

Looking Behind the Trucks – Conversation Patterns and the Implications of Curation in Twitter



1. Introduction

Team Convoy's experience of hacking the Freedom Convoy data sets led us to consider Twitter curation algorithms. We needed to learn more about the biases in this process.

The research found evidence of bias in three areas:

- ❖ Recency Bias
- ❖ Popularity Bias
- ❖ Exposure Bias

We also found that Twitter's curation algorithms can distort the nature and flow of the conversations. We concluded "It was no longer about the trucks." The idea of algorithm curation by Twitter resonated with our experiences of the data and helped us formulate the research questions.

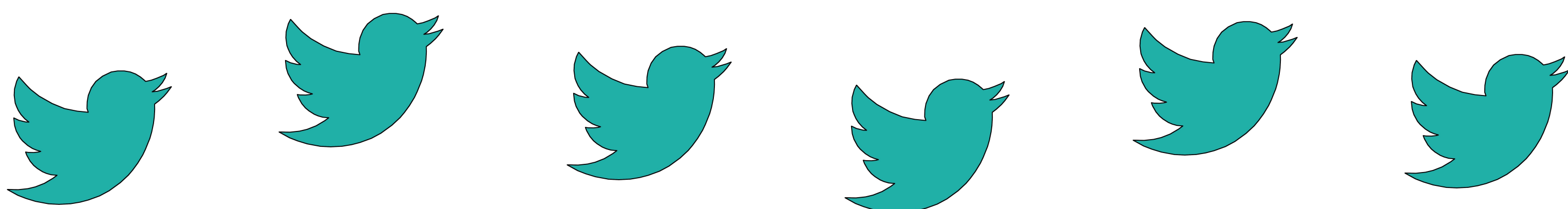
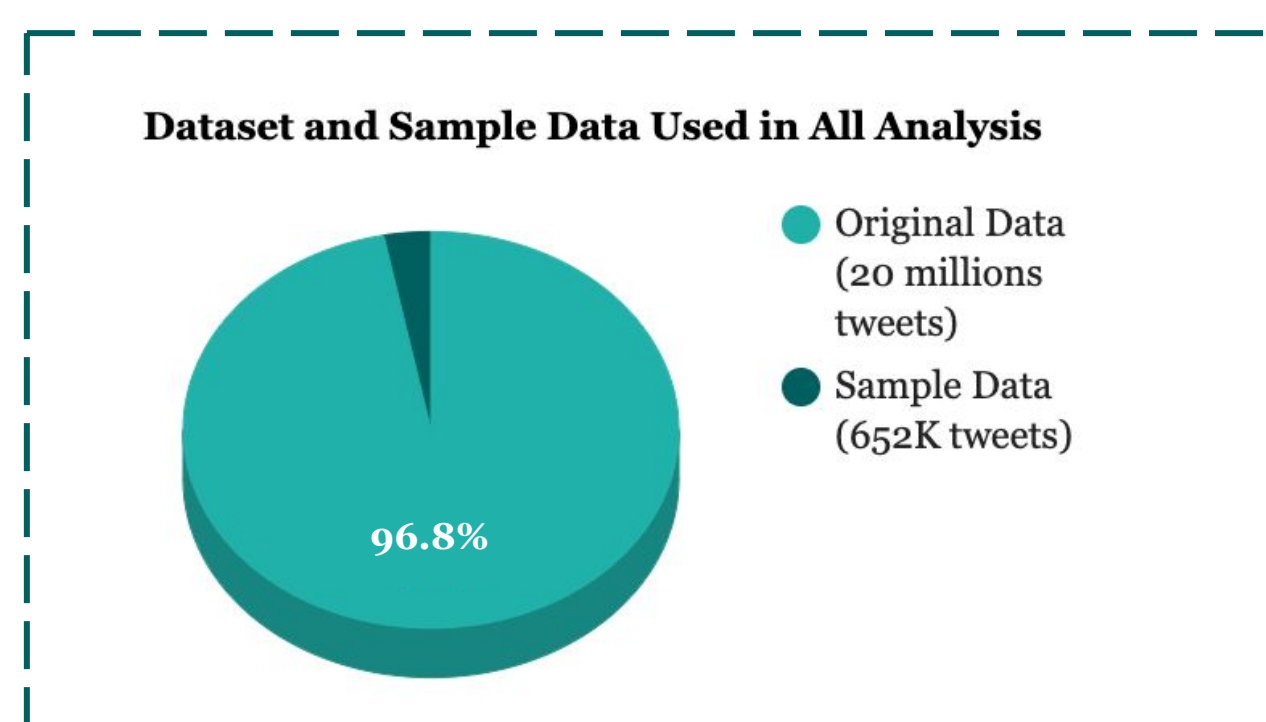
2. Research Questions

