A diachronic corpus-based analysis of COVID-19 topic modeling, sentiment, metaphor frames in media discourse: a case study of Hong Kong

232703M Yolanda Guan
Supervisor: Dr. Thomas Chan
2024 10 30

OUTLINE

1. Introduction
2. Literature review
   - Topic modeling
   - Sentiment analysis
   - Metaphor analysis
3. Research Questions
4. Methodology
   - Toipc modelng
   - Sentiment analysis
   - Metaphor analysis
5. Results and discussion
   - Toipc modelng
   - Sentiment analysis
   - Metaphor analysis
6. References

Introduction

Introduction

In the past five years, the COVID-19 pandemic has had a huge and lasting impact on the public's health, economy, and society.
Many experts and scholars in various fields are studying the impact of COVID-19 on various aspects of society.
Metaphor analysis is crucial in various discourses, as it is not only a rhetorical device but also a tool for expressing complex ideas using relevant terminology.
In the context of online news discourse, metaphors can evoke public emotions, shape citizens' perceptions, and guide the interpretation of social phenomena (Lakoff&Johnson, 2008; Pragglejaz Group, 2007; Steen et al., 2010; Ahrens & Zeng, 2022).
Sentiment analysis is useful and effective in news texts. This study aims to investigate the role of topic modeling, sentiment and metaphors in addressing the COVID-19 issues in the news discourse in

Introduction

Understanding the functions of topic modeling, sentiment and metaphors in news discourse can establish COVID-19 perception.

focuses on the relationship between sentiment analysis and words associated with metaphors.

It contributes to the field with practical implications by providing a detailed topic modeling, sentiment and metaphor analysis of COVID-19 discourse in the context of Hong Kong.

Literature Review

Literature review

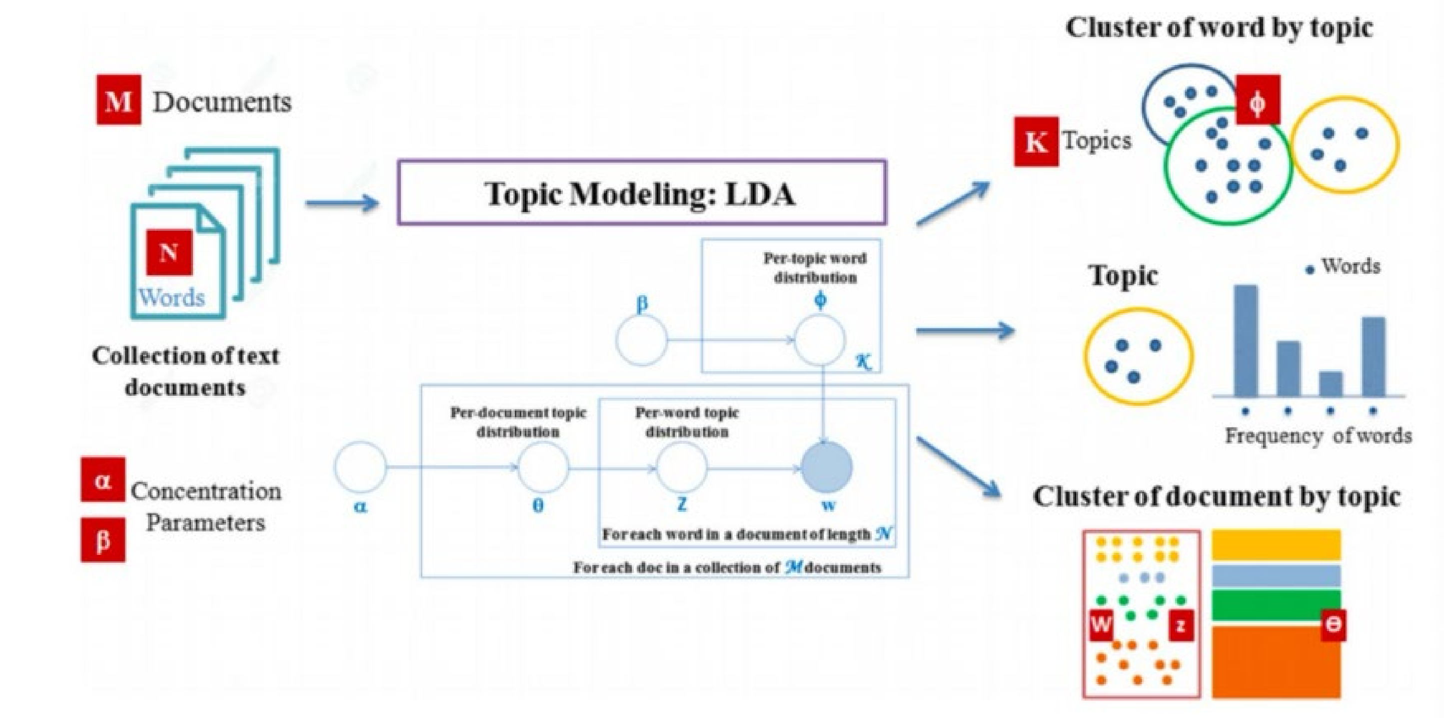
In today's era of information explosion, textual data is ubiquitous in our digital lives, ranging from comments on social media to product reviews on e-commerce platforms, and extending to articles in academic journals and reports from news media (Jelodar et al., 2019).

