

Association for Information Systems

## AIS Electronic Library (AISeL)

---

Proceedings of the 2021 Pre-ICIS SIGDSA  
Symposium

Special Interest Group on Decision Support and  
Analytics (SIGDSA)

---

Winter 12-12-2021

### **What Facebook Messages Told Us About How We Handled Disaster Management during the COVID-19 Pandemic?**

Wenting Jiang

Ashish Gupta

Kenan Xiao

Xiao Qin

Follow this and additional works at: <https://aisel.aisnet.org/sigdsa2021>

---

This material is brought to you by the Special Interest Group on Decision Support and Analytics (SIGDSA) at AIS Electronic Library (AISeL). It has been accepted for inclusion in Proceedings of the 2021 Pre-ICIS SIGDSA Symposium by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# **What Facebook Messages Told Us About How We Handled Disaster Management during the COVID-19 Pandemic?**

*Research-in-Progress*

**Wenting Jiang**  
Auburn University  
wzj0027@auburn.edu

**Kenan Xiao**  
Auburn University  
kzx0010@auburn.edu

**Ashish Gupta**  
Auburn University  
azg0074@auburn.edu

**Xiao Qin**  
Auburn University  
xqin@auburn.edu

## **Abstract**

As COVID-19 continues, social media platforms such as Facebook have become an increasingly important tool for communication and information sharing for public and government agencies. The generic disaster management cycle (mitigation, preparedness, response, and recovery) provides systematic guidance to the public and government agencies to respond to the crisis and suggest appropriate measures for different disaster stages. This study examines various trending topics and themes during the COVID-19 outbreak. Using this generic disaster management cycle as our guiding framework, we examine news topics' evolution during the COVID-19 pandemic on Facebook during two of the four phases. Latent Dirichlet Allocation (LDA) is used for topic modeling to identify topics and themes.

## **Keywords**

Disaster Management, COVID-19, Latent Dirichlet Allocation, Text Analytics, Natural Language Processing, Thematic Deduction

## **Introduction**

We have recently seen exponential growth in COVID-19 cases, with approximately 254 million positive cases and 5.11 million mortalities globally reported in November 2021. This rise in COVID-19 cases is also signified by increased social media usage. Due to social distancing and stay-at-home orders during the COVID-19 pandemic, people's reliance on social media to acquire information as the COVID-19 outbreak has also increased continues (Rovetta and Bhagavathula 2020; Tsao et al. 2021). Recent reports on news consumption through social media revealed similar trends, with Facebook having the most newsreaders across all other social media-based news platforms (Shearer and Mitchell 2021). This has also led to other problems such as infodemic (Gupta et al. 2021)

Social media is more than a tool for getting daily news. More importantly, it has been harnessed as crucial information dissemination and emergency management tool that can assist in reducing disaster risk, crisis response, disaster preparations, distributing real-time warnings and alerts, providing emergency services and assistance, and monitoring evolving situations, etc. (Alexander 2014; Xiao et al. 2015). For example, social media was employed to deliver emergency services during floods in Queensland and Victoria (Bird et al. 2012), and respond to the crisis, and communicate with the public 2015 Nepal Earthquake (Subba and Bui 2017) and when Hurricane Harvey hit Texas and Louisiana in 2017 (Liu et al. 2018).

As the pandemic progresses through various stages, the nature and content of the information shared across social media platforms also evolve. Furthermore, it will be insightful to investigate the elements of connectedness in these discussions as network-based characteristics embedded in such media postings may exhibit dynamic behavior. For example, as topics get added to the media posting, some old topics might fade out as COVID-19 shifts from one phase to another. The connections among certain topics might become less or stronger with time. Capturing the dynamics behind the news topics network will provide increased awareness about the frequency of connectedness among various news topics and may serve to guide the government in identifying topics that may help mitigate disaster risk or improve crisis response.

The disaster management cycle empowers the public and government agencies to respond to the crisis in an informed way by taking appropriate actions before, during, and after disasters (Adams 2002). The disaster management cycle consists of four phases: mitigation, preparedness, response, and recovery (Khan et al., 2008). During mitigation, activities that minimize the effects of the disaster are conducted, preparedness involves response planning, response phase includes undertaking the efforts to minimize the impact, and recovery refers to initiatives that attempt to return the community to normal. Taking the right actions and making the correction decisions at each disaster management phase can better prepare the public to respond to a disaster (Adams, 2002). Social media is believed to be helpful in mitigation, preparedness, response, and recovery phases to enhance crisis communications and reduce vulnerability in the community (Chan 2014; Joseph et al. 2018).