Main – How were Twitter conversations around Convoy initiated, developed, and manipulated?

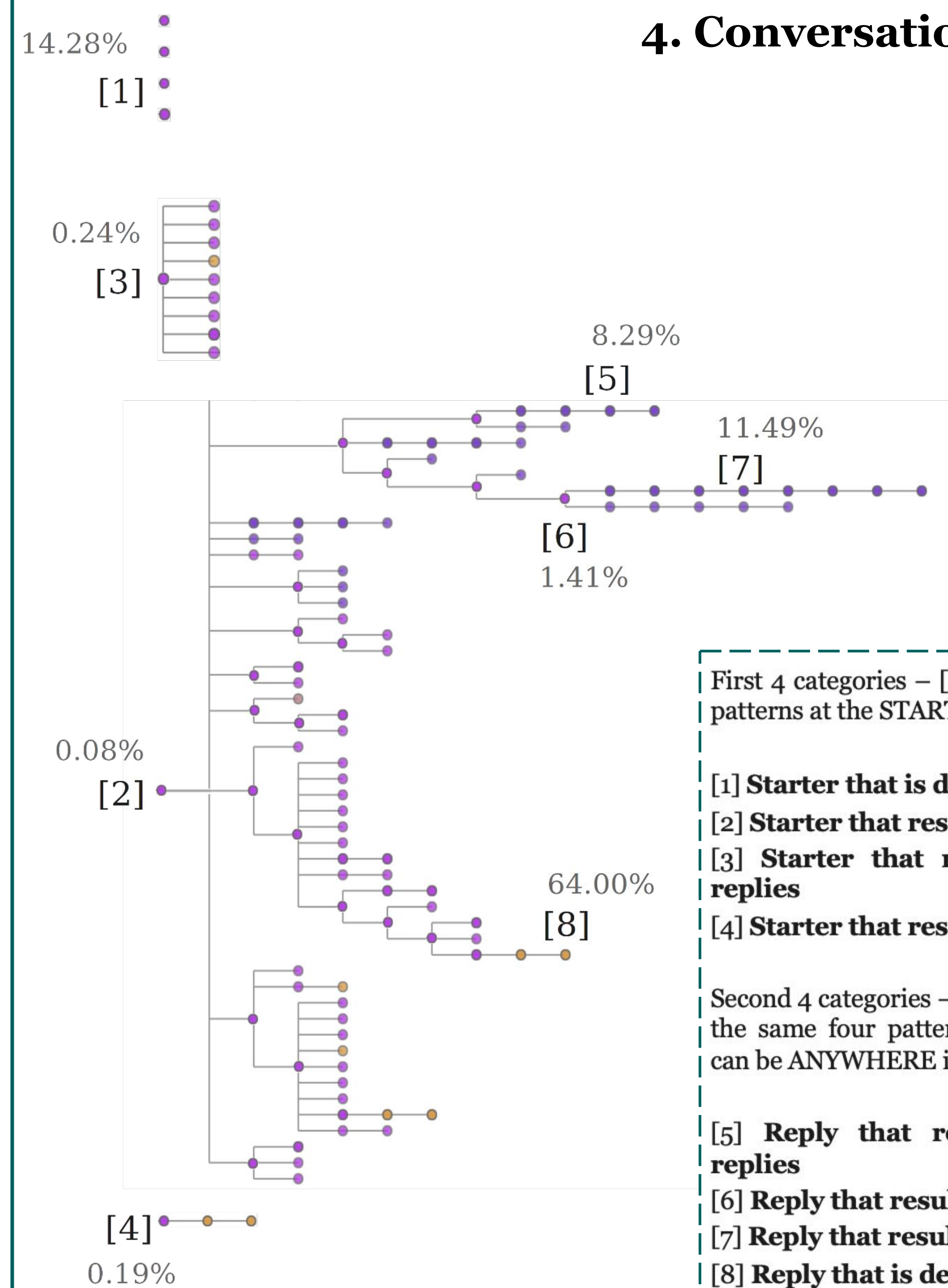
- ❖ Can we identify conversation patterns in the Canadian freedom convoy discussions on Twitter?
- ❖ What is the role of the platform and its' algorithmic curation of tweets?
- ❖ What kind of manipulation can be detected?

