Semantic Analysis of Misinformation Policy using BERTopic

INFO 5731: Computational Methods of Information Systems

March 10, 2024

Sai Koushik Valluri
Section-020
11547937
saikoushikvalluri@my.unt.edu

Srikanth Nampally
Section-020
11592092
srikanthnampally@my.unt.edu

Mohana Srinitha Shaga
Section-022
11637773
shagamohanasrinitha@my.unt.edu

Abhigna Dara
Section-020
11643029
AbhignaDara@my.unt.edu

ACM Reference Format:

Sai Koushik Valluri, Srikanth Nampally, Mohana Srinitha Shaga, and Abhigna Dara. 2024. Semantic Analysis of Misinformation Policy using BERTopic
INFO 5731: Computational Methods of Information Systems
March 10, 2024. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

CONTENTS

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

1.1 Background and Significance

Dr. Haihua Chen's "Semantic Analysis of Misinformation Policy using BERTopic" project focuses on exploring the governance models of misinformation used by various governments, with a special emphasis on their regulatory activities. The concept for this project derives from the worldwide complexity and multifaceted nature of misinformation/disinformation governance, which involves a wide range of parties with various roles. The value rests in knowing how different governments regulate various components of disinformation, as well as the importance of this research in comparison to previous studies. This project aims to offer light on governance models, policy aspects, and stakeholder networks addressing misinformation.

1.2 Research Question

RQ1: What are some of the current governance models of misinformation adopted by different governments? **RQ2:** What are the characteristics of the misinformation policy arena? **RQ3:** How have the misinformation-related topics/issues and the network of different stakeholders evolved over time?

1.3 Research Purpose

The research questions concentrate on governance forms, policy aspects, and stakeholder evolution in misinformation regulation. The major purpose is to do semantic analysis on acquired policy documents by extracting keywords, entities, and modeling topics.

1.4 Research Methods

To answer the research questions, we will use a comprehensive methodology that includes: **Document Collection:** Organizing a carefully selected set of policy documents from various nations. **Keyword Extraction:** Analyze these documents to extract relevant keywords that reveal each policy's core focus. **Entity extraction:** The process of identifying and categorizing significant entities involved in misinformation regulation. **Topic Modeling (BERTopic):** Using advanced natural language processing techniques, specifically BERTopic, to gain a detailed knowledge of misinformation-related issues.

2 Related work

2.1 Cross-platform social dynamics: an analysis of ChatGPT and COVID-19 vaccine conversations

Through an analysis of over 12 million postings on social media sites like Facebook, Twitter, and Reddit on the release of ChatGPT and conversations about the COVID-19 vaccine, this study investigates the effects of social media on the dissemination of knowledge and public discourse [1].

2.2 Analyzing qanon on twitter in context of us elections 2020: Analysis of user messages and profiles using vader and bert topic modeling

The study focuses on QAnon references in over 12 million tweets that were sent during the 2020 US Presidential Elections. According to analysis, QAnon talk is primarily pro-Trump and does not precisely follow geographic political leanings. Users frequently identify as conservative and nationalist, as seen by the terms they employ in their profiles [2].

2.3 Racist framing through stigmatized naming: A topical and geo-locational analysis of# chinavirus and# chinesevirus on Twitter

The purpose of this study is to examine how social media, including hashtags like Chinavirus and Chinesevirus, contributed to the propagation of racism and xenophobia during the COVID-19 outbreak. It examines the dissemination of racist speech and identifies influential figures through the use of geo-locational analysis and BERTopic modeling [3].

2.4 Uncovering discussion groups on claims of election fraud from twitter

Using network modeling, the study examines Twitter's impact on conversations around 2020 U.S. election fraud, elucidating user demographics and subjects. It efficiently pinpoints 25 important themes and separate groups while providing insights into the mechanics of the dissemination of fraud claims [4].

2.5 Reactions to science communication: discovering social network topics using word embeddings and semantic knowledge

Using Twitter conversations from doctors, researchers, science communicators, and research institutions over a two-year period, this study examines how the public views and responds to scientific information given on social media during the COVID-19 epidemic. The study creates a way for quickly detecting pertinent social media comments by utilizing topic modeling and machine learning techniques [5].

