This accepted version of the article may differ from the final published version.
This is an Accepted Manuscript for *Disaster Medicine and Public Health Preparedness*
as part of the Cambridge Coronavirus Collection
DOI: 10.1017/dmp.2022.207

Enhancing Artificial Intelligence for Twitter-based Public Discourse on Food Security During the COVID-19 Pandemic

Nina M. Martin, PhD[1,*], Lisa Poirier, MHS[2], Andrew J. Rosenblum, MSPH[3,4], Melissa M. Reznar, PhD, MPH[5], Joel Gittelsohn, PhD[6], Daniel J. Barnett, MD, MPH[7]

[1]Assistant Scientist, Johns Hopkins Bloomberg School of Public Health, Department of International Health, Human Nutrition Program, Address: 615 N. Wolfe St., Baltimore, MD, 21205, Email : nmarti38@jhu.edu
[2]Research Associate, Johns Hopkins Bloomberg School of Public Health, Department of International Health, 615 N Wolfe St, Baltimore, MD, 21205, Email: lpoirie4@jhu.edu
[3]Law Student, American University Washington College of Law, 4300 Nebraska Ave. NW, Washington DC 20016
[4]Part Time Lecturer, Krieger School of Arts & Sciences, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218, Email: arosenblum@jhu.edu
[5]Associate Professor, Oakland University School of Health Sciences, 433 Meadow Brook Road Rochester, MI 48309, Phone : 248-364-8668, Email: reznar@oakland.edu
[6]Professor, Johns Hopkins Bloomberg School of Public Health, Department of International Health, Center for Human Nutrition, 615 N Wolfe St, Baltimore, MD, 21205. Phone: 410-955-3927 Email: jgittel1@jh.edu
[7]Associate Professor, Johns Hopkins Bloomberg School of Public Health, Department of Environmental Health and Engineering, 615 N. Wolfe Street, Room E7036, Baltimore, Maryland 21205, Phone: 410-502-0591, Fax: 410-955-0617, Email: dbarnet4@jhu.edu

***Corresponding author:** Email : nmarti38@jhu.edu

Short Running Title: Public Discourse on Food Security During COVID-19 Pandemic

Sources of support that require acknowledgment: Bloomberg American Health Initiative

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

ABSTRACT

Objective: Food security during public health emergencies relies on situational awareness of needs and resources. Artificial intelligence (AI) has revolutionized situational awareness during crises, allowing the allocation of resources to needs through machine learning algorithms. Limited research exists monitoring Twitter for changes in the food security-related public discourse during the COVID-19 pandemic. We aim to address that gap with AI by classifying food security topics on Twitter and showing topic frequency per day.

Methods: Tweets were scraped from Twitter from January 2020 through December 2021 using food security keywords. Latent Dirichlet Allocation (LDA) topic modeling was performed, followed by time-series analyses on topic frequency per day.

Results: 237,107 tweets were scraped and classified into topics, including food needs and resources, emergency preparedness and response, and mental/physical health. After the WHO's pandemic declaration, there were relative increases in topic density per day regarding food pantries, food banks, economic and food security crises, essential services, and emergency preparedness advice. Threats to food security in Tigray emerged in 2021.

Conclusions: AI is a powerful yet underused tool to monitor food insecurity on social media. Machine learning tools to improve emergency response should be prioritized, along with measurement of impact. Further food insecurity word patterns testing, as generated by this research, with supervised machine learning models can accelerate the uptake of these tools by policymakers and aid organizations.

Key words: Food Security, Machine Learning, Topic Modeling, Twitter, Natural Language Processing

Abbreviations: U.S. (United States)), SNAP (Supplemental Nutrition Assistance Program), GPS (Global Positioning System), WIC (SNAP for Women, Infants, and Children), Application Programming Interface (API), Latent Dirichlet Allocation (LDA), artificial intelligence (AI).

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

1. INTRODUCTION

Food insecurity has surged globally during the COVID-19 pandemic.[1] Worldwide, food assistance and distribution remains one of the most basic needs, both in emergency and non-emergency settings. In the United States (U.S.), rates of food insecurity generally have continue to increase, particularly among lower income adults.[2–4] This trend has been exacerbated by the COVID-19 pandemic, which caused an unprecedented surge in food insecurity nationwide.[6–8] Since the beginning of the pandemic, an unprecedented number of food banks experienced supply shortages and many were forced to close.[9] Food pantries have traditionally provided critical nutritional resources for clients, especially during disasters in already food insecure populations; however, the magnitude and duration of the COVID-19 pandemic has exposed critical limitations amid unparalleled needs.

Simultaneously, as other fields have applied insights from COVID-19 to mature their real-time situational awareness technologies, these efforts have been notably lacking for food distribution agencies.[5] Notably, other public health fields have employed social listening,[6,7] or the monitoring of social media networks in real time, for trends in topics, sentiments, misinformation, and emerging issues.[8] During the COVID-19 pandemic, public health researchers used social listening to spot and correct COVID-19 misinformation, understand opinions on COVID-19 vaccinations and masks, report emergencies, and track symptoms and emerging cases. In other fields, social listening has been used to combat the *Infodemic*, or the overwhelming amount of information—trustworthy and untrustworthy—that is now found on online platforms.[7,9,10] Both the World Health Organization (WHO) and the United States Centers for Disease Control and Prevention have recently placed an emphasis on further developing these tools in the last couple of years.[7,9–11] The WHO in particular is developing tools to advance real-time social listening that promotes an adaptive public health response and can be iteratively improved as new information is collected.[6]

A strength of these social listening strategies lies in the near ubiquity of social media and digital communications platforms. This level of engagement between peers can drive individual and group-level behavior change.[12] That engagement can in turn be used to shape behaviors via

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

digital 'networking' that reinforces or drives behaviors. To date, these strategies have shown a wide range of successful uses, from diabetes management in adolescents[13] and prevention,[14] to weight reduction[15] and increasing consumer spending through group discounts.[16] Based on these previous findings, scalable digital strategies are urgently needed to promote food assistance use and management.

