

Health Organizations and Public Interaction on Instagram: A Study of Topics, Emotions, and Sentiment

Abigail Paradise Vit, Avi Magid

Submitted to: JMIR Infodemiology
on: December 25, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	36
Figures	37
Figure 1.....	38
Figure 2.....	39
Figure 3.....	40
Figure 4.....	41
Figure 5.....	42
Figure 6.....	43
Figure 7.....	44
Figure 8.....	45
Figure 9.....	46
Figure 10	47
Multimedia Appendixes	48
Multimedia Appendix 1.....	49
Multimedia Appendix 2.....	49

Health Organizations and Public Interaction on Instagram: A Study of Topics, Emotions, and Sentiment

Abigail Paradise Vit¹; Avi Magid^{2,3}

¹Emek Yezreel The Max Stern Emek Yezreel College Emek Yezreel IL

²Management Rambam Healthcare Campus Haifa IL

³Department of International Health Maastricht University Maastricht NL

Corresponding Author:

Abigail Paradise Vit

Emek Yezreel

The Max Stern Emek Yezreel College

Department of Information Systems

Emek Yezreel

IL

Abstract

Background: Social media is an essential tool of Health organizations, who use this platform to deliver evidence-based information, educate the public, dispel myths, and promote a healthier, more informed society.

Objective: To categorize health organizations' content on social media into topics, examine public engagement, sentiment and emotional responses to these topics and identify gap in fear between health organizations' messages and public response.

Methods: Real data was collected from the official Instagram accounts of health organizations worldwide. The BERTopic algorithm for topic modeling was used to categorize health organizations' posts into distinct topics. For each identified topic, we analyzed the engagement metrics (number of comments and likes) of posts categorized under the same topic, calculating the average engagement received. Additionally, we examined the sentiment and emotional content of both posts and responses within the same topic, providing insights into the distributions of sentiment and emotions for each topic. Particular attention was paid to the presence of the fear emotion in the posts and responses.

Results: A total of 6,082 random posts and 82,982 comments were collected from the official Instagram accounts of eight health organizations. The study reveals that topics related to COVID-19, vaccines, and humanitarian crises (such as the Ukraine conflict and the war in Gaza) generated the highest engagement. Our sentiment analysis of the responses to health organizations' posts showed that topics related to vaccines and Monkeypox generated the highest percentage of negative responses. Fear was the dominant emotion expressed in the posts text, while the public's responses showed more varied emotions, with anger notably high in discussions around vaccines. Gaps were observed between the level of fear conveyed in posts published by health organizations and the in the fear conveyed in public's responses to such posts especially regarding mask-wearing during COVID-19 and the influenza vaccine.

Conclusions: This study underscore the importance of transparent communication that consider the emotional and sentiment-driven responses of the public on social particularly regarding vaccines. Understanding the psychological and social dynamics associated with public interaction with health information online can help health organizations achieve public health goals, fostering trust and efficiency, countering misinformation, and promoting informed health behavior.

(JMIR Preprints 25/12/2024:70576)

DOI: <https://doi.org/10.2196/preprints.70576>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to the public.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>, I will be able to make my accepted manuscript PDF available to anyone.

No. Please do not make my accepted manuscript PDF available to anyone.



Original Manuscript

Original Paper

Health Organizations and Public Interaction on Instagram: A Study of Topics, Emotions, and Sentiment

Abstract

Background: Social media is an essential tool of Health organizations, who use this platform to deliver evidence-based information, educate the public, dispel myths, and promote a healthier, more informed society.

Objective: To categorize health organizations' content on social media into topics, examine public engagement, sentiment and emotional responses to these topics and identify gap in fear between health organizations' messages and public response.

Methods: Real data was collected from the official Instagram accounts of health organizations worldwide. The BERTopic algorithm for topic modeling was used to categorize health organizations' posts into distinct topics. For each identified topic, we analyzed the engagement metrics (number of comments and likes) of posts categorized under the same topic, calculating the average engagement received. We examined the sentiment and emotional content of both posts and responses within the same topic, providing insights into the distributions of sentiment and emotions for each topic. Particular attention was paid to the presence of the fear emotion in the posts and responses. Additionally, a linguistic analysis and an analysis of sentiments and emotions over time were conducted.

Results: A total of 6,082 posts and 82,982 comments were collected from the official Instagram accounts of eight health organizations. The study reveals that topics related to COVID-19, vaccines, and humanitarian crises (such as the Ukraine conflict and the war in Gaza) generated the highest engagement. Our sentiment analysis of the responses to health organizations' posts showed that topics related to vaccines and Monkeypox generated the highest percentage of negative responses. Fear was the dominant emotion expressed in the posts text, while the public's responses showed more varied emotions, with anger notably high in discussions around vaccines. Gaps were observed between the level of fear conveyed in posts published by health organizations and in the fear conveyed in public's responses to such posts especially regarding mask-wearing during COVID-19 and the influenza vaccine.

Conclusions: This study underscores the importance of transparent communication that consider the emotional and sentiment-driven responses of the public on social media particularly regarding vaccines. Understanding the psychological and social dynamics associated with public interaction with health information online can help health organizations achieve public health goals, fostering trust, countering misinformation, and promoting informed health behavior.

Keywords: Emotion Analysis; Fear; Health Communication; Healthcare; Instagram; Official Health Organizations; Sentiment Analysis; Social Media; Vaccines

Introduction

Social networks are used by billions of people around the world, making them an effective platform for reaching a wide range of audiences. Health organizations, such as the World Health Organization (WHO), use social networks to disseminate important health-related information [1-5]; provide real-

time updates, news, and emergency guidelines [6]; and promote awareness of diseases [7] and mental health [8]. Health organizations also promote vaccination compliance by disseminating information regarding vaccination importance, safety, and disease severity [9].

In contrast to the content published by social network users unaffiliated with health organizations, who can spread a large amount of inaccurate, false, and misleading information about health-related issues [10,11], the health-related content published on social networks by official health organizations enables the public to have access to reliable and useful information [12-14]. Therefore, health organizations are important actors in social media as reliable sources, providing evidence-based and authoritative information [15]. Through their efforts, these organizations educate the public, dispel myths, and help the community become healthier and more informed [16].

In addition to serving as a means of disseminating health-related information, social networks provide the public with the opportunity to participate in discussions and conversations with health organizations [17] by posting comments, sharing posts and liking posts. Members of the public can also ask questions about the health issue being discussed, and the organizations that disseminate the information can respond. The public's active participation can enhance individuals' understanding of health-related content [18, 19] and can foster public trust in and appreciation for science [20]. In addition, health organizations are able to pinpoint concerns related to specific health topics, gain insight into public opinion [21-25][17-21], and identify topics that result in misinformation [26].

However, health-related messages disseminated by health organizations can provoke negative public reactions [27]. When combined with contradictory messages from unaffiliated social media users, this can lead to undesired health behaviors, such as vaccine hesitancy, noncompliance with health directives, and diminished trust in the reliability of health organizations [28-30].

Understanding public emotions in response to information disseminated by health organizations on social media is important for assessing the effectiveness of health communication strategies [31]. Sentiment and emotion analysis are widely utilized for determining sentiment polarity and detecting specific emotions expressed in textual data [32]. Sentiment analysis categorizes a text as positive, negative, or neutral, while emotion analysis identifies specific emotions expressed within a text [32]. Fundamental emotions such as happiness, sadness, anger, disgust, surprise, and fear can be detected, along with more nuanced emotions like confusion and trust [33,34]. Among these emotions, fear is an important factor in health communication, as it can influence public perception, engagement, and behavioral responses [35].

Various theoretical models, such as the Extended Parallel Process Model (EPPM) [36], have explored how fear is utilized in public health messaging to encourage protective behaviors. This model suggests that individuals respond to fear-based messages depending on their perceived threat level and their sense of efficacy in managing the threat. According to this model, when perceived risk is high and the message also offers a clear solution, individuals are more likely to engage positively and adopt protective behaviors. However, if either condition is not met, responses may be negative, leading to outcomes such as avoidance or denial [36]. Therefore, fear is a crucial factor to consider in public health communication.

Despite the need to examine how the public emotionally responds to information shared by health organizations on social media, studies examining the topics communicated by these organizations and the corresponding responses from the public are sparse. The aim of this study is to analyze and characterize the content disseminated by health organizations on social media into topics, as well as social media users' responses to this content and the engagement, sentiment and emotions induced by

this content. Additionally, we aim to identify gaps in fear between health organizations' messages and public responses.

Related

work

In this section, we provide an overview of various studies associated with health-related content, the topic modeling and sentiment analysis of such content.

Sentiment analysis of health-related social media content

Many studies examined the sentiment and public opinion surrounding the COVID-19 vaccine on Twitter [21,23,24,37-53]. For example, an analysis of Twitter data was conducted by Niu et al [24] to examine public opinion and sentiment before and during the administration of the COVID-19 vaccines in Japan. They found that negative sentiment toward the vaccines dominated positive sentiment in Japan and concerns about side effects may have outweighed fears of infection at the beginning of the vaccination process.

Numerous studies leveraged machine learning techniques to classify tweets as positive, negative, or neutral sentiment towards vaccines, enabling the identification of vaccine hesitancy among communities and social media users [35,54-68]. The majority of these studies collected data from Twitter using keywords or hashtags related to vaccinations. Chakraborty et al [56] used deep learning to analyze 226,668 COVID-19 tweets from Dec 2019–May 2020, achieving 81% accuracy. Most tweets showed positive or neutral sentiment, while highly retweeted posts were predominantly neutral or negative.

Topic modeling for health-related social media content

Several studies have utilized topic modeling to examine health-related discussions on social media, focusing on topics such as blood donation [69], cancer-related content [70], and vaccine-related conversations [71]. Paul and Dredze [72] proposed the ailment topic aspect model to identify health topics on Twitter. Analyzing 144 million tweets, they identified 13 topics linked to seasonal influenza, allergies, temporal surveillance, and obesity-related geographic data in the U.S.

Seltzer et al. [73] analyzed 500 Instagram images tagged #zika from May–August 2016, analyzing them by sentiment, content, and engagement. 299 images focused on health and 193 on public interest. Sentiments and emotion analysis revealed that fear and negative emotions were linked to Zika transmission and response uncertainty. The study highlighted Instagram's value in understanding public sentiment and addressing gaps in health communication. Also in Instagram, Muralidhara et al [74] collected 96,426 Instagram posts between September and October 2016 using 269 health-related hashtags. Polylingual topic modeling approach was used to identify 47 health-related topics spanning ten broad categories: acute illness, alternative medicine, chronic illness and pain, diet, exercise, health care and medicine, mental health, musculoskeletal health and dermatology, sleep, and substance abuse. Kim et al [75] analyzed 96,302 Instagram photos and 513,694 comments with antivaccination hashtags, focusing on photo features, engagement and sentiment. Most photos (52.24%) were categorized as "text." "Food" and "plant" photos received the most positive comments, while "text" photos, despite high engagement, received less positive responses.

Other studies focused on dividing vaccine content on social networks into topics [40,41,43,45,52,53,76]. The study by Kwok et al [76] examined tweets of Australian users regarding

COVID-19 vaccination on Twitter. Using a latent Dirichlet allocation (LDA) topic model, they identified three commonly discussed topics: attitudes towards COVID-19 and vaccination, advocacy for infection control measures against COVID-19, and misconceptions and complaints regarding COVID-19. Also using LDA, Lyu et al [40] analyzed COVID-19 vaccine discussions on Twitter, identifying 16 topics grouped into five themes. Vaccination opinions were the most discussed topic. Emotion analysis showed trust as the dominant emotion, followed by anticipation, fear, and sadness. Additionally, Chandrasekaran et al [43] used the CorEx topic modeling algorithm to examine COVID-19 vaccine-related tweets. The authors identified 16 topics in the COVID-19 vaccination tweets, which were grouped into six broader themes. Most tweets regarding COVID-19 vaccination centered on vaccine policy, vaccine hesitancy, and postvaccination symptoms and side effects.

Analysis of healthcare providers' content and public responses on social media

Several studies have examined content published on social networks by healthcare providers. Among them, Kim and Kim [77] analyzed 1,545 Instagram photos published by the Centers for Disease Control and Prevention (CDC) and public comments using Microsoft Azure Cognitive Services. It was found that text and people type of photos dominated, but larger faces and gaudy features reduced engagement. Happiness and neutral emotions in comments were negatively correlated with interaction levels. Pinto et al [78] analyzed 632 Instagram posts from Portugal's National Health Service (NHS) and Brazil's Ministry of Health (MH) in 2019, mapping 53 topics for NHS and 63 for MH. NHS emphasized healthy eating and blood donation, while MH focused on vaccination campaigns, dengue prevention, and HIV awareness.

Mello et al [79] examined how the WHO and the CDC used Instagram in 2020 to communicate COVID-19 risks by analyzing 726 posts. Focusing on threat and efficacy messaging, cues to action, and credibility cues. According to the findings, efficacy messages, such as those promoting preventive behaviors, were more prevalent, while threat messages addressing the susceptibility and severity of COVID-19 were less common. The study concluded that enhancing credibility cues, visuals, tailoring content to diverse audiences, and leveraging Instagram's interactive features, public health communication could improve engagement, trust, and impact.

Vaghefi et al [80] analyzed healthcare providers messages on Twitter from May 2018 to May 2019 using machine learning, including BERT-based models, to classify tweets as professional communications or health-related information, further categorizing them as fear-based or hope-based. Results showed that fear-based messages were widely shared by the public but were less effective at motivating constructive health actions, while hope-based messages resonated more with healthcare providers. While this study [80] examined healthcare providers' content and the fear evoked by their messages on Twitter, our study focuses on categorizing healthcare providers' posts into distinct topics, analyzing engagement metrics, sentiment, and emotional content within each topic, and identifying the gaps between the fear expressed in messages from health organizations and the fear observed in public responses to such content. In contrast, Vaghefi et al [80] did not analyze the public comments to measure fear in relation to healthcare providers' messages, instead focusing on public interactions, such as retweets and replies, to study information propagation.

Methods

In this section, we describe the proposed methodology. The four main phases of the methodology are illustrated in Figure 1) defining the targeted health organizations, 2) gathering data from the selected

health organization accounts, 3) performing topic modeling on health organization content, and 4) analyzing the sentiment and emotion associated with each topic and the public's response to it. Detailed descriptions of each phase are provided in the subsections that follow.

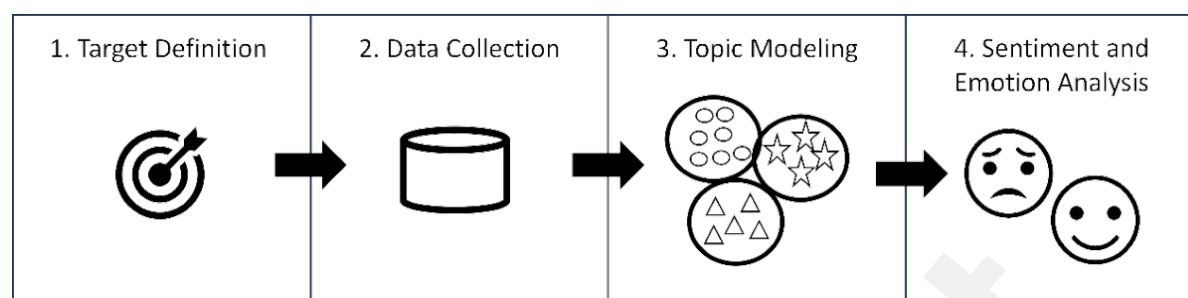


Figure 1 - An overview of the methodology's main phases.

Target Definition

The first phase of our methodology is to define the group of health organizations to be examined. These organizations were selected based on their status as official government entities or government-supported health agencies responsible for public health policies, research, and regulations, ensuring that the study focuses on authoritative sources [81-88]. Additionally, a minimum threshold of 10,000 followers was set for health organization accounts to ensure sufficient activity and public engagement.

A brief description of each organization is provided below:

1. The World Health Organization (WHO) - The WHO is the United Nations agency tasked with connecting nations around the world in order to promote health, keep the world safe, and serve the vulnerable [81]. The organization's official Instagram account is named WHO.
2. Department of Health and Human Services (HHS) – The Department of HHS aims to enhance the health and well-being of the residents or citizens of the U.S, by providing effective health and human services and fostering sound, sustained advances in the sciences underlying medicine, public health, and social services [82]. HHSgov is the official Instagram account of the HHS.
3. The Office of Minority Health (OMH) - The Office of Minority Health is part of the HHS Services dedicated to improving the health of racial and ethnic minorities. The OMH fulfills its commitment to improving the health of racial and ethnic minorities in large part by developing health policies and programs that help eliminate health disparities [83]. MINORITYHEALTH is the official Instagram account of the OMH.
4. National Institutes of Health (NIH) - The National Institutes of Health is the leading federal agency in the U.S responsible for conducting and supporting medical research. The NIH, which is part of the HHS, is one of the world's most prominent centers for medical research. Its mission is to enhance human health by advancing research across a wide range of scientific disciplines [84]. The organization's official Instagram account is named NIHgov.
5. The National Institute of Mental Health (NIMH) - The National Institute of Mental Health (NIMH) is the U.S agency responsible for research on mental health. Its primary objective is to understand, treat, and prevent mental illness by conducting basic and clinical research. The NIMH is one of the 27 institutes and centers that comprise the NIH, which is part of the HHS. [85]. NIMHgov is the official Instagram account of the NIMH.
6. The Centers for Disease Control and Prevention (CDC) – The CDC is a science-based, data-driven organization that leads the United States' efforts to protect the public's health. The CDC is one of the major components of the U.S. Department of Health and Human Services, and it aims to protect the residents of the U.S from health, safety, and security threats, both foreign and domestic. [86]. CDCgov is the official Instagram account of the CDC.

