Sentiment Analysis and Topic Detection on Post-Pandemic Healthcare Challenges: A Comparative Study of Twitter Data in the US and Indonesia

George Morris William Tangka[1], Ibrena Reghuella Chrisanti*[2], Jacquline Waworundeng[3], Raissa Camilla Maringka[4], Green Arther Sandag[5]
[1,3,4,5]Universitas Klabat; Jl. Arnold Mononutu, Airmadidi, Minahasa Utara, Sulawesi Utara,
[2]PT. Vfirst Komunikasi Indonesia; Jl. Jenderal Sudirman No.Kav 86 Lt. 23, Karet Tengsin, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta
[1]Information System, [2]Information Management, [3,4,5]Informatics
e-mail: [1]gtangka@unklab.ac.id, *[2]**ibrena.chrisanti@vfirst.com**, [3]jacquline.morlav@.unklab.ac.id, [4]raissam@unklab.ac.id, [5]greensandag@unklab.ac.id

Abstract

This study examines public sentiment and key topics in Twitter discussions regarding the COVID-19 vaccine and the Omicron variant in the US and Indonesia. The importance of this research lies in understanding people's changing views on vaccination, especially in light of new virus variants. Using sentiment analysis with VADER and topic modeling with Latent Dirichlet Allocation (LDA), this research analyzes 637,367 tweets from the US and 91,679 tweets from Indonesia collected over two months from January 21 to February 21, 2022. The results reveal that US discussions on vaccines are predominantly positive, while those on Omicron are mostly negative. In contrast, discussions in Indonesia are largely neutral, followed by positive sentiment. Additionally, five main topics were identified for each country, with the US showing a broader range of vaccine-related discussions. These findings suggest that while the vaccine is seen as a source of hope in both countries, factors such as literacy, socioeconomic status, and education contribute to negative sentiment and vaccine resistance.

Keywords— Sentiment Analysis, Topic Modeling, Social Media, COVID-19 Vaccine, Omicron

1. INTRODUCTION

The global outbreak of COVID-19, caused by the SARS-CoV-2 virus, has significantly disrupted societies worldwide since its first detection in December 2019 in Wuhan, China. COVID-19 has rapidly spread across borders, causing severe health and economic consequences. The virus can manifest in a variety of symptoms, ranging from mild conditions like sore throat and nasal congestion to severe respiratory distress and death. It is primarily transmitted through small droplets expelled during coughs, sneezes, or even talking, which can land on surfaces and be transferred to individuals who touch those surfaces and then their eyes, nose, or mouth. The World Health Organization (WHO) reported, by February 2022, over 386 million confirmed cases and nearly 5.7 million deaths globally, underscoring the pandemic's vast scale and impact on public health. As the pandemic continues to affect millions of people, the emergence of multiple virus variants has posed additional challenges. The evolution of the virus has resulted in strains such as the Alpha, Beta, and Delta variants, each with varying transmissibility, severity, and immune escape characteristics. The Omicron variant (B.1.1.529), first identified in late 2021, has become a particular point of concern due to its ability to evade immunity provided by earlier variants and vaccines. The Omicron variant has led to a surge in cases globally, further complicating efforts to control the pandemic [1, 2].

1.1. Literature Review

Vaccination has been one of the most effective tools in combating the spread of the virus and mitigating the severity of illness. The United States, the country with the highest number of Covid-19 cases, launched its vaccination campaign in December 2020. Initially, there was a high rate of vaccine acceptance, with the U.S. Food and Drug Administration (FDA) emergency authorizing vaccines from companies like Pfizer-BioNTech and Moderna. Early studies showed that the vaccines had an effectiveness rate above 85% in preventing severe disease and hospitalization. However, as the virus continued to mutate and new variants like Omicron emerged, vaccine effectiveness was diminished, leading to a decline in public confidence and an increase in vaccine hesitancy. By 2021, vaccine acceptance rates in the U.S. showed a downward trend. A survey conducted by Szilagyi et al. [3] revealed a decline in willingness to get vaccinated, dropping from 75% to 56% over the year. Factors contributing to this decline include misinformation about vaccine safety, political polarization, and distrust in government actions. Research by Mondal, Sinharoy, & Su [4] further indicated that sociodemographic factors such as age, education, ethnicity, and family income play significant roles in vaccine acceptance.

In Indonesia, the situation mirrored the U.S. to some extent. At the beginning of the pandemic, Indonesia had one of the highest vaccination acceptance rates, with about 93.3% of the population willing to receive the vaccine. However, as misinformation spread, particularly concerning vaccine side effects and government policies, public confidence started to wane. By the end of 2021, a survey revealed that only 65% of the population still supported vaccination, showing a significant drop in vaccine acceptance. Additionally, religious beliefs and the influence of local communities played a critical role in shaping people's views toward vaccination. With the ongoing changes in the virus's transmissibility and the introduction of new variants, governments worldwide faced the dual challenge of maintaining public trust while implementing effective control measures [5]. Social distancing, mask mandates, and lockdowns became the new norms, but the reliance on technology for communication and work increased. Social media platforms, particularly Twitter, saw an unprecedented surge in usage as people turned to these platforms for updates, social connection, and emotional support. While social media provided a valuable space for information exchange, it also became a breeding ground for misinformation, conspiracy theories, and public anxiety, further complicating the global response to the pandemic [6].

1.2. Context and Motivation

Twitter, with its massive user base, became a crucial platform for public discourse, especially in the U.S. and Indonesia. In the U.S., Twitter reported 77.75 million active users by November 2021, making it one of the leading platforms for sharing opinions and sentiments. The platform has proven to be a rich source for sentiment analysis, as it provides real-time public reactions to pressing issues, including the COVID-19 vaccine and the Omicron variant. Studies conducted by Saud et al. [7] found that, in Indonesia, people showed a positive attitude towards social media, using it to connect with family and friends and seek information on health topics. However, concerns were also raised about the mental health implications of prolonged social media use, with feelings of anxiety, distrust, and frustration becoming increasingly prevalent as the pandemic continued [6].

1.3. Research Objectives

The purpose of this study is to identify public sentiment in the United States and Indonesia towards the COVID-19 vaccine and the Omicron variant, as well as the most frequently discussed topics related to these issues. This research is particularly relevant given the ongoing emergence of new Coronavirus variants, especially Omicron. The United States, with its high number of cases and deaths, was the first country to roll out a vaccine program and has the highest proportion of Twitter users globally. Conversely, Indonesia initially displayed high public acceptance of vaccination during the first wave of Covid-19. However, acceptance rates decreased by approximately 30%, influenced by conspiracy theories, misinformation, doubts, and religious beliefs [5]. Given these contrasting contexts, understanding public attitudes toward vaccines and emerging variants is critical for policymakers and health organizations to design effective strategies and interventions. This study seeks to address key questions regarding public attitudes towards vaccines and emerging variants, such as Omicron, in two contrasting contexts, namely the United States and Indonesia. Specifically, it aims to determine the most dominant sentiment polarity—whether positive, negative, or neutral—expressed in the vaccine- and Omicron-related discussions within these countries. Furthermore, it explores the primary topics or entities that emerge as the top five focal points in these discussions, shedding light on who or what drives public discourse in each context.

2. RESEARCH METHODS

This study follows a structured research framework to analyze public sentiment and identify prevalent topics in Twitter discussions related to COVID-19 vaccines and the Omicron variant. The framework consists of four main stages: data collection, data pre-processing, sentiment analysis, and topic modeling. Each step is designed to ensure the data is properly prepared, analyzed, and interpreted to provide insights into public opinions across two countries, the United States and Indonesia.

For sentiment analysis, VADER (Valence Aware Dictionary and sEntiment Reasoner), a lexicon-based tool specifically optimized for social media text. VADER has been shown to perform well on short, informal texts like tweets, where the language is often highly contextual, contains slang, and may involve emoticons or hashtags. Furthermore, VADER is well-suited for handling sentiment expressed in varied formats and offers a fast, efficient, and easy-to-implement solution for analyzing sentiment at scale. We believe that VADER's accuracy in detecting sentiment nuances in social media data makes it ideal for this research, particularly in understanding public sentiment toward COVID-19 vaccines and the Omicron variant [8, 9].

To identify prevalent topics and themes in the discussions, Latent Dirichlet Allocation (LDA), is a well-established method for topic modeling. LDA is particularly effective for uncovering hidden thematic structures in large text corpora. Given the large volume of tweets related to COVID-19 vaccines and the Omicron variant, LDA helps reveal the underlying topics without requiring prior labeling of data. The choice of LDA is further supported by its ability to handle high-dimensional data, such as the diverse range of discussions on Twitter. LDA has been widely used in social media analysis due to its interpretability and effectiveness in extracting coherent topics from unstructured data. In this study, it allows us to understand the diverse range of conversations around Covid-19 and the Omicron variant, providing insights into the public's concerns, interests, and discourse [10, 11].