A large body of researchers examined various trending topics and themes during the COVID-19 outbreak and pointed to the plethora of information related to COVID-19 in areas such as disease statistics, the nature of the disease, preventions, protection methods, treatments, and diagnostic procedures that have propagated rapidly on various social media platforms across different countries (Abd-Alrazaq et al. 2020; Bastani and Bahrami 2020; Chandrasekaran et al. 2020; Hung et al. 2020; Li et al. 2020a; Li et al. 2020b; Liu et al. 2020; Medford et al. 2020; Singh et al. 2020; Thelwall and Thelwall 2020; Xue et al. 2020; Zhao and Xu 2020). While these studies have identified different topics and themes during the pandemic, several research opportunities exist. For example, (1) majority of the studies have focused on the pandemic's early stages, without paying attention to the long-time development timeline of COVID-19, (2) the dynamics behind news topics have not been studied, which implies that studies that treat news topics as continuously evolving and interconnected, rather than static and isolated from each other into their research scope, are lacking (3) these studies have primarily concentrated on Twitter and Sina Weibo (the Chinese equivalent of Twitter), however, the topics and themes regarding COVID-19 information that are published on Facebook have not been investigated yet, and (4) there is a paucity of research focus on how the public and government agencies used social media to conduct disaster management during the COVID-19 outbreak.

To fill these research gaps, we propose to study the evolution of news topics during the COVID-19 pandemic on Facebook across progressively different timeframes based on four different phases of the pandemic cycle. These four phases are based on the 7-day average new reported cases of COVID-19 originating from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Specifically, the following

research questions are studied. (1) What COVID-19 topics were discussed in Facebook during different phases from January 2020 to September 2021? (2) How do these topics change from one phase to another (e.g., the relationship between these topics, most influential topics, the lifetime of topics, etc.)? (3) Whether Facebook discussion reflect each disaster management phase (mitigation, preparedness, response, recovery)?

This study makes two specific contributions. First, we use the disaster management cycle to understand and explain Facebook messages published from January 1, 2020, to September 10, 2021. Specifically, we examined various themes that emerge from social media postings during different phases that confirm the disaster management cycle. Second, we provide guidance for the government to identify topics that can be used more effectively to reduce disaster risk and respond to crises.

## **Method**

Natural Language Processing (NLP) techniques such as Latent Dirichlet Allocation (LDA) is used to compare and contrast the COVID-19 related Facebook messages that emerged in different phases based on 7-day average new reported cases originating from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. The text analytics process used in this study, consisting of data acquisition, data preprocessing, and topic modeling.

### ***Data Acquisition***

Facebook messages were collected using CrowdTangle (CrowdTangleTeam 2021) via keywords: coronavirus, covid, covid19, covid 19, covid-19, wuhancoronavirus, wuhanvirus, 2019nCOV, SARS-CoV-2 during the period from January 1, 2020, to September 10, 2021. CrowdTangle is a Facebook-owned tool to track public content across social media.

### ***Study Period***

In this study, we slit the timeframe of COVID-19 developments in the United States into four phases based on the COVID-19 7-day averages of newly reported cases reported by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Phase One starts on January 1, 2020, and ends on September 11, 2020. The 7-day average was relatively low during this period, and we retrieved 7,185,722 Facebook messages. Phase Two starts on September 12, 2020, and ends on January 9, 2021. The 7-day average is climbing up with a steep slope during this period, and we retrieved 3,950,761 Facebook messages. Phase Three starts on January 10, 2021, and ends on July 5, 2021. The 7-day average went down during this period, and we retrieved 4,251,392 Facebook messages. Phase Four starts on July 6, 2021, and ends on September 10, 2021. The 7-day average started climbing up again during this period, and we retrieved 546,091 Facebook messages.

### ***Data Preprocessing***

To get data ready for topic modeling, we used the Spacy package, an open-source Python-based library for NLP (Honnibal and Montani 2017). After removing duplicated Facebook messages, we performed various preprocessing activities such as: tokenizing each word, converting each word into lower cases, expanding contraction words (e.g., you've, I'd, it's, etc.), removing hashtags content, removing URL, eliminating stop words (e.g., the, a, is are, etc.) that carry very little useful information, removing punctuations, replacing words with their synonyms (for example, new coronavirus was replaced with novel coronavirus), and lemmatizing each token.