7. National Health Service (NHS) - The National Health Service was established as the public healthcare system of the United Kingdom. It is one of the largest and most comprehensive healthcare systems in the world. The NHS provides various types of services, including mental health services, general practitioners (GPs), hospitals, and treatment facilities. The NHS is dedicated to improving the overall health of the population by promoting public health initiatives, health education, vaccination programs, and disease prevention campaigns [87]. The organization's official Instagram account is named NHS.
8. The U.S Food and Drug Administration (FDA) - The FDA is a governmental regulatory agency responsible for protecting public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices [88]. The organization's official Instagram account is named FDA.

It's important to note that, since we selected health organizations with a high number of followers, the resulting sample skewed toward U.S.-based organizations. The majority of accounts originated from the United States, with one from the UK and one global organization (WHO). This concentration may affect the generalizability of the findings, as both the published content and the public responses are likely influenced by U.S.-specific health priorities and cultural context.

Data Collection

The data collection process and analysis was approved by our institution's review board: the Emek Yezreel College Ethical Review Board (approval number 2023-81 YVC EMEK).

We chose Instagram as the social media platform. We chose Instagram for its visual and interactive nature, which allows health organizations to share information and engage with the public effectively [89]. We searched for the official Instagram accounts of healthcare organization using Instagram's search box. For each organization, Table 1 lists the name of each official Instagram account, along with the total number of posts published since the account's creation and the number of followers, both recorded as of August 2024. As can be seen in the table, the CDC, WHO, and OMH are the organizations that publish the most posts. The WHO's account has the largest number of followers.

Table 1 – Statistics for the health organizations' Instagram accounts, ranked by number of followers (highest to lowest).

Organization name	Instagram account name	Number of published posts	Number of followers
CDC	CDCGOV	6,359	2.5M
WHO	WHO	3,893	12M
NHS	NHS	695	564K
HHS	HHSGOV	3,269	202K
FDA	FDA	801	122K
NIMH	NIMHGOV	674	64.1K
NIH	NIHGOV	1,867	279K
OMH	MINORITYHEALTH	3,909	15.4K

In order to collect the health organizations' Instagram posts, we connected with the Instagram API using the RapidAPI website. This website is a large API hub that allowed to connect with tens of thousands of public REST APIs over the Internet.

The posts were collected between April 7, 2017 and November 17, 2023. Each Instagram post

includes the publication date, the number of likes, the number of comments, text, and a photo. For each post, we collected the comments, including the publication date, the comment text, and the number of likes. A total of 6,082 posts and 82,982 comments were collected using the Instagram API. All retrieved posts and comments were included in the analysis. Table 2 presents the relevant statistics for the collected posts and comments for each health organization's Instagram account. We analyzed all the posts that the API allowed us to retrieve, without applying selection criteria or filtering specific posts. Comments were collected only from the original posts, excluding replies to comments.

Table 2 – Statistics for the collected posts and comments of each health organization's Instagram account.

Name of health organization	Number of collected posts	Number of collected comments	Average number of likes
CDC	4,298	58,471	3,133.940
WHO	527	15,448	26,612.700
NHS	277	3,826	2,021.760
HHS	116	896	265.584
FDA	85	565	289.166
NIMH	301	1,096	250.281
NIH	253	1,987	1,208.972
OMH	225	603	41.476

Topic Modeling

Text from the posts published by health organization accounts was analyzed using BERTopic [90]. Note that we removed the names of the health organizations and their abbreviations from the text in order to avoid creating topics around each organization. Stop words were also removed from the text.

We used the C_v coherence score to evaluate the quality and interpretability of the topics produced by the model. This measure combines the indirect cosine measure with normalized pointwise mutual information (NPMI) and a boolean sliding window [91]. The coherence score indicates how closely related and coherent the words are in a topic. Using the average coherence metric, which is the average of the coherence metrics within each topic, we measured the model's ability to generate coherent and meaningful topics. This score enabled us to evaluate the overall effectiveness of the model in producing topics with a high degree of semantic similarity.

We performed the six steps of BERTopic to analyze the posts:

1. Embedding tweets - In this step, the text in the posts is converted into numerical representations using a sentence-transformers model named All-MiniLM-L6-v2 [92].
2. Dimensionality reduction - Uniform Manifold Approximation and Projection (UMAP) [93] is used to reduce the dimensionality of the embedded text, with the following parameters: *_neighbors = 15, n_components = 3, min_dist = 0, metric = 'cosine.'*
3. Cluster tweets - The text is grouped into clusters using the HDBSCAN density-based clustering technique [94], with the following parameters: *min_cluster_size=40, metric="euclidean", cluster_selection_method='eom.'*
4. Word frequency analysis in clusters - The frequency of each word in each cluster is determined at

the cluster level.

5. Topic representation - To represent the topics in the Instagram posts and responses, TF-IDF (term frequency-inverse document frequency) was adapted to work on a cluster or topic level instead of a tweet level. A new TF-IDF representation is used called c-TF-IDF (class-based term frequency-inverse document frequency).
6. Outlier reduction - HDBSCAN identifies texts that are outliers, meaning they do not belong to any of the established topics. To address this, we calculate the c-TF-IDF representation for each outlier text and compare its cosine similarity with the c-TF-IDF representations of the existing topics. By associating outlier texts with the closest matching topic based on similarity, we minimize the number of texts classified as outliers.

Sentiment and Emotion Analysis

Having categorized the health organizations' posts into topics, we analyzed the emotions and sentiment of each post and comment associated with a particular topic. Sentiment analysis was conducted using *distilbert-base-multilingual-cased-sentiments-student* which achieved an average accuracy of 0.808 on test set [95]. This is a distilled version of a zero-shot classification pipeline trained on the Multilingual Sentiment dataset. Zero-shot classification is a machine learning technique that allows models to classify data into categories they have never encountered during training without requiring labeled examples to be provided. It accomplishes this by leveraging contextual understanding and semantic relationships between seen and unseen classes, often by utilizing embeddings or natural language models [96]. In this case, a larger "teacher" model, MoritzLaurer/mDeBERTa-v3-base-mnli-xnli was used to train a smaller "student" model, *distilbert-base-multilingual-cased*. Using this distillation process, the student model maintains high classification performance while being more efficient and lightweight. According to the training log, the student model achieved an impressive agreement rate of 88.29% with its teacher model.

Emotion analysis of six basic emotions (fear, anger, disgust, joy, sadness, surprise) was conducted using a fine-tuned checkpoint of the DistilRoBERTa-base model called *j-hartmann/emotion-english-distilroberta-base*. The model, which was trained on six diverse datasets [97]. The model was trained on a balanced subset from several datasets of nearly 20k observations in total. 80% of this balanced subset is used for training and 20% for evaluation. The evaluation accuracy was 66%.

To ensure the accuracy of the models in identifying sentiments and emotions, we randomly selected 100 posts and comments. We manually classified them based on their dominant sentiment and dominant emotion. The results showed that the sentiment model achieved 92% accuracy, while the emotion model demonstrated 84% accuracy. Considering sentiment models predict binary classification, and emotion models face greater complexity due to emotions' multidimensional nature, the results are logical. Therefore, applying the models to the data is expected to provide sufficiently reliable outcomes.

Each post and posts' comments received a sentiment score of positive, negative, or neutral, as well as an emotion score of fear, anger, disgust, joy, sadness, or surprise.

For each topic, we calculated the following:

1. Average number of comments: the average number of comments for all posts in the topic.
2. Average number of likes: the average number of likes for all posts in the topic.
3. Average post sentiment scores: the average positive, negative, and neutral sentiment scores for all posts in the topic.
4. Average post emotion scores: the average fear, anger, disgust, joy, sadness, and surprise scores

for all posts in the topic.

5. Average comment sentiment scores: the average positive, negative, and neutral sentiment scores for all post comments in the topic.
6. Average comment emotion scores: Average fear, anger, disgust, joy, sadness, and surprise scores were calculated for all post comments in the topic.
7. Gap: The difference between the average post fear score to the average comment fear score.

Linguistic Analysis

The objective of this analysis was to identify the most significant phrases used by health organizations in posts that resonated more positively with the public (i.e., associated with higher positive sentiment in comments) compared to those that elicited more negative reactions (higher negative sentiment in comments).

For this purpose, we calculated the average sentiment score (positive and negative) of all comments related to each post. Posts were then categorized based on the dominant sentiment (positive or negative) derived from these average sentiment scores in comments, resulting in two groups: posts with predominantly positive comments and posts with predominantly negative comments.

For each group of posts, we extracted the top 50 most significant phrases, including single words (unigrams), two-word phrases (bigrams), three-word phrases (trigrams), and four-word phrases (fourgrams). The preprocessing involved removing hashtags from the text, eliminating stopwords, and calculating term importance using TF-IDF. TF-IDF estimates the significance of terms within a group of posts based on their frequency within each post as compared to the frequency across all posts in the group. The cumulative TF-IDF scores for each term were calculated by summing across all posts in the group, enabling the identification of the most significant phrases for each group of posts.

Time Analysis

In the time analysis section, we examined the average positive and negative sentiment in both comments and posts over the years. Additionally, we analyzed the average levels of various emotions, including fear and anger as expressed in both posts and comments over the period of time collected. Further, we decided to calculate additional emotions of trust, disappointment, and confusion as well. The purpose of including these emotions was to extend the analysis beyond the six basic emotions, providing a broader range of emotional insights.

To calculate trust, we utilized the *ayoubkirouane/BERT-Emotions-Classifier* model [98], a fine-tuned BERT-based model designed for multi-label emotion classification. This model was trained on the *sem_eval_2018_task_1* dataset. This model includes a wide range of emotions such as anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, and trust.

For confusion and disappointment emotions, we used the *SamLowe/roberta-base-go_emotions* model [99], which is trained on the *go_emotions* dataset for multi-label classification. The *go_emotions* dataset, based on Reddit data, contains 28 emotion labels. The model achieved high performance in identifying confusion (with an accuracy of 0.972) and disappointment (with an accuracy of 0.974).

As part of the temporal analysis, we incorporated weekly COVID-19 mortality data in the U.S using publicly available dataset published by the National Center for Health Statistics [100]. The analysis included calculating a "COVID-19 death ratio" which represents the proportion of U.S COVID-19-related deaths to the total number of U.S deaths. We applied a rolling average smoothing technique

across a 10-week window to reduce noise and variability in the data, allowing us to identify clearer trends and patterns over time. This approach allowed us to observe trends in mortality data alongside our sentiment and emotional analyses, providing valuable insights into the alignment between public responses and real-world outcomes.

Results

Topic Modeling Results

Using BERTopic, 36 topics were identified for the posts of the eight health organization Instagram accounts. The average coherence score was 0.7374. A health expert reviewed the list of topics and suggested that we combine related topics.

The following topics were therefore combined: seven topics related to COVID-19 were grouped together, two topics related to vaccines for children were grouped together, two topics related to monkeypox were grouped together, and two topics related to booster vaccines were combined. Combining the topics resulted in a total of 25 topics, with an average coherence score of 0.7298. The health expert assigned a representative name to each topic. The topic names were assigned by analyzing the 10 most significant words for each topic, as determined by the c-TF-IDF scores (see the methods section for more information about c-TF-IDF) from the BERTopic algorithm and cross-referencing them with example posts related to that topic. Each topic was automatically assigned a unique number starting from 0, following the default numbering convention used by the BERTopic algorithm, and Table 3 lists the topics, indicating their number, name, and number of posts. We present in Multimedia Appendix 1 the five words with the highest c-TF-IDF for each topic.

It is important to note that the words may include hashtags, which is why some words appear without spaces, such as Mentalhealth or Publichealth.

Table 3 – The number, name, and size of each topic.

Topic number	Topic name	Topic size
0	COVID-19	1,446
1	Mental Health	752
2	Children and Vaccines	496
3	COVID-19 Booster Vaccine	288
4	Pregnant	246
5	Foodborne	235
6	Research	226
7	Community Health	223
8	Flu	216
9	Cancer	198
10	Monkeypox	197
11	Cardiovascular Diseases	193
12	Public Health	183
13	Climate	148
14	Sepsis	141
15	Masks	127
16	Vector	122
17	Health Equity	108
18	Antibiotics	108
19	Ebola	92
20	Humanitarian	91

21	Smoking	85
22	RSV	59
23	Sun Damage	52
24	Noise Damage	50

Engagement Analysis Results

Our dataset includes the following engagement information for each post: the number of likes and comments provided by users.

We calculated the average number of comments and likes for all posts in each topic. Figure 2 presents the average number of comments received per topic, while Figure 3 displays the average number of likes. Among the topics, those with the highest user engagement based on the average number of post comments were humanitarian issues, masks, and COVID-19. In terms of engagement measured by the average number of post likes, the leading topics were humanitarian issues, masks, and cancer. Both figures also showed increased engagement for topics related to vaccines and COVID-19.

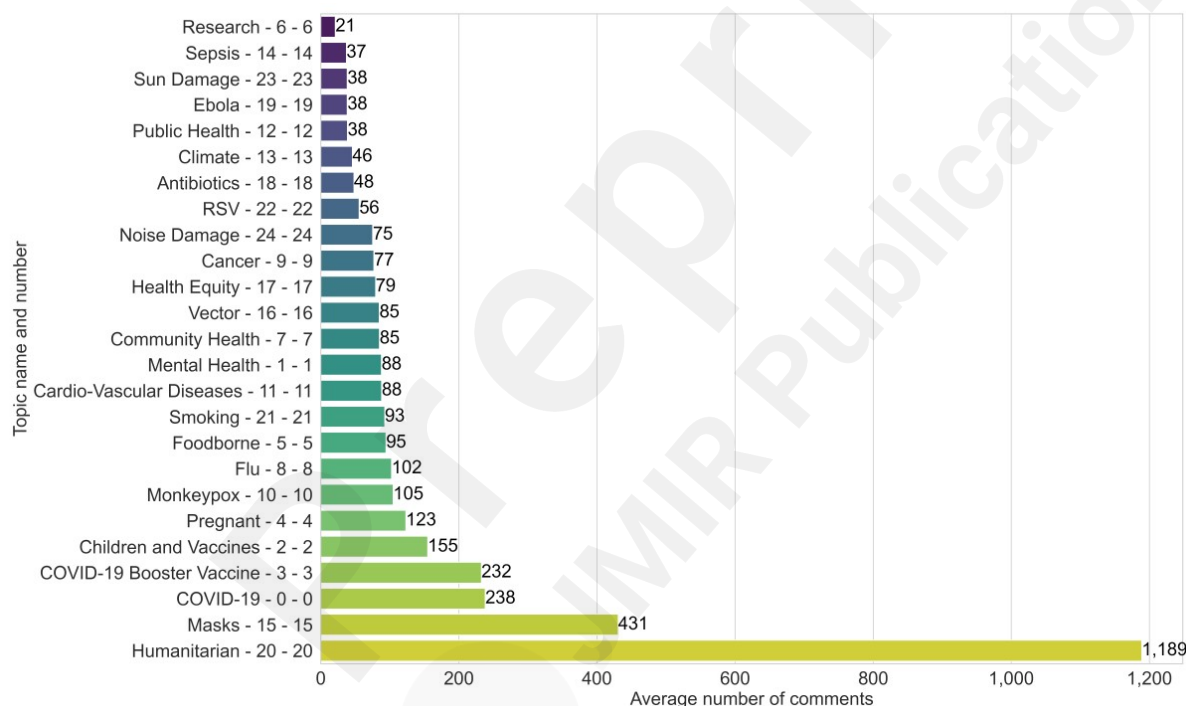


Figure 2 – Average number of post comments per topic.