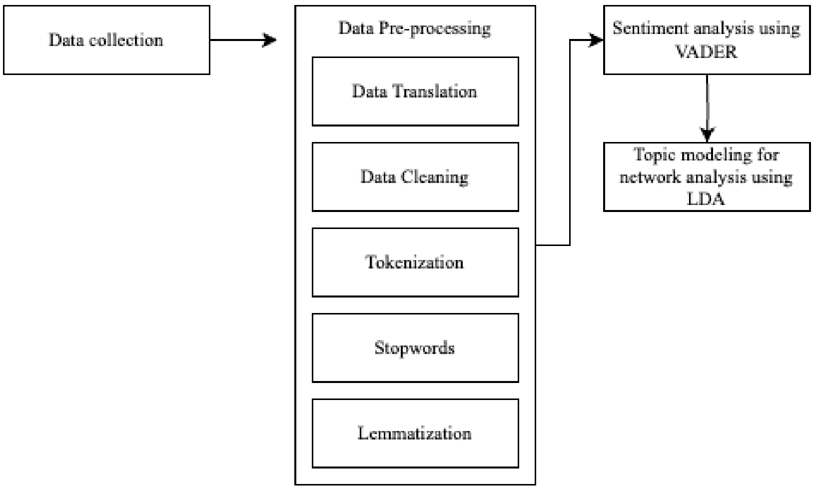

Figure 1. Research Framework

2.1. Data Collection

The primary data for this research was collected from Twitter using the Twitter Intelligence Tool (twint), which facilitates the retrieval of publicly available tweets; considering is a famous microblogging and social network platform that enables users to share, get information and interact with others around the world. Therefore, many scholars use Twitter information as their source of data in conducting research, and it was proved by the development of applications and researchers in various domains such as Humanitarian Assistance and Disaster Relief, Health, Internet Cyber, and Education [12, 13, 14].

In addition, a big increase in daily active users was counted for about 166 million since the earlier time since coronavirus outbreak Twitter provides a place for people to stay connected, informed and share information, help, and support each other in the middle of the world pandemic where people was not allowed to do physical interaction [15].

The data was collected over two months, from January 21 to February 21, 2022, using specific keywords related to COVID-19 vaccines and the Omicron variant. Keywords for the United States included terms like “Covid-19 vaccine”, “vaccine”, “Omicron”, “Pfizer”, “AstraZeneca”, and “Moderna”. For Indonesia, the keywords were “vaksin”, “corona vaksin”, “Sinovac”, “AstraZeneca”, and “omicron”. The dataset was selected for this study because it represents a critical period during the COVID-19 pandemic when public discourse surrounding vaccines and the Omicron variant was particularly intense. This period coincided with the rapid global spread of the Omicron variant, which became the dominant strain and raised widespread concerns about its impact on vaccine effectiveness and healthcare systems. It also marked significant developments in vaccination campaigns, including booster rollouts and updated public health guidelines in response to Omicron. These events triggered heightened social media activity, providing a rich dataset to analyze public sentiment and discussions. Additionally, the selected timeframe aligns with significant events in both the United States and Indonesia, such as booster shot campaigns in the US and government-led public health measures in Indonesia, making it suitable for comparative analysis.

A total of 637,367 tweets were collected from the United States, while 91,679 tweets were retrieved from Indonesia. These tweets, which represent discussions surrounding vaccination and the Omicron variant, offer valuable insights into public sentiment and discourse during a critical phase of the pandemic. Although the data reflects historical discussions, the findings remain highly relevant in understanding the evolution of public attitudes towards vaccination campaigns, which can inform strategies for addressing hesitancy and misinformation regarding new infectious diseases, such as [insert contemporary disease or issue]. Additionally, analyzing this dataset provides a foundation for exploring the long-term effects of pandemic communication strategies on public trust in healthcare systems, particularly in the context of mental health and resilience in a post-pandemic world.

2.2. Data Pre-processing

Data pre-processing is a crucial step in preparing the raw data for analysis. Given that the dataset consists of tweets, which often contain informal words that may include noise, pre-processing ensures that the text is cleaned, standardized, and ready for sentiment analysis and topic modeling. This section outlines the pre-processing steps performed on the collected Twitter data.

1. *Translation*: Tweets in Indonesian were translated into English to standardize the dataset for analysis.
2. *Cleaning*: Irrelevant columns and noise, such as user handles, hashtags, symbols, URLs, and non-ASCII characters, were removed.
3. *Tokenization*: The text data was split into individual tokens (words) to facilitate further analysis.
4. *Stopword Removal*: Commonly used but uninformative words, like "the," "is," and "and," were removed from the dataset to focus on meaningful words.
5. *Lemmatization*: Words were reduced to their base forms (e.g., "running" to "run") to ensure

that variations of words were treated as the same token.

2.3. Sentiment Analysis

The sentiment of each tweet was classified into three categories: positive, neutral, and negative using the Valence Aware Dictionary and sEntiment Reasoner (VADER). VADER is a lexicon-based sentiment analysis tool specifically designed for social media texts, which takes into account the intensity of emotions in the text. The sentiment score for each tweet is calculated on a scale from -1 to 1, where -1 represents negative sentiment, 1 represents positive sentiment, and 0 represents neutral sentiment. To classify tweets into sentiment categories, the following thresholds were applied based on VADER's standard recommendations, the default threshold of the machine learning algorithm. VADER was chosen for its effectiveness in handling short and informal text data, which is characteristic of tweets. Additionally, VADER's ability to process negations, emoticons, and punctuation makes it particularly suitable for analyzing the tone of social media posts related to the vaccine and Omicron variant [16, 17].

2.4. Topic Modeling

The next step in the methodology was to identify the underlying topics discussed in the tweets. This was achieved using Latent Dirichlet Allocation (LDA), a probabilistic model commonly used for topic modeling, as shown in Figure 2.

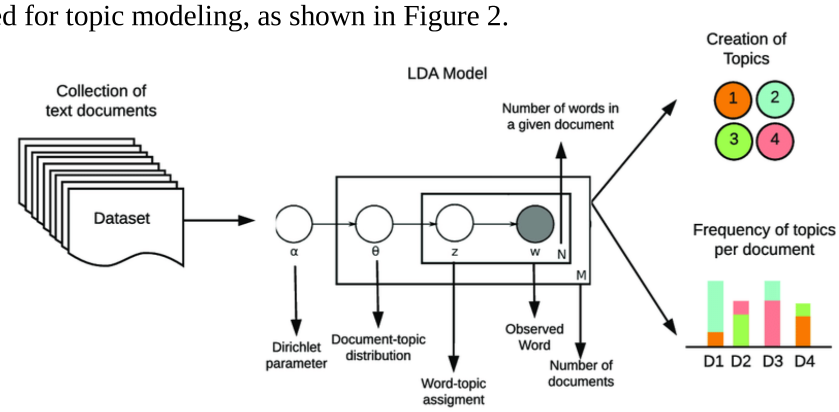

Figure 2. Latent Dirichlet Allocation Architecture [18]

Figure 2 explains that each tweet is a mixture of topics and that each topic is represented by a set of words. The LDA model was trained on the cleaned and pre-processed data, and a grid search was conducted to determine the optimal number of topics. Grid search is an exhaustive search technique used to identify the best combination of parameters for a model by systematically testing all possible values within a specified range. In this study, the grid search involved varying the number of topics between 5 and 30, with increments of 1 [18].

The model identified five key topics that represented the main themes of the vaccine and Omicron-related discussions in both countries. After identifying the topics, a social network analysis was conducted to examine the relationships between words within each topic. Social network analysis provides a graphical representation of connections between words, offering insights into the structural relationships and co-occurrence patterns in the data. To construct the network graphs, the top 8-10 keywords for each topic (as determined by their relevance scores in the LDA model) were selected to form the nodes of the graph. Edges between these nodes were created based on the co-occurrence frequency of words within the same document (tweet), with stronger co-occurrence frequencies resulting in higher edge weights. The frequency of a word's appearance in the dataset determined the size of the node, and the strength of the relationship between words was represented by the thickness of the edges. This analysis helped to identify the central themes of public discourse in each country and provide insights into the most discussed issues with COVID-19 vaccines and the Omicron variant [17, 18, 19].

3. RESULT AND DISCUSSION

This section presents the results of sentiment analysis and topic modeling applied to Twitter data related to the COVID-19 vaccine and the Omicron variant in the United States and Indonesia. The findings are discussed with the public sentiment in each country and the key topics that emerged from the Twitter discussions.

3.1. Sentiment Analysis

The sentiment analysis was conducted using the VADER (Valence Aware Dictionary and sEntiment Reasoner), which classifies tweets as positive, neutral, or negative based on the content. The analysis was performed on 637,367 tweets from the United States and 91,679 tweets from Indonesia over two months [20].

