

Project 3 Q&A

DSC680

Matthew Fikes

1. If disinformation groups were already removed, how are there still groups to analyze?

Only the largest group was removed, the others were only quarantined. They still allow posts but do not show up in search results. Users also find a place in other communities. Since they can be created by any user, it is difficult to keep subreddits down permanently.

2. Are there other subreddits that may also warrant a closer look?

Crawling the posting activity of users who posted to disinformation subreddits did show a few other groups that may also be spreading disinformation – CoronavirusCirclejerk, ChurchofCOVID, and LockdownCriticalLeft showed up as a common link between users.

3. Do some users stand out for the frequency of posting?

While Reddit administrators suggest I leave out individual user information out of my report to respect privacy and prevent witch-hunting, there were a small handful of users that would post far more frequently than the rest. The top five most prolific users had an average of 153 posts, and the 2nd five average only 47.

4. Why look at community moderator activity?

The job of moderator is unpaid and can require a significant amount of time depending on the size of the subreddit and how many other moderators there are to handle content. It would make sense for these individuals to have a strong investment in the idea of anti-vaccination and a good place to start for looking at activity.

5. Why is a group related to the stock market associated with all of the disinformation moderators?

The “why” is still unclear but there seems to be links between conservative political subreddits, investing/stock subreddits, and antivaccine subreddits.

6. Was there anything surprising in the data?

I was surprised to find more that seemed related to disinformation that I had not heard of. I use the site regularly and had not seen their content. I would suggest that means they are filtering even non-quarantined communities in listings, but I have seen the targeted subreddits recently.

7. Why limit activity to the top 10 communities?

There are a few reasons I decided to limit this. For one, the network graph diagram was difficult to comprehend with too many nodes. It also serves as a method to find user focus. Some users may interact with dozens of communities, but if the interactions are 100:1 then chances are the community does not strongly align with the user.

8. Are there any recognizable patterns to the posting of disinformation?

At this point I have not looked into specific posting times to see if there are certain hours within the day where posting activity is higher. This could be a future project, along with collecting the content of specific posts to measure which are disinformation compared to general posts.

9. How much disinformation is spread in a day?

As I mentioned in the previous question, I did not look at specific content of posts or timings. I can say that the counts for top posts of the day have the following totals:

LockdownSkepticism = 43

TakeTheJab = 13

DebateVaccines = 31

conspiracy = 100 (maximum result of API)

10. What can be done about this problem?

This is not a simple solution. Attempts could certainly be made to scrape the content of posts and comments or flag users who post links from certain blacklisted domains. The issue here is motivating the company to do that when they have already shown reluctance to address it.