A more advanced tool for social listening is machine learning. Machine learning algorithms, a subset of artificial intelligence, can analyze a body of texts, such as tweets or news media posts, and categorize messages into topics or themes, known as topic modelling.[17] Additionally, other machine learning algorithms can categorize texts based on their positive or negative sentiments.[18] By combining topic modelling with sentiment analysis, public health researchers can gain rapid insight into the current state of thinking and feeling of large populations of the general public—or population segments thereof. [19]

To date, these technologies have seen limited application in the food assistance sector. While there are studies generally focused on classifying needs and resource messages, and predicting needs, few studies have examined food-specific conversations.[20,21] Past challenges in this applied research context have included limited accessibility to machine learning tools outside of computer science. That limitation makes it difficult to adapt algorithms and define parameters that relate to complex and multivariate public health concepts like food security. However, the past five years have seen a marked increase in the use of these technologies, as well as an increase in open access code and software packages that simplify previously arcane knowledge.

These technologies will help the food assistance sector improve interconnectivity between food banks and those in need, offer cost-effective solutions to overcome funding barriers, and provide stable scaling of a robust distribution network. This is particularly relevant for the scale of the COVID-19 pandemic and potential future disasters that affect lower-income populations.[5,22]

Technological innovation represents a key opportunity to address the disconnect between the need and availability of supplemental nutrition. A recent scoping review of digital tools for emergency food assistance found that the food security sector is lagging behind others in terms

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

of use and research on modern digital technologies to improve situational awareness during disasters.[5] As a follow up to that research, this study explores using artificial intelligence—specifically the machine learning algorithm Latent Dirichlet Allocation (LDA) model—to mine Twitter for relevant tweets and analyze the topics within conversations. The aims of this study were to:

1. Develop and test a Twitter query string to pull all tweets related to food insecurity and food assistance from January 2020 to December 2021.
2. Determine topics within the corpus of pulled tweets by evaluating and iteratively improving an LDA topic model.
3. Perform a time series analysis on frequency of each topic per day on Twitter from 2020 to 2021.

2. METHODS

To assess tweets pertaining to food security on Twitter, a unique search string is required to effectively retrieve relevant Twitter messages. The information retrieval query to the Twitter Application Program Interface (API) was built using a three-prong process: 1) a Scoping literature review; 2) compilation of a table of words used by other studies; and 3) expert stakeholder input to the iterative list. Ten food security stakeholders (2 academics, 5 food policy actors, 3 emergency preparedness experts) provided input into a social media search string regarding food security.

A. Data collection

Once the unique string was created, a Python[23] based script was utilized to run a daily pull from the Twitter API from January 1, 2020 through December 31, 2021 (Table 1). The start date was chosen to allow for analysis of Tweets in the months prior to the WHO pandemic declaration on March 11, 2020. Tweets from January 1, 2020 to March 10, 2020 were used as a comparison group, by topic frequency, for the pre-pandemic phase. The methods to request and use the Twitter API have been described previously.[24] Results were collected, analyzed and visualized using Python and a Google Colaboratory[25] notebook. All Google Colaboratory notebooks are available on GitHub.[26] Historic tweets were collected through use of a Twitter (San Francisco,

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

California USA) developer account on the Twitter for Academic Research by the corresponding author.[27]

B. Data Preprocessing

After exploratory data analysis, the data was cleaned through deduplication of tweets and removal of retweets, mentions, and URL's. Emojis and other non-characters, extra space, and punctuation were removed, and all letters were converted to lowercase. Stop words, e.g., the/in/or, were removed. Word, bigram, and trigram frequency per document and the entire corpus was calculated and visualized using re,[28] numpy,[29] pandas,[30] gensim,[31] WordCloud,[32] spacy,[33] matplotlib,[34] PIL.[35]

C. Topic Modeling

Unsupervised machine learning, namely Latent Dirichlet allocation (LDA), was used to examine data for patterns. This is a commonly used statistical-based model for exploratory data analysis of previously unseen unstructured textual data.[17,19,36,37] LDA topic modeling is preferred to traditional qualitative approaches because it summarizes textual data at a scale beyond the capacity of human annotation.[36] By deriving a probabilistic clustering, or latent topic distributions, LDA topic models can identify patterns, themes, and structures of the Tweets texts; it enables efficient classification of large bodies of data based on patterns and features.

Topic modeling had been widely used to gain a descriptive understating of unstructured Twitter big data in social science research [12].

Using the Latent Dirichlet Allocation (LDA) algorithm, 22 topics in the corpus[26] were analyzed as described by Sharma.[38]

In order to select the optimal number of topics to analyze from the corpus, coherence score, alpha, and beta model hyperparameters were calculated on 75% of the corpus as described by Kapadia.[39]

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

To iteratively improve and understand the 22 generated topics, a series of metrics and visualizations were considered including dominant topics, representative sentences for each topic, distribution of dominant topics within each document, visualization of the most important words that make up each topic, and word cloud images. To understand the importance (or weight) of the keywords in the topics, frequency and density of top topic words, calculations were made using collections and matplotlib packages. If there were words that had higher frequency than weight in multiple topics, this indicated that they were most likely not important to the topic and were added to the list of stop words.[39] To visualize which topics are most discussed in each and all documents, pandas and matplotlib were used to calculate both the number of documents by dominant topic and by topic weightage. Finally, the pyLDAvis package produced a dynamic intertopic distance map (via multidimensional scaling) including a list of the top 30 most salient terms overall and 30 most relevant terms per topic.[38]

D. Time series analysis of topics

To determine the change in the nature or content of food security-related Tweets from before the WHO pandemic declaration on March 11, 2020 (i.e., January 2020 – March 10, 2020) versus after, a time series analysis of the number of posts per day per topic was calculated. The Pandas[30] python package was used to calculate a timeseries analysis of the number of posts per day. The Calplot[40] package was used to create a heatmap of the results.