The results, shown in Figure 3, present the overall sentiment proportions for the COVID-19 vaccine and Omicron-related tweets in the United States. For the vaccine-related discussions, positive sentiment dominates at 37%, followed by 36% negative sentiment, and 27% of tweets are neutral. In contrast, Omicron-related discussions are mainly dominated by negative sentiment at 36%, with 32% each for both positive and neutral sentiment.

Figure 4 illustrates the sentiment proportions for both vaccine and Omicron-related discussions in Indonesia. In both cases, neutral sentiment is the most prevalent, accounting for 64% of vaccines and 50% for Omicron. Positive sentiment follows, by making up 25% of vaccine-related discussions and 31% of Omicron-related tweets. The remaining portions are filled with negative sentiment, representing 11% for vaccines and 19% for Omicron.

The sentiment results in Figures 3 and 4 answer the first research question of this study. Vaccine-related discussions in the US are dominated by positive sentiment, while Omicron-related discussions in the US are mainly negative. In contrast, both vaccine and Omicron-related discussions in Indonesia are dominated by neutral sentiment. This indicates some similarities in public sentiment towards the vaccine and Omicron discussions. In both countries, most people reacted neutrally or positively to the vaccine, likely due to widespread vaccination, easy access to vaccines, and informative government approaches via Twitter that keep the public updated. However, there is a noticeable difference in reactions to Omicron. In the US, negative sentiment is more prevalent, possibly due to concerns about the potential for lockdowns if the new variant spreads. On the other hand, in Indonesia, the virus does not seem to be perceived as a serious threat, and there is more confidence in handling it, leading to lower negative sentiment compared to the US.

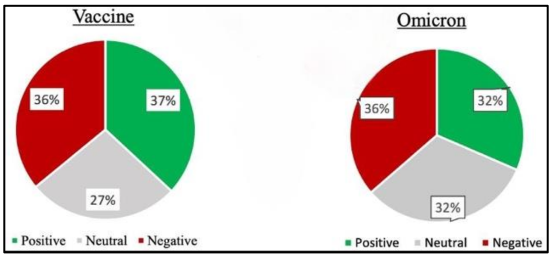

Figure 3. US Covid-19 Vaccine and Omicron Sentiment

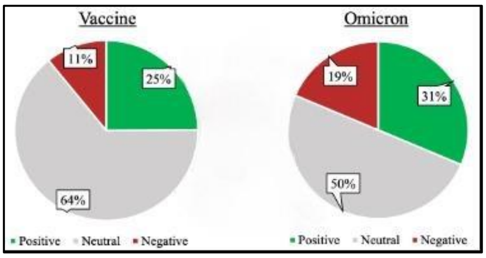

Figure 4. Indonesia Covid-19 Vaccine and Omicron Sentiment

3.2. Topic Modeling and Social Network Analysis

To identify the optimal number of topics for the LDA model, a grid search was performed within the range of 5 to 30 topics, with the best result at 5 topics, based on the perplexity score. Perplexity measures how well a model can handle new data, with a lower value indicating better performance [20]. The lowest perplexity score of 817 was achieved with 5 topics. Thus, this study selected 5 topics for each dataset, consisting of 5 to 10 keywords per topic, deemed most relevant and interpretable based on the data.

After determining the topics and identifying the top keywords, a social network graph was created to analyze the importance and relationships between the words in each topic. In the graph, each word is represented as a node, with its size reflecting its frequency, while the connecting lines indicate the relationships between words.

1. *COVID-19 Vaccine-Related Discussion in the United States*

To answer the second research question, Figure 5 shows the 5 topics related to the Covid-19 vaccine in the US, with the following keywords: **Booster Availability**: available, today, year, clinic, month, health, shot, child, dose, booster; **Vaccine Type**: sign, study, omicron, site, available, mRNA, trial, testing, data, test; **Canada's Trucker**: trucker, health, people, rate, Canada, risk, anti, mask, health, mandate; **Psychological Impact**: good, know, think, right, year, need, want, make, people, like; **Stop Mask**: stop, like, mask, don't, virus, work, know, getting, vaccinated, people.

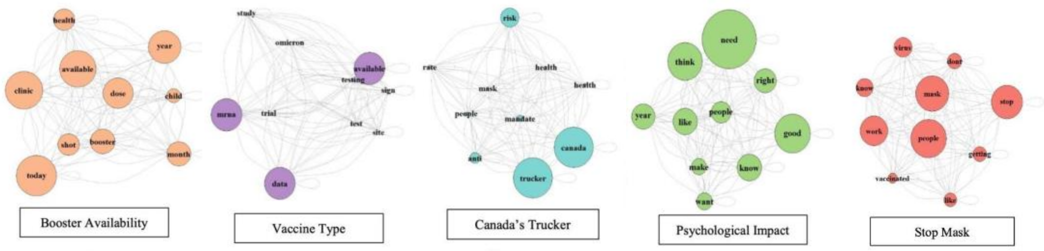

Figure 5. US Covid-19 Vaccine Network

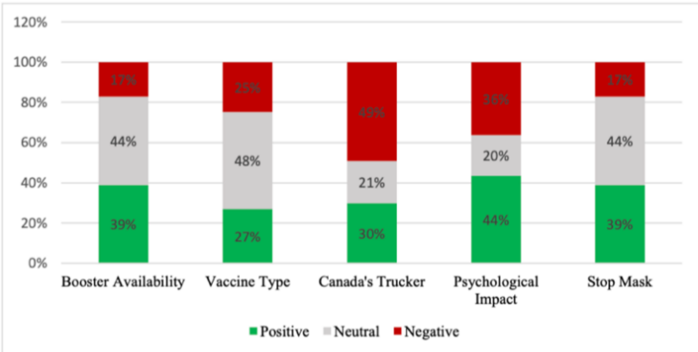

Figure 6. US Vaccine Topic Sentiment

This study incorporates network centrality measures to better understand the network analysis. Degree centrality identifies the importance of a node based on its number of connections. Betweenness centrality captures the role of nodes in connecting others. Closeness centrality measures the shortest path between nodes, indicating how easily a node connects with others. Eigenvector centrality assesses the influence of a node based on its connections to high-degree nodes in the network [19].

Figure 5 shows the Booster Availability network, with nodes such as available, clinic, and today being frequently mentioned in discussions about booster availability in the US. Table 4 presents the centrality measures for this topic. All nodes in this network share the same degree of centrality, indicating equal importance in the network. The node month plays a key role in betweenness centrality, connecting other nodes, while health stands out in closeness centrality as the most accessible node. The node with the highest eigenvector centrality is a booster, reflecting its central role in the network due to its connection to other influential nodes. Regarding sentiment, 44% of the discussions on Booster Availability were neutral, 39% were positive, and 17% were negative, indicating overall support with some concerns.

Table 1. US Covid-19 Vaccine Network Centrality for Booster Availability

topic 1	available	today	year	clinic	month	health	shot	child	dose	booster
Degree	18	18	18	18	18	18	18	18	18	18

Betweenness	0	3	0	1	11	5	0	6	0	0
Closeness	0.0003	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003	0.0004	0.0003	0.0002
Eigen	0.6151	0.4176	0.4750	0.5553	0.5694	0.5074	0.6685	0.4734	0.8524	1.0000

Topic 2: Vaccine Type in Figure 5 highlights the most mentioned words, including available, mRNA, and delta, with Table 2 showing the top eight nodes with the highest degree of centrality: study, omicron, available, mRNA, trial, testing, data, and test. Although a sign does not have the highest centrality, it plays a crucial role in connecting other nodes due to its high betweenness and closeness centrality. The node testing has the highest eigenvector centrality, indicating its strong influence due to connections with key nodes. Sentiment analysis of discussions about Vaccine Type in the US shows 48% neutral, 27% positive, and 25% negative sentiment. This likely reflects people's lack of familiarity with the different vaccine development approaches, resulting in mostly neutral reactions (see Figure 6).

The vaccine development discussion centers around three approaches: using whole viruses, parts of viruses, and genetic material (e.g., BioNTech Pfizer and Moderna). Genetic-based vaccines, being newer, raise concerns about their effectiveness and safety, which are frequently discussed. Additionally, the topic overlaps with conversations about COVID-19 testing appointments, reflecting public interest in easily accessible vaccine and test information [21].

Table 2. US Covid-19 Vaccine Network Centrality for Vaccine Type

topic 2	**sign**	**study**	**omicron**	**site**	**available**	**mRNA**	**trial**	**testing**	**data**	**test**
Degree	17	18	18	17	18	18	18	18	18	18
Betweenness	25	5	0	3	0	5	0	0	2	0
Closeness	0.0147	0.0098	0.0087	0.0097	0.0100	0.0123	0.0092	0.0049	0.0135	0.0081
Eigen	0.9114	0.0177	0.0651	0.9719	0.9698	0.0195	0.0371	1.0000	0.0327	0.9728

Topic 3: Canada's Trucker in Figure 5 is dominated by the words trucker and Canada. Table 3 shows that all nodes in this topic have an equal degree of centrality, indicating their equal importance in the network. However, the trucker has the highest betweenness and closeness centrality, while the mandate stands out with a high eigenvector score, reflecting its strong influence.

