of Shares for Each Topic

Topic	No. of Post	Average No. of Share
Distrust in Pharmaceutical Companies	12	51.83
Health Risks and Side Effects of Vaccines	15	49.73
Freedom of Choice and Mandates	10	40.86
Misinformation and Conspiracy Theories	6	39.67
Alternative Medicine and Natural Immunity	7	35.90



Topic Modeling : Top 7 Posts with the Most Shares

Rank	Number of Shares	Post ID	User ID	Topic
1	239	329	239	Freedom of Choice and Mandates
2	234	304	89	Health Risks and Side Effects of Vaccines
3	168	347	125	Misinformation and Conspiracy Theories
4	159	318	81	Distrust in Pharmaceutical Companies
5	157	320	75	Distrust in Pharmaceutical Companies
6	157	342	12	Alternative Medicine and Natural Immunity
7	154	306	81	Health Risks and Side Effects of Vaccines



Interventions

Intervention: Creator has significant fewer follower numbers

User ID	Edges	Post ID	Status
81	154	318	Creator
239	230		Sharer
125	212		Sharer

Indication:

- The post gained traction due to its content, rather than the poster's influence.
- Sharers play a significant role

Interventions:

- Content Monitoring and Early Detection:
 - Tracking and flag anti-vax content that begins gaining traction
 - Provide real-time corrections
- Sharer-Centric Counter-Messaging:
 - Addressing common anti-vax myths
 - Spread accurate content debunking anti-vax posts
- Behavioral Nudge to Sharers:
 - Incentivize users to share verified, and pro-vaccination content

Intervention: Creator has significant higher follower numbers

User ID	Edges	Post ID	Status
12	259	342	Creator
256	126		Sharer
194	114		Sharer

Indication:

- The post's virality is driven primarily by the poster's influence
- The influencer's authority gives credibility to the anti-vax related topics

Interventions:

- Influencer Engagement and Collaboration:
 - Approach these influential creators with evidence
 - Counteract misinformation from anti-vax influencers
- Deprioritizing Harmful Content:
 - Collaborate with the platform to deprioritize anti-vax content
 - Restrict for influential accounts spreading misinformation

Intervention: Creator and sharers have similar follower numbers

User ID	Edges	Post ID	Status
239	230	329	Creator
125	212		Sharer
89	195		Sharer

Indication:

- The post may reflect an organic spread from key influencers networks.

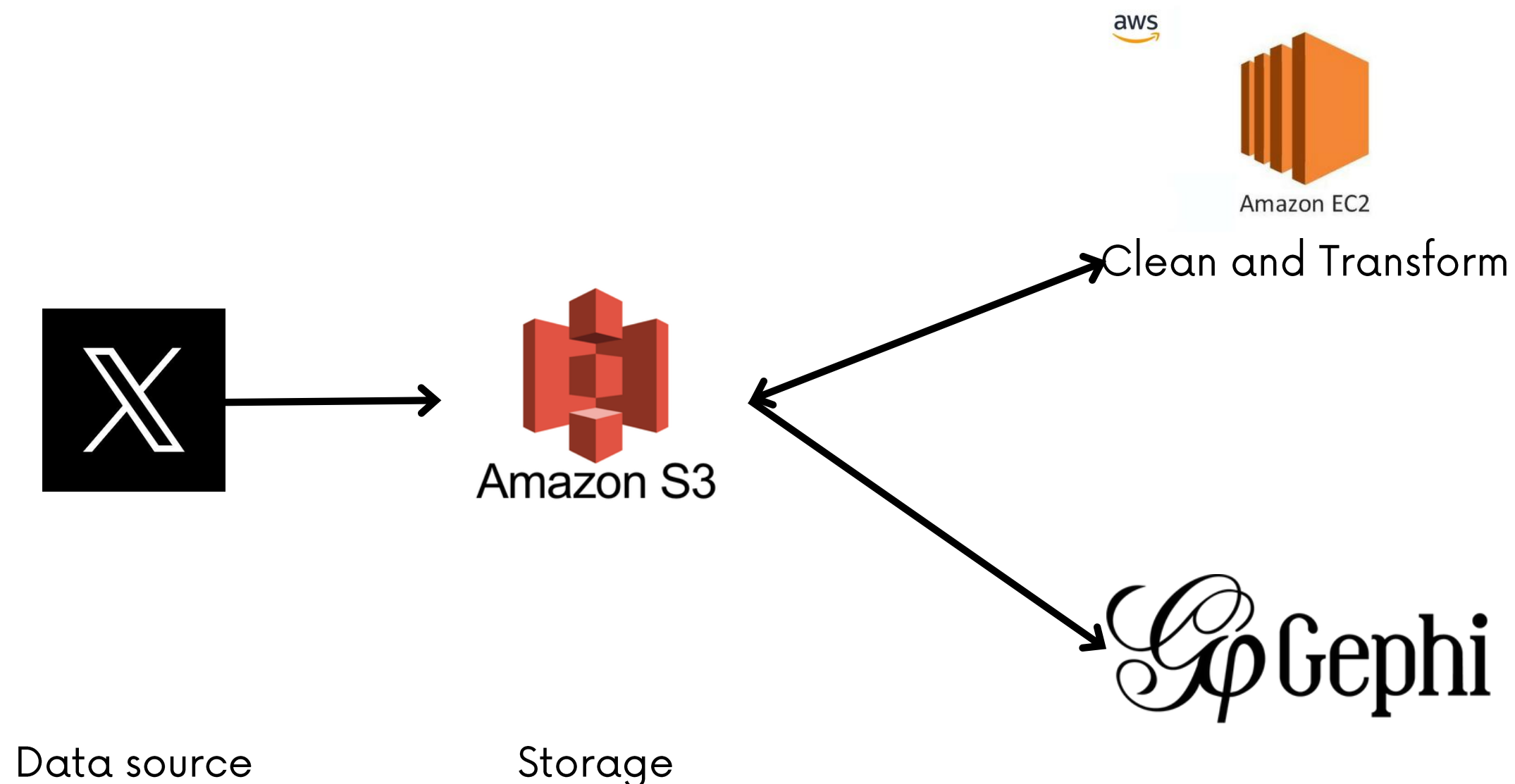
Interventions:

- Leverage Micro-Influencers:
 - Have high engagement rates and trusted relationships with their followers
 - Partner with micro-influencers
 - Equip them with ready-to-share
- Viral Challenges and User-Generated Content:
 - Encourage mass participation, creating a snowball effect
 - Leverage the creativity of users makes the campaign feel authentic
 - Launch a hashtag campaign e.g., #VaxForLife
- Gamified Online Engagement:
 - Tap into the competitive and reward-seeking behavior of users
 - Interactive tools make education fun and viral.
 - Create short quizzes on vaccine myths and facts
 - Partner with platforms to gamify to offer digital rewards.

Future work

Limitation Discussion

- Mock-Up Data:
 - 300 accounts followed almost every account
- On-Premise:
 - Difficult to process streaming data with large scale



Thank you