3. RESULTS

3.1 Stakeholder-derived food security search string

Based on stakeholder input, the final search terms were: "food emergency OR food crisis OR food aid OR food need OR food corona OR food rona OR food assistance OR food distribution OR food distribute OR food free OR WIC OR food security OR food insecurity OR food hungry OR #freefood OR #mealservice OR food pantry OR food bank OR covid food."

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

3.2 Exploratory Data Analysis

The search returned a total of 116,500 tweets from 2020 and 120,607 tweets from 2021 (Figure 1 and Supplementary Table 1). Of the tweets from 2020, there were 919,317 likes, 17,572 quotes, 72,655 replies, and 215,398 retweets. In 2021, there were 1,038,339 likes, 23,036 quotes, 81,430 replies, and 294,122 retweets. The top five platforms for posting to Twitter for both years are Twitter for iPhone, Twitter for Android, the Twitter Web App, Twitter for iPad, and the OLIO food sharing app (Figure 1). Each OLIO app post includes the hashtags #free, #UnitedKingdom, #food, and #foodwaste. These hashtags were in the top 10 hashtags for both years (Supplementary Table 3). Other popular hashtags included, #COVID19, #TigrayGenocide, #StopTigrayFamine, and #fastfoodies.

During 2020, presidential candidates Donald Trump and Joe Biden were in the top ten mentions list (Supplementary Table 3). Other government officials included @senatemajldr, Speaker Pelosi, and the White House. @Pulte is the family name of a company and philanthropist. In 2021, top mentions included the United Nations, POTUS, Secretary Blinken, Pulte, EU Commission, and the World Food Programme.

The most liked and most retweeted tweet of 2020 was, “Here is a grown man complaining about the free food at an event he's not going to https://t.co/c0wQTfdL71” (Figure 1 and Supplementary Table 4). For 2021, the most liked and retweeted tweet was, “I saw this on FB and it's the best food pantry donation list I've ever seen. Usually posts like this either end up shaming the donor for what they do or don't give, or shaming the receiver for what they do or don't choose to eat. As a former pantry employee, I endorse! https://t.co/GdvAkLHUDY.”

3.3 Topic Model Evaluation

The results of the parameter and hyperparameter (coherence score, alpha, and beta) are presented in Supplementary Figure 1. 22 topics were chosen as the optimal topic number with a coherence value of .514. While there were topic numbers with a slightly higher coherence score, these topics either included repeat topics or missed important sub-topics upon examination. Alpha and beta hyperparameters were also optimized at .91.

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

3.4 Topic Model Results

LDA topic modeling returned the top 30 terms per topic (Table 2). Each topic was labeled based on these terms as well as representative tweets. Table 6 shows the categories, topics, tweet counts per category and topic, and at least one representative tweet for each topic. The categories, described in the subsequent sub-sections, included: (1) food assistance, needs, and resources, (2) emergency preparedness and response, (3) mental and physical health, (4) stay-at-home life, and (5) entertainment.

I. Food Assistance, Needs and Resources

Food assistance, needs and resources as a category comprised the dominant category with 31.2% of all tweets (n = 73,924). The United Kingdom-based smartphone application OLIO returned the following prominent tweets in categories related to food sharing: free, foodwaste, meat (n = 21,392). All posts in this category included the hashtags #foodwaste or #free, plus a location. A representative tweet in this category is: "bagels plain 5 pack 14 april donated by Tesco in #London #UnitedKingdom https://t.co/2KitI4TQ4w #foodwaste #free." Other topics in this category include food pantries (family, donate, support; n = 14,987); government assistance (assistance, stamp, program; n = 11,491); food systems (security, grow, system; n = 11,247); pet food (food, buy, feed; n = 8,815); and food banks (food, week, bank; n = 5,992).

II. Emergency Preparedness and Response

Tweets relating to emergency preparedness and response comprised 30.7% (n = 72,783) of all posts. The most dominant topic in this category related to emergency aid following the unrest in Tigray, Ethiopia in 2021, resulting in tweets regarding food insecurity and calls for aid. For example, one user posted, "@GlobalGsts @WHO @DrTedros @OCHA_Ethiopia @GlobalFund @WFP @gavi @ICRC @ICRC_Africa @ifrc @IFRCAfrica @UNDP @UNICEF @UNICEFEthiopia @AfricaCDC Please help the Tigrians with basic humanitarian assistance, such as food water and medicine. Politics will never end." Other types of emergencies related to individual financial crises (money, pay, buy; n = 12,561); COVID-19-related food insecurity (insecurity, covid, crisis; n = 10,787); family needs (people, child, care; 9,672); essential services

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

(work, covid, home; 9,170); advice for emergency preparedness (food, water, emergency; n = 8,421); and values (food, people, thing; n = 7,090).