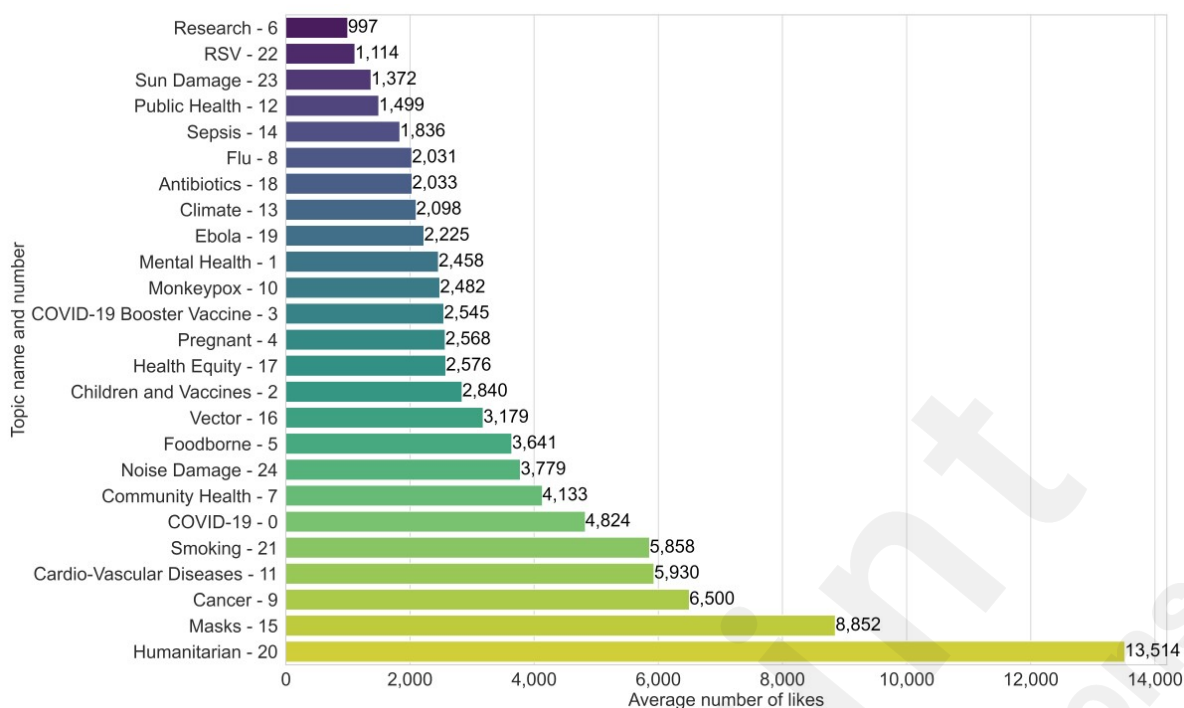


Figure 3 - Average number of post likes per topic.

Sentiment Analysis Results

We analyzed the sentiment in the text of the posts and comments and calculated the sentiment scores of the posts and comments as described in the methods section.

Figures 4 and 5 present the average sentiment scores for health organizations' posts and comments, respectively. In both figures, we selected only the 15 largest topics (those containing the greatest number of posts). In Figure 4, it can be seen that there are certain topics, such as sepsis, climate, and foodborne, in the health organizations' posts where negative sentiment dominates all other sentiment. However, in Figure 5, which presents the scores for the comments to the health organizations' posts, we see that there are more topics where negative sentiment is dominant than in Figure 4, with the greatest negative sentiment found in the comments for posts about vaccines, specifically for posts in the COVID-19, children and vaccines, booster vaccines, and monkeypox topics.

While certain topics show a dominant positive sentiment in the posts published by health organizations (as illustrated in Figure 4), this does not necessarily reflect how the public responds to those posts. The dominance of positive sentiment in these topics is based on the average sentiment scores, where the positive score was higher than both the negative and neutral scores. This indicates that these subjects were presented mostly positively by the organizations. However, when examining the sentiment expressed in user comments (Figure 5), we observe a contrasting trend. In topics such as COVID-19, children and vaccines, and the COVID-19 booster vaccine, the comments exhibit predominantly negative sentiment, even though the organizations framed these topics positively. In these cases, the average negative score in the comments was higher than both the positive and neutral scores. In other words, while the organizations try to communicate these topics in a positive light, the public's reaction to them was largely negative.

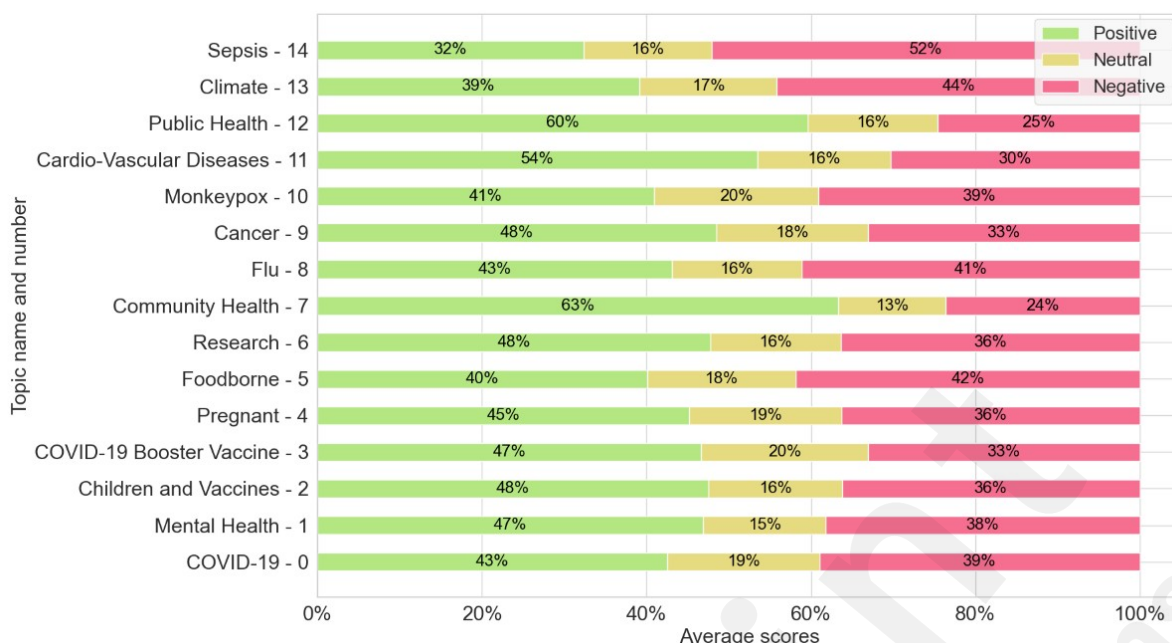


Figure 4 –

Posts' sentiment scores (average positive, neutral, negative) for each topic.

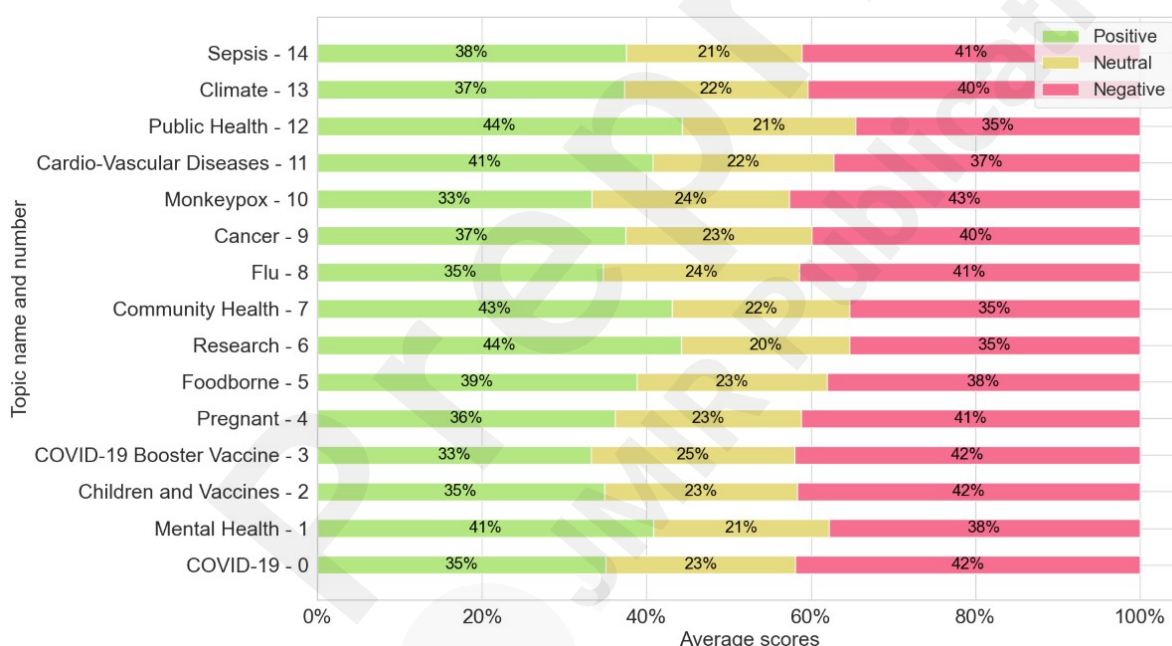


Figure 5 –

Comments' sentiment scores (average positive, neutral, negative) for each topic.

Emotion Analysis Results

We analyzed the emotion in the text of the posts and comments and calculated the emotion scores of the posts and comments as described in the methods section.

The emotions present were anger, disgust, fear, joy, neutral, sadness, and surprise. Figure 6 presents the average emotion scores for the posts, and Figure 7 displays the average emotion scores for the comments. In both figures, we selected only the 15 largest topics (those containing the greatest number of posts).

As seen in Figure 6, the dominant emotion in all topics is fear, and its scores are higher than those of all the other emotions in the posts. In other words, the text in the health organization posts is characterized by a very high level of fear. Among the topics with the highest fear emotion were

sepsis, monkeypox, cancer, and flu.

However, when examining Figure 7, fear is not dominant, and instead a neutral emotion is, which suggests that there are gaps in terms of what health organizations perceive to be issues that should be more feared and how the public perceives this and shows less fear in its reactions. When looking at other emotions, in Figure 7 we can see which topic raises the most anger, namely vaccines. The COVID-19 booster vaccine topic has the highest level of anger. In terms of sadness, cancer and sepsis are the topics that evoked the strongest feelings of sadness.

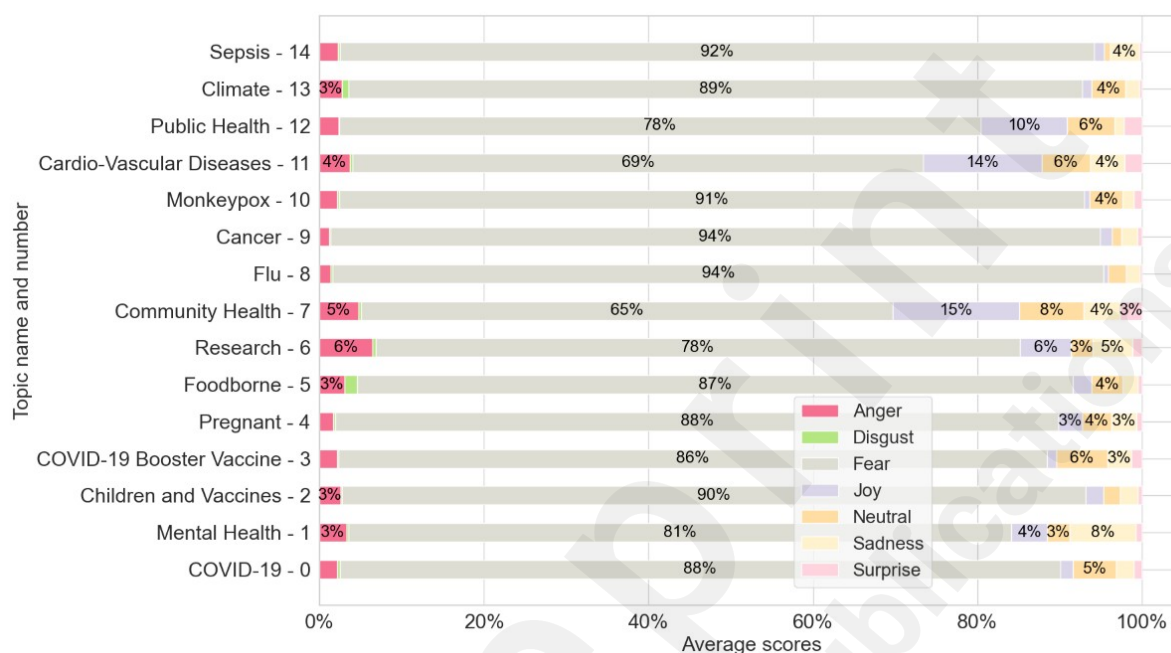


Figure 6 – Posts' emotion scores (average anger, disgust, fear, joy, neutral, sadness, and surprise) for each topic.

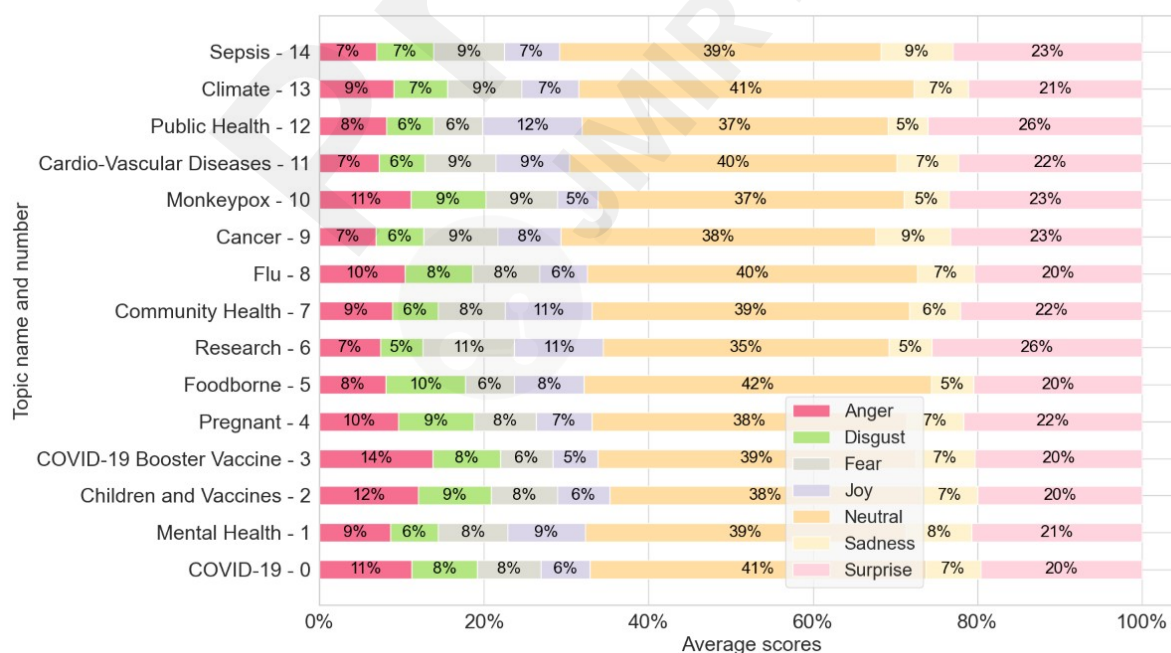


Figure 7 – Comments' emotion scores (average anger, disgust, fear, joy, neutral, sadness, surprise) for each topic.

Gap Analysis Results

Given the difference observed between the high average fear in the health organizations' posts compared to the very low fear in the public's responses to the posts, we examined the topics with the largest gap between the two average fears as revealed in the emotion analysis. For each topic, Table 4 presents the average fear in the posts of health organizations, the average fear in the comments, and the difference between the two. As can be seen from the table, the topics with the highest gap are masks and flu.

Table 3 - Each topic's average fear in posts, comments, and gap.

Topic name	Topic number	Average fear in posts	Average fear in comments	Gap
Masks	15	0.958	0.092	0.881
Flu	8	0.938	0.066	0.857
Vector	16	0.931	0.101	0.846
Cancer	9	0.936	0.062	0.846
Sepsis	14	0.917	0.106	0.831
Children and Vaccines	2	0.904	0.086	0.823
Noise Damage	24	0.913	0.084	0.821
Monkeypox	10	0.906	0.061	0.819
Ebola	19	0.898	0.085	0.812
Foodborne	5	0.870	0.077	0.810
Pregnant	4	0.879	0.086	0.803
Climate	13	0.893	0.090	0.803
COVID-19	0	0.876	0.060	0.799
COVID-19 Booster Vaccine	3	0.861	0.086	0.798
RSV	22	0.897	0.087	0.796
Health Equity	17	0.815	0.091	0.754
Antibiotics	18	0.835	0.081	0.751
Sun Damage	23	0.800	0.081	0.734
Mental Health	1	0.806	0.111	0.722
Public Health	12	0.780	0.060	0.720
Smoking	21	0.777	0.076	0.715
Research	6	0.783	0.064	0.672
Humanitarian	20	0.724	0.081	0.618
Cardiovascular Diseases	11	0.694	0.084	0.608
Community Health	7	0.646	0.077	0.565

Linguistic Analysis Results

The results are presented in Multimedia Appendix 2. It contains the top 50 phrases for unigrams, bigrams, trigrams, and four-grams in posts receiving predominantly positive or negative responses, along with their TF-IDF scores.

Results revealed that phrases associated with positive public responses promote public health awareness, vaccination benefits, and preventative measures. As an example, the phrases that were only included in the top 50 phrases of the health organizations' posts that received positive responses included "stay healthy" "awareness month", "signs symptoms", "raise awareness", "save lives", "help slow spread", "health care provider", and "better health better understanding". In contrast, phrases associated with negative sentiment related to vaccination efforts, outcomes, policy mandates, and

health monitoring with phrases as "data tracker", "vaccinated covid", "dose vaccine", "severe illness hospitalization", "illness hospitalization death", and "covid 19 vaccine booster".

