Theme Song: Welcome Back

Coreen Mullen

Xinyi Miao

Evans Etrue Howard

REDWARN

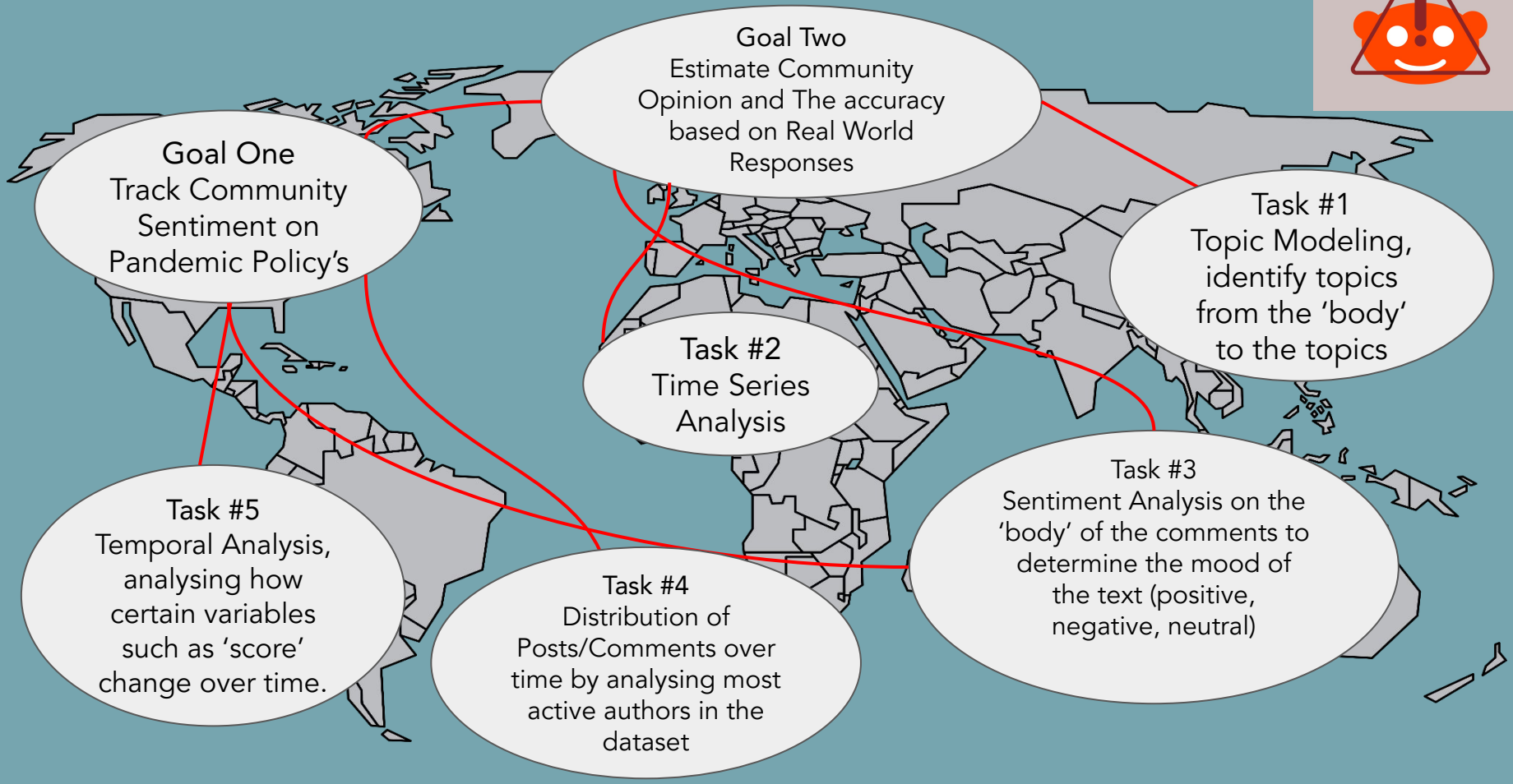
Mentors

- Emily Javan
- Oluwasegun Ibrahim
- Lydia Fletcher

Qimora Mason

Ahmad Samyono

REDWARN (Reddit Data for Early Warning and Response to Pandemics)



Project Plan

Evans

- Github Lead
- Code Collaboration
- Data Analysis
- Task #3 :

- Analyse frequency of posts
- Distribution of posts over time by authors
(5hrs)

Qimora

- Code collaboration
- Analysing data
- Task #1

- Clean “text” column
-Handle Missing Values
(5hrs)

Ahmad

- Analysing Data
- Code Collaboration
- Task #4

-How semantics change over time
-Study how the semantics changed after policy changed
(5hrs)

Xinyi

- Code Collaboration
- Visualisation
- #Task #5

-Prepare text for topic modeling
- Topic correlation
(5hrs)

Coreen

- Poster and Slide creation
- Code Collaboration
- Analysing the Data
- #Task 2

-Aggregate Sentiment scores
(5hrs)

Hackers



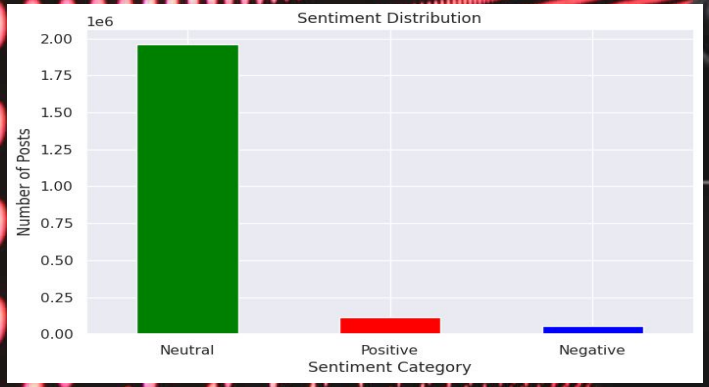


Mentors



Technology Used:
Google Colab.
Python and R.

Bottlenecks/Issues:
We had to use Eureka due to the size of the data, we pre-processed it then uploaded it to our Google Colab.



Status Updates

Task #1: Topic Modeling - 5 hours.
Task #2: Time Series Analysis - Finished.
Task #3: Sentiment Analysis - Finished.
Task #4: Distribution of Comments/Posts Over Time - 5 hours.
Task #5: Temporal Analysis - 5 hours.

—Analyzing the results is our next major task.
—Combining Time Series Analysis + Topic Modeling.

REDWARN (Reddit Data for Early Warning and Response to Pandemics)