The text can be book chapters, emails, news, journal articles, and any type of unstructured text (Jelodar et al., 2019).

Latent Dirichlet Allocation (LDA) is an unsupervised machine learning technique widely used in the field of natural language processing, aimed at uncovering the hidden semantic structures within text documents, known as "topics."

In LDA, a "topic" is understood as a pattern that associates words in the vocabulary with their frequency of occurrence in documents.

Through topic modeling, researchers can identify key concepts in large collections of documents and classify and annotate documents accordingly, providing valuable insights into large corpora, including the content of


M Documents
N
Words
Collection of text documents
α Concentration Parameters
β
Topic Modeling: LDA
Per-topic word distribution
β
ϕ
K
Per-document topic distribution
Per-word topic distribution
α
θ
Z
w
For each word in a document of length N
For each doc in a collection of M documents
Cluster of word by topic
K Topics
ϕ
Topic
Words
Frequency of words
Cluster of document by topic
W
z
θ

Literature review

Sentiment analysis is a sophisticated process that includes the examination of positive and negative emotions. It also represents an intellectual endeavor to extract feelings and sentiments from textual data (Bhardwaj et al., 2024; Saberi & Saad, 2017).

This intricate procedure employs natural language processing, text analysis, and statistical methods to assess human emotions and categorize them into various classifications.

Sentiment analysis systems are increasingly being utilized across numerous economic and social domains, as opinions and perspectives are fundamental to nearly all human activities.

Furthermore, sentiment analysis serves as a crucial component for comprehensive investigations in several real-world applications, such as forecasting stock market trends (Khan et al., 2022) and monitoring mental health conditions (Benrouba & Boudour, 2023).

These insights offer a panoramic view of the populace across diverse social and economic sectors, enabling relevant organizations to implement reforms accordingly.

Research gap

Reviewing the past research achievements, we find that most studies have focused on analyzing the characteristics of specific stages of COVID-19 pandemic news, while comprehensive studies covering the entire pandemic period are relatively few.

In the fields of topic modeling, sentiment analysis, and metaphor analysis, in-depth exploration of the corpus throughout the entire COVID-19 period will help us obtain a comprehensive analysis of COVID-19 news coverage.

Furthermore, few studies have explored the interrelationships between these three analytical methods. Therefore, this study conducts a detailed discussion of these three methods, hoping to contribute to metaphor research in the field of journalism.

Research Questions

Research Questions

1. How do COVID-19 topics in Hong Kong news discourse change over time?

2. How do news writers' attitudes towards COVID-19 pandemic and measures change over time?

3. How does use of COVID-19 metaphors in Hong Kong news discourse change over time?

Methodology

Corpus creation

This research examines news articles from the online Hong Kong Free Press website at https://hongkongfp.com/comment-analysis/.