### ***Topic Modeling***

In order to extract topics that had been discussed in each phase of the COVID-19 pandemic, we first use the Bag of Words (BOW) approach to convert each bigram into vectors. For example, in the sentence "the sun will rise tomorrow", the bigrams are: "the sun ", "sun will ", "will rise ", "rise tomorrow ". We used these bigrams as our unit of analysis to capture the context within the documents (Wallach 2006). Second, we harnessed Latent Dirichlet Allocation (LDA), a popular unsupervised generative statistical topic modeling

technique used for discovering latent topics from a collection of the corpus (Blei et al. 2003, Jelodar et al. 2019).

## Results

### ***Phase One (January 1, 2020 – September 11, 2020)***

Four topics were extracted after running LDA for Phase One. The percentage of tokens for four topics is very close, ranging from 24% to 27%, with Topic 1 having the highest percentage and Topic 4 having the lowest percentage. Topic 1 consists of four themes in terms of the coronavirus pandemic: business impact, children safety, support and relief programs, and how information is being shared on social media. Topic 2 concentrates on four themes: epicenters, combating the virus, effects/impact of the virus, and government agencies/stakeholders. Topic 3 focuses on statistical trends and awareness, mitigation and prevention strategies, and healthcare agencies. Topic 4 covers four themes: care and support, awareness and patient literacy, preventions, and COVID-19 challenges.

### ***Phase Two (September 12, 2020 – January 9, 2021)***

Five topics were extracted after running LDA for Phase One. The percentage of tokens for four topics is very close, ranging from 15.9% to 24.6%, with Topic 1 having the highest percentage and Topic 5 having the lowest percentage. Topic 1 covers three themes: holiday season during COVID-19, challenges during COVID-19, and social media/information sharing. Topic 2 concentrates on four themes: government agencies, epicenters, COVID-19 second wave, and the impact of the virus. Topic 3 focuses on two themes: the spread of virus and mitigation and prevention strategies. Topic 4 has three themes: preventions of COVID-19, statistical trends and awareness, and healthcare agencies. Topic 5 consists of three themes: business impact, stakeholders, and support and relief programs.

## Conclusion

The topics related to mitigation and preparedness were more prevalent in the early phases of COVID-19, and response-related and recovery-related discussions were more common in the later phases.

## References

- Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., and Shah, Z. 2020. "Top Concerns of Tweeters During the Covid-19 Pandemic: Inveillance Study," *Journal of medical Internet research* (22:4), p. e19016.
- Adams, J. 2002. *Environmental Health in Emergencies and Disasters: A Practical Guide*. World health organization.
- Alexander, D. E. 2014. "Social Media in Disaster Risk Reduction and Crisis Management," *Science and engineering ethics* (20:3), pp. 717-733.
- Bastani, P., and Bahrami, M. A. 2020. "Covid-19 Related Misinformation on Social Media: A Qualitative Study from Iran," *Journal of medical Internet research*.
- Bird, D., Ling, M., and Haynes, K. 2012. "Flooding Facebook-the Use of Social Media During the Queensland and Victorian Floods," *Australian Journal of Emergency Management, The* (27:1), pp. 27-33.
- Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. "Latent Dirichlet Allocation," *the Journal of machine Learning research* (3), pp. 993-1022.
- Chan, J. C. 2014. "The Role of Social Media in Crisis Preparedness, Response and Recovery By,").
- Chandrasekaran, R., Mehta, V., Valkunde, T., and Moustakas, E. 2020. "Topics, Trends, and Sentiments of Tweets About the Covid-19 Pandemic: Temporal Inveillance Study," *Journal of medical Internet research* (22:10), p. e22624.
- CrowdTangleTeam. 2021. "Crowdtangle. Facebook, Menlo Park, California, United States. List Id: <https://Apps.Crowdtangle.Com/Nlpnetwork/Lists/1558852>."
- Gupta, A. Li, A., Farnoush, A., Jiang, W. 2021. Understanding Patterns of COVID Infodemic: A Systematic and Pragmatic Approach to Curb Fake News, *Journal of Business Research*,
- Honnibal, M., and Montani, I. 2017. "Spacy 2: Natural Language Understanding with Bloom Embeddings, Convolutional Neural Networks and Incremental Parsing," *To appear* (7:1), pp. 411-420.