3. Data

Twitter data are extracted using strings related to the Freedom Convoy, e.g. #TruckersForFreedom, #GoHomeTruckers, #NoVaccineMandate, #OttawaSiege, #CanadaTruckers, #OttawaOccupiers, #freedomconvoy22, #TrudeauTyranny, #ResignTrudeau, Canadian Trucking Alliance, PM Trudeau, etc.



4. Conversation Patterns



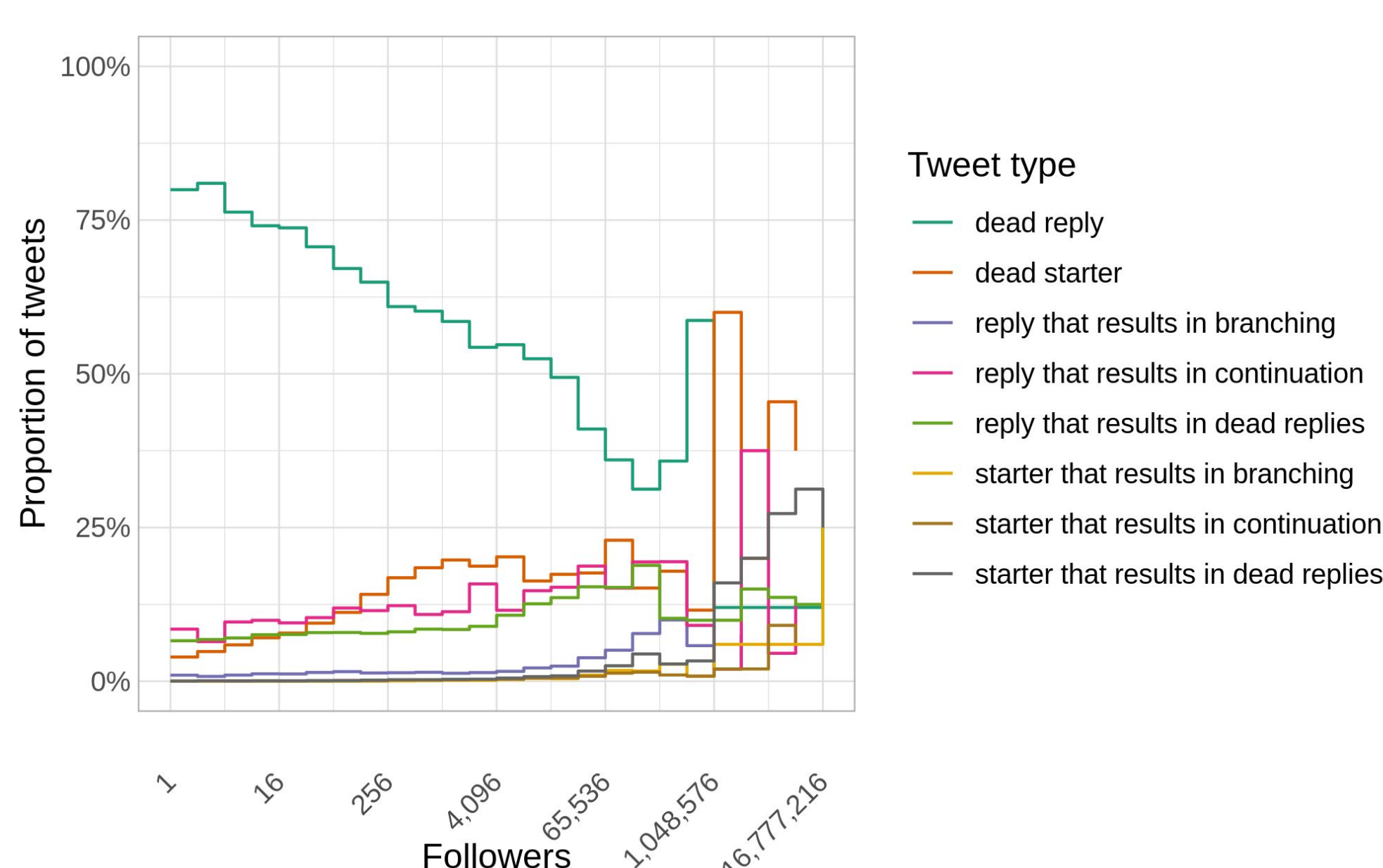
First 4 categories – [1], [2], [3], [4] – contain patterns at the START of the conversation:

- [1] Starter that is dead
- [2] Starter that results in branching
- [3] Starter that results in only dead replies
- [4] Starter that results in continuation

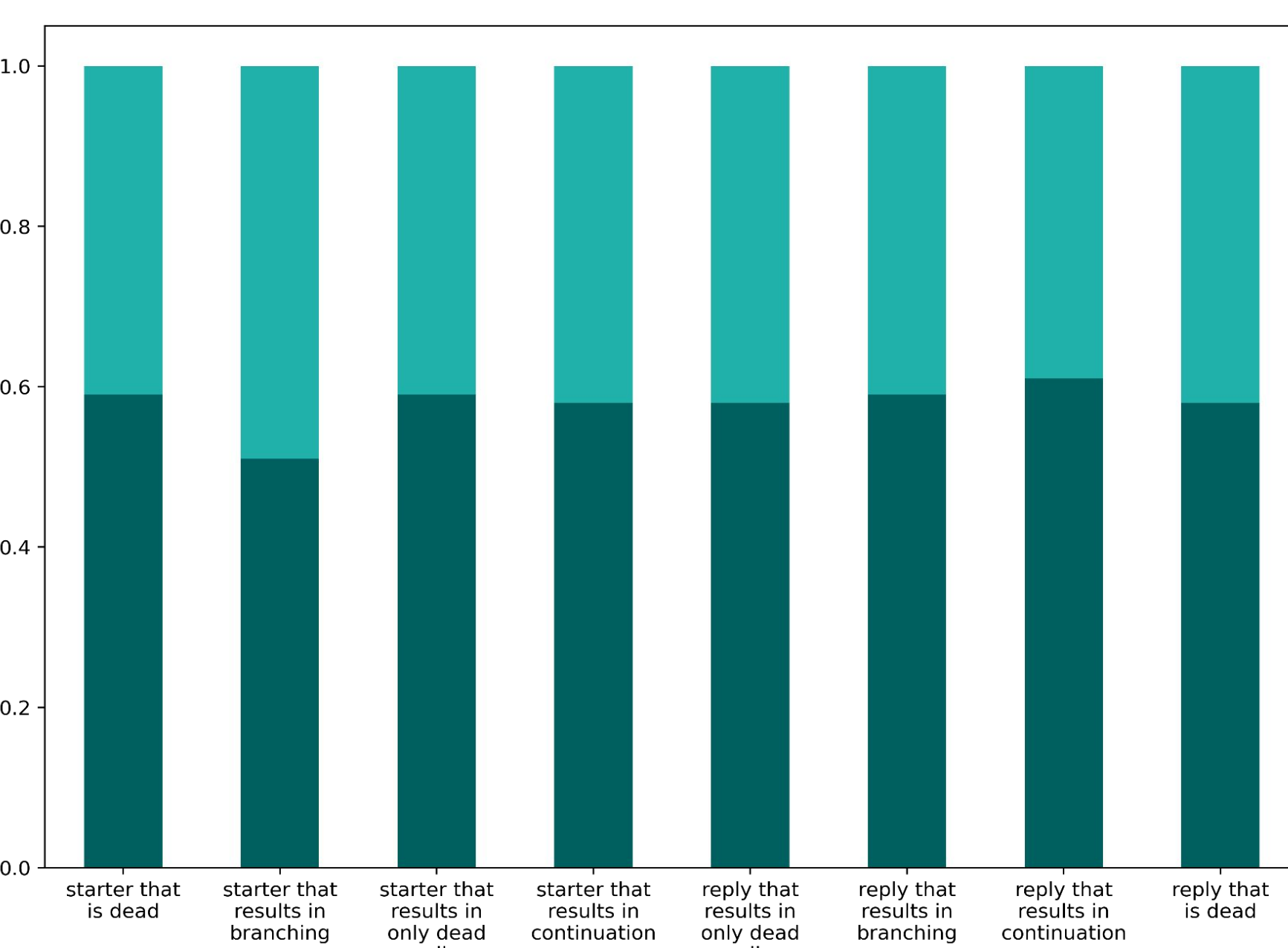
Second 4 categories – [5], [6], [7], [8] – mimic the same four patterns, but tweet's position can be ANYWHERE in the conversation:

- [5] Reply that results in only dead replies
- [6] Reply that results in branching
- [7] Reply that results in continuation
- [8] Reply that is dead

6. User Analysis

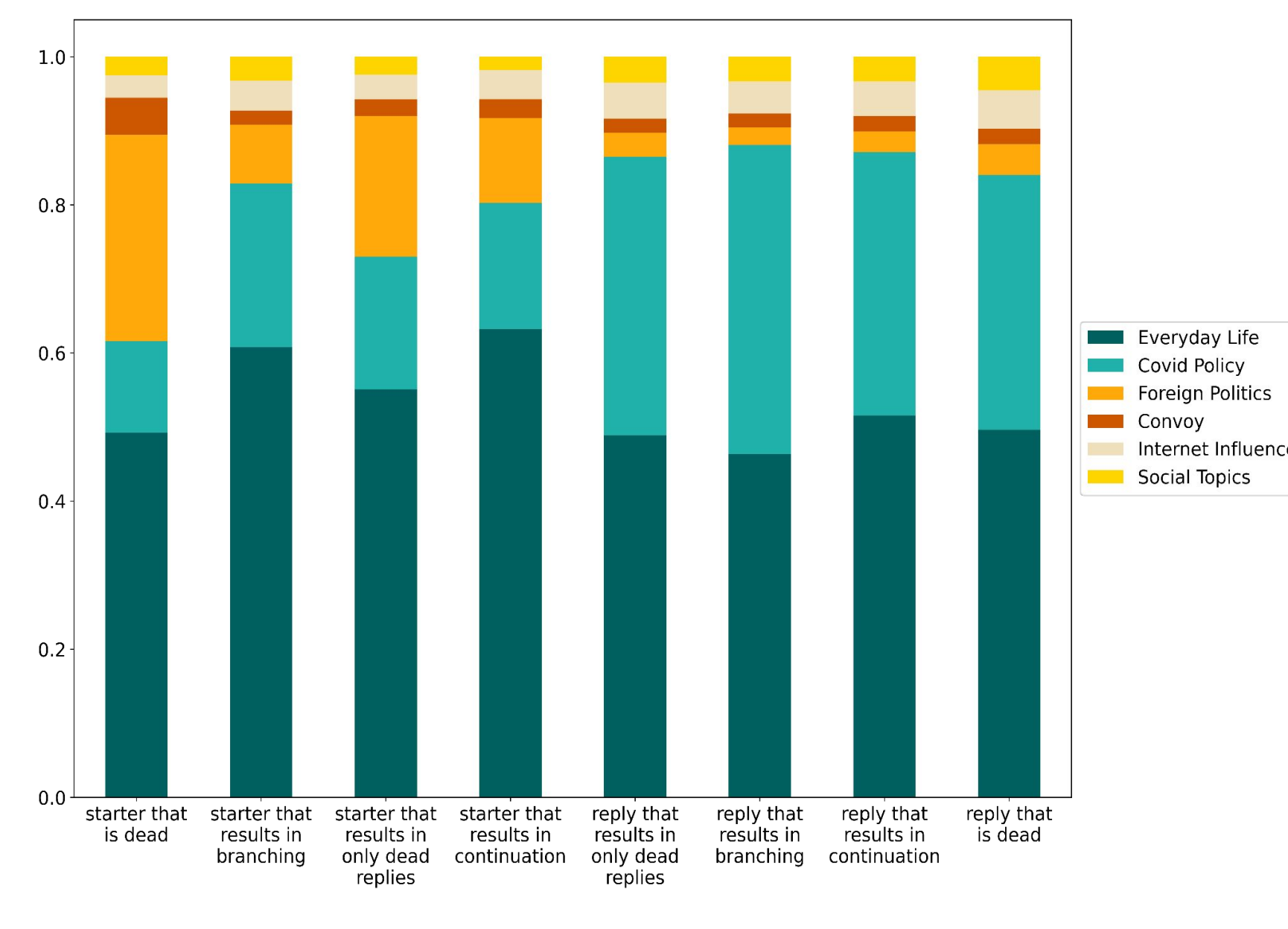


Tweet category distribution analysis across different users groups with numbers of followers.



Bot and non-bot distribution analysis across different Tweet categories.

5. Text Analysis



Top 20 - the users who have received the most replies which have not lead to further discussions

- ❖ **Publicists and Professionals** (8) Professional people with some celebrity status promoting their own brands as sole traders in the online world
- ❖ **Media Organisations & Journalists** (2) They cover issues as expected in mainstream media: Politics, Crimes against humanity, Anti Trump, Support for Ukraine Charity
- ❖ **Professional Bloggers** (4) US/Australia Bloggers who promote diverse views from Republican Conservative to Futurist perspectives, with diverse interests from 'dark polls' to Stock Markey trades!
- ❖ **Ordinary Users** (6) Retirees, Housewives, Dog Lovers with their own interests and opinions on day-to-day affairs. Two are Canadian men. This is the only visible evidence of Convoy. Both do not support the Truckers and reveal that without vitriol. But one is very measured, uses retweets and makes no overt reference to the Truckers or downtown activities



From the topic modeling we can clearly see that the most prevalent topics in the twitter conversations for that time were everyday life and Covid policies. The conversations revolved around Covid and the pandemic appear to have attracted more interaction and generated more conversations.

7. Conclusion and Future Steps

We identified 8 tweet structures to analyze twitter discussions. These structures show some variety in the conversation starting topics, although on average, the average sentiment and lexical diversity show little difference between the categories, likewise the difference in the bot distributions of users is minimal.

The findings call for further investigation by combining different user and tweet characteristics to understand who starts successful discussions and how do the discussions change and develop through time. Twitter conversations are dependent of how the platform is built which makes it more important for researchers to describe what users cannot see.