III. Mental and Physical Health

Topics related to mental and physical health compromised 15.8% (n = 37,432) of tweets. Users posted about well-being and nutrition (eat, feel, hungry; n = 13,077); food as self-medication (food, day, today; n = 9,141); food as love (food, give, love; n= 9,082); and recommendations (food, time, year; n = 6,131). Thousands of users posted about using food as a mechanism to feel better during the pandemic (e.g., food as self-medication). For example, "Okay, I just need to start on some food. That might help distract me from whatever this day is throwing at me." Food as love refers to people gifting food as an act of kindness and/or love. For example, "@CielPha52115605 "I'll always be here love" no matter what they go through, he will always be there to give her hugs and buy her food when shes [sic] hungry, let her rest when shes [sic] tired." Recommendation posts include content like: "@streets @thewhicha @ScooterRiderFMX @InsulateLove £90 from dyas -don't get the expensive one unless u can afford it -you don't need to plan -you put the ingredients in the night before -set the timer and come home to delicious healthy food kept warm and sealed on slow cook for hrs."

IV. Stay-At-Home Life

Stay-At-Home Life (n = 31,823, 13.4%) represents tweets pertaining to daily life eating and other types of habits while people were forced to stay at home during the early phases of the pandemic. Of these tweets, messages about hunger were most abundant (eat, feel, hungry; n = 13,117). Hunger in this case does not refer to food insecurity in the literal sense. For example, one user shared, "I'm cold. I'm tired. I'm hungry. Therefore cuddles, a movie and some food delivery sound like a perfect date [right now]." Other topics in this category include food delivery (food, order, call; n = 9,395) and daily eating habits (food, back, leave; n = 9,311).

V. Entertainment

The entertainment category (n = 21,145, 8.9%) includes two topics, free drinks and entertainment (free, drink, watch; n = 12,102) and free/cheap food (food, good, free; n = 9,043).

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

There were 21,145 tweets assigned to this category, representing 8.9% of tweets. Examples of tweets from this category includes, “TACO TUESDAY AT SOHO LOUNGE ITS GOIN UP #FREEFOOD https://t.co/F4jZn7craX” and “anyone else thinks food tastes better when it's free?”

3.5 Time Series Results

The results of the time series analysis were plotted in a heatmap (Figure 2) for January 1, 2020 to December 31, 2021. Five topics clearly showed a sharp increase in posts per day in March 2020. This included food pantries, food banks, individual financial crisis, food insecurity crisis, essential services, and advice for emergency preparedness. Other topics showed a sharp decrease in the number of posts per day, including food sharing via the OLIO app, well-being and nutrition, hunger generally, and free drinks and entertainment. The topics of food systems, daily eating habits, and recommendations remained consistently high over the two-year period. The topic regarding aid for Tigray increased sharply in 2021, especially in the second half of the year.

4. DISCUSSION

The COVID-19 pandemic has upended nearly all facets of society, including food security. The combination of big data and artificial intelligence has greatly improved situational awareness, but its potential in food security remains underused. This is the first study to explore the food security conversation on Twitter throughout the COVID-19 pandemic using big data and AI. This study also highlights the utility of machine learning to detect needs and resources on Twitter as well as other discourse related to food security. Several salient findings remained relevant throughout the first two-years of the pandemic: the continuous need and requests for philanthropy (e.g., accounts like Pulte), modifications of daily eating habits particularly during periods of social distancing, food systems discussions (e.g., farming, technology, policies, practices), and the ever-present need and utilization of food banks and pantries. In addition, this study was successful at achieving its three aims.

First, this study improved upon previous work on food security topic modeling that used too broad or too narrow searches to retrieve all relevant food security tweets.[41,42] The current study

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

succeeded in defining first a search string, based on a literature review and stakeholder input. Second, this study defined twelve topics and related top 30 salient terms for retrieving food security tweets. The other topics need additional follow up analysis to determine if tweets are related to food insecurity or not. For example, while some tweets in the mental and physical health category are clearly related to food insecurity, others need more context to determine relevance. Identifying key words and parameters that can differentiate between food security and insecurity tweets, as has been done in this study, are vital to future research aiming for real-time analysis of food needs and resources. This labeled data set will be used as the basis for future supervised learning algorithms.

While this progress on food security is valuable, future work can be improved upon by taking guidance from other fields that have made progress on identifying needs and resources on microblogs, even though they are not explicitly focused on food.[42–48] Recent studies have improved performance of classifying needs and resources through the use of key terms mapped to crisis scenarios and the general detection of needs and resources.[49] This includes crisis-related lexicons, e.g., EMTerms[50] and CrisisLex[51], that contain over 7,000 terms used in Twitter to describe various crises. A specific lexicon related to food security terms could be beneficial in disaster and non-disaster settings.

The identification of 12 topics and associated key words that were clearly related to food security will be used as the base data set for additional research. Future work should aim to supplement the irregular or infrequently collected real-time national and global level surveys on food security. This study also shows that Twitter is an excellent source of information on the public's opinion and knowledge the entire disaster lifecycle, including preparedness, response, mitigation, and resilience. Other work supports that social media is an important venue for big data and AI analysis regarding the disaster lifecycle—and that more such research can improve real-time response and preparedness.[52–59]

Food intake has long been associated with mental health and well-being,[60–63] so it is unsurprising that the third most popular tweet category involved these themes. During the COVID-19 pandemic, other studies have shown that lock-down periods were associated vicious cycles of

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

increased food insecurity and worse mental health outcomes—like depression and anxiety disorder—and in turn, these disorders lead to more unhealthy eating.[64–68] Observations from the current study supported this cycle of food insecurity, poor mental health, and increased unhealthy eating via increased frequency of tweets about fast food, home food delivery as well as isolation, tiredness and anxiety. However, further analysis is needed to distinguish if posts in this category are truly a reflection of food insecurity or not. For example, users posting about being hungry may or may not be food insecure. Nevertheless, our research supports results found elsewhere using surveys and in-depth interviews[66,68] that the declaration of the pandemic and stay-at-home orders had a profound impact on food insecurity as well as mental health, physical health, and daily life.