- Hung, M., Lauren, E., Hon, E. S., Birmingham, W. C., Xu, J., Su, S., Hon, S. D., Park, J., Dang, P., and Lipsky, M. S. 2020. "Social Network Analysis of Covid-19 Sentiments: Application of Artificial Intelligence," *Journal of medical Internet research* (22:8), p. e22590.
- Jagarlamudi, J., Daumé III, H., and Udupa, R. 2012. "Incorporating Lexical Priors into Topic Models," *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 204-213.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., and Zhao, L. 2019. "Latent Dirichlet Allocation (Lda) and Topic Modeling: Models, Applications, a Survey," *Multimedia Tools and Applications* (78:11), pp. 15169-15211.
- Joseph, J. K., Dev, K. A., Pradeepkumar, A., and Mohan, M. 2018. "Big Data Analytics and Social Media in Disaster Management," in *Integrating Disaster Science and Management*. Elsevier, pp. 287-294.
- Khan, H., Vasilescu, L. G., and Khan, A. 2008. "Disaster Management Cycle-a Theoretical Approach," *Journal of Management and Marketing* (6:1), pp. 43-50.
- Li, J., Xu, Q., Cuomo, R., Purushothaman, V., and Mackey, T. 2020a. "Data Mining and Content Analysis of the Chinese Social Media Platform Weibo During the Early Covid-19 Outbreak: Retrospective Observational Infeveillance Study," *JMIR Public Health and Surveillance* (6:2), p. e18700.
- Li, L., Zhang, Q., Wang, X., Zhang, J., Wang, T., Gao, T.-L., Duan, W., Tsoi, K. K.-f., and Wang, F.-Y. 2020b. "Characterizing the Propagation of Situational Information in Social Media During Covid-19 Epidemic: A Case Study on Weibo," *IEEE Transactions on Computational Social Systems* (7:2), pp. 556-562.
- Liu, Q., Zheng, Z., Zheng, J., Chen, Q., Liu, G., Chen, S., Chu, B., Zhu, H., Akinwunmi, B., and Huang, J. 2020. "Health Communication through News Media During the Early Stage of the Covid-19 Outbreak in China: Digital Topic Modeling Approach," *Journal of Medical Internet Research* (22:4), p. e19118.
- Liu, W., Lai, C.-H., and Xu, W. W. 2018. "Tweeting About Emergency: A Semantic Network Analysis of Government Organizations' Social Media Messaging During Hurricane Harvey," *Public relations review* (44:5), pp. 807-819.
- Medford, R. J., Saleh, S. N., Sumarsono, A., Perl, T. M., and Lehmann, C. U. 2020. "An" Infodemic": Leveraging High-Volume Twitter Data to Understand Public Sentiment for the Covid-19 Outbreak," *medRxiv*).
- Rovetta, A., and Bhagavathula, A. S. 2020. "Covid-19-Related Web Search Behaviors and Infodemic Attitudes in Italy: Infodemiological Study," *JMIR public health and surveillance* (6:2), p. e19374.
- Shearer, E., and Mitchell, A. 2021. "News Use across Social Media Platforms in 2020," Pew Research Center.
- Singh, L., Bansal, S., Bode, L., Budak, C., Chi, G., Kawintiranon, K., Padden, C., Vanarsdall, R., Vraga, E., and Wang, Y. 2020. "A First Look at Covid-19 Information and Misinformation Sharing on Twitter," *arXiv preprint arXiv:2003.13907*).
- Subba, R., and Bui, T. 2017. "Online Convergence Behavior, Social Media Communications and Crisis Response: An Empirical Study of the 2015 Nepal Earthquake Police Twitter Project," *Proceedings of the 50th hawaii international conference on system sciences*.
- Thelwall, M., and Thelwall, S. 2020. "A Thematic Analysis of Highly Retweeted Early Covid-19 Tweets: Consensus, Information, Dissent, and Lockdown Life,").
- Tsao, S.-F., Chen, H., Tisseverasinghe, T., Yang, Y., Li, L., and Butt, Z. A. 2021. "What Social Media Told Us in the Time of Covid-19: A Scoping Review," *The Lancet Digital Health*).
- Wallach, H. M. 2006. "Topic Modeling: Beyond Bag-of-Words," *Proceedings of the 23rd international conference on Machine learning*, pp. 977-984.
- Xiao, Y., Huang, Q., and Wu, K. 2015. "Understanding Social Media Data for Disaster Management," *Natural hazards* (79:3), pp. 1663-1679.
- Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., and Zhu, T. 2020. "Twitter Discussions and Emotions About the Covid-19 Pandemic: Machine Learning Approach," *Journal of medical Internet research* (22:11), p. e20550.
- Zhao, Y., and Xu, H. 2020. "Chinese Public Attention to Covid-19 Epidemic: Based on Social Media," *medRxiv*).