2.6 Health-related misinformation and public governance of COVID-19 in South Africa

The study by Kariuki et al. titled "Health-related misinformation and public governance of COVID-19 in South Africa" discusses the spread of false information about health during the pandemic and how it affected public governance in that country. The results highlight a link between the propagation of false information and problems with institutional coherence and government coordination [6].

2.7 Social Noise on Social Media and Users Perception of Global Warming

The article focuses on the sentiment analysis and topic modeling to contextually analyze Twitter data pertaining to "global warming" from 2010 to 2011. It seeks to evaluate how the general population views global warming and the influence of false information, sometimes known as "social noise." The majority of tweets, according to the results, reflect opinions rather than empirical data [7].

2.8 Examining the Prevailing Negative Sentiments Surrounding Measles Vaccination: Unsupervised Deep Learning of Twitter Posts from 2017 to 2022

This study examined the unfavorable opinions on the measles vaccine on Twitter over a five-year period using cutting-edge machine learning techniques, finding notable negative opinions from distinct people. Four primary themes were identified using thematic analysis and BERT-based analysis [8].

2.9 In the Spotlight: The Russian Government's Use of Official Twitter Accounts to Influence Discussions About its War in Ukraine

The article analyses how Russian government tweets in the English language have influenced discussions on the conflict in Ukraine by depicting Russia as a peace-seeking nation and holding the "Kiev Regime" accountable for starting the conflict. The significance that state-sponsored social media plays in spreading misinformation during geopolitical crises is highlighted by this study [9].

2.10 Artificial intelligence–enabled analysis of statin-related topics and sentiments on social media

An AI technique was used in this qualitative study to categorize social media information about statins into discussion-worthy categories. Six theme categories were then created from these 100 topics. An AI method like this may be used to examine vast amounts of recent social media data and provide insights regarding how the general public feels about statins [10].

3 Methodology

In this project, we will use the following methodology 1. Data Collection. 2. Data Preprocessing. 3. Keyword Extraction. 4. Entity Extraction. 5. Topic Modelling. 6. Topic Evolution. 7. Visualization. 8. Validation and Evaluation. 11. Insights and Recommendations.

4 Data Collection and Cleaning Plan

4.1 Data Collection

We will collect misinformation policy documents from different sources like government websites and academic databases. These documents will contain laws, guidelines, regulations, official statements and reports related to misinformation. We will collect data from different countries. We will also develop a search criteria to extract documents that are closely related to research questions.

4.2 Cleaning Plan

Collected data will be cleansed and pre-processed to improve the performance of the analysis. We will perform text normalization, tokenization, stopword removal and lemmatization/stemming. Quality checks will be performed to identify inconsistencies and missing information in the data. We will remove duplicate documents to avoid redundancy.

5 Experiment and Data Analysis Plan

5.1 Model Development

Keywords will be extracted from preprocessed documents using methods like TF-IDF, RAKE as part of Keyword Extraction. Entity Extraction/ Entity Recognition will be used to identify entities which are relevant in misinformation governance i.e. Identify government bodies that are responsible to draft regulations related to misinformation, recognize major technology platforms such as Facebook, Google, Twitter, Reddit and Instagram which are involved in content generation. As part of Topic Modelling, we will be using Bertopic algorithm to cluster documents into topics and identify the topic class.

5.2 Evaluation Metrics

Using metrics like TF-IDF score we will evaluate the relevance and importance of extracted keywords. Accuracy of the entity recognition can be validated using precision, recall and F1-Score. We will evaluate the coherence and interpretability of topics using metrics like topic coherence scores/ manual inspection of topic keywords.

5.3 Data Analysis

We will analyze extracted keywords to identify the related topics and important terms which are related to misinformation policies. We will explore the roles and relationships of identified entities. We have to investigate the topics generated by BERTopic.

5.4 visualization

Work clouds will be used to visualize the most frequent words used. Network graphs can help in displaying the relationship between entities involved in misinformation governance. We will visualize topic clusters and their interrelations using techniques such as t-SNE or PCA for dimensionality reduction.