Additionally, these categories may also reflect a growing demand among public health professionals and stakeholders for updating the concept of 'food security' to 'nutrition security.'[69–71] Nutrition security is an emerging theme in public health that highlights that improving health outcomes can only be achieved by providing nutritious food, not just any food. The results of this study can be used to guide stakeholders in their efforts to target getting nutritious food to those in need by correlating the need to the frequency of posts.

There were relative differences in topic density, as shown by the time series analysis, on Twitter before (January 1-March 10, 2020) and after the WHO declaration of the COVID-19 pandemic (March 11 – December 31, 2021). Based on our analysis of Twitter-based social discourse on food security during COVID-19, topics that increased in frequency included emerging threats to food security (Tigray, food shortages, corruption/warehouse food hoarding, crop threats), fast food and home delivery, aid organizations and donation requests, individual and large-scale economic crises, stimulus checks, stay at home life, and food pantry advertisements. In concordance with these increased discussion points, related accounts also became more engaged in the discussion, including the World Food Programme, the United Nations and the President of the United States (i.e., President Biden). Detecting real-time trends in public food insecurity discourse can fill the gaps of survey-based data, i.e., the USDA's Household Food Security Survey, which had its last update in 2020.

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

As the pandemic moved from an initial disruption of daily life to an extended response and concurrent recovery, novelty aspects of staying at home, entertainment to abate pandemic boredom, and political overtones dissipated. These findings were noted as politicians and ideas such as free food and drinks occurred less frequently in identified tweets.

LIMITATIONS

This study is subject to several limitations. The complexity of the two main concepts – food assistance and detecting needs and resources – creates many permutations of possibly relevant tweets. While the search was built iteratively and with multiple stakeholder's input, it is possible a subset of relevant tweets was missed. However, more analysis is needed to determine if tweets classified as 'Entertainment,' Mental Health and Well-being,' and 'Daily Eating Habits' are related to food insecurity. For example, users mentioning hunger is not necessarily an indication of food insecurity and the context of such posts must be examined to determine relevancy. Additionally, while our query was not tailored to English language posts, our search terms may have only promoted the retrieval of English language tweets to the exclusion of other languages. Thus, there could be valuable, non-English language posts that were missed. Future work can improve on these limitations by using supervised learning approaches with word embeddings and/or crisis specific features that have been used to in disaster settings to classify tweets into informative versus non-informative and/or, more specifically, into the six needs and resource classes as defined by UNOCHA (United Nations Office for the Coordination of Humanitarian Affairs)[72] and resources such as CrisisLex[51] and EMTerms[50].

While Twitter is a popular site for public discourse, it may not capture some percentage, minor or major, that other social media platforms, such as Instagram, TikTok, or Facebook, may cover. Additionally, our study did not exclude tweets that did not pertain to food security. All topics were analyzed for relevance to the study themes and were deemed related to food assistance; however, some of the results may have been biased by non-relevant tweets. This deliberate choice to seek broader terms sought to better categorize the available evidence, and no extraneous topics appeared in the top 10 words of each topic. By consensus, the topic themes were generated with confidence.

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

Unsupervised learning algorithms, such as the LDA algorithm used for topic modeling, are known to have limitations, such as lower accuracy of assigning tweets to the correct topic.[37,39,73] For exploratory data analysis, observations, and preliminary detection of topics in an unfamiliar corpus, it is acceptable to use LDA modeling despite these known limitations.[38,39] Future work could take the results of this topic modeling analysis as the basis for feature selection for more accurate, supervised learning algorithms.

Despite these limitations, this study underscores timely gaps in emergency preparedness interventions and response while defining the scope of existing public discourse on food assistance during the first two years of the pandemic.

CONCLUSIONS

As the pandemic enters year three, real-time big data collection and analysis with artificial intelligence remains important. This study demonstrated how Twitter and artificial intelligence can be leveraged to increase preparedness and response to food insecurity in future disasters in an all-hazards environment. These new solutions to for real-time detection of food needs and resources point to the pressing need for the development of stabile food security programs.[9]

This study is the first to develop a query string for food security public discourse during the COVID-19 pandemic. Future research could utilize the topic key words associated with food security, as the basis for supervised learning models for detecting real-time food needs, resources, and emerging threats to food security. Other fields have improved their real-time analyses by feeding their base topic model results into more advanced supervised learning algorithms and feature development.[41]Current platforms for showing food security information are lacking in real-time data, modern user interfaces, and actionable or usable data for policymakers and aid organization to make real-time resource allocation decisions.

As reflected in this study, there are many policy and practice implications. The results of this study show the pressing need to understand resource allocation challenges during emergencies while enhancing public discourse-informed situational awareness and optimizing food insecurity-response strategies. This model can also be useful for understanding emerging threats to food security. For example, this model unexpectedly focused in on the impact of Tigray on

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

public discourse and concern for food security. The topic modeling was also useful in identifying the economic implications of Covid on global, national, and individual financial crises as well as supply and demand and supply chain issues. Stakeholders can also use this modeling to understand public opinion on food security policies and programs.

As global disasters continue to increase due to climate change and political instability, the development and evaluation of tools that address all disaster phases– prevention, preparedness, response, and recovery—are urgently needed. Machine learning tools to improve emergency response should be prioritized, along with rigorous qualitative and quantitative measurement of impact. Further testing of the generated search string with supervised machine learning models as identified by this study will maximize the efficacy and uptake of these promising digital food security tools by policymakers and aid organizations.