Unlike previous topics, sentiment towards this topic in US vaccine-related discussions is largely negative at 49%, followed by 30% positive and 21% neutral. This reflects the negative reactions tied to the truckers' protest against vaccine and mask mandates, particularly the protest in Canada in February, where truckers opposed vaccination requirements for cross-border travel. The protest, which lasted for three weeks, escalated into a broader opposition to pandemic restrictions and the Canadian government's measures [22, 23], leading to widespread negative sentiment on Twitter.

Table 3. US Covid-19 Vaccine Network Centrality for Canada's Trucker

topic 3	**truck**	**health**	**people**	**rate**	**Canada**	**risk**	**anti**	**mask**	**health**	**mandate**
Degree	18	18	18	18	18	18	18	18	18	18
Betweenness	23	0	0	0	3	17	0	0	0	0
Closeness	0.0015	0.0010	0.0005	0.0010	0.0012	0.0015	0.0009	0.0006	0.0006	0.0010
Eigen	0.5308	0.2796	0.4113	0.2045	0.4676	0.1352	0.4114	0.6882	0.2796	1.0000

Topic 4 in Figure 5 is dominated by words such as need, think, and good, which reflect the Psychological Impact of the pandemic. Table 4 shows that all nodes contribute equally to the network, with "good" playing a key role in connecting other nodes due to its high betweenness and closeness centrality. The node "people" is the most influential, with strong connections to high-centrality nodes.

Unlike the other topics, this one has the highest positive sentiment at 44%, followed by 36% negative and 20% neutral sentiment. This indicates that the pandemic's impact is not just physical but also psychological. Research shows that hopefulness and optimism are key factors in coping with health anxiety, which is influenced by factors like hopelessness, psychological resilience, and exposure to COVID-19-related news. Managing mental health can be supported by sharing positive information and promoting activities that encourage self-control and optimism [23, 24]. The positive sentiment in this topic reflects people's optimism for a better year, with many social events, family gatherings, and work resuming. However, some individuals continue to struggle with maintaining their mental health.

Table 4. US Covid-19 Vaccine Network Centrality for Psychological Impact

topic 4	good	know	think	right	year	need	want	make	people	like
Degree	18	18	18	18	18	18	18	18	18	18
Betweenness	16	0	0	0	0	0	0	0	0	0
Closeness	0.0007	0.0004	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0002	0.0003
Eigen	0.3390	0.5665	0.5552	0.5124	0.4769	0.5359	0.6336	0.5568	1.0000	0.8054

Figure 5 shows the social network for Stop Mask, highlighting keywords like people, mask, stop, and work. Table 5 shows that, like the previous topic, all nodes in this network share an equal degree of centrality, with the virus acting as a node that helps connect other members. The most powerful node is people, with the highest influence.

Sentiment analysis reveals that 44% of discussions on this topic are neutral, 39% are positive, and 17% are negative. This reflects resistance to both mask mandates and vaccines, influenced by factors like political polarization and conspiracy theories. Studies show that people's decisions on health measures are often shaped more by social identity than by scientific evidence, leading to misinformation [24]. Recent studies also show that Republicans tend to have a lower acceptance of pandemic regulations compared to Democrats, leading to higher death rates in Republican-governed areas. The premature lifting of mask mandates in the US confused, with some states continuing mandates while others dropped them [25, 26].

Table 5. US Covid-19 Vaccine Network Centrality for Stop Mask

topic 5	stop	like	mask	don't	virus	work	know	getting	vaccinated	people
Degree	18	18	18	18	18	18	18	18	18	18
Betweenness	0	0	0	0	4	0	0	0	0	0
Closeness	0.0003	0.0002	0.0002	0.0002	0.0003	0.0002	0.0002	0.0002	0.0002	0.0001
Eigen	0.4415	0.4579	0.5475	0.5383	0.3957	0.5468	0.5678	0.5766	0.7354	1.0000

2. *Omicron-Related Discussion in the United States*

After analyzing the US Vaccine dataset, the researcher began analyzing the US Omicron dataset. As described earlier, the number of topics was determined using a grid search, resulting in 5 topics, each containing 5 to 10 keywords. These topics answer the second research question of the study and are as follows: **New Subvariant**: spread, wave, expert, contagious, despite, school, case, surge, mask, variant; **Booster Effectiveness**: effective, data, study, death, people, immunity, booster, infection, variant, delta; **Covid-19 Test Accuracy**: test, cold, think, going, don’t, work, know, like, people; **Significant Wave**: number, state, news, peak, death, surge, variant, wave, health, case; **Treatment for Patients**: delta, Canada, contra, plus, virus, work, treatment, antibody, variant.

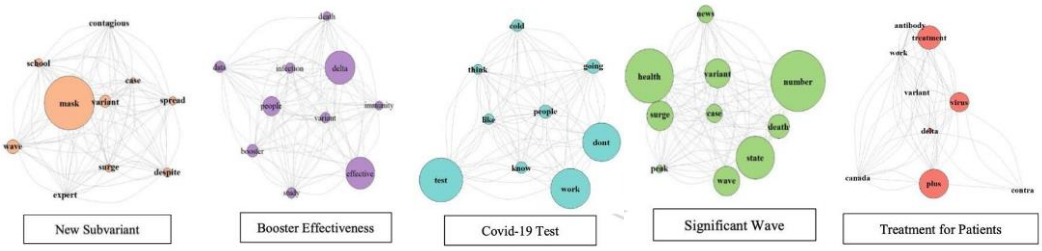

Figure 7. US Omicron Network

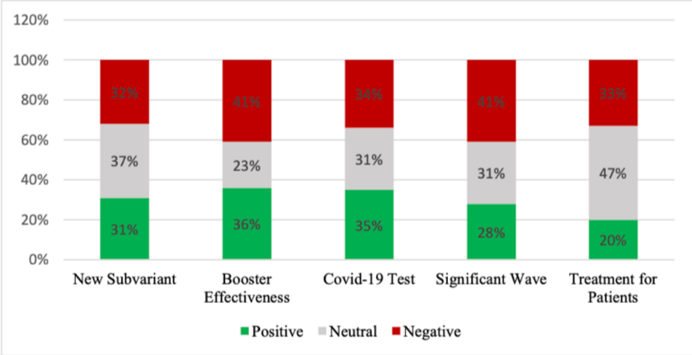

Figure 8. US Omicron Sentiment

An in-depth analysis of network centrality and sentiment was conducted for each topic in the Omicron-related discussions. The first topic, New Subvariant, shown in Figure 7, highlights the most mentioned words, with "mask" being the most frequently discussed. Table 6 shows that all words in this topic contribute equally to degree centrality, with "expert" and "despite" serving as key nodes for connectivity. "Case" is the most influential node due to its high centrality.

Sentiment analysis for this topic, shown in Figure 8, reveals nearly equal distribution across the three sentiments: 37% neutral, 32% negative, and 31% positive. The dominance of neutral sentiment suggests that, despite the availability of vaccines and the concept of a "new normal," the emergence of new Omicron subvariants remains a significant concern. New variants like BA.2 and BA.2.12.1, which are more transmissible, accounted for 80.6% of new cases in New York as of April 2022 [26].

Table 6. US Omicron Network Centrality for New Subvariant

topic 1	spread	wave	expert	contagious	despite	school	case	surge	mask	variant
degree	18	18	18	18	18	18	18	18	18	18
betweenness	0	5	19.5	6.5	21.5	0	0	0	0	0
closeness	0.0037	0.0065	0.0081	0.0072	0.0081	0.0047	0.0038	0.0038	0.0020	0.0050
eigen	0.4798	0.2286	0.3700	0.5708	0.3462	0.3575	0.6213	0.4917	0.5658	0.0000

Figure 7 depicts the network for Topic 2: Booster Effectiveness, where the most discussed words on Twitter in the US related to the Omicron virus are effective and delta. Table 7 shows that all nodes in this network have a high degree of centrality, indicating their importance. The effective node plays a key role in connecting other nodes, with high betweenness and closeness centrality, while the variant is the most influential node in the network.

Sentiment analysis, as shown in Figure 8, reveals that 41% of the posts are negative, followed by 36% positive and 23% neutral. The negative sentiment likely stems from concerns about vaccine effectiveness, as stated by the CDC. With the emergence of the Delta and Omicron variants, vaccine protection against hospitalization and death decreased. However, booster shots have proven to provide higher protection, particularly for those over 50 years old [27, 28]. The ongoing discussion about vaccine and booster effectiveness reflects growing concerns about the evolving virus. Some people question whether the vaccine or booster can effectively prevent the virus, especially those who perceive the virus as mild and view the vaccine as unsafe [28, 29].