Time Analysis Results

Figures 8-10 illustrate sentiment and emotional trends in posts and comments from 2018 to 2023 along with the COVID-19 death ratio. The numbers 1-5 in the figures represent five significant milestones during the COVID-19 pandemic. These include the first report of COVID-19 in late 2019 (1), the declaration of COVID-19 as a global pandemic by the WHO in March 2020 (2); the administration of the first COVID-19 vaccine in December 2020 in the United Kingdom (3), the peak of the Omicron wave in early 2022 (4), and the WHO's declaration in May 2023 that COVID-19 was no longer a global health emergency (5).

Figure 8 presents the average positive and negative sentiment expressed in posts and comments over time. The negative sentiment in comments increased over time and then began to decrease after the peak of the Omicron wave (4). While positive sentiment in comments shows a similar but opposite trend as it decreases and then increases after the peak of the Omicron wave (4). Regarding the sentiment of posts, it appears that positivity in posts increases slightly over the years, while negativity decreases. Regarding the COVID-19 death ratio, it peaks align with increased negative sentiment in comments.

Figure 9 explores the average levels of emotions—fear, trust, disappointment, anger, and confusion—in posts. Fear consistently dominated posts and remained relatively steady until it declined alongside the significant decrease in the death ratio. Trust, anger, and disappointment remained relatively steady throughout the years. While confusion gradually increased following the pandemic declaration (2).

Figure 10 shows the average levels of emotions—fear, trust, disappointment, anger, and confusion—in comments over the same period. Emotions of fear and trust remained relatively steady throughout the years. Anger and disappointment increase around death ratio peaks and, overall, show a general upward trend over the years. However, anger begins to decline slightly after the Omicron peak and the subsequent decrease in the death ratio.

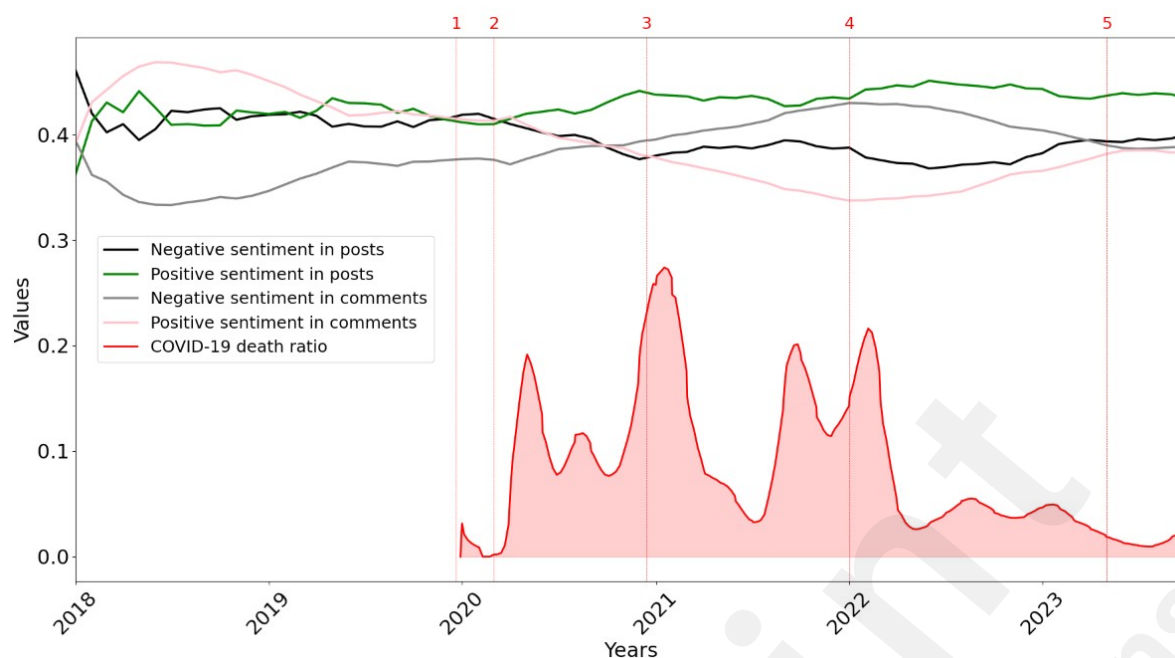


Figure 8 – Sentiment trends (average positive and negative scores) in posts and comments over the years, combined with the COVID-19 death ratio.

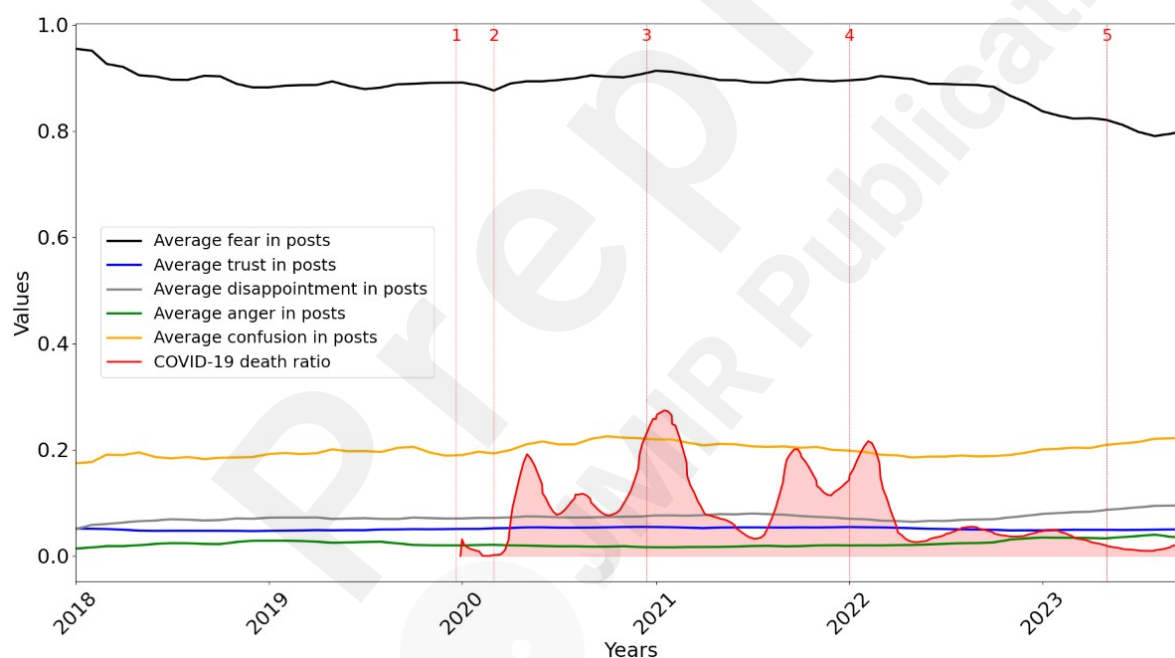


Figure 9 – Emotion trends (average scores for trust, fear, anger, confusion, and disappointment) in posts over the years, combined with the COVID-19 death ratio.

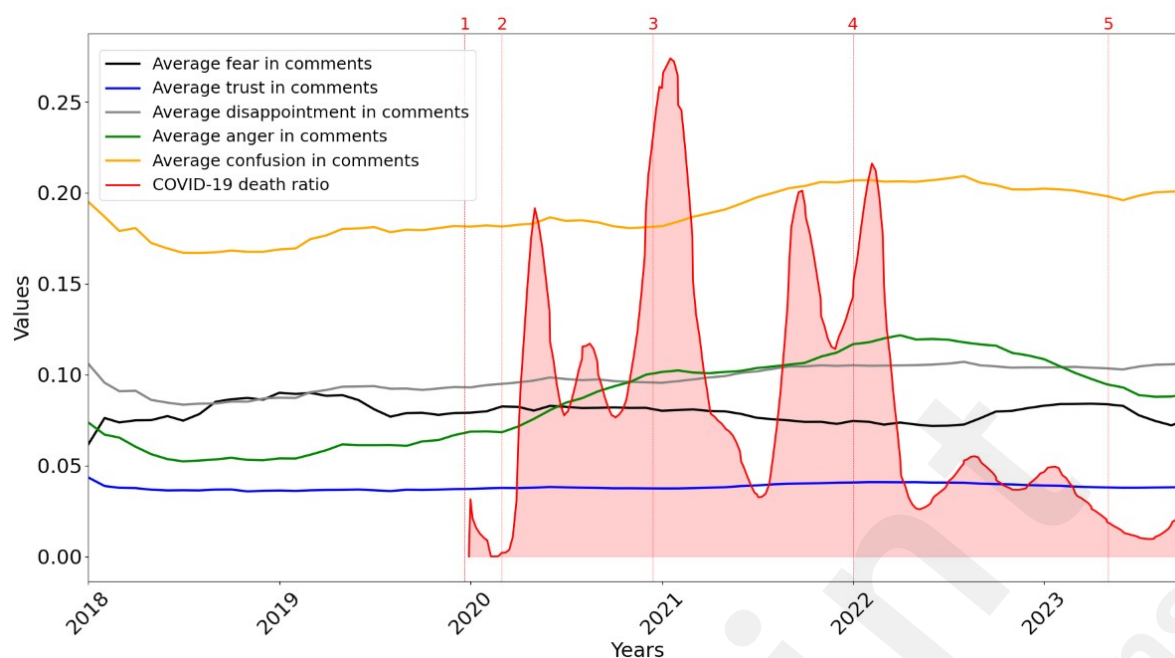


Figure 10 – Emotion trends (average scores for trust, fear, anger, confusion, and disappointment) in comments over the years, combined with the COVID-19 death ratio.

Discussion

The findings of this study should be interpreted with consideration of the dataset's geographic composition. As the majority of the health organizations analyzed are based in the U.S., the topics emphasized may reflect U.S.-specific health priorities, and the patterns of public engagement, sentiment and emotional responses to these topics may also be shaped by U.S. cultural and social contexts. The selected health organizations have a high number of followers on the platform, hold substantial influence, and reach audiences beyond national boundaries. Still, the framing and intensity of reactions observed in this study may not be fully generalizable to health organizations in other countries or regions, particularly because the characteristics of the users engaging with the content (such as their geographic location or demographics) are unknown.

Our findings on engagement analysis show that humanitarian issues, including the Ukraine war and the war in Gaza and the COVID-19 pandemic received the greatest response from the public. This supports prior research which has shown that those disasters, including humanitarian crises such as nature-related crises and wars are a major concern of the public and policy makers [101]. In recent years, conflicts such as the ongoing Ukraine war and the war in Gaza have had significant impact on international stability, affecting economies, migration patterns, and security measures, particularly in Europe [102]. These conflicts have not only disrupted local societies but have also prompted re-evaluation of humanitarian aid and intervention strategies globally. The connection between humanitarian conflicts and adverse health outcomes is well-documented. Studies have shown increased incidence of mental health disorders, infectious disease outbreaks, and chronic health conditions in conflict-affected areas [103]. As a result, health organizations worldwide have recognized the urgent need to develop comprehensive strategies for emergency preparedness. This includes enhancing public health response capabilities, improving pre-hospital care, and integrating disaster medicine in public health frameworks to better address humanitarian crises [104,105].

The COVID-19 pandemic has fundamentally altered global perceptions and practices across various sectors, revealing both strengths and weaknesses in systems and policies worldwide [106]. During the pandemic, the public turned to health organization websites and social media channels to receive

timely updates on infection rates, government guidelines, and evolving safety protocols [107,108]. Trust in organizations such as the WHO and CDC was crucial, as these bodies communicated essential information, adapting advice as new data emerged about the virus's spread and impact. The results of this study further highlight that a significant portion of posts shared by health organizations focused on the topic of COVID-19. Additionally, these posts generated substantial public engagement, demonstrating the public's heightened interest and concern regarding COVID-19-related information. Given the profound impact of COVID-19, the public remains extremely interested in these insights, recognizing that understanding what was done well and what failed can better prepare societies for future health crises. As such, the ongoing dialogue among scientists, policymakers, and the public about these findings continues to shape post-pandemic recovery strategies and fortify global health preparedness [109].

Our sentiment and emotion analysis of responses to health organizations' posts showed that the highest percentage of negative scores was in topics related to vaccines and Monkeypox. Moreover, our results show that the highest level of anger was observed in topics related to the COVID-19 booster and children's vaccination in general. The high anger and negative sentiment can be attributed to the vaccine hesitancy.

Vaccine hesitancy and anti-vaccine movements have long posed significant challenges to public health [110], particularly in communities with historically low trust in government and pharmaceutical institutions [111]. The analysis of topics associated with high anger emotions, such as COVID-19 vaccines, can be enriched by examining the emotional characteristics and motivations of specific user groups with vaccine hesitancy. For instance, women are generally more likely to exhibit vaccine hesitancy than men [112]. Alternatively, older individuals tend to display lower levels of hesitancy, likely due to their awareness of their vulnerability to COVID-19 complications, which may alleviate their fear or anger [112]. Vaccine hesitancy is also more prevalent among individuals with lower economic security [112]. Trust also shapes emotional and cognitive characteristics. Higher trust in healthcare providers, scientists, and global health organizations like the WHO correlates with reduced vaccine hesitancy [112]. Conversely, in some contexts, higher trust in religious leaders is linked to increased hesitancy [112].

The COVID-19 pandemic has increased skepticism toward routine immunizations and increased hesitancy even among those previously compliant with vaccination schedules. The rapid development of COVID-19 vaccines, coupled with widely publicized adverse effects, heightened fears and deepened mistrust [113]. Additionally, misinformation and conflicting narratives on social media further fueled fear, anger, and uncertainty, particularly in communities already skeptical of public health authorities. Our findings reflect these trends, showing increases in public negativity, anger, and disappointment over time.

Our linguistic analysis highlights that messages emphasizing safety, prevention, and public health benefits tend to elicit positive sentiments, while messages emphasizing vaccination efforts, outcomes, policy mandates, and health monitoring tend to elicit negative sentiments. This underscores the need for health organizations to refine communication strategies, focusing on clear, trust-building messaging to address concerns and counteract negative sentiments effectively.

Our analysis also reveals a gap between the fear conveyed in posts by health organizations and the public's responses to these posts, particularly regarding the influenza vaccine and face masks during the pandemic. While health organizations intended to convey urgency, public comments did not reflect the same fear. Addressing skepticism and ambivalence about influenza vaccine and mask usage is critical. Face masks are a critical tool in pandemic response [114], enhancing public

understanding of their effectiveness is essential for improving adherence and compliance during future outbreaks.

The difference in fear responses between the influenza vaccine and face masks can be understood by examining the interplay of social, psychological, and cultural factors [114]. One factor is psychological reactance, which occurs when individuals perceive mandates, such as mask requirements, as threats to their autonomy. This perception, combined with beliefs that masks are ineffective and an aversion to being forced to wear them, can trigger anger and counter-arguments. Individuals with strong psychological reactance are particularly likely to exhibit these responses, reinforcing and intensifying their anti-mask attitudes [115, 116]. Such resistance is further exacerbated by personality traits, which not only strengthen anti-mask sentiments but also link to broader vaccine skepticism, exaggeration of COVID-19 risks, and resistance to social distancing, often influenced by political conservatism [117].

Social and cultural dynamics also shape mask perceptions. For example, in many Western societies, masks are seen as extraordinary artifacts associated with emergency, while in Asian cultures, they are normalized as part of daily life [118]. In collectivist cultures, mask-wearing aligns with a sense of duty to protect the community [116]. Additionally, stigma, appearance concerns, and fears of being perceived as overly cautious further hinder acceptance [119].

Policy decisions and public health campaigns significantly influence perceptions. During the COVID-19 pandemic, inconsistent messaging from health authorities and varying guidance on mask use between countries and organizations contributed to public confusion and skepticism regarding their efficacy [120], despite substantial scientific evidence supporting their role in reducing transmission rates [120,121]. Therefore, consistent messaging and targeted communication strategies should focus on populations with anti-mask perceptions, aiming to reduce stigmas and address the underlying factors driving these attitudes.

Regarding the influenza vaccine, historically, the influenza vaccine has engendered significant hesitancy and objection from segments of the general public, a trend particularly evident among healthcare workers, who are critical in promoting vaccination [122]. Factors contributing to this hesitancy include misconceptions about vaccine efficacy, fears of adverse effects, and a perceived lack of urgency surrounding influenza, particularly when compared to more severe illnesses such as COVID-19 [123,124]. The emergence of the COVID-19 pandemic exacerbated this situation. During the pandemic, public focus shifted intensely toward COVID-19 vaccination campaigns, leading to a diversion of attention and resources from influenza vaccination efforts [125]. The flood of information regarding COVID-19 vaccines overshadowed the long-standing influenza vaccine campaigns. As a result, many individuals prioritize the COVID-19 vaccine over the seasonal flu vaccine [125]. Data indicating that compliance with the influenza vaccine plummeted in several countries during 2023 and 2024 compared to pre-pandemic years highlight the ongoing challenges faced by public health authorities [125].

Despite the availability of the influenza vaccine, social media misinformation and evolving health narratives regarding influenza and COVID-19 have led to an atmosphere of uncertainty. For instance, some individuals may mistakenly believe that if COVID-19 variants pose health risks, the influenza virus may be less significant or that acquiring one vaccine negates the need for others. This shift necessitates a reassessment of public health strategies to re-engage communities with the importance of influenza vaccination. Health organizations must develop targeted messaging that addresses misconceptions, enhances understanding of the influenza virus's potential impact, and reinforces the protective benefits of vaccination even during periods dominated by attention to other diseases.