ACKNOWLEDGEMENTS

We would like to thank the Bloomberg American Health Initiative for funding support.

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

REFERENCES

1. HungerMap LIVE. World Food Programme. Published 2018. Accessed August 18, 2021. https://hungermap.wfp.org/

2. Coleman-Jensen A. Household Food Security in the United States in 2020. :2.

3. Pooler JA, Srinivasan M, Miller Z, Mian P. Prevalence and Risk Factors for Food Insecurity Among Low-Income US Military Veterans. *Public Health Rep*. Published online 2021. doi:10.1177/0033354920974662

4. Singh DR, Sunuwar DR, Shah SK, Sah LK, Karki K, Sah RK. Food insecurity during COVID-19 pandemic: A genuine concern for people from disadvantaged community and low-income families in Province 2 of Nepal. *PLoS One*. 2021;16(7):e0254954. doi:10.1371/journal.pone.0254954

5. Martin NM, Rex S, Barnett D, et al. Digital tools for emergency food assistance: a scoping review. *Disaster medicine and public health preparedness*. Published online September 2021.

6. Mahajan A, Czerniak C, Lamichhane J, et al. Advances in real-time social listening for an adaptive public health response: WHO's EARS platform. *Eur J Public Health*. 2021;31(Suppl 3):ckab164.501. doi:10.1093/eurpub/ckab164.501

7. 4th Virtual WHO Infodemic Management Conference: Advances in Social Listening for Public Health. Accessed November 11, 2021. https://www.who.int/news-room/events/detail/2021/05/04/default-calendar/4th-virtual-who-infodemic-management-conference-advances-in-social-listening-for-public-health

8. The Keys to Social Listening | NYU School of Global Public Health. Accessed November 11, 2021. https://publichealth.nyu.edu/events-news/news/2020/12/17/keys-social-listening

9. Purnat TD, Vacca P, Czerniak C, et al. Infodemic Signal Detection During the COVID-19 Pandemic: Development of a Methodology for Identifying Potential Information Voids in Online Conversations. *JMIR Infodemiology*. 2021;1(1):e30971. doi:10.2196/30971

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

10. Purnat TD, Vacca P, Burzo S, et al. WHO Digital Intelligence Analysis for Tracking Narratives and Information Voids in the COVID-19 Infodemic. *Stud Health Technol Inform*. 2021;281:989-993. doi:10.3233/SHTI210326

11. Social Listening and Monitoring Tools. :9.

12. Dunn EE, Gainforth HL, Robertson-Wilson JE. Behavior change techniques in mobile applications for sedentary behavior. *Digit Health*. 2018;4:2055207618785798. doi:10.1177/2055207618785798

13. Sepah SC, Jiang L, Peters AL. Translating the Diabetes Prevention Program into an Online Social Network: Validation against CDC Standards. *Diabetes Educ*. 2014;40(4):435-443. doi:10.1177/0145721714531339

14. Castensøe-Seidenfaden P, Husted GR, Jensen AK, et al. Testing a Smartphone App (Young with Diabetes) to Improve Self-Management of Diabetes Over 12 Months: Randomized Controlled Trial. *JMIR Mhealth Uhealth*. 2018;6(6):e141. doi:10.2196/mhealth.9487

15. Willis EA, Szabo-Reed AN, Ptomey LT, et al. Distance learning strategies for weight management utilizing online social networks versus group phone conference call. *Obes Sci Pract*. 2017;3(2):134-142. doi:10.1002/osp4.96

16. Dholakia U, Tsabar G. A Startup's Experience with Running a Groupon Promotion. In: ; 2011. doi:10.2139/SSRN.1828003

17. Topic Modeling: An Introduction. MonkeyLearn Blog. Published September 26, 2019. Accessed October 27, 2021. https://monkeylearn.com/blog/introduction-to-topic-modeling/

18. Abdalla G, Ozyurt F. Sentiment Analysis of Fast Food Companies With Deep Learning Models. *COMPUTER JOURNAL*. 2021;64(3):383-390. doi:10.1093/comjnl/bxaa131

19. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. *PLOS ONE*. 2020;15(9):e0239441. doi:10.1371/journal.pone.0239441

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

20. Martin NM, Barnett DJ, Poirier L, Sundermeir SM, Reznar MM, Gittelsohn J. Moving Food Assistance into the Digital Age: A Scoping Review. *International Journal of Environmental Research and Public Health*. 2022;19(3):1328. doi:10.3390/ijerph19031328

21. Martin NM, Barnett DJ, Sedoc J, et al. Harnessing Artificial Intelligence to Improve Food Security: A Scoping Review of Machine Learning Tools. *In Review*. Published online March 22, 2022.