Table 7. US Omicron Network Centrality for Booster Effectiveness

topic 2	effective	data	study	death	people	immunity	booster	infection	variant	delta
degree	18	18	18	18	18	18	18	18	18	18
betweenness	12	3	0	5	0	0	0	0	0	0
closeness	0.0007	0.0006	0.0006	0.0007	0.0004	0.0005	0.0005	0.0004	0.0002	0.0003
eigen	0.3992	0.4257	0.4399	0.6263	0.6911	0.6002	0.5286	0.7158	1.0000	0.9729

Figure 7 shows that Topic 3: Covid-19 Test in Omicron-related discussions is dominated by the terms test, work, and don’t. Table 8 presents the network centrality for this topic, where all nodes share equal importance. The test node serves as a mediator, connecting other nodes through the shortest path. Sentiment analysis for this topic reveals a fairly even distribution: 35% positive, 34% negative, and 31% neutral as shown in Figure 8.

Testing is essential for determining Covid-19 infection status. There are two primary types of COVID-19 tests: diagnostic tests (including Polymerase Chain Reaction (PCR) and rapid tests) and antibody tests (currently not used for diagnosing active infections) [29]. Self-testing kits are widely available, providing quicker results. However, a study of 225 adults and children in the US found that self-tests have lower sensitivity and may require retesting after 1-2 days. PCR tests, which offer higher accuracy, require a visit to a testing site, where medical professionals administer the test, and results are typically available within a day [30].

Table 8. US Omicron Network Centrality for Covid-19 Test

topic 3	test	cold	think	going	don't	work	know	like	people
degree	16	16	16	16	16	16	16	16	16
betweenness	12.5	2	0	0	0	0	0	0	0
closeness	0.0013	0.0011	0.0010	0.0011	0.0009	0.0008	0.0007	0.0006	0.0005
eigen	0.3013	0.4336	0.4839	0.4648	0.6907	0.6826	0.7944	0.7912	1.0000

Figure 7 shows that Topic 4: Significant Wave in Omicron-related discussions is dominated by keywords like number, health, and state. Table 9 indicates that all nodes in the network have the same degree of centrality, with news acting as a bridge between other nodes. Peak serves as the shortest path for most nodes to connect, while case is the most powerful node.

Figure 8, sentiment analysis reveals 41% negative, 31% neutral, and 28% positive sentiment. Research on COVID-19 infection and transmission rates in the US shows that during the Omicron period, transmission was 6-8 times higher than during the Delta period, making Omicron more infectious, though less severe, with lower hospitalization rates [31]. This surge in infections likely contributed to the high negative sentiment expressed on Twitter.

Table 9. US Omicron Network Centrality for Covid-19 Test

topic 4	number	state	news	peak	death	surge	variant	wave	health	case
degree	18	18	18	18	18	18	18	18	18	18
betweenness	1	0	19	4.5	0	0	0	1.5	0	0
closeness	0.0011	0.0011	0.0013	0.0009	0.0008	0.0007	0.0006	0.0009	0.0007	0.0005
eigen	0.3928	0.4477	0.2593	0.4432	0.6130	0.5369	0.7843	0.4847	0.5517	1.0000

Result in Figure 7 shows the social network for Topic 5: Treatment for Patients, where keywords like Plus, Virus, and Treatment dominate the discussions in the US. Unlike other topics, only the Delta and Virus nodes have the highest degree of centrality. The node Canada serves as the shortest path connecting other nodes, while Antibody is the most influential node. Sentiment analysis reveals 47% neutral, 33% negative, and 20% positive sentiment. The discussion primarily focuses on Delta and Delta Plus variants and treatments. Due to the severe disease caused by these variants, antibody infusions were less effective, leading to more negative reactions [32].

Table 10. US Omicron Network Centrality for Covid-19 Test

topic 5	delta	canada	contra	plus	virus	work	treatment	antibody	variant
degree	16	14	5	14	16	16	14	13	14
betweenness	0	26	0	0	0	0	0	0	0
closeness	0.0094	0.0156	0.0120	0.0091	0.0078	0.0128	0.0131	0.0111	0.0034
eigen	0.3312	0.0562	0.0179	0.1502	0.2834	0.2812	0.9169	1.0000	0.8349

3. *COVID-19 Vaccine Related Discussion in Indonesia*

This analysis also includes a social network analysis of COVID-19 vaccine-related discussions in Indonesia, answering the second research question by identifying five key topics. The first topic, **Vaccine for Child**, is associated with keywords like aged, complete, country, need, good, government, child, said, receive, and year. The second topic, **Safe Development**, includes remember, look, news, data, number, development, want, safe, police, and people. The third topic, **Booster Safety**, is linked to belief, study, need, come, protection, make shot, time, and booster. The fourth topic, **Vaccine Coverage**, features keywords such as million, effective, used, yesterday, ministry, government, virus, type, dosage, and health. The final topic, **Side Effect**, includes know, second, really, body, like, fever, loading, effect, dose, and booster.

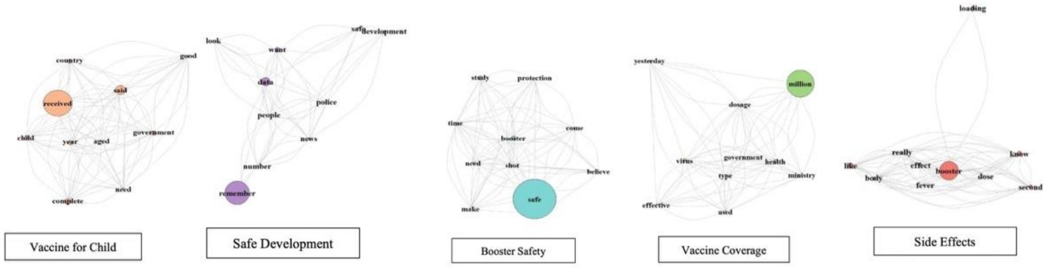

Figure 9. Indonesia Covid-19 Vaccine Social Network

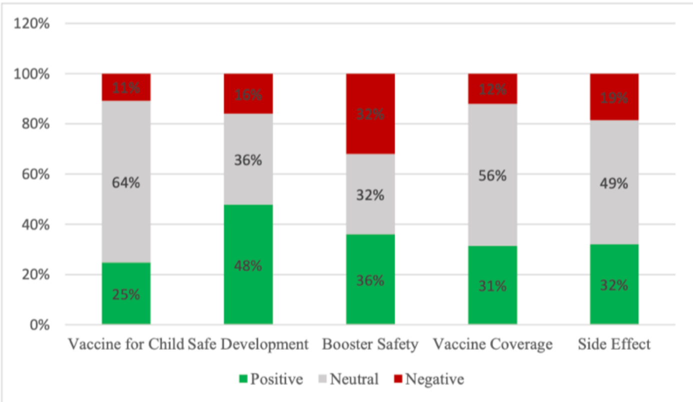

Figure 10. Indonesia Covid-19 Vaccine Sentiment

Figure 9 shows the social network for the Topic: Vaccine for Child, where the term receives dominates the discussion. Table 11 highlights the centrality metrics, showing that government, said, and year have the highest degree of centrality. Complete serves as the key mediator node with the highest betweenness and closeness scores, while year holds the strongest power in the network. Sentiment analysis, as shown in Figure 10, reveals that 64% of the discussion is neutral, 25% is positive, and 11% is negative.

Vaccination plays a crucial role in achieving herd immunity during Covid-19. As of May 2022, 96.13% of people received the first dose, 80.34% the second dose, and 21.80% the booster [33]. However, 70% coverage is needed for herd immunity with the third dose. A study in Indonesia on parents' knowledge, attitude, and practice regarding the COVID-19 vaccine showed a positive response, particularly among parents with higher education and occupation levels. This aligns with the current research, where most people responded neutrally, likely due to limited information on vaccination programs for children. Despite this, many remain optimistic about the vaccine, while a small portion still rejects it [34].

Table 11. Indonesia Covid-19 Vaccine Network Centrality for Vaccine for Child

topic 1	aged	complete	country	need	good	government	child	said	receive	year
degree	15	16	16	16	15	18	17	18	17	18
betweenness	0	9.3333	0	1.5	1.6416	0	0	1	0	0
closeness	0.0333	0.0588	0.0370	0.0400	0.0666	0.0500	0.0500	0.0454	0.0370	0.0476
eigen	0.6378	0.4149	0.3927	0.4453	0.0914	0.3113	0.6691	0.4219	0.6779	1.0000

Figure 9 shows the social network for the Safe Development topic, where remember dominates the discussion. Table 12 reveals that, surprisingly, among ten members, only people have the highest degree of centrality. News acts as a connector for most nodes, while safe is the most influential node, connecting to those with high centrality. Two other members have the shortest paths, helping other nodes link.

Sentiment for this topic differs from the previous one, with 48% positive, 36% neutral, and 16% negative, which can be seen in Figure 10. The optimistic sentiment stems from people's belief in the safety and readiness of the vaccine development. The vaccine must achieve 50% efficacy to be approved and undergo safety trials, which have shown that COVID-19 vaccines reduce infections and severity, and prevent death [35]. This aligns with the positive views held by most people in Indonesia.