A total of 3,352 articles related to COVID-19 were collected using a web crawler, implemented in Python with the Requests library.

A web crawler, also known as a web spider, web robot, or web chaser, is a program or script designed to automatically gather information from the internet according to specific rules.

Data cleaning and pre-processing

To begin with, the marked data is segmented, as the dataset utilized in this study consists of English textual information.
English text employs spaces as delimiters, making it relatively straightforward to distinguish words; thus, explicit text segmentation is not performed here.
Certain terms within the text may lack substantive meaning, such as onomatopoeia and articles.
Commonly encountered meaningless words in English texts—referred to as stop words—include "a," "an," and "and."
The scope of stop words varies depending on the specific domain of the text being processed; therefore, it is often necessary to customize stop word sets to enhance processing quality.

LDA topic modeling

Firstly, a corpus containing multiple news articles was collected and organized. To ensure the validity of the data, strict preprocessing steps were implemented. This stage involved removing meaningless characters such as punctuation marks and numbers, as well as performing word segmentation and eliminating stop words.

For the selection of stop words, both common English stop word libraries and characteristics specific to news texts were utilized. Additionally, certain vocabulary (e.g., "COVID-19," "Hong Kong," etc.) was customized to further optimize the model's performance. The resulting preprocessed text is more concise, helping to reduce noise interference in topic analysis.

Next, a Bag-of-Words model was employed to convert the text into feature vectors for input into the LDA model. During the construction of the Bag-of-Words model, thresholds for maximum and minimum document frequency were established to filter out overly frequent or rare words.

LDA topic modeling

The Coherence Score is a metric used to evaluate the quality of topics generated by topic modeling algorithms, particularly Latent Dirichlet Allocation (LDA) and other probabilistic models. It provides a measure of how interpretable and coherent the topics are, which is crucial for understanding and using the results of topic modeling effectively.

$$\text{Coherence} = \frac{1}{N}\sum_{i=1}^{N} \text{NPMI}(w_i, w_j)$$

Perplexity is a measure of the performance of a language model or a topic model, reflecting the model's predictive ability on test data. The lower the perplexity, the more accurate the model's prediction and the better the clustering effect.

The formula for calculatin

$$\text{perplexity}(D) = \exp\left(-\frac{1}{N}\sum_{d=1}^{M}\sum_{w\in d}\log p(w)\right)$$

Sentiment analysis

The TextBlob tool for sentiment analysis provides, in addition to the polarity score, a subjectivity score. The subjectivity score ranges from [0.0, 1.0], where 0.0 indicates a very objective (factual) stance, and 1.0 indicates a very subjective (opinionated) stance (De Smedt & Daelemans, 2012; Wicke & Bolognesi 2021).

Assuming that the non-linear correlation between time and polarity fits better ($R2$: 0.356, $p<0.001$, f-statistic=26.81) and explains 35.6% of the variance. It is worth noting that higher polynomials provide better fitting, but they cannot serve our survey to determine a simple trend and can overfit our data.

Metaphor analysis

MIPVU

The metaphor identification tool used in this study is MIPVU developed by Steen et al. (2010). The main steps are as follows:

(1).After determining the basic meaning of each keyword with help of the Oxford English Dictionary, read the text carefully to determine the contextual meaning of these keywords;

(2). Determine whether there is a difference between the basic meaning and contextual meaning of these words;

(3). Investigate whether there is a similarity relationship between referents of basic meaning and contextual meaning of these words.

If so, this word is a "metaphor-bearing word".

By determining the semantic domain to which the metaphor-bearing word belongs, that is, the source domain in the conceptual metaphor, the mapping relationship between the source domain and the target domain in the metaphor model is constructed.

(4). During the whole process, Antconc Software, Python, and Wmatrix are used to assist MIPVU in obtaining more accurate metaphor analysis results.