To summarize, in this study, we analyzed the quality of disseminated messages along two dimensions: public responses to topics and the emotions and sentiments expressed in those responses. Mapping these reactions to messages posted by health organizations is important for evaluating public engagement with specific health-related issues. Identifying emotional gaps can also help assess the effectiveness of health-related messages, revealing potential discrepancies between the importance health organizations assign to certain topics and the public's perceived importance.

Our findings highlight gaps in fear responses regarding the influenza vaccine and wearing face masks during the COVID-19 underscore the need for transparent communication from health authorities and comprehensive education campaigns to address misconceptions and reassure the public. Understanding the psychological and social factors driving vaccine hesitancy post-pandemic is essential for tailoring effective public health strategies. Efforts must focus on fostering trust through consistent and clear messaging, transparency about vaccine development processes, and open forums for addressing public concerns. By doing so, it will be possible to mitigate misinformation, reduce fear, and enhance public compliance and improve vaccine uptake across populations.

Limitations

Our study may have some limitations. We collected data from Instagram, and therefore, our results and conclusions derive from the posts and interaction on this social media platform. Our data did not include specific information about the users who commented on the health organization posts. However, according to general information about Instagram users by Statista [126], Instagram had two billion monthly active users in 2024, with India leading the platform's user base at approximately 360 million, followed by the United States with 169 million and Brazil with 134 million users. Younger users dominate the platform, with the 18-24 age group being the largest demographic, followed by the 25-34 age group. Participation declines significantly among older age groups, particularly those aged 55 and above, who represent only a small fraction of the audience. Moreover, the majority of users in the 18-34 age group are men, while the majority of users over the age of 34 are women.

The absence of detailed user characteristics in our dataset may limit the generalizability of our findings, as biases could arise due to age, gender, educational level, or other demographic factors that influence engagement and sentiment patterns. Certain user groups may be overrepresented or underrepresented in the data, potentially shaping the interpretation of public responses. Future research should aim to systematically characterize respondents' profiles, leveraging available metadata or incorporating external surveys to gain a clearer understanding of the audience engaging with health-related content.

The study also has limitations regarding the influence of social media algorithms on post visibility and engagement. The social media algorithms prioritize content based on user interactions, interests, and platform-specific ranking mechanisms, potentially introducing bias in public sentiment patterns. Nevertheless, the decision as to whether to engage and how to respond is left to the users. Therefore, although engagement data may not fully represent the broader public, it still offers insights into those actively participating in discussions. To mitigate bias, we ensured a diverse dataset by including a substantial number of posts from multiple health organizations.

It is important to acknowledge that the models used in this study may have limitations. Although BERTopic is an effective topic modeling technique, it assumes that each document (in our case, each post) is associated with a single dominant topic, which does not always reflect reality. As posts can discuss a variety of interconnected topics, it is often difficult to classify them accurately under a

single theme. Furthermore, topic separation relies on clustering techniques, which may not always produce clear or optimal topic divisions, resulting in the merging of distinct topics or fragmentation of related discussions, reducing interpretability.

Similarly, the sentiment and emotion models may inherit biases from their training datasets, influencing detection. These models are not always fully accurate and often struggle to capture context-dependent sentiment shifts, irony, and implicit emotional expressions, which may result in misinterpretations.

In addition, as previously mentioned, the health organizations in this study is predominantly U.S.-based, with limited representation from other countries. This concentration may influence the topics represented in the dataset, reflecting U.S.-specific health priorities and cultural dynamics, as well as the public sentiment, emotional responses, and engagement observed in relation to those topics. Future research could incorporate data from a more geographically diverse set of health organizations, enabling cross-cultural comparisons and providing a broader understanding of global public health communication and audience responses.

Conclusions

This study demonstrates the value of our methodology in assessing public responses and emotions expressed regarding health-related messages. By identifying emotional gaps, particularly fear, we were able to uncover discrepancies between the fear health organizations assign to issues and the fear that the public expresses in response. The greatest gaps were revealed with regard to influenza vaccines and face masks during COVID-19.

These findings emphasize the need for transparent communication and trust-building strategies that consider the emotional and sentiment-driven responses of the public on social media. By understanding the psychological and social dynamics of public interaction with health information, particularly regarding vaccine resistance, organizations can support public health goals, foster trust and efficient engagement, counter misinformation, and encourage informed health behaviors.

In future research, we plan to apply our topic modeling and sentiment and emotion analysis approach on other social media platforms to gain a more comprehensive view of the public's response to the posts of health organizations across various digital platforms. Additionally, we aim to extend our analysis beyond social media to offline settings, such as newspapers and television campaigns. Since public responses are essential to our method, we will also incorporate surveys to assess audience reactions and engagement with these campaigns. Investigating how fear-based messaging and public emotional responses manifest in both online and offline environments will provide deeper insights into the broader impact of health communication strategies.

In addition, we plan to investigate misinformation within the comments on health organization posts and identify the topics where inaccuracies are prevalent.

Data Availability

The datasets generated and analyzed during this study are not publicly available, because they contain sensitive information that could compromise the privacy of social media users. The transformed data are, however, available upon reasonable request from the authors.

Authors' Contributions

APV contributed to the study conceptualization, software, methodology, data curation, formal analysis, visualization, and writing. AM contributed to the study design, methodology, and writing. All authors approved the final version of this manuscript.

Conflicts of Interest

None declared.

Abbreviations

API: Application Programming Interface

c-TF-IDF: Class-based term frequency-inverse document frequency

CDC: Centers for Disease Control

FDA: Food and Drug Administration (US)

HHS: Department of Health and Human Services

LDA: Latent Dirichlet allocation

NHS: National Health Service

NIH: National Institutes of Health

NIMH: National Institute of Mental Health

OMH: Office of Minority Health

TF-IDF: Term frequency-inverse document frequency

WHO: World Health Organization

Multimedia Appendix 1

The number, name, size, and top-5 most significant words of each topic.

Multimedia Appendix 2

Top-50 phrases for unigrams, bigrams, trigrams, and four-grams in posts receiving predominantly positive or negative responses with corresponding TF-IDF scores.

References

1. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. *Lancet*. 2010;376(9748):1261-1271. doi:10.1016/S0140-6736(10)60809-4.
2. Gatewood J, Monks SL, Singletary CR, Vidrascu E, Moore JB. Practice brief report: Social media in public health—strategies to distill, package, and disseminate public health research. *J Public Health Manag Pract*. 2020;26(5):489. doi:10.1097/PHH.0000000000001096.
3. Chen J, Wang Y. Social media use for health purposes: systematic review. *J Med Internet Res*. 2021;23(5):e17917. doi:10.2196/17917.
4. Tseng TS, Gonzalez G. Social media and types with their current applications in public health and healthcare. *Effective Use of Social Media in Public Health*. 2023;3-22. doi:10.1016/B978-0-323-95630-7.00007-X.
5. Jeyaraman M, Ramasubramanian S, Kumar S, Jeyaraman N, Selvaraj P, Nallakumarasamy A, et al. Multifaceted role of social media in healthcare: opportunities, challenges, and the need for quality control. *Cureus*. 2023;15(5). doi:10.7759/cureus.39111
6. Ramanadhan S, Mendez SR, Rao M, Viswanath K. Social media use by community-based organizations conducting health promotion: a content analysis. *BMC Public Health*. 2013;13(1):1-10. doi:10.1186/1471-2458-13-1129.
7. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: example of cancer awareness via Twitter. *JMIR Public Health Surveill*. 2016;2(1):e5205. doi:10.2196/publichealth.5205.
8. Latha K, Meena KS, Pravitha MR, Dasgupta M, Chaturvedi SK. Effective use of social media platforms for promotion of mental health awareness. *J Educ Health Promot*. 2020;9.

doi: 10.4103/jehp.jehp_90_20.

9. Anwar A, Malik M, Raees V, Anwar M, Anwar A. Role of mass media and public health communications in COVID-19 vaccination. *Med Res Arch*. 2023;11(2). doi: 10.7759/cureus.10453.
10. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med*. 2019;240:112552. doi:10.1016/j.socscimed.2019.112552.
11. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res*. 2021;23(1):e17187. doi:10.2196/17187.
12. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, Fagen MC. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. *Health Promot Pract*. 2012;13(2):159-164. doi:10.1177/1524839911433467.
13. Edney S, Bogomolova S, Ryan J, Olds T, Sanders I, Maher C. Creating engaging health promotion campaigns on social media: observations and lessons from Fitbit and Garmin. *J Med Internet Res*. 2018;20(12):e10911. doi:10.2196/10911.
14. Afful-Dadzie E, Afful-Dadzie A, Egala SB. Social media in health communication: A literature review of information quality. *Health Inf Manag J*. 2021;52(1):3-17. doi:10.1177/1833358321992683.
15. Kington RS, Arnesen S, Chou WYS, Curry SJ, Lazer D, Villarruel AM. Identifying credible sources of health information in social media: principles and attributes. *NAM Perspect*. 2021;10:31478. doi: 10.31478/202107a
16. Eysenbach G. Credibility of health information and digital media: new perspectives and implications for youth. *Digit Media Learn*. 2007;9(2):123. doi: 10.1162/dmal.9780262562324.123
17. Heldman AB, Schindelar J, Weaver JB. Social media engagement and public health communication: implications for public health organizations being truly “social.” *Public Health Rev*. 2013;35:1-18. doi:10.1007/BF03391698.
18. Aydin G. Improving health and efficiency with strategic social media use in health organizations: a critical review of the status quo. *Multidimensional Perspectives and Global Analysis of Universal Health Coverage*. 2020:309-336. doi: 10.4018/978-1-7998-8960-1.ch006.
19. Ghahramani A, Prokofieva M, de Courten MP. Health promotion campaigns using social media: association rules mining and co-occurrence network analysis of Twitter hashtags. *BMC Public Health*. 2025;25(1):67. doi: 10.1186/s12889-024-21255-5
20. De Freitas C, Silva S, Maia T, Amorim M. Public participation in health data governance: a scoping review. *Eur J Public Health*. 2019;29(Suppl 4):ckz185-647. doi: 10.1093/eurpub/ckz185.647
21. Tavoschi L, Quattrone F, D’Andrea E, Ducange P, Vabanesi M, Marcelloni F, Lopalco PL. Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. *Hum Vaccin Immunother*. 2020;16(5):1062-1069. doi: 10.1080/21645515.2020.1714311.
22. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. *J Med Internet Res*. 2021;23(4):e26627. doi:10.2196/26627.
23. Hu T, Wang S, Luo W, Zhang M, Huang X, Yan Y, et al. Revealing public opinion towards COVID-19 vaccines with Twitter data in the United States: spatiotemporal perspective. *J Med Internet Res*. 2021;23(9):e30854. doi: 10.2196/30854.
24. Niu Q, Liu J, Kato M, Shinohara Y, Matsumura N, Aoyama T, Nagai-Tanima M. Public

- opinion and sentiment before and at the beginning of COVID-19 vaccinations in Japan: Twitter analysis. *JMIR Infodemiol.* 2022;2(1):e32335. doi:10.2196/32335.
25. Saleh SN, McDonald SA, Basit MA, Kumar S, Arasaratnam RJ, Perl TM, et al. Public perception of COVID-19 vaccines through analysis of Twitter content and users. *Vaccine.* 2023. doi: 10.1016/j.vaccine.2023.06.058.
 26. Finset A, Bosworth H, Butow P, Gulbrandsen P, Hulsman RL, Pieterse AH, et al. Effective health communication—a key factor in fighting the COVID-19 pandemic. *Patient Educ Couns.* 2020;103(5):873. doi:10.1016/j.pec.2020.03.027.
 27. Rovetta A. World Health Organization is losing online credibility towards health-sensitive topics: infodemiological analysis of Facebook users' reactions. *Health Promot Perspect.* 2022;12(4):367. doi:10.34172/hpp.2022.48.
 28. Gemenis K. COVID-19 conspiracy beliefs and vaccination intentions among social media users. *Stat Polit Policy.* 2022;13(3):279-296. doi: 10.1515/spp-2022-0005.
 29. Lee J, Bissell K. Correcting vaccine misinformation on social media: the inadvertent effects of repeating misinformation within such corrections on COVID-19 vaccine misperceptions. *Curr Psychol.* 2024;1-13. doi: 10.1177/14614448231169697.
 30. Ayaz H, Celik MH, Koytak HZ, Yanik IE. Exploring vaccine hesitancy in digital public discourse: from tribal polarization to socio-economic disparities. *PLoS One.* 2024;19(11):e0308122. doi:10.1371/journal.pone.0308122.
 31. Kite J, Chan L, MacKay K, Corbett L, Reyes-Marcelino G, Nguyen B, Freeman B. A model of social media effects in public health communication campaigns: systematic review. *J Med Internet Res.* 2023;25:e46345. doi: 10.2196/46345.
 32. Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. *Soc Netw Anal Min.* 2021;11(1):81. doi: 10.1007/s13278-021-00776-6.
 33. Demszky D, Movshovitz-Attias D, Ko J, Cowen A, Nemade G, Ravi S. GoEmotions: A dataset of fine-grained emotions. *arXiv preprint.* 2020;arXiv:2005.00547.
 34. Chutia T, Baruah N. A review on emotion detection by using deep learning techniques. *Artif Intell Rev.* 2024;57(8):203. doi:s10462-024-10831-1
 35. Chen L, Yang X, Fu L, Liu X, Yuan C. Using the extended parallel process model to examine the nature and impact of breast cancer prevention information on mobile-based social media: content analysis. *JMIR Mhealth Uhealth.* 2019;7(6):e13987. doi: 10.2196/13987.
 36. Witte K. Putting the fear back into fear appeals: The extended parallel process model. *Commun Monogr.* 1992;59(4):329-349. doi: 10.1080/03637759209376276.
 37. Raghupathi V, Ren J, Raghupathi W. Studying public perception about vaccination: a sentiment analysis of tweets. *Int J Environ Res Public Health.* 2020;17(10):3464. doi:10.3390/ijerph17103464.
 38. Ansari MTJ, Khan NA. Worldwide COVID-19 vaccines sentiment analysis through Twitter content. *Electron J Gen Med.* 2021;18(6):em324. doi: 10.29333/ejgm/11316
 39. Liu S, Liu J. Public attitudes toward COVID-19 vaccines on English-language Twitter: a sentiment analysis. *Vaccine.* 2021;39(39):5499-5505. doi: 10.1016/j.vaccine.2021.08.058.
 40. Lyu JC, Han EL, Luli GK. COVID-19 vaccine-related discussion on Twitter: topic modeling and sentiment analysis. *J Med Internet Res.* 2021;23(6):e24435. doi:10.2196/24435.
 41. Monselise M, Chang CH, Ferreira G, Yang R, Yang CC. Topics and sentiments of public concerns regarding COVID-19 vaccines: social media trend analysis. *J Med Internet Res.* 2021;23(10):e30765. doi:10.2196/30765.
 42. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. *Int J Infect Dis.* 2021;108:256-262. doi:10.1016/j.ijid.2021.05.059.
 43. Chandrasekaran R, Desai R, Shah H, Kumar V, Moustakas E. Examining public sentiments and attitudes toward COVID-19 vaccination: infoveillance study using Twitter posts. *JMIR*

- Infodemiol. 2022;2(1):e33909. doi:10.2196/33909.
44. Chen N, Chen X, Pang J. A multilingual dataset of COVID-19 vaccination attitudes on Twitter. *Data Brief*. 2022;44:108503. doi:10.1016/j.dib.2022.108503.
 45. Huangfu L, Mo Y, Zhang P, Zeng DD, He S. COVID-19 vaccine tweets after vaccine rollout: sentiment-based topic modeling. *J Med Internet Res*. 2022;24(2):e31726. doi:10.2196/31726.
 46. Nguyen A, Catalan-Matamoros D. Anti-vaccine discourse on social media: an exploratory audit of negative tweets about vaccines and their posters. *Vaccines*. 2022;10(12):2067. doi:10.3390/vaccines10122067.
 47. Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. *Postgrad Med J*. 2022;98(1161):544-550. doi:10.1136/postgradmedj-2021-140685.
 48. Mir AA, Sevukan R. Sentiment analysis of Indian tweets about COVID-19 vaccines. *J Inf Sci*. 2022;48(6):814-827. doi:10.1177/01655515221118049.
 49. Reshi AA, Rustam F, Aljedaani W, Shafi S, Alhossan A, Alrabiah Z, et al. COVID-19 vaccination-related sentiments analysis: a case study using worldwide Twitter dataset. *Healthcare (Basel)*. 2022;10(3):411. doi:10.3390/healthcare10030411.
 50. Stracqualursi L, Agati P. COVID-19 vaccines in Italian public opinion: identifying key issues using Twitter and natural language processing. *PLoS One*. 2022;17(11):e0277394. doi:10.1371/journal.pone.0277394.
 51. Yiannakoulis N, Darlington JC, Slavik CE, Benjamin G. Negative COVID-19 vaccine information on Twitter: content analysis. *JMIR Infodemiol*. 2022;2(2):e38485. doi:10.2196/38485.
 52. Zulfiker MS, Kabir N, Biswas AA, Zulfiker S, Uddin MS. Analyzing the public sentiment on COVID-19 vaccination in social media: Bangladesh context. *Array*. 2022;15:100204. doi:10.1016/j.array.2022.100204.
 53. Xu H, Liu R, Luo Z, Xu M. COVID-19 vaccine sensing: sentiment analysis and subject distillation from Twitter data. *Telemat Inform Rep*. 2022;8:100016. doi:10.1016/j.teler.2022.100016.
 54. Rustam F, Khalid M, Aslam W, Rupapara V, Mehmood A, Choi GS. A performance comparison of supervised machine learning models for COVID-19 tweets sentiment analysis. *PLoS One*. 2021;16(2):e0245909. doi:10.1371/journal.pone.0245909.
 55. Blankenship EB, Goff ME, Yin J, Tse ZTH, Fu KW, Liang H, et al. Sentiment, contents, and retweets: a study of two vaccine-related Twitter datasets. *Perm J*. 2018;22. doi: 10.7812/TPP/17-138.
 56. Chakraborty K, Bhatia S, Bhattacharyya S, Platos J, Bag R, Hassanien AE. Sentiment analysis of COVID-19 tweets by deep learning classifiers—A study to show how popularity is affecting accuracy in social media. *Appl Soft Comput*. 2020;97:106754. doi:10.1016/j.asoc.2020.106754.
 57. Argyris YA, Monu K, Tan PN, Aarts C, Jiang F, Wiseley KA. Using machine learning to compare provaccine and antivaccine discourse among the public on social media: algorithm development study. *JMIR Public Health Surveill*. 2021;7(6):e23105. doi:10.2196/23105.
 58. Aygün I, Kaya B, Kaya M. Aspect-based Twitter sentiment analysis on vaccination and vaccine types in COVID-19 pandemic with deep learning. *IEEE J Biomed Health Inform*. 2021;26(5):2360-2369. doi: 10.1109/JBHI.2021.3133103.
 59. Basiri ME, Nemati S, Abdar M, Asadi S, Acharrya UR. A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. *Knowl-Based Syst*. 2021;228:107242. doi:10.1016/j.knosys.2021.107242.
 60. Kaur H, Ahsaan SU, Alankar B, Chang V. A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. *Inf Syst Front*. 2021;1-13. doi: 10.1007/s10796-

021-10135-7.