Table 12. Indonesia Covid-19 Vaccine Network Centrality for Safe Development

topic 2	**remember**	**look**	**news**	**data**	**number**	**development**	**want**	**safe**	**police**	**people**
degree	4	8	11	11	10	6	12	8	10	16
betweenness	0.5333	4.2833	7.1666	0	0.25	0.5	1.085	1.0666	4.8333	2.5
closeness	0.0500	0.0714	0.0666	0.0555	0.0555	0.0555	0.0714	0.0500	0.0625	0.0526
eigen	0.2216	0.0201	0.0719	0.1071	0.2520	0.9956	0.0840	1.0000	0.0774	0.1587

The social network for Booster Safety, where safety dominates the discussion is shown in Figure 9. Table 13 reveals that the booster is the most powerful node, with the highest degree of centrality and eigenvector centrality. Come has the highest betweenness score, while believe serves as the shortest path to connect other nodes.

Based on Figure 10, sentiment analysis shows 36% positive, with 32% neutral and 32% negative sentiment. The positive sentiment is driven by high vaccine penetration in Indonesia, with many believing that the booster provides added protection. However, some people remain skeptical for various reasons. To achieve herd immunity, the government needs to strengthen policies and education related to vaccination programs and coverage [36].

Table 13. Indonesia Covid-19 Vaccine Network Centrality for Booster Safety

topic 3	**believe**	**study**	**need**	**come**	**protection**	**make**	**shot**	**time**	**safe**	**booster**
degree	13	12	15	15	14	15	16	17	15	18
betweenness	1.2476	3.8928	0	6.6785	4.0119	0	1.9047	0.6666	3	0
closeness	0.0344	0.0322	0.0227	0.0333	0.0322	0.0243	0.0312	0.0312	0.0322	0.0067
eigen	0.1869	0.2623	0.5294	0.2760	0.2614	0.1743	0.5953	0.3726	0.3106	1.0000

Figure 9 shows the social network for Vaccine Coverage, where the discussion is dominated by the word million. Table 14 reveals that health has the highest degree of centrality and is the most powerful node, connecting with other high-centrality nodes. Effectively holds the highest betweenness, while ministry serves as the node with the shortest path to connect other nodes.

Sentiment analysis in Figure 10 shows that 56% of people reacted neutrally, followed by 31% positive and 12% negative. In Indonesia, with its diverse population, the government has successfully achieved over 80% vaccination coverage for the second dose, including health workers, the elderly, the public workforce, and children [33]. Most reactions are neutral, though some people are optimistic, believing that widespread vaccination will bring an end to the pandemic.

Table 14. Indonesia Covid-19 Vaccine Network Centrality for Vaccine Coverage

topic 4	**million**	**effective**	**used**	**yesterday**	**ministry**	**government**	**virus**	**type**	**dosage**	**health**
degree	11	12	14	10	12	15	15	14	15	18
betweenness	1.1666	9.7873	0.25	7.5126	6.5611	0	2.4301	1.25	3.9079	1.5
closeness	0.0588	0.0666	0.0526	0.0666	0.0714	0.0434	0.0588	0.0625	0.0714	0.0555
eigen	0.1929	0.1260	0.6297	0.0368	0.8827	0.7444	0.2274	0.7249	0.3537	1.0000

The social network for the Side Effect topic, where booster dominates the discussion can be seen in Figure 9. Table 15 reveals that booster has the highest degree of centrality, betweenness, and is the most powerful node in the network, with second being the closest node. Figure 10, shows the sentiment analysis shows that 49% of the discussion is neutral, followed by 32% positive and 19% negative. For those who took the vaccine, side effects (both short-term and long-term) vary due to factors like gender, geography, and marital status. A survey of 311 respondents in Indonesia found that short-term side effects were more common, including muscle pain, headache, and nausea [37]. Despite the side effects, only 21.8% of the population in Indonesia has received the booster, and most discussions remain neutral, with some expressing optimism about the booster's effectiveness.

Table 15. Indonesia Covid-19 Vaccine Network Centrality for Side Effects

topic 5	know	second	really	body	like	fever	loading	effect	dose	booster
degree	16	16	16	16	16	16	2	16	16	18
betweenness	1.5	1.55	0	1.2	7	0	0	0	0	8
closeness	0.0075	0.0077	0.0060	0.0078	0.0070	0.0064	0.0026	0.0026	0.0044	0.0026
eigen	0.1417	0.2557	0.2020	0.2878	0.2300	0.4381	0.0011	0.5344	0.8130	1.0000

4. *Omicron-Related Discussion in Indonesia*

Figure 11 shows the social network for Indonesia regarding the Omicron virus known as the latest variant of the Covid-19 virus. These topics are as follows: **Health Protocol**: need, high, protocol, wave, number, increase, health, booster, case; **Government Measure**: level, availability, face, news, city, country, president, case, surge, government; **Health Service**: world, free, minister, said, come, delta, ministry health, variant; **Safety Extension**: doctor, suppress, community, year, government, variant, abroad, prevent spread; and **Self- Management**: said, positive, home, want, really, virus, time, like, people, delta.

Table 16 indicates that all nodes share a similar degree of centrality, while need connects nodes most effectively, and protocol provides the shortest path for linking other nodes. The sentiment analysis for this topic is 50% neutral, 31% positive, and 19% negative as shown in Figure 12. The dominance of neutral sentiment is due to government posts about COVID-19 protocols, which are widely shared on social media. A study by Muslih et al. [38] shows that Indonesian citizens are knowledgeable about COVID-19 and generally respond positively to health protocols like staying home, social distancing, and mask-wearing, with education and occupation levels influencing their attitudes.

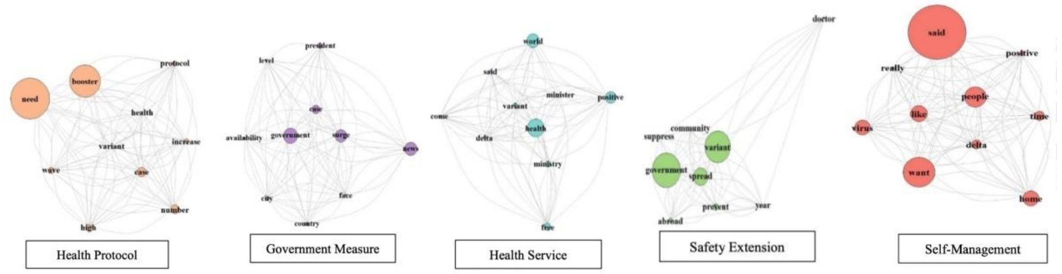

Figure 11. Indonesia Omicron Social Network

Table 16. Indonesia Omicron Network Centrality for Health Protocol

topic 1	need	high	protocol	wave	number	increase	variant	health	booster	case
degree	18	18	18	18	18	18	18	18	3.8333	0
betweenness	8	3.1666	2.8166	0	0	0	0	3.3333	0	0
closeness	0.0068	0.0074	0.0090	0.0079	0.0073	0.0066	0.0025	0.0032	0.0071	0.0040
eigen	0.2231	0.3378	0.6626	0.2977	0.5196	0.7104	0.9123	1.0000	0.3708	0.9946

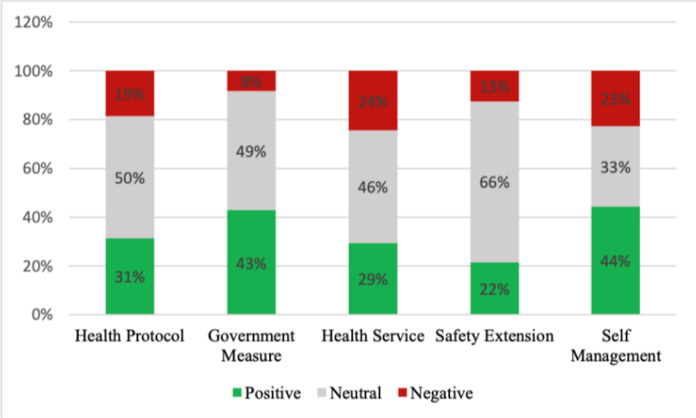

Figure 12. Indonesia Omicron Topic Sentiment

Figure 11 shows the social network for the Government Measure topic, with government, case, surge, and news dominating the discussion. Table 17 highlights that five members (city, county, case, surge, and government) have the highest degree of centrality, with government being the most influential, connecting strongly to other central nodes. Availability and country act as mediators, with the country being the closest node. Figure 12 shows the sentiment for this topic is 49% neutral, 43% positive, and 8% negative. The neutral sentiment is due to government tweets about policies, similar to the first topic. A study by Muslih et al. [38] shows that Indonesian citizens, especially women, have higher knowledge of COVID-19, which is linked to a positive attitude towards government measures, resulting in better public response to virus containment policies.