Metaphor analysis

Source domain verification

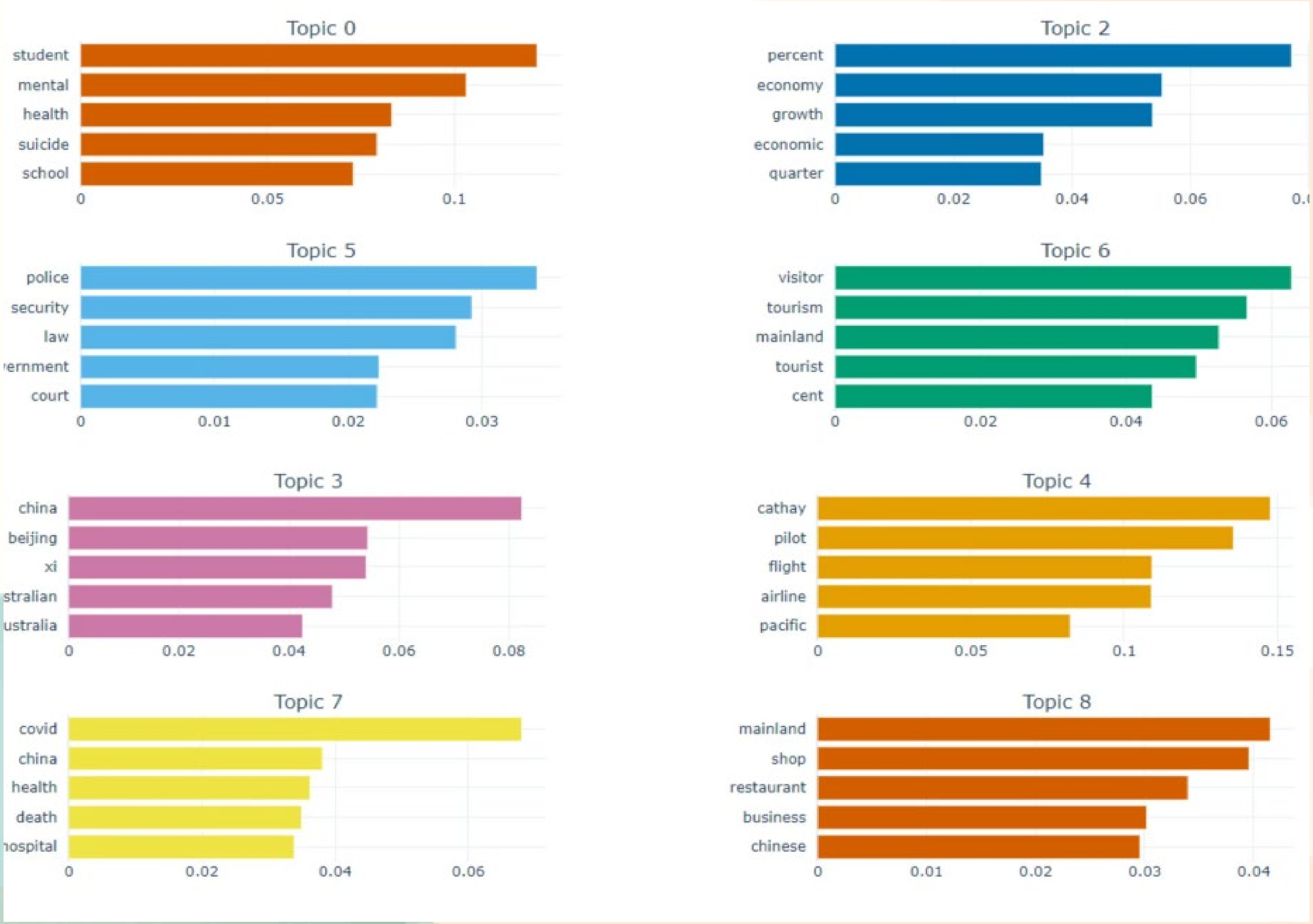
After obtaining all potential metaphorical keywords in the corpus, these keywords were verified to see if they can indeed be classified under the source domain (Ahrens & Jiang, 2020). The source domain verification procedure proposed in the previous study were strictly followed. the meaning and category of the keyword in WordNet-SUMO(Suggested Upper Merged Ontology, http://ontology.teknowledge.com) or the dictionary were determined and were made sure that they match.