61. Villavicencio C, Macrohon JJ, Inbaraj XA, Jeng JH, Hsieh JG. Twitter sentiment analysis towards COVID-19 vaccines in the Philippines using naïve Bayes. *Information*. 2021;12(5):204. doi:10.3390/info12050204.
62. Bari A, Heymann M, Cohen RJ, Zhao R, Szabo L, Apas Vasandani S, et al. Exploring coronavirus disease 2019 vaccine hesitancy on Twitter using sentiment analysis and natural language processing algorithms. *Clin Infect Dis*. 2022;74(Suppl_3):e4-e9. doi: 10.1093/cid/ciac141.
63. Islam MB, Hasibunnahar S, Shukla PK, Rawat P, Dange J. Twitter opinion mining on COVID-19 vaccinations by machine learning presence. *Proceedings of Third Doctoral Symposium on Computational Intelligence: DoSCI 2022*. Singapore: Springer Nature Singapore; 2022. p. 37-55. doi: 10.1007/978-981-19-3148-2_4
64. Maulana FI, Adi PDP, Lestari D, Purnomo A, Prihatin SY. Twitter data sentiment analysis of COVID-19 vaccination using machine learning. *Proceedings of the 2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. IEEE; 2022. p. 582-587. doi: 10.1109/ISRITI56927.2022.10053035.
65. Nyawa S, Tchuente D, Fosso-Wamba S. COVID-19 vaccine hesitancy: a social media analysis using deep learning. *Ann Oper Res*. 2022;1-39. doi: 10.1007/s10479-022-04792-3.
66. Shahzad A, Zafar B, Ali N, Jamil U, Alghadhbhan AJ, Assam M, et al. COVID-19 vaccines related user's response categorization using machine learning techniques. *Computation*. 2022;10(8):141. doi:10.3390/computation10080141.
67. Albahli S, Nawaz M. TSM-CV: Twitter sentiment analysis for COVID-19 vaccines using deep learning. *Electronics*. 2023;12(15):3372. doi:10.3390/electronics12153372.
68. Biri I, Kucuktas UT, Uysal F, Hardalac F. Forecasting the future popularity of the anti-vax narrative on Twitter with machine learning. *J Supercomput*. 2023;1-31. doi: 10.1007/s11227-023-05567-8.
69. Ramondt S, Kerkhof P, Merz EM. Blood donation narratives on social media: a topic modeling study. *Transfus Med Rev*. 2022;36(1):58-65. doi: 10.1016/j.tmr.2021.10.001.
70. Chen L, Wang P, Ma X, Wang X. Cancer communication and user engagement on Chinese social media: content analysis and topic modeling study. *J Med Internet Res*. 2021;23(11):e26310. doi:10.2196/26310.
71. Jiang S, Wang P, Liu PL, Ngien A, Wu X. Social media communication about HPV vaccine in China: a study using topic modeling and survey. *Health Commun*. 2023;38(5):935-946. doi: 10.1080/10410236.2021.1983338.
72. Paul MJ, Dredze M. Discovering health topics in social media using topic models. *PLoS One*. 2014;9(8):e103408. doi:10.1371/journal.pone.0103408.
73. Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. *Public Health*. 2017;150:170-175. doi:10.1016/j.puhe.2017.07.015.
74. Muralidhara S, Paul MJ. #Healthy selfies: exploration of health topics on Instagram. *JMIR Public Health Surveill*. 2018;4(2):e10150. doi:10.2196/10150.
75. Kim Y, Song D, Lee YJ. #Antivaccination on Instagram: a computational analysis of hashtag activism through photos and public responses. *Int J Environ Res Public Health*. 2020;17(20):7550. doi:10.3390/ijerph17207550.
76. Kwok SWH, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: machine learning analysis. *J Med Internet Res*. 2021;23(5):e26953. doi:10.2196/26953.
77. Kim Y, Kim JH. Using photos for public health communication: A computational analysis of the Centers for Disease Control and Prevention Instagram photos and public responses. *Health Inform J*. 2020;26(3):2159-2180. doi: 10.1177/1460458219896673.
78. Pinto PA, Antunes MJL, Almeida AMP. Public Health on Instagram: An analysis of health

- promotion strategies of Portugal and Brazil. *Procedia Comput Sci.* 2021;181:231-238. doi: 10.1016/j.procs.2021.01.142.
79. Mello S, Glowacki E, Fuentes I, Seabolt J. Communicating COVID-19 risk on Instagram: A content analysis of official public health messaging during the first year of the pandemic. *J Health Commun.* 2023;28(1):38-52. doi: 10.1080/10810730.2023.2175278.
80. Vaghefi MS, Beheshti N, Jain H. Dissemination of health messages in online social network: A study of healthcare providers' content generation and dissemination on Twitter. *Inf Manag.* 2024;61(2):103925. doi:10.1016/j.im.2023.103925.
81. World Health Organization. Available from: <https://www.who.int/about/> [accessed December 5, 2024].
82. U.S. Department of Health and Human Services. Available from: <https://www.hhs.gov/about/index.html> [accessed December 5, 2024].
83. The Office of Minority Health. Available from: <https://minorityhealth.hhs.gov/about-office-minority-health> [accessed December 5, 2024].
84. The National Institutes of Health. Available from: <https://www.nih.gov/about-nih> [accessed December 5, 2024].
85. The National Institute of Mental Health. Available from: <https://www.nimh.nih.gov/about> [accessed December 5, 2024].
86. Centers for Disease Control and Prevention. Available from: <https://www.cdc.gov/> [accessed December 5, 2024].
87. National Health Service. Available from: <https://www.nhs.uk/nhs-services/> [accessed December 5, 2024].
88. The U.S Food and Drug Administration. Available from: <https://www.fda.gov/> [accessed December 5, 2024].
89. Pinto PA, Antunes MJL, Almeida AMP. Instagram as a communication tool in public health: a systematic review. *Proceedings of the 2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*. June 2020; Seville, Spain. Piscataway, NJ: IEEE; 2020. doi: 10.23919/CISTI49556.2020.9140809.
90. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint*. 2022. Available from: <https://arxiv.org/abs/2203.05794>.
91. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. *Proc ACM Int Conf Web Search Data Min.* February 2015; Shanghai, China. New York, NY: ACM; 2015. doi: 10.1145/2684822.2685324.
92. Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. *arXiv preprint*. 2019. Available from: <https://arxiv.org/abs/1908.10084>.
93. McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. *arXiv preprint*. 2018. Available from: <https://arxiv.org/abs/1802.03426>.
94. McInnes L, Healy J, Astels S. HDBSCAN: hierarchical density based clustering. *J Open Source Softw.* 2017;2(11):205. doi:10.21105/joss.00205.
95. Yuan LX. distilbert-base-multilingual-cased-sentiments-student. 2023. Available from: <https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>.
96. Xian Y, Akata Z, Sharma G, Nguyen Q, Hein M, Schiele B. Latent embeddings for zero-shot classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. June 2016; Las Vegas, NV, USA. Piscataway, NJ: IEEE; 2016.
97. Hartmann J. Emotion english distilroberta-base. 2022. Available from: <https://huggingface.co/j-hartmann/emotion-english-distilroberta-base>.
98. Kirouane A. Emotion english distilroberta-base. 2024. Available from: <https://huggingface.co/ayoubkirouane/BERT-Emotions-Classfier>.
99. Lowe S. Emotion english distilroberta-base. 2024. Available from:

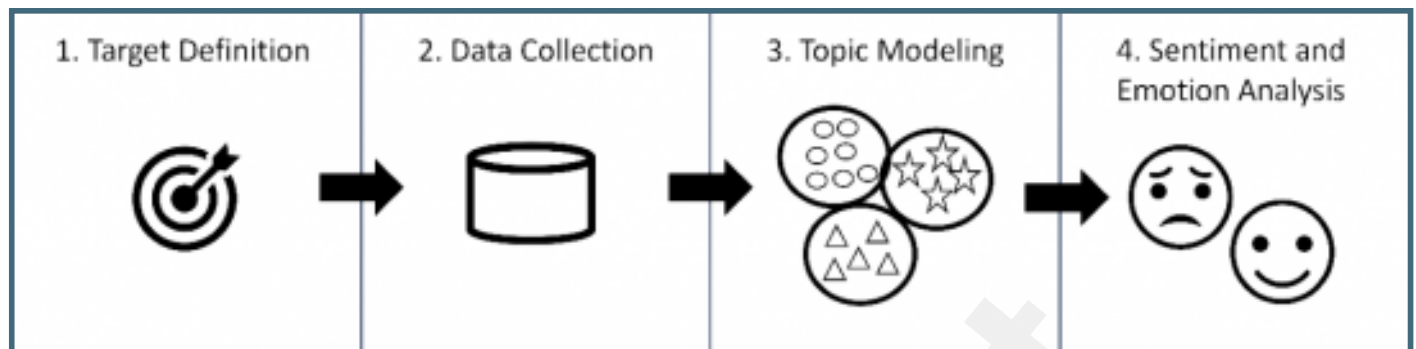
- https://huggingface.co/SamLowe/roberta-base-go_emotions.
100. National Center for Health. Provisional COVID-19 Death Counts by Week Ending Date and State. 2025. <https://data.cdc.gov/d/r8kw-7aab> [accessed Jan 22, 2025].
 101. Sandifer PA, Walker AH. Enhancing disaster resilience by reducing stress-associated health impacts. *Front Public Health*. 2018;6:373. doi:10.3389/fpubh.2018.00373.
 102. Colombo A, Magri P. Europe in the Age of Insecurity: ISPI Report 2024. 1st ed. Milan: Ledizioni; 2024. ISBN: 9791256000968.
 103. Garry S, Checchi F. Armed conflict and public health: into the 21st century. *J Public Health*. 2020;42(3):e287-e298. doi:10.1093/pubmed/fdz095.
 104. Sofuoglu Z, Jaeger B, Sofuoglu T, Kamyabi L, Bianchi S. Determining the emergency medical services innovation needs through a coordination and support action. *Open Res Eur*. 2024;4:264. doi: 10.12688/openreseurope.18449.1.
 105. Lee JM, Jansen R, Sanderson KE, Guerra F, Keller-Olaman S, Murti M, et al. Public health emergency preparedness for infectious disease emergencies: a scoping review of recent evidence. *BMC Public Health*. 2023;23(1):420. doi: 10.1186/s12889-023-15313-7.
 106. Lombatti S, Magid A, Davidovitch N, Middleton J, Sheek-Hussein M, Lopes H, et al. The COVID-19 pandemic as a public health teacher—the lessons we must learn. *Public Health Rev*. 2024;45:1607232. doi:10.3389/phrs.2024.1607232.
 107. Al-Dmour H, Salman A, Abuhashesh M, Al-Dmour R. Influence of social media platforms on public health protection against the COVID-19 pandemic via the mediating effects of public health awareness and behavioral changes: integrated model. *J Med Internet Res*. 2020;22(8):e19996. PMID:32750004 doi:10.2196/19996.
 108. Abuhashesh MY, Al-Dmour H, Masa'deh RE, Salman A, Al-Dmour R, Boguszewicz-Kreft M, AlAmaireh QN. The role of social media in raising public health awareness during the pandemic COVID-19: an international comparative study. *Informatics*. 2021;8(4):80. doi:10.3390/informatics8040080.
 109. Codd M, Barros H, Davidovitch N, Razum O, Mabhala M, Kostoulas P, et al. ASPHER statement: a new public health curriculum for a “new normal.” *Public Health Rev*. 2023;44:1606539. doi:10.3389/phrs.2023.1606539.
 110. Mir EA. The lingering challenge: addressing vaccine hesitancy in a post-pandemic world. *Cent Eur J Public Health*. 2023;31(4):296-299. doi: 10.21101/cejph.a7593.
 111. Chen X, Lee W, Lin F. Infodemic, institutional trust, and COVID-19 vaccine hesitancy: a cross-national survey. *Int J Environ Res Public Health*. 2022;19(13):8033. doi: 10.3390/ijerph19138033
 112. Rozek LS, Jones P, Menon A, Hicken A, Apsley S, King EJ. Understanding vaccine hesitancy in the context of COVID-19: the role of trust and confidence in a seventeen-country survey. *Int J Public Health*. 2021;66:636255. doi: 10.3389/ijph.2021.636255
 113. Carpiano RM, Callaghan T, DiResta R, Brewer NT, Clinton C, Galvani AP, et al. Confronting the evolution and expansion of anti-vaccine activism in the USA in the COVID-19 era. *Lancet*. 2023;401(10380):967-970. doi:10.1016/S0140-6736(23)00440-8.
 114. Schönweitz F, Eichinger J, Kuiper JM, Ongolly F, Spahl W, Prainsack B, Zimmermann BM. The social meanings of artifacts: face masks in the COVID-19 pandemic. *Front Public Health*. 2022;10:829904. doi: 10.3389/fpubh.2022.829904
 115. Taylor S, Asmundson GJ. Negative attitudes about facemasks during the COVID-19 pandemic: the dual importance of perceived ineffectiveness and psychological reactance. *PLoS One*. 2021;16(2):e0246317. doi: 10.1371/journal.pone.0246317
 116. Shim JM. Mask sociology as a way of action theory: voices of the face mask among social individuals in the COVID-19 masquerade in South Korea. *PLoS One*. 2023;18(11):e0293758. doi: 10.1371/journal.pone.0293758
 117. Rosenfeld DL, Balcetis E, Bastian B, Berkman ET, Bosson JK, Brannon TN, et al.

- Psychological science in the wake of COVID-19: social, methodological, and metascientific considerations. *Perspect Psychol Sci*. 2022;17(2):311-333. doi: 10.1177/1745691621999374
118. Syed Q, Sopwith W, Regan M, Bellis MA. Behind the mask: journey through an epidemic—some observations of contrasting public health responses to SARS. *J Epidemiol Community Health*. 2003;57(11):855-856. doi: 10.1136/jech.57.11.855.
119. Buregyeya E, Mitchell EM, Rutebemberwa E, Colebunders R, Criel B, Kiguli J, Nuwaha F. Acceptability of masking and patient separation to control nosocomial tuberculosis in Uganda: a qualitative study. *J Public Health*. 2012;20:599-606. doi: 10.1007/s10389-012-0503-1.
120. Unlu A, Truong S, Tammi T, Lohiniva AL. Exploring political mistrust in pandemic risk communication: mixed-method study using social media data analysis. *J Med Internet Res*. 2023;25:e50199. doi:10.2196/50199.
121. Boulos L, Curran JA, Gallant A, Wong H, Johnson C, Delahunty-Pike A, et al. Effectiveness of face masks for reducing transmission of SARS-CoV-2: a rapid systematic review. *Philos Trans R Soc A*. 2023;381(2257):20230133. doi:10.1098/rsta.2023.0133.
122. Kisa S, Kisa A. A comprehensive analysis of COVID-19 misinformation, public health impacts, and communication strategies: scoping review. *J Med Internet Res*. 2024;26:e56931. doi:10.2196/56931.
123. Kaur M, Coppeta L, Olesen OF. Vaccine hesitancy among healthcare workers in Europe: a systematic review. *Vaccines*. 2023;11(11):1657. doi:10.3390/vaccines11111657.
124. Magid A, Hussein K, Dabaja-Younis H, Szwarcwort-Cohen M, Almog R, Mekel M, et al. Knowledge of the serological response to the third BNT162b2 vaccination may influence compliance of healthcare workers to booster dose. *Antibodies*. 2024;13(3):63. doi:10.3390/antib13030063.
125. Walkowiak MP, Domaradzki J, Walkowiak D. Are we facing a tsunami of vaccine hesitancy or outdated pandemic policy in times of omicron? Analyzing changes of COVID-19 vaccination trends in Poland. *Vaccines*. 2023;11(6):1065. doi:10.3390/vaccines11061065.
126. statista. Instagram - statistics & facts. 2024. <https://www.statista.com/topics/1882/instagram/> [accessed Jan 22, 2025].