Table 17. Indonesia Omicron Network Centrality for Government Measure

topic 2	level	availability	face	news	city	country	president	case	surge	government
degree	17	14	18	16	18	18	17	18	18	18
betweenness	0	8	5	2	1.6666	8	5.3333	0	0	0
closeness	0.0128	0.0147	0.0151	0.0140	0.0140	0.0161	0.0149	0.0070	0.0082	0.0077
eigen	0.2911	0.5755	0.3791	0.0921	0.3908	0.2757	0.3782	0.6803	0.5578	1.0000

The social network for Health Service, where health, world, and positive dominate the discussion is shown in Figure 11. Table 18 indicates that five members (positive, said, come, delta, health, and variant) share equal significance in the network, with health being the most influential due to its connections with other high-centrality nodes. Free serves as a mediator and the closest node. The sentiment for this topic in Figure 12 indicates 46% neutral, 29% positive, and 24% negative. The Indonesian government announced a free telemedicine service for COVID-19 patients in five major cities, allowing consultations without visiting a hospital [39]. However, the limited availability of this service led to neutral reactions from most people, as many did not have access to or experience with it. Despite free telemedicine, paid services are still widely used, and factors like pricing, education, and geographic limitations influence people's willingness to use telehealth services [34].

Table 18. Indonesia Omicron Network Centrality for Health Service

topic 3	world	free	minister	positive	said	come	delta	ministry	health	variant
degree	15	16	15	18	18	18	18	16	18	18
betweenness	4.8333	23.5	0	6	0	0	0	0	0	0
closeness	0.0277	0.0303	0.0172	0.0270	0.0243	0.0196	0.0196	0.0149	0.0130	0.0088
eigen	0.2414	0.1175	0.4474	0.1625	0.3732	0.1746	0.5372	0.6510	1.0000	0.9189

The fourth topic, Safety Extension, is dominated by variant and spread, with the spread being the most powerful node as shown in Figure 11. Table 19 shows Broad has the highest betweenness and, along with doctor, contributes as the closest node. The sentiment as in Figure

12, indicates 66% neutral, 29% positive, and 13% negative. To address the surge in travel during Ramadan and Eid, the government implemented safety extensions, particularly in Java and Bali, requiring travelers to follow regulations regardless of transportation mode. This policy also applies to international travel [40]. The discussion on this topic was largely driven by government updates and support from Indonesian medical organizations to prevent further infections.

Table 19. Indonesia Omicron Network Centrality for Safety Extension

topic 4	doctor	suppress	community	year	government	variant	abroad	prevent	spread
degree	8	9	13	12	14	15	14	14	15
betweenness	9.038	1.138	0	4.7904	0	0	9.9619	0	0
closeness	0.0555	0.0526	0.0312	0.0476	0.0303	0.0227	0.0555	0.0303	0.0384
eigen	0.0107	0.5285	0.3812	0.2737	0.6208	0.5855	0.3871	0.7073	1.0000

Figure 11 shows the social network for Self-Management, where said and wants dominate the discussion. Table 20 reveals that all nodes share equal importance, with the home having the highest betweenness and closeness, while people are the most powerful node. Unlike the other topics, this one is primarily positive (44%), followed by 33% neutral and 23% negative as shown in Figure 12. The discussion mainly involves individuals sharing their experiences after being infected or in contact with someone who tested positive for COVID-19, many of whom managed symptoms at home. This topic is linked to the availability of telehealth services, where people consult doctors from home [34]. Additionally, the high vaccination rate in Indonesia has reduced disease severity and hospitalizations [33].

Table 20. Indonesia Omicron Network Centrality for Self-Management

topic 5	said	positive	home	want	really	virus	time	like	people	delta
degree	18	18	18	18	18	18	18	18	18	18
betweenness	5	0	24.5	0	0	1	0	0	0.0045	0
closeness	0.0078	0.0055	0.0092	0.0043	0.0052	0.0064	0.0044	0.0055	0.0045	0.0058
eigen	0.5212	0.5789	0.3449	0.6566	0.7024	0.7044	0.7388	0.8964	1.0000	0.9551

4. CONCLUSION

This study analyzed public sentiment and topic discussions related to COVID-19 vaccines and Omicron in the US and Indonesia using Twitter data. Sentiment analysis was conducted with VADER, topic modeling with LDA, and social network analysis (SNA) to examine relationships between key terms. The dataset included 637,367 tweets from the United States and 92,679 tweets from Indonesia. In the US, vaccine discussions were dominated by positive sentiment (37%), followed by negative (36%) and neutral (27%), while Omicron-related discussions were more negative (36%). In Indonesia, neutral sentiment was most prevalent (30-64%), followed by positive (25-31%) and negative (11-19%). Neutral sentiment reflected government and news posts, while positive sentiment showed optimism about vaccines and government efforts. Negative sentiment in the US stemmed largely from concerns about vaccine effectiveness and government policies. The differences in sentiment were also reflected in the topic modeling results. In the US, vaccine-related discussions focused on issues like booster availability, vaccine type, psychological impacts, and the Canada Trucker protests. Omicron-related discussions highlighted concerns about the virus's contagiousness and vaccine effectiveness. Meanwhile, in Indonesia, discussions centered on vaccine safety, coverage, and side effects, with Omicron-related discussions emphasizing health protocols, government measures, and self-management strategies. These thematic differences underscore how public sentiment and discourse were shaped by the unique cultural, social, and governmental contexts in each country.

The findings of this study not only offer insights into how pandemic-related discourse continues to influence public sentiment but also highlight opportunities for addressing ongoing health challenges. In the US, the mix of optimism and doubts about vaccines and booster effectiveness suggests a need for targeted campaigns to rebuild trust and counter misinformation. In Indonesia, the optimism surrounding vaccines reflects the success of government communication, but concerns about booster safety and health services point to areas where further engagement and education are needed. Building on these insights, several actionable recommendations can be proposed for health organizations and governments. First, tailored vaccine literacy campaigns should address the specific concerns of each population. In the US, this means focusing on misinformation about vaccine efficacy and booster safety, while in Indonesia, emphasis should be placed on improving confidence in booster programs and transparent communication about vaccine side effects. Second, social media platforms should be leveraged not only to monitor public sentiment but also to engage directly with users, providing accurate information and addressing misconceptions in real-time. Third, governments should integrate mental health resources into pandemic responses, offering support for individuals dealing with vaccine-related anxiety or skepticism. Finally, these findings highlight the need for governments to strengthen preparedness for future health crises by developing communication protocols that prioritize transparency, timely updates, and collaboration with trusted community leaders.

While this study examines historical data, its findings remain highly relevant in the current context. Lessons learned from the COVID-19 pandemic, such as managing public sentiment and combating misinformation, can be applied to current and emerging health crises, including the rise of new infectious diseases and post-pandemic mental health challenges. Additionally, this study contributes to understanding the long-term societal changes caused by the pandemic, including shifts in public trust, reliance on digital platforms for information, and evolving attitudes toward health measures. This study, however, is not without limitations. First, the analysis relies solely on Twitter data, which may not fully capture the perspectives of populations with limited access to or use of social media. This could result in a bias towards the views of more digitally connected users, potentially underrepresenting rural or older demographics. Second, the sentiment analysis using VADER may have limitations in accurately interpreting complex emotions, sarcasm, or cultural nuances, especially in multilingual datasets such as those from Indonesia. Third, the focus on historical data (January to February 2022) may limit the applicability of findings to rapidly evolving public health contexts, underscoring the need for further studies incorporating more recent data. Lastly, while this study provides valuable insights into public sentiment and topic discussions, it does not address causal relationships or evaluate the effectiveness of specific public health interventions.

Future research should build on these findings to explore how public sentiment and discourse evolve in response to new health challenges. For example, examining the psychological impacts of the pandemic and the role of trust in healthcare systems could provide valuable insights for addressing vaccine hesitancy and improving health outcomes. Additionally, studies like those analyzing sentiments toward online learning during the pandemic reveal that public attitudes toward Covid-19-related issues vary significantly based on the context of discussion, with certain topics, such as online education, showing higher levels of negative sentiment due to challenges like accessibility and fatigue [41]. By connecting past pandemic responses to future challenges, this study demonstrates the value of leveraging social media data and machine learning techniques to inform public health strategies, foster trust, and enhance resilience in the face of future crises.

REFERENCES

[1] W.H.O, “Corona Virus Disease,” WHO, 2021. [Online]. Available: https://www.who.int/health-topics/coronavirus#tab=tab_1..

[2] W.H.O, “Tracking SARS-CoV-2 Variants,” WHO, 2022. [Online]. Available: https://www.who.int/activities/tracking-SARS-CoV-2-variants..

[3] P. G. Szilagyi, K. Thomas, M. D. Shah, N. Vizueta, Y. Cui, S. Vangala, and A. Kapteyn,, “ "National trends in the US public’s likelihood of getting a COVID-19 vaccine—April 1 to December 8, 2020,",” *JAMA*, vol. 325, no. 4, pp. 396-398, 2021.