6 Task assignment and timeline

6.1 Task Assignment

In each phase of the project, we will assign four distinct methods for accomplishing a task, ensuring diversity in approach and maximizing insights. Specifically, for keyword extraction and entity extraction, each team member will employ one of four varied methods, resulting in four unique outputs for each task. These outputs will then serve as inputs for topic modeling using BERTopic, allowing us to compare the outcomes effectively.

6.2 Project Timeline

Week Number	Week Start Date	Week End Date	Project Tasks
1	3/3/2024	3/9/2024	Project Proposal
2	3/10/2024	3/16/2024	Literature Review
3	3/17/2024	3/23/2024	Keyword Extraction
4	3/24/2024	3/30/2024	Entity Extraction
5	3/31/2024	4/6/2024	Topic Modelling
6	4/7/2024	4/13/2024	Optimising
7	4/14/2024	4/20/2024	Research Questions
8	4/21/2024	4/27/2024	Documentation
9	4/28/2024	5/4/2024	Presentation
10	5/5/2024	5/11/2024	Reporting

References

[1] Shayan Alipour, Alessandro Galeazzi, Emanuele Sangiorgio, Michele Avalle, Ljubisa Bojic, Matteo Cinelli, and Walter Quattrociocchi. 2024. Cross-platform social dynamics: an analysis of ChatGPT and COVID-19 vaccine conversations. *Scientific Reports* 14, 1 (2024), 2789.

[2] Ahmed Anwar, Haider Ilyas, Ussama Yaqub, and Salma Zaman. 2021. Analyzing qanon on twitter in context of us elections 2020: Analysis of user messages and profiles using vader and bert topic modeling. In *DG. O2021: The 22nd Annual International Conference on Digital Government Research*. 82–88.

[3] Miyoung Chong and Haihua Chen. 2021. Racist framing through stigmatized naming: A topical and geo-locational analysis of# chinavirus and# chinesevirus on Twitter. *Proceedings of the association for information science and technology* 58, 1 (2021), 70–79.

[4] Jose Martins da Rosa Jr, Renan Saldanha Linhares, Carlos Henrique Gomes Ferreira, Gabriel P Nobre, Fabricio Murai, and Jussara M Almeida. 2022. Uncovering discussion groups on claims of election fraud from twitter. In *International Conference on Social Informatics*. Springer, 320–336.

[5] Bernardo Cerqueira de Lima, Renata Maria Abrantes Baracho, Thomas Mandl, and Patricia Baracho Porto. 2023. Reactions to science communication: discovering social network topics using word embeddings and semantic knowledge. *Social Network Analysis and Mining* 13, 1 (2023), 119.

[6] Paul Kariuki, Lizzy Oluwatoyin Ofusori, Maria Lauda Goyayi, and Prabhakar Rontala Subramaniam. 2023. Health-related misinformation and public governance of COVID-19 in South Africa. *Digital Policy, Regulation and Governance* 25, 1 (2023), 58–74.

[7] Nayana Pampapura Madali, Manar Alsaid, and Suliman Hawamdeh. 2023. Social Noise on Social Media and Users Perception of Global Warming. *Journal of Information & Knowledge Management* 22, 06 (2023), 2350050.

[8] Qin Xiang Ng, Yu Qing Jolene Teo, Chee Yu Kiew, Bryant Po-Yuen Lim, Yu Liang Lim, and Tau Ming Liew. 2023. Examining the Prevailing Negative Sentiments Surrounding Measles Vaccination: Unsupervised Deep Learning of Twitter Posts from 2017 to 2022. *Cyberpsychology, Behavior, and Social Networking* 26, 8 (2023), 621–630.

[9] Benjamin Shultz. 2023. In the Spotlight: The Russian Government's Use of Official Twitter Accounts to Influence Discussions About its War in Ukraine. In *Proceedings of the 2nd ACM International Workshop on Multimedia AI against Disinformation*. 45–51.

[10] Sulaiman Somani, Marieke Meija van Buchem, Ashish Sarraju, Tina Hernandez-Boussard, and Fatima Rodriguez. 2023. Artificial intelligence–enabled analysis of statin-related topics and sentiments on social media. *JAMA Network Open* 6, 4 (2023), e239747–e239747.