22. Barnett DJ, Rosenblum AJ, Strauss-Riggs K, Kirsch TD. Readying for a Post-COVID-19 World: The Case for Concurrent Pandemic Disaster Response and Recovery Efforts in Public Health. *J Public Health Manag Pract*. 2020;26(4):310-313. doi:10.1097/PHH.0000000000001199

23. Welcome to Python.org. Accessed September 27, 2021. https://www.python.org/

24. Edward A. An Extensive Guide to collecting tweets from Twitter API v2 for academic research using Python 3. Medium. Published June 17, 2021. Accessed October 20, 2021. https://towardsdatascience.com/an-extensive-guide-to-collecting-tweets-from-twitter-api-v2-for-academic-research-using-python-3-518fcb71df2a

25. Google Colaboratory. Accessed September 27, 2021. https://colab.research.google.com/notebooks/intro.ipynb?utm_source=scs-index

26. Martin N. *Github Repository for LDA Topic Modeling Code*.; 2021. Accessed September 27, 2021. https://github.com/drscicomm/food-security/blob/a70b06f68c4e36f6a4a49bab75f765061a226f1e/LDA_viz_C_20210924.ipynb

27. Twitter API for Academic Research | Products. Accessed February 11, 2022. https://developer.twitter.com/en/products/twitter-api/academic-research

28. re — Regular expression operations — Python 3.9.7 documentation. Accessed October 4, 2021. https://docs.python.org/3/library/re.html

29. NumPy. Accessed October 4, 2021. https://numpy.org/

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

30. pandas - Python Data Analysis Library. Accessed October 4, 2021. https://pandas.pydata.org/

31. NLP Gensim Tutorial - Complete Guide For Beginners - GeeksforGeeks. Accessed October 4, 2021. https://www.geeksforgeeks.org/nlp-gensim-tutorial-complete-guide-for-beginners/

32. Python Word Clouds: How to Create a Word Cloud. DataCamp Community. Published November 8, 2019. Accessed October 4, 2021. https://www.datacamp.com/community/tutorials/wordcloud-python

33. spaCy · Industrial-strength Natural Language Processing in Python. Accessed October 4, 2021. https://spacy.io/

34. Matplotlib: Python plotting — Matplotlib 3.4.3 documentation. Accessed October 4, 2021. https://matplotlib.org/

35. Pillow. Accessed October 4, 2021. https://pillow.readthedocs.io/en/stable/index.html

36. Topic modeling visualization - How to present results of LDA model? | ML+. Machine Learning Plus. Published December 4, 2018. Accessed October 27, 2021. https://www.machinelearningplus.com/nlp/topic-modeling-visualization-how-to-present-results-lda-models/

37. Brunila M, Zhao R, Mircea A, Lumley S, Sieber R. Bridging the gap between supervised classification and unsupervised topic modelling for social-media assisted crisis management. In: Association for Computational Linguistics (ACL); 2021:33-49.

38. Sharma H. Topic Model Visualization using pyLDAvis. Medium. Published June 5, 2021. Accessed October 4, 2021. https://towardsdatascience.com/topic-model-visualization-using-pyldavis-fecd7c18fbf6

39. Kapadia S. Evaluate Topic Models: Latent Dirichlet Allocation (LDA). Medium. Published December 29, 2020. Accessed September 27, 2021. https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

40. Calendar heatmaps from Pandas time series data — calplot 0.1.7.3 documentation. Accessed January 24, 2022. https://calplot.readthedocs.io/en/latest/

41. Lukyamuzi A, Ngubiri J, Okori W. Tracking food insecurity from tweets using data mining techniques. In: *Proc Int Conf Software Eng*. IEEE Computer Society; 2018:27-34. doi:10.1145/3195528.3195531

42. Basu M, Ghosh K, Das S, Dey R, Bandyopadhyay S, Ghosh S. Identifying Post-Disaster Resource Needs and Availabilities from Microblogs. In: *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*. ACM; 2017:427-430. doi:10.1145/3110025.3110036

43. Madichetty S, M S. A stacked convolutional neural network for detecting the resource tweets during a disaster. *Multimed Tools Appl*. Published online September 25, 2020:1-23. doi:10.1007/s11042-020-09873-8

44. Behl S, Rao A, Aggarwal S, Chadha S, Pannu H. Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. *INTERNATIONAL JOURNAL OF DISASTER RISK REDUCTION*. 2021;55. doi:10.1016/j.ijdrr.2021.102101

45. Aipe A, Mukuntha NS, Ekbal A, Kurohashi S. Deep learning approach towards multi-label classification of crisis related tweets. In: Vol 2018-May. Information Systems for Crisis Response and Management, ISCRAM; 2018:705-717.

46. Basu M, Shandilya A, Khosla P, Ghosh K, Ghosh S. Extracting Resource Needs and Availabilities From Microblogs for Aiding Post-Disaster Relief Operations. *IEEE Transactions on Computational Social Systems*. 2019;6(3):604-618. doi:10.1109/TCSS.2019.2914179

47. Madichetty S, Muthukumarasamy S, Jayadev P. Multi-modal classification of Twitter data during disasters for humanitarian response. *J Ambient Intell Human Comput*. Published online January 3, 2021. doi:10.1007/s12652-020-02791-5

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

48. Madichetty S, Sridevi M. Classifying informative and non-informative tweets from the twitter by adapting image features during disaster. *MULTIMEDIA TOOLS AND APPLICATIONS*. 2020;79(39-40):28901-28923. doi:10.1007/s11042-020-09343-1

49. Basu M, Bit SD, Ghosh S. Utilizing microblogs for optimized real-time resource allocation in post-disaster scenarios. *Soc Netw Anal Min*. 2021;12(1):15. doi:10.1007/s13278-021-00841-0

50. Temnikova I. EMTerms 1.0: A terminological resource for crisis tweets. *ISCRAM 2015 Conference Proceedings - 12th International Conference on Information Systems for Crisis Response and Management*. 2015;2015-January:147-157.

51. Olteanu A. CrisisLex: A lexicon for collecting and filtering Microblogged communications in crises. *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*. Published online 2014:376-385.