[4] P. Mondal, A. Sinharoy, and L. Su, “Sociodemographic predictors of COVID-19 vaccine acceptance: a nationwide US-based survey study,” *Public Health*, vol. 198, pp. 252–259, 2021.

[5] R. Triwardani, “Indonesian officials and media fight vaccine hesitancy, misinformation,” *Asian Politics & Policy*, vol. 13, no. 4, pp. 635–639, Oct. 2021.

[6] H. Yin, S. Yang, and J. Li, “Detecting Topic and Sentiment Dynamics Due to COVID-19 Pandemic Using Social Media,” *Lecture Notes in Computer Science*, vol. 12447, Springer, pp. 610–623, 2020.

[7] M. Saud, M. Mashud, and R. Ida, “Usage of social media during the pandemic: Seeking support and awareness about COVID-19 through social media platforms,” *Journal of Public Affairs*, vol. 20, no. 4, pp. e2417, Nov. 2020.

[8] C. Çılgın, M. Baş, H. Bilgehan, and C. Ünal, “Covid-19 Salgını Esnasında VADER ile Twitter Duygu Analizi,” *AJIT-e: Academic Journal of Information Technology*, vol. 13, pp. 72-89, 2022.

[9] V. D. Chaithra, “Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments,” *International Journal of Electrical and Computer Engineering (IJECE)*, 2019.

[10] L. Hong and B. D. Davison, “Empirical study of topic modeling in Twitter,” in *Proceedings of the First Workshop on Social Media Analytics*, vol. 10, pp. 80-88, 2010.

[11] J. W. Mohr and P. Bogdanov, “Introduction—Topic models: What they are and why they matter,” *Poetics*, vol. 41, no. 6, pp. 545-569, 2013.

[12] B. Jongman, J. Wagemaker, B. Revilla Romero, and E. Coughlan De Perez, “Early Flood Detection for Rapid Humanitarian Response: Harnessing Near Real-Time Satellite and Twitter Signals,” *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, 2015.

[13] L. Sinnenberg, A. M. Buttenheim, K. Padrez, C. Mancheno, L. Ungar, and R. M. Merchant, “Twitter as a Tool for Health Research: A Systematic Review,” *American Journal of Public Health*, vol. 107, no. 1, pp. e1–e8, Jan. 2017..

[14] M. A. Al-Garadi, M. R. Hussain, N. Khan, G. Murtaza, H. F. Nweke, and I. Ali, “Predicting Cyberbullying on Social Media in the Big Data Era Using Machine Learning Algorithms: Review of Literature and Open Challenges,” IEEE Access, vol. 7, pp. 70701 - 70718, 2019.

[15] W Q. Wong and J. Skillings, “Twitter's user growth soars amid coronavirus, but uncertainty remains,” *CNET*, Apr. 30, 2020.

[16] C. Hutto and E. Gilbert, “VADER: A Parsimonious Rule Based Model for Sentiment Analysis of Social Media Text.,” in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014.

[17] M. B. David, “Probabilistic topics models,” *Communications of the ACM*, vol. 55, pp. 77-84, 2012.

[18] S. K. Sahoo, S. K. Dash, and S. K. Rath, “Phase-based Cepstral features for Automatic Speech Emotion Recognition of Low Resource Indian languages,” *Computer Speech & Language*, vol. 73, pp. 101-312, 2022.

[19] S. P. Borgatti, “Centrality and Network Flow,” *Social Networks*, vol. 21, no. 1, pp. 55-71, 2005.

[20] M. S. Kaiser, A. Bandyopadhyay, and M. Mahmud, “ Normalized Approach to Find Optimal Number of Topics in Latent Dirichlet Allocation (LDA),” in *Proceedings of the International Conference on Trends in Computational and Cognitive Engineering*, pp. 273–281, 2020.

[21] N. Pardi, M. J. Hogan, F. W. Porter, and D. Weissman, “mRNA vaccines a new era in vacciology,” *Nature Reviews Drug Discovery*, vol. 17, no. 4, pp. 261-279, 2018.

[22] BBC, “Canada Protests: Police arrest leaders id trucker convoy,” BBC, 2021. [Online]. Available: https://www.bbc.com/news/world-us-canada-60420470..

[23] Ö. K. Kızılkurt, A. Yılmaz, C. O. Noyan, and N. Dilbaz, “Health anxiety during the early phases pg COVID-19 pamdemic in Turkey and its relationship with postpandemic attitudes hopelemess and psychhological resilience,” *Perspectives in Psychiatric Care*, vol. 57, no. 1, pp. 399-407, 2021.

[24] T. Bolsen and R. Palm, “Politicization and COVID-19 vaccine resistance in the U.S,” *Progress in Molecular Biology and Translational Science*, vol. 188, no. 1, pp. 81-100, 2022.

[25] O. Dyer, “Covid-19: Mask mandates fall across US against public health,” *BMJ (British Medical Journal* , vol. 276, 2022.

[26] New York State Department of Health, “New York State Department of Health Announces Emergence of Recently Identified, Highly Contagious Omicron Subvariants in New York and Urges Continued Vigilance Against COVID-19,” Albany, NY, USA, Press Release, Apr. 13, 2022.

[27] A. G. Johnson et al., “COVID-19 Incidence and Death Rates Among Unvaccinated and Fully Vaccinated Adults with and Without Booster Doses During Periods of Delta and Omicron Variant Emergence,” *MMWR Morbidity and Mortality Weekly Report, vol*, vol. 71, no. 4, pp. 132-138, 2022.

[28] L. C. Karlsson et al., “ Fearing the disease or the vaccine: The case of COVID-19,” *Personality of Individual Differences*, vol. 172, 2021.

[29] FDA, “COVID-19 Test Basics,” 2022. [Online]. Available: https://www.fda.gov/consumers/consumer-updates/covid-19-test-basics..

[30] V. T. Chu et al., “Comparison of Home Antigen Testing With RT-PCR and Viral Culture During the Course of SARS-CoV-2 Infection,” *JAMA Internal Medicine*, 2022.

[31] L. Wang, N. A. Berger, D. C. Kaelber, P. B. Davis, N. D. Volkow, and R. Xu, “COVID infection rates, clinical outcomes, and racial/ethnic and gender disparities before and after Omicron emerged in the US,” *MedRxiv*, 2022.

[32] J. C. O'Horo et al., “Effectiveness of Monoclonal Antibodies in Preventing Severe COVID-19 With Emergence of the Delta Variant,” *Mayo Clinic Proceedings*, vol. 97, no. 2, pp. 327-332, 2022.

[33] Kemenkes, “Vaksinasi COVID-19 Nasional,” Kemenkes, 2022. [Online]. Available: https://vaksin.kemkes.go.id/#/vaccines..

[34] R. K. Sinuraya, A. S. W. Kusuma, Z. E. Pardoel, M. J. Postma, and A. A. Suwantika, “Parents' Knowledge, Attitude, and Practice on Childhood Vaccination During the COVID-19 Pandemic in Indonesia,” *Patient Prefer Adherence*,, vol. 16, pp. 105-112, 2022.

[35] A. Ditamy et al., “The Effectiveness, Side Effects, and Implementation Between Variation of COVID-19 Vaccines in Indonesia,” in *Proceedings of the 3rd Tarumanagara International Conference on the Applications of Social Sciences and Humanities (TICASH 2021)*, 2021.

[36] J H. Juliani, K. C. S. Wibawa, and Solechan, “COVID-19 Vaccine Policy as an Effort to Achieve National Herd Immunity in Indonesia,” *Pakistan Journal of Medical & Health Sciences*, vol. 16, no. 3, pp. 492-494, 2022.

[37] J. Wibowo et al., “Factors associated with side effects of COVID-19 vaccine in Indonesia,” *Clinical and Experimental Vaccine Research*, vol. 11, no. 1, pp. 85-95, 2022.

[38] M. Muslih, H. D. Susanti, Y. A. Rias, and M. H. Chung, “Knowledge, Attitude, and Practice of Indonesian Residents toward COVID-19: A Cross-Sectional Survey,” *International Journal of Enviromental Resource Public Health*, vol. 18, no. 9, 2021.

[39] Setkab, “Health Ministry Expands Telemedicine Services to Jakarta’s Satelit Cities,” 2021. [Online]. Available: https://setkab.go.id/en/health-ministry-expands-telemedicine-services-to-jakartas-satellite-cities/..

[40] Directorate General of Immigration, “Information on immigration regulations during the Covid-19 Pandemic,” 2022. [Online]. Available: https://www.imigrasi.go.id/en/covid19/covid19-1/.

[41] J. M. S. Waworundeng, G. A. Sandag, R. A. Sahulata, and G. D. Rellely, “Sentiment Analysis of Online Lectures Tweets using Naïve Bayes Classifier,” *Cogito Smart Journal*, vol. 8, no. 2, pp. 341-384, 2022.