Inter-coder reliability

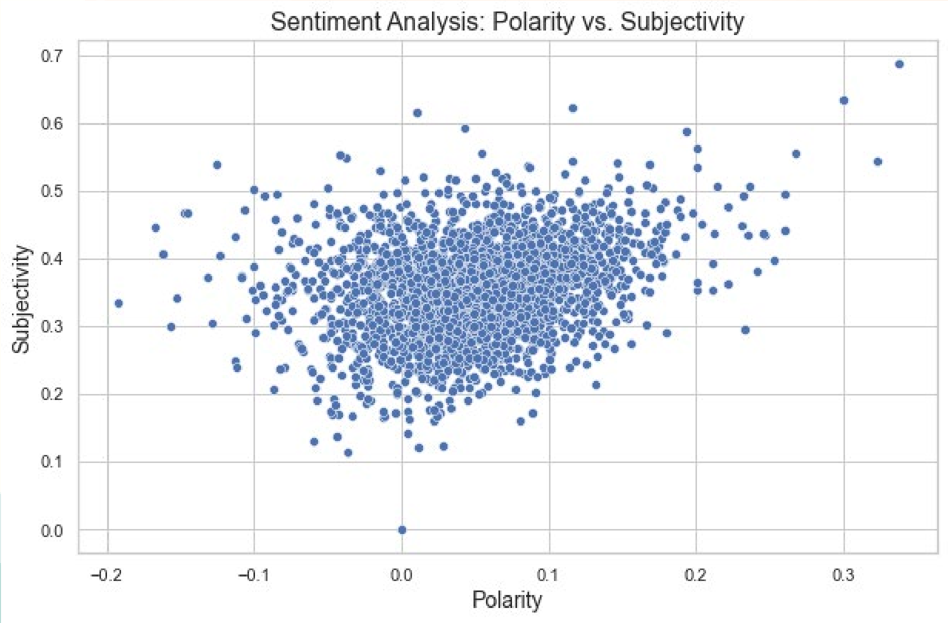
Two coders in linguistics finished the coding, such as the categarization of metaphorical instances and identification of source domains. They had been trained before they could independently analyze the data. They discussed and resolved ambiguous cases after reliability testing. The subset of each reliability analysis varies between 10% and 25% of the entire data set (Wimmer & Dominick, 2011). Cohen' s kappa coefficient was used in this study to ensure the objectivity and accuracy of metaphorical analysis in this research.

Results and discussion

china
chinese
medium
official
party
government
twitter
report
police
government
arrest
protest
journalist
press
public
china
beijing
trump
taiwan
court
chinese
beijing
chinese
country
china
government
trump
china
virus
country
government
outbreak
chinese
coronavirus
government
wuhan
china
day
outbreak
coronavirus
mask
virus
business
government
student
mask
government
test
social distancing
police
assembly
vigil
ban
law
macau
national security
police
law
government
pro democracy
strike
virus
china
taiwan
vaccine
china
health
test
covid
pandemic
hotel
trump
election
biden
china
vote
government
test
infection
quarantine
china
government
public
restaurant
taiwan
china
virus
government
chinese
government
vaccination
infection
vaccine
china
government
health
virus
outbreak
coronavirus


Topic 0
student
mental
health
suicide
school
0
0.05
0.1
Topic 2
percent
economy
growth
economic
quarter
0
0.02
0.04
0.06
Topic 5
police
security
law
ernment
court
0
0.01
0.02
0.03
Topic 6
visitor
tourism
mainland
tourist
cent
0
0.02
0.04
0.06
Topic 3
china
beijing
xi
stralian
ustralia
0
0.02
0.04
0.06
0.08
Topic 4
cathay
pilot
flight
airline
pacific
0
0.05
0.1
0.15
Topic 7
covid
china
health
death
ospital
0
0.02
0.04
0.06
Topic 8
mainland
shop
restaurant
business
chinese
0
0.01
0.02
0.03
0.04

Sentiment analysis



Sentiment analysis

When analyzing the above Figure, the first thing is the relationship between emotional Polarity and subjectivity. The figure shows the distribution of a large number of text data points, with the horizontal axis representing emotional Polarity and the vertical axis representing subjectivity.