52. Acerbo FS, Rossi C. Filtering informative tweets during emergencies: A machine learning approach. In: Association for Computing Machinery; 2017:1-6. doi:10.1145/3152896.3152897

53. Chy AN, Siddiqua UA, Aono M. A Neural Network Model to Identify the Crisis-related Actionable Informative Tweets for Disaster Management. In: Institute of Electrical and Electronics Engineers Inc.; 2020:390-393. doi:10.1109/TENSYMP50017.2020.9230679

54. Gata W, Amsury F, Wardhani NK, Sugiyarto I, Sulistyowati DN, Saputra I. Informative Tweet Classification of the Earthquake Disaster Situation in Indonesia. In: Institute of Electrical and Electronics Engineers Inc.; 2019. doi:10.1109/ICCED46541.2019.9161135

55. Kumar A, Singh J, Dwivedi Y, Rana N. A deep multi-modal neural network for informative Twitter content classification during emergencies. *ANNALS OF OPERATIONS RESEARCH*. doi:10.1007/s10479-020-03514-x

56. Madichetty S, Sridevi M. Detecting Informative Tweets during Disaster using Deep Neural Networks. In: Institute of Electrical and Electronics Engineers Inc.; 2019:709-713. doi:10.1109/COMSNETS.2019.8711095

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

57. Neppalli VK, Caragea C, Caragea D. Deep neural networks versus naive bayes classifiers for identifying informative tweets during disasters. In: Vol 2018-May. Information Systems for Crisis Response and Management, ISCRAM; 2018:677-686.

58. Sreenivasulu M, Sridevi M. Mining informative words from the tweets for detecting the resources during disaster. In: Vol 10682 LNAI. Springer Verlag; 2017:348-358. doi:10.1007/978-3-319-71928-3_33

59. Roy S, Mishra S, Matam R. Classification and Summarization for Informative Tweets. In: Institute of Electrical and Electronics Engineers Inc.; 2020. doi:10.1109/SCEECS48394.2020.128

60. Whitaker RC, Phillips SM, Orzol SM. Food Insecurity and the Risks of Depression and Anxiety in Mothers and Behavior Problems in their Preschool-Aged Children. *Pediatrics*. 2006;118(3):e859-e868. doi:10.1542/peds.2006-0239

61. Laraia BA, Siega-Riz AM, Gundersen C, Dole N. Psychosocial Factors and Socioeconomic Indicators Are Associated with Household Food Insecurity among Pregnant Women. *The Journal of Nutrition*. 2006;136(1):177-182. doi:10.1093/jn/136.1.177

62. Heflin CM, Siefert K, Williams DR. Food insufficiency and women's mental health: Findings from a 3-year panel of welfare recipients. *Social Science & Medicine*. 2005;61(9):1971-1982. doi:10.1016/j.socscimed.2005.04.014

63. Leung CW, Epel ES, Willett WC, Rimm EB, Laraia BA. Household Food Insecurity Is Positively Associated with Depression among Low-Income Supplemental Nutrition Assistance Program Participants and Income-Eligible Nonparticipants. *The Journal of Nutrition*. 2015;145(3):622-627. doi:10.3945/jn.114.199414

64. Erokhin V, Gao T. Impacts of COVID-19 on Trade and Economic Aspects of Food Security: Evidence from 45 Developing Countries. *International Journal of Environmental Research and Public Health*. 2020;17(16):5775. doi:10.3390/ijerph17165775

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

65. Carroll N, Sadowski A, Laila A, et al. The Impact of COVID-19 on Health Behavior, Stress, Financial and Food Security among Middle to High Income Canadian Families with Young Children. *Nutrients*. 2020;12(8):2352. doi:10.3390/nu12082352

66. Fang D, Thomsen MR, Nayga RM. The association between food insecurity and mental health during the COVID-19 pandemic. *BMC Public Health*. 2021;21(1):1-8. doi:10.1186/s12889-021-10631-0

67. Aguiar A, Pinto M, Duarte R. The bad, the ugly and the monster behind the mirror - Food insecurity, mental health and socio-economic determinants. *Journal of Psychosomatic Research*. 2022;154:110727. doi:10.1016/j.jpsychores.2022.110727

68. Immediate impact of stay-at-home orders to control COVID-19 transmission on socioeconomic conditions, food insecurity, mental health, and intimate partner violence in Bangladeshi women and their families: an interrupted time series - ClinicalKey. Accessed March 2, 2022. https://www-clinicalkey-com.proxy1.library.jhu.edu/#!/content/playContent/1-s2.0-S2214109X20303661?returnurl=https:%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2214109X20303661%3Fshowall%3Dtrue&referrer=

69. Mozaffarian D, Fleischhacker S, Andrés JR. Prioritizing Nutrition Security in the US. *JAMA*. 2021;325(16):1605-1606. doi:10.1001/jama.2021.1915

70. Perspective | Food security isn't enough. Anti-hunger experts say the focus should be on nutrition security. *Washington Post*. https://www.washingtonpost.com/lifestyle/wellness/food-security-nutrition-usda-school-andres/2021/06/10/8375e7c6-c951-11eb-81b1-34796c7393af_story.html. Accessed March 2, 2022.

71. Ingram J. Nutrition security is more than food security. *Nat Food*. 2020;1(1):2-2. doi:10.1038/s43016-019-0002-4

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press

72. Bose A, Tandon S, Basu M, Ghosh S. Social Media for Post-Disaster Relief: Mapping Needs and Availabilities to UNOCHA Resource Classes. In: Association for Computing Machinery; 2022:294-298. doi:10.1145/3491003.3493236

73. Kersten J, Bongard J, Klan F. Combining supervised and unsupervised learning to detect and semantically aggregate crisis-related twitter content. In: Vol 2021-May. Information Systems for Crisis Response and Management, ISCRAM; 2021:744-754.

https://doi.org/10.1017/dmp.2022.207 Published online by Cambridge University Press