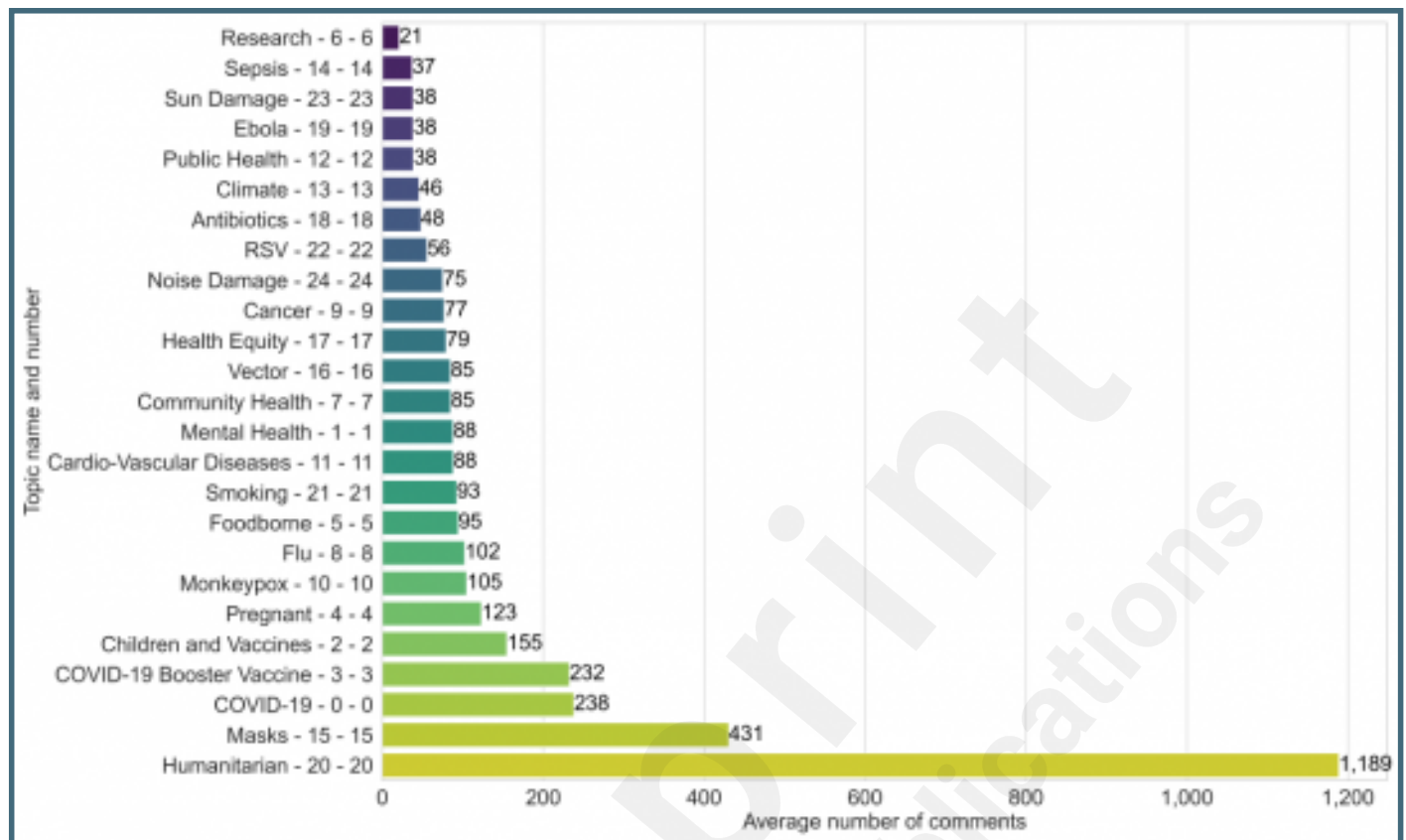
Supplementary Files

Figures

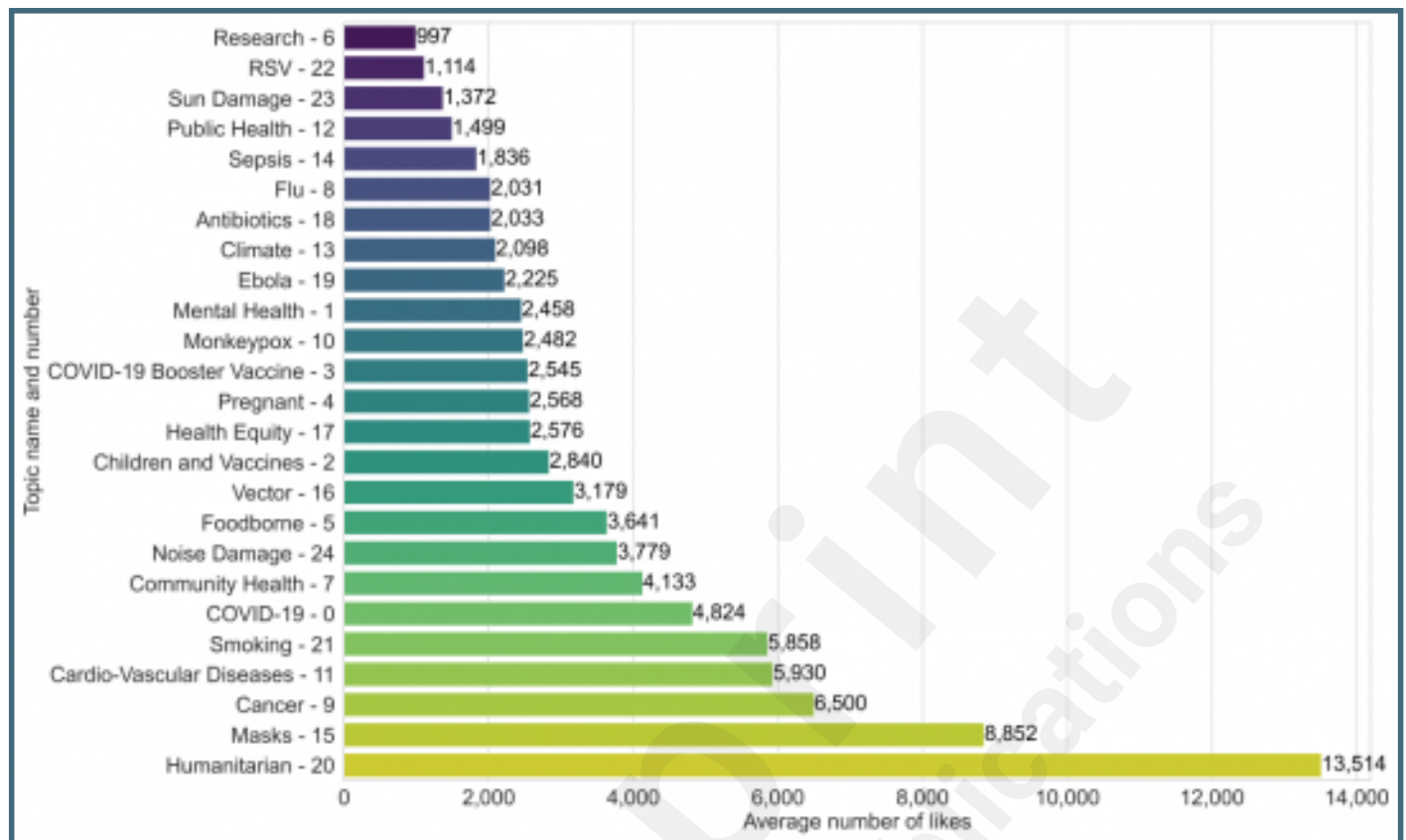
An overview of the methodology's main phases.



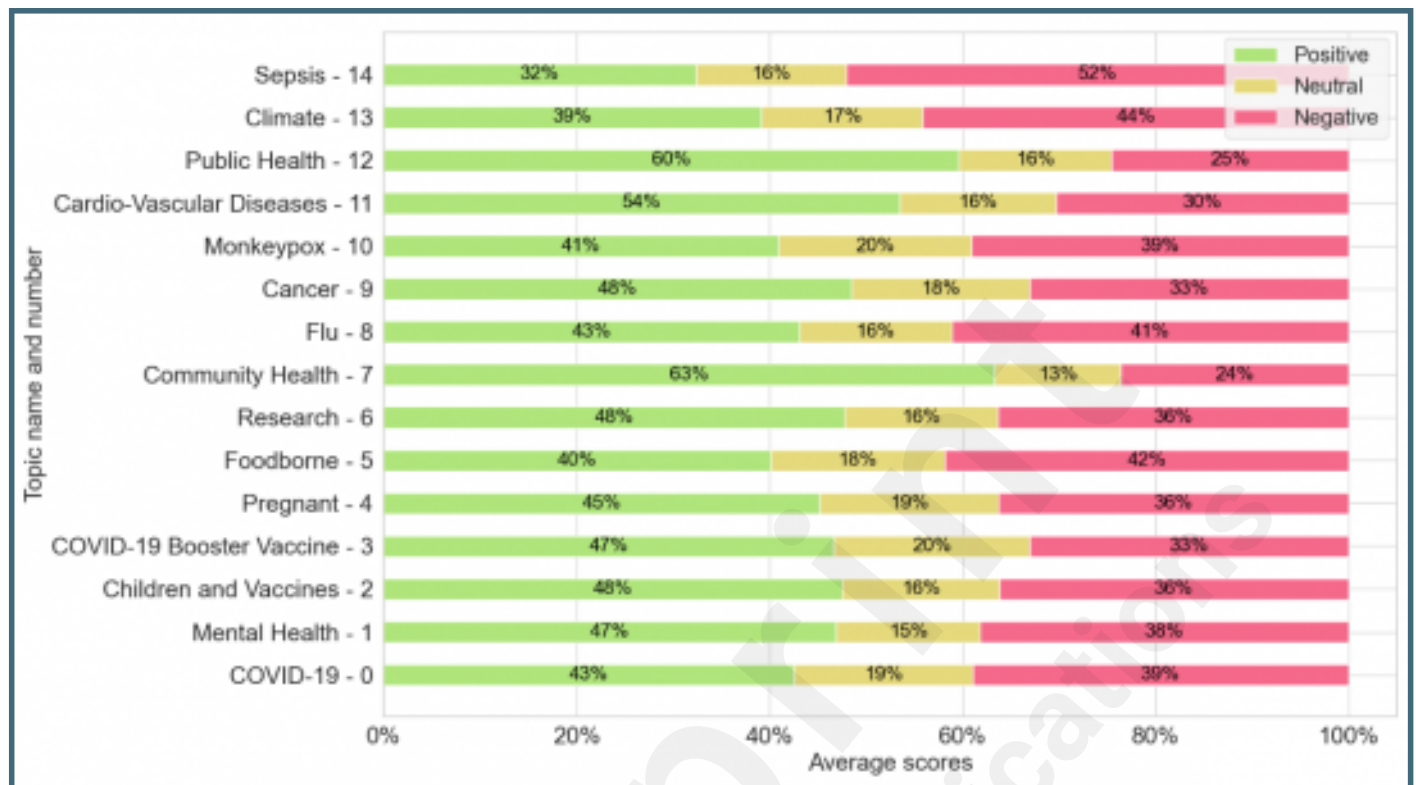
Average number of post comments per topic.



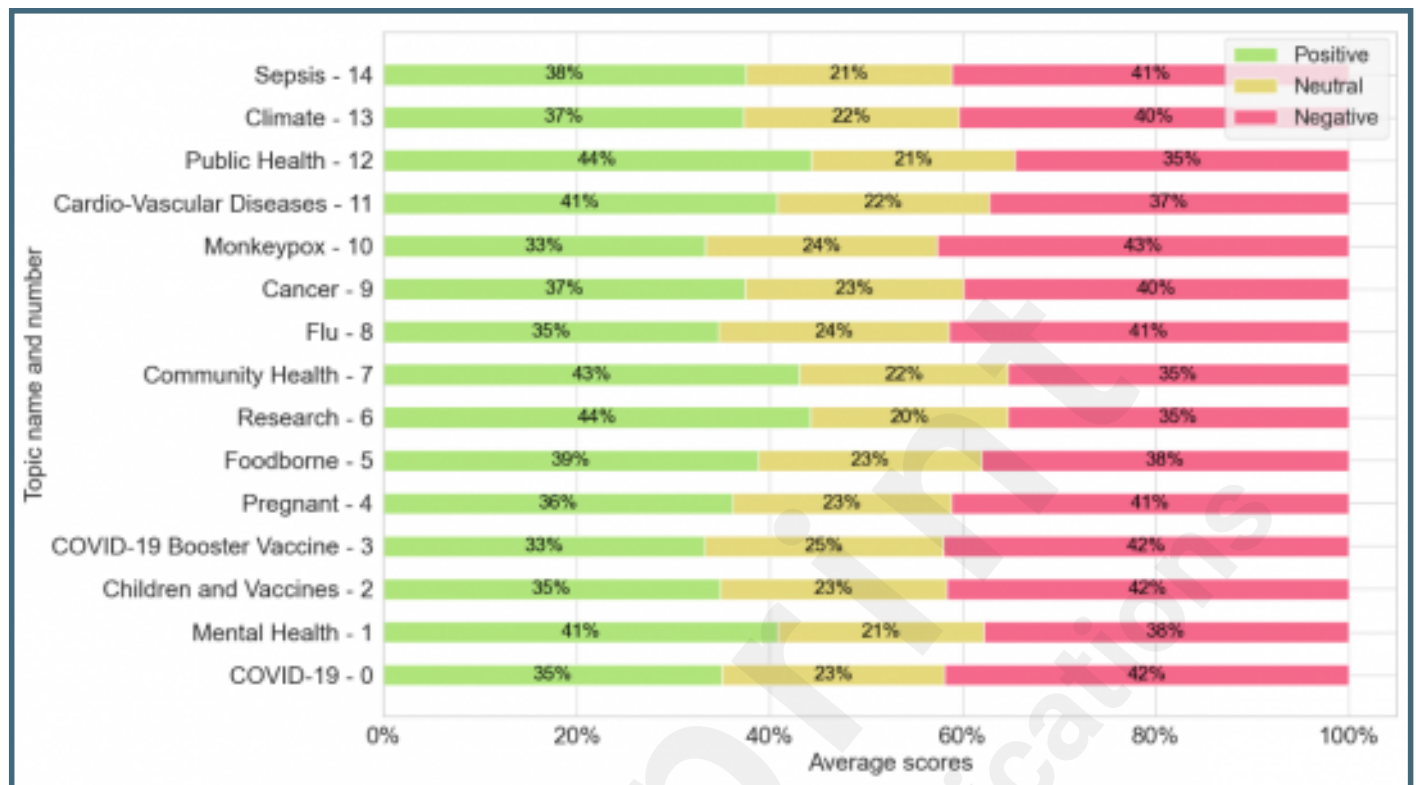
Average number of post likes per topic.



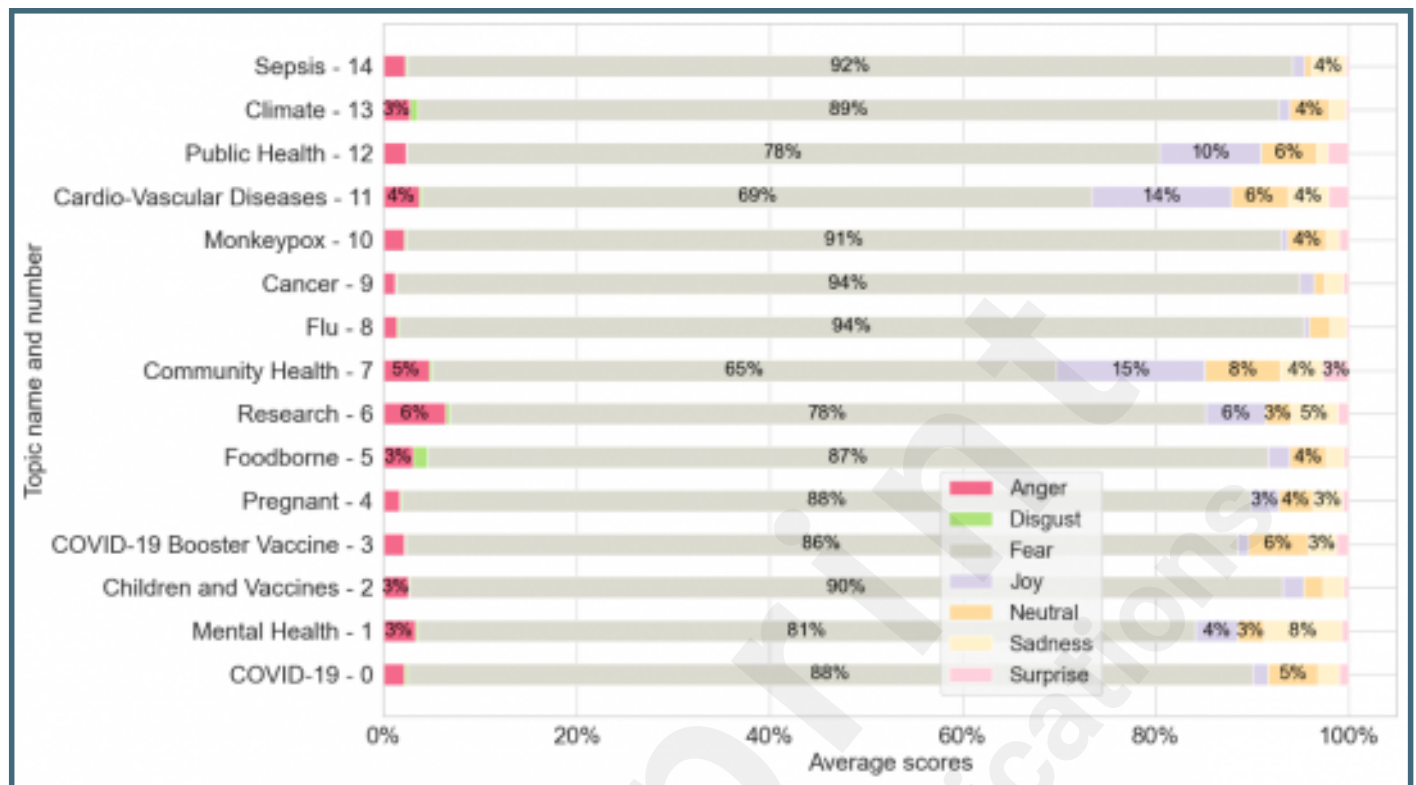
Posts' sentiment scores (average positive, neutral, negative) for each topic.



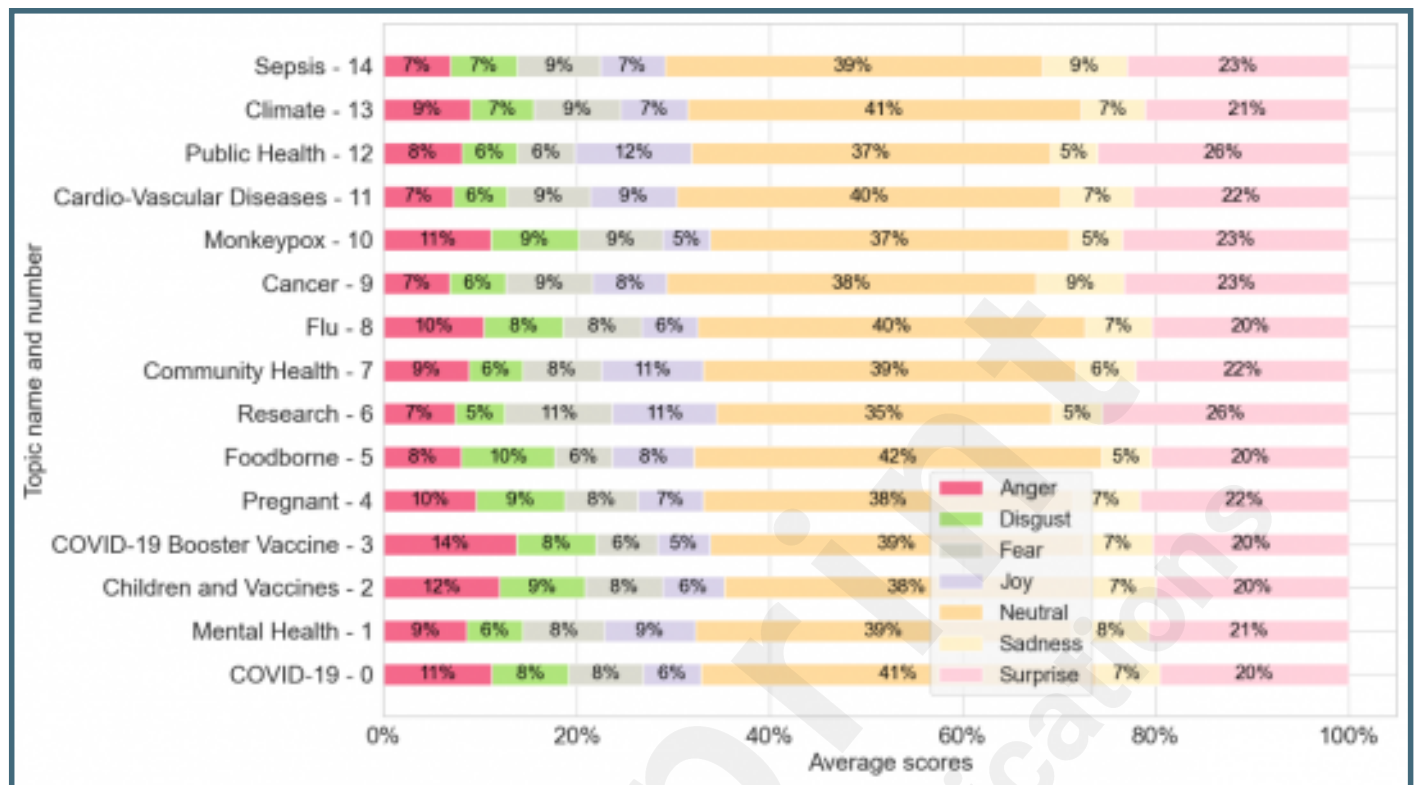
Comments' sentiment scores (average positive, neutral, negative) for each topic.



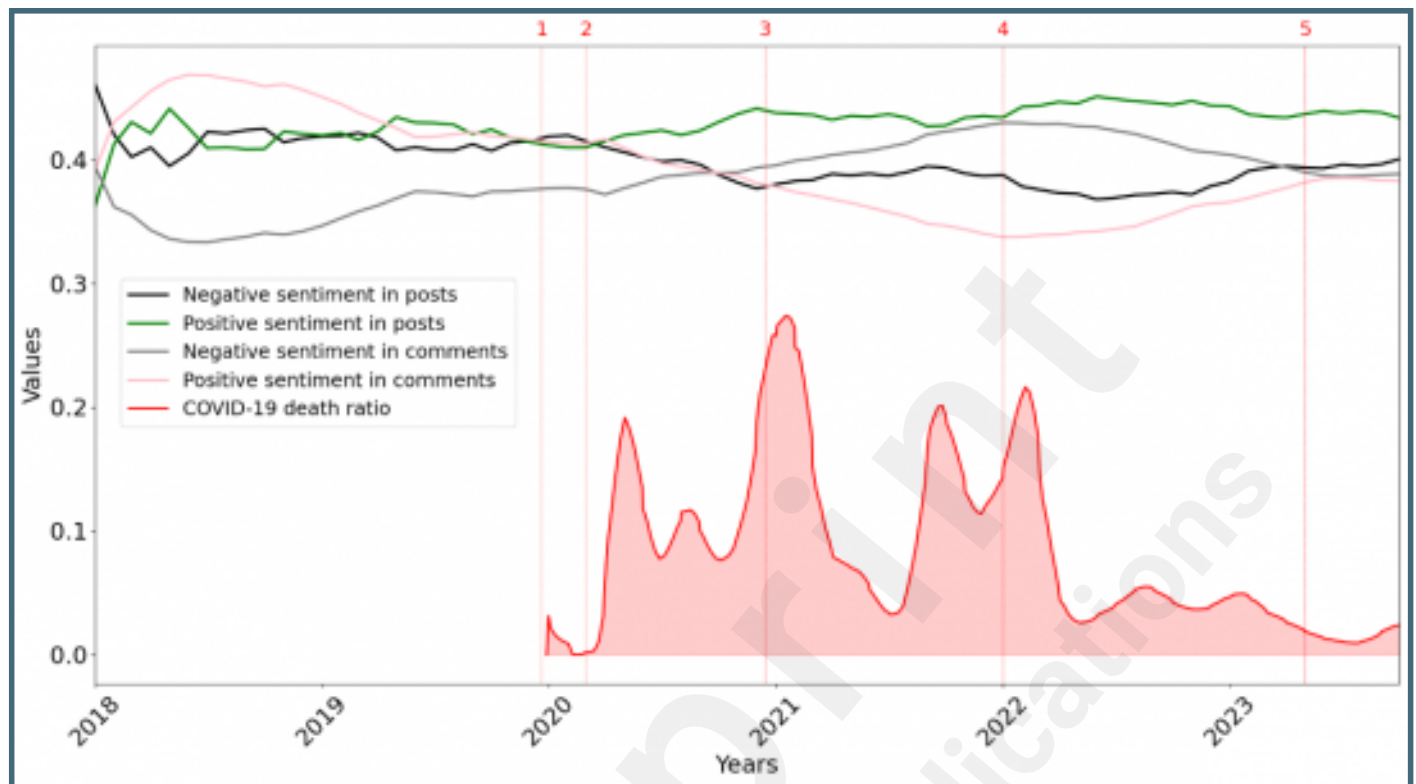
Posts' emotion scores (average anger, disgust, fear, joy, neutral, sadness, and surprise) for each topic.



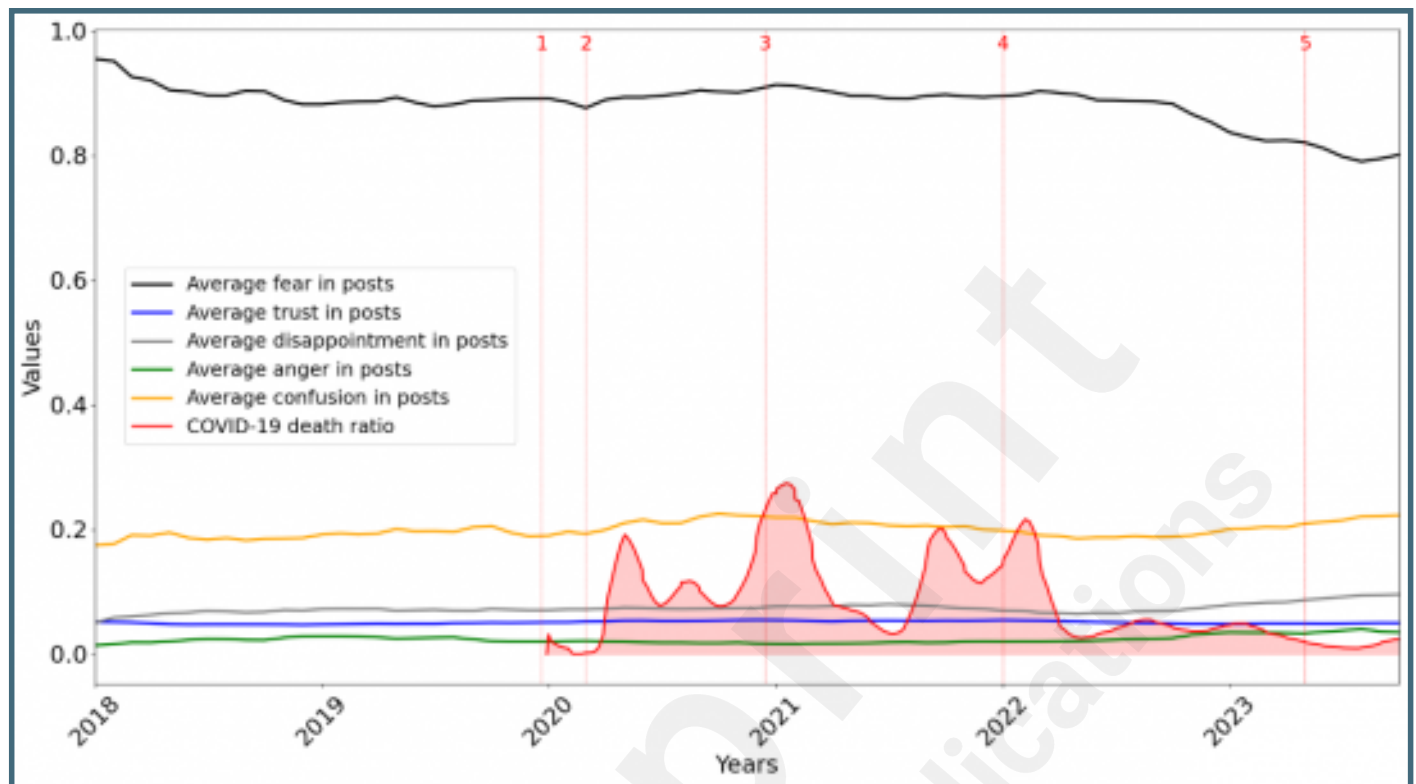
Comments' emotion scores (average anger, disgust, fear, joy, neutral, sadness, surprise) for each topic.



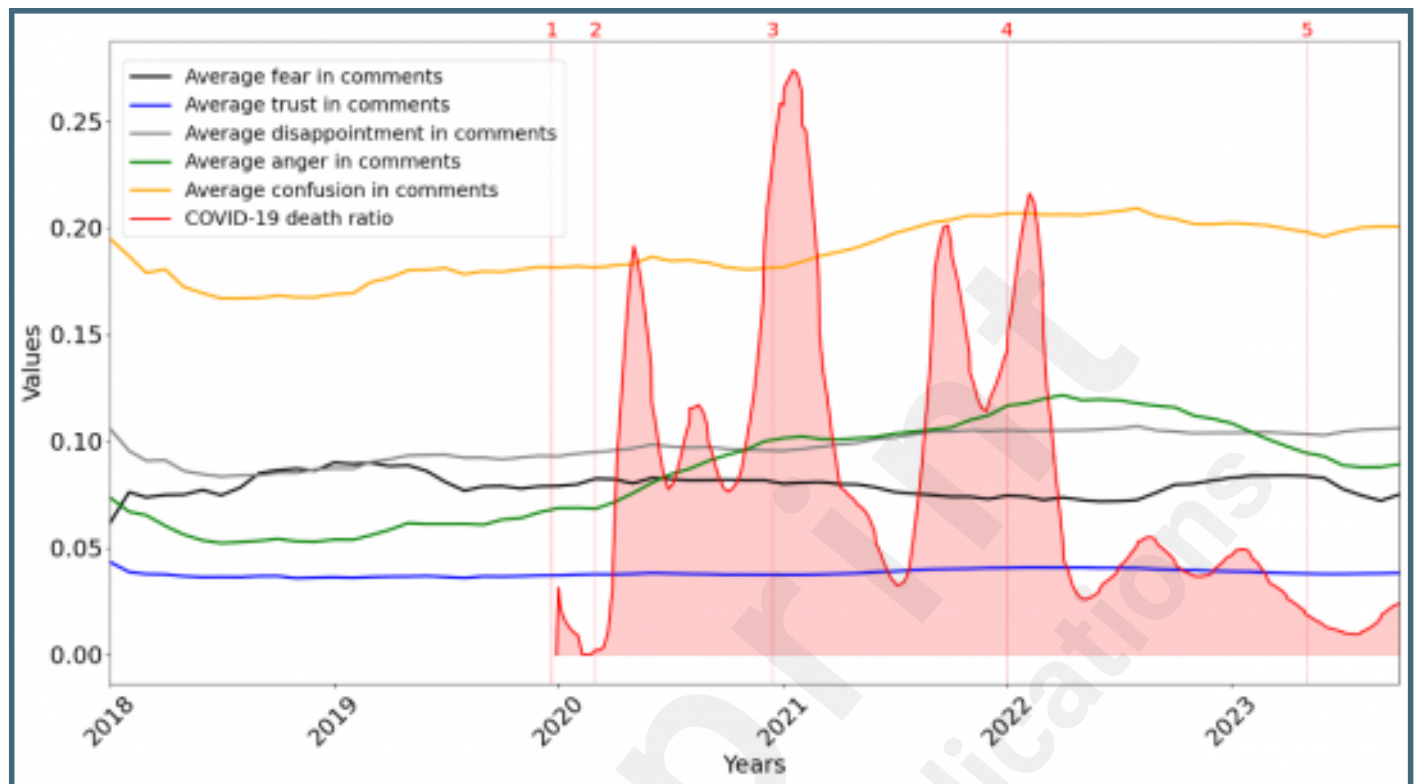
Sentiment trends (average positive and negative scores) in posts and comments over the years, combined with the COVID-19 death ratio.



Emotion trends (average scores for trust, fear, anger, confusion, and disappointment) in posts over the years, combined with the COVID-19 death ratio.



Emotion trends (average scores for trust, fear, anger, confusion, and disappointment) in comments over the years, combined with the COVID-19 death ratio.



Multimedia Appendixes

The number, name, size, and top-5 most significant words of each topic.

URL: <http://asset.jmir.pub/assets/ce52db0686b7314ad518d93786c736b1.docx>

Top-50 phrases for unigrams, bigrams, trigrams, and four-grams in posts receiving predominantly positive or negative responses with corresponding TF-IDF scores.

URL: <http://asset.jmir.pub/assets/64bcaa3cb86c7b68cc3dd96e19e1e874.xlsx>