From the overall distribution, most of the data points are concentrated in the range of Polarity from -0.2 to 0.4, and the value of Subjectivity is roughly between 0.2 and 0.5.

This distribution indicates that the vast majority of texts have neutral or slightly positive emotional tendencies, while the subjectivity of these texts

Sentiment analysis

Overall, this scatter plot provides us with preliminary insights into the distribution of textual emotions and subjectivity, emphasizing the neutral tendency of textual emotions and their relatively low subjective expression.

In the subsequent analysis, it is possible to further explore how the emotions and subjectivity of these texts change in different contexts, in order to obtain a more comprehensive understanding.

Sentiment analysis

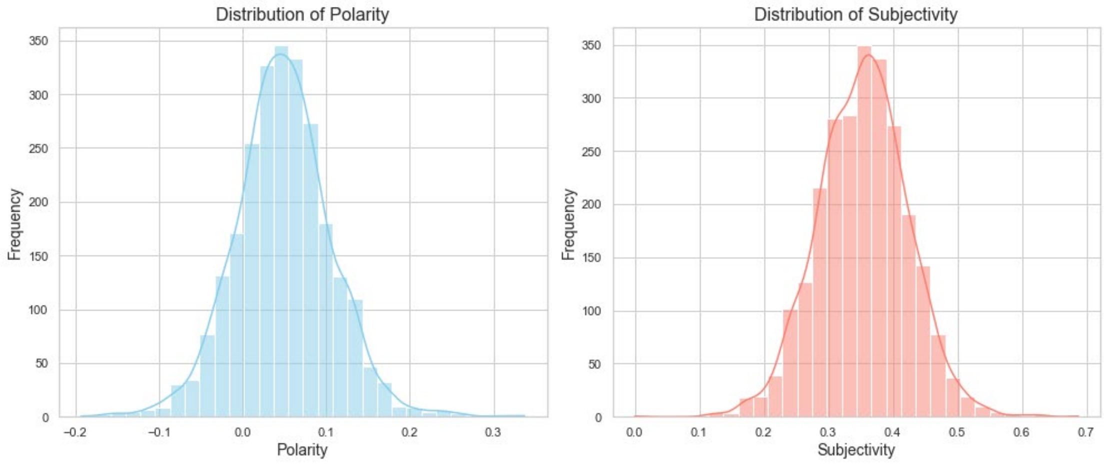
R^2 value is 0.00798

R-squared (R-squared), also known as the coefficient of determination, represents the proportion of the total variability in the dependent variable that's explained by the independent variables in a regression model. The range of R-squared values is from 0 to 1, with values closer to 1 indicating a better fit of the model.

In sentiment analysis, R-squared is a statistical measure used to assess the degree to which a model fits the data. It represents the ratio of the variability in the predicted sentiment scores explained by the model to the total variability in the actual sentiment scores. A high R-squared value implies that the model can explain a significant portion of the variability in the data, while a low R-squared value suggests that the model's predictive power is weak.

Based on search results, we can see that in some applications of sentiment analysis, the R-squared value can be very high, close to 1, indicating that the model can For example, in a provided example, a model's R-squared value reached 0.90000000, demonstrating a high level of accuracy in predicting sentiment. However, in another case, the R-squared value was very low, almost 0, indicating that the model's predictive capabilities are very limited. These values show that R-squared is an important metric for evaluating the

Sentiment analysis



Sentiment analysis

It reveals the distribution of sentiment polarity and subjectivity, which provides important insights.
Firstly, the sentiment polarity histogram on the left reveals that the distribution of data roughly follows a normal distribution pattern.
The majority of the text's Polarity values are concentrated around 0, slightly biased towards the positive side.
This indicates that emotional expressions in the analyzed text samples tend to be neutral or slightly positive, while the number of texts with negative emotions is relatively small.
This distribution may reflect the balance and neutrality of the text content, possibly due to the emphasis on objectivity in writing and avoiding strong emotional tendencies.

Sentiment analysis

The subjective distribution map on the right also shows a shape close to normal distribution, but its peak is slightly shifted to the right, concentrated between 0.3 and 0.4.

This means that the subjectivity of most texts is in a moderately low range, indicating that these texts are more based on facts, and although there is a certain degree of subjective expression, overall, they still maintain relative objectivity.

It is worth noting that the number of texts with extreme subjectivity or extreme objectivity is relatively small, further indicating the moderate expression of subjectivity in the content of the text

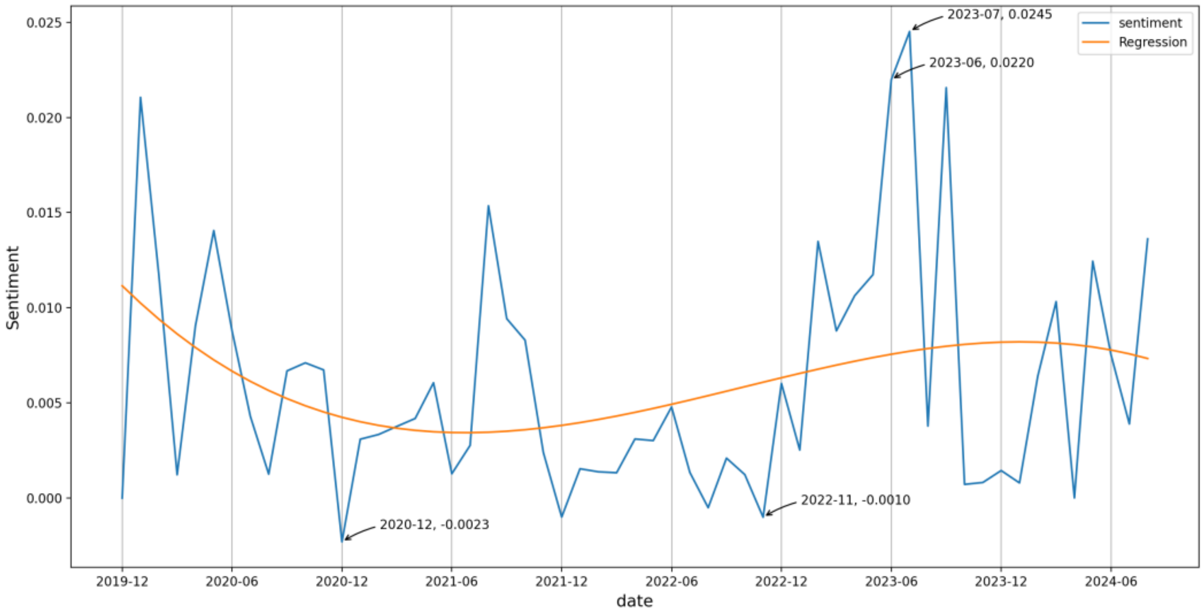
Sentiment analysis

Comparing the two distribution maps, the analyzed text exhibits a tendency that most texts are emotionally mild and biased towards neutrality while also being moderately subjective in expression. This feature may be related to the type of text, especially if these texts are essentially objective content such as news reports or academic articles. So, the gentleness of emotions and the objectivity of expression complement each other, making the overall text present a stable and impartial style.

Overall, these distribution maps provide us with important information about text emotions and expression styles, indicating that balance and neutrality are dominant features in these texts.

In further research, it can be explored whether these features have consistency across different categories of text, or whether there are

Sentiment analysis



Sentiment analysis

On July 10 2023, the booking volume of Hong Kong's tourist tickets exceeded expectations by giving away free tickets, and the tourism of Hong Kong was revitalized and promoted again after the pandemic. In the meanwhile, to boost the city's morale and economy, the government has launched an HK $20 million "Happy Hong Kong" campaign. These measures by the Hong Kong authorities demonstrate that they are determined to revive economy and thus also show optimism and positive feelings. For example:

Airline HK Express will launch its free flight giveaway at 10. 30 am on Tuesday, with round-trip tickets to destinations in Japan, South Korea, Thailand, Vietