# The right opinion? Anti-vaccine conspiracies on Reddit

*Maheep Tripathi, Sanchi Singh*

## Introduction

The COVID19 pandemic has seen the rise of a parallel 'infodemic' where a slew of misinformation surrounding the origins of the virus and potential cures for it have run amok. Misinformation that has been circulated and shared primarily through social media platforms.

As protests against lockdowns, masking and other safety measures erupted across countries, more insidious perhaps has been the rise in anti-vaccine rhetoric which has the potential to undercut the efforts of governments and public health officials to curtail the spread of the virus.

In a bid to understand how conspiracy theories in anti-vaccine online forums take hold we use topic modelling to show how anti-vaccine concerns and rhetoric shifts within our chosen sample subreddit r/conspiracy as the pandemic progresses.

## Theory and Hypotheses

**H1: The pandemic has led to an increase in consumption of anti-vaccine content.**
The theory of discursive polarisation stipulates that the tendency of like-minded individuals engaged in discussion with one another to fortify their pre-existing views and to move towards more extreme points of view in the general direction in which they are already tending our first hypothesis should hold.

**H2: As time goes by anti-vaccine rhetoric and concerns shift**
The theory of motivated reasoning states that motivations are desired end-states that individuals want to achieve. So misinformation or conspiracy theories might evolve or change over time but the anti-vaccine attitude, belief, stance remains.

## Analysis

Fig.1: Weekly comment frequency

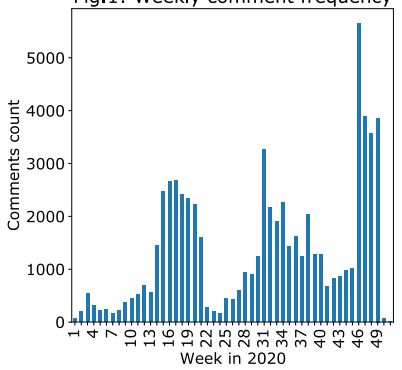


Fig.2: Weekly comment share per cluster

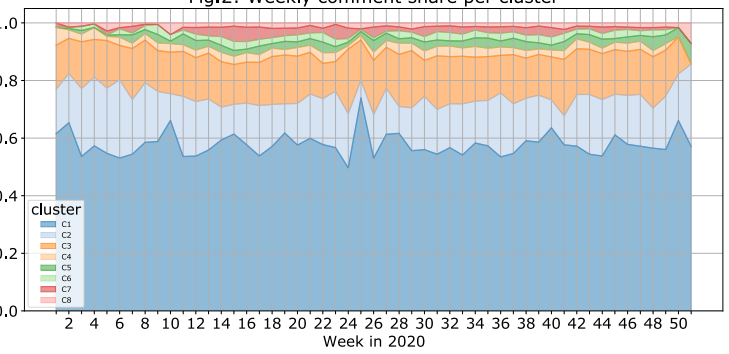


Fig 3: Mean topic loading per cluster

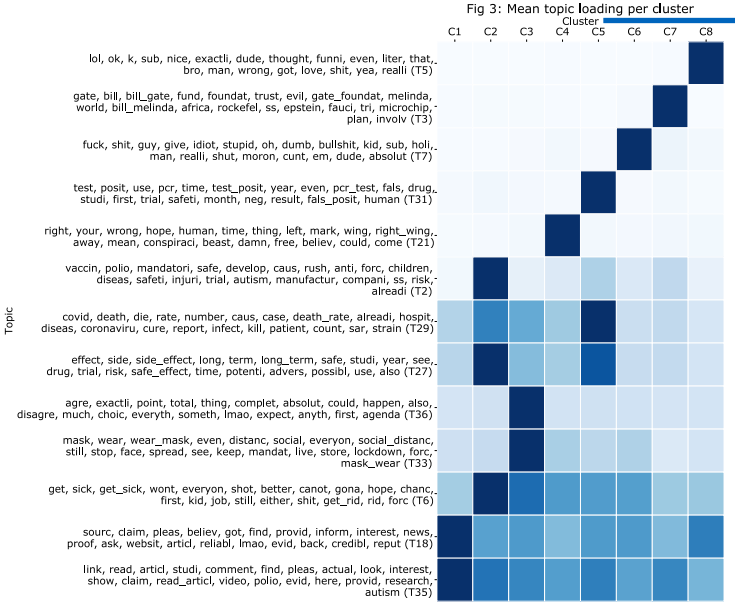


Table 1: Cluster descriptors

| # | Label | Representative Comment |
|---|---|---|
| C1 | Skeptics | I have no idea what the truth is. What I DO know is that it isn't as it has been portrayed by the media. I'm neither confirming NOR denying the existence of COVID19. No way for me to know, or anyone for that matter. |
| C2 | Side-effects | Some vaccines and pills can work, some can kill you. Not everyone responds to medicine the same or has the same exact immune system. Humans are the guinea pigs for big pharma. And they are profit driven. |
| C3 | Political | Political ploy to get trump out of office. AOC literally tweeted this. |
| C4 | Rights | He's violating human rights, goes against Nuremberg, it's also racist to be targeting the Third World like this. |
| C5 | Safety | Semantics. It has never undergone a double-blind three phase testing process like drugs. We don't know anything about its safety or efficacy. |
| C6 | Indignant | There is ZERO fucking need for a vaccine, for a virus that 99.98% of people naturally develop antibodies for. Bill Gates is a fucking psychopath and he needs to be locked up for his crimes against humanity |
| C7 | Bill Gates | The entire plandemic is a scam to allow Bill Gates and other globalist criminals to cash in. |
| C8 | Troll | China doing something to help the rest of the world and not gain anything themselves? LOL okay. |

## Methods

**1 data**
- r/conspiracy submissions containing 'vaccine' from Pushshift query
- collect comments on each submission with PRAW, Jan '20 - Dec '20

**2 pre-processing**
- clean bots, deleted comments
- clean URLs, mentions, emojis
- expand contractions, fix repeating letters
- tokenise (lowercase & remove punctuation)
- remove stopwords
- lemmatize

**3 topic modelling**
- join and add common bigrams
- bag-of-words vector representation
- normalise for tf-idf representation
- NMF model with num_topics = 36

**4 clustering**
- check for latent clusters in document-topic vector space
- k-means clustering on topic loadings with num_clusters = 8
- validate clusters manually

## Results

**H1 is false**. Fig.1 does not show a monotonic increase in comment frequency. Our theory only accounts for rise in consumption of anti-vac content, not fluctuations.

**H2 is false.** Fig.2 shows share of comments by each cluster remained stable over time. There was no discernible shift in anti-vac rhetoric on r/conspiracy.

**8 clusters identified.** Tbl.1 shows the label and representative comment chosen manually for each cluster. Fig.3 shows the mean topic loadings of each cluster on a subset of the most variable topics.

*Refs: [1] Klein C, Clutton P and Polito V (2018) Topic Modeling Reveals Distinct Interests within an Online Conspiracy Forum. Front. Psychol. 9:189 [2] Kunda, Z (1990) The Case for Motivated Reasoning. American Psychological Association 108 (3). pp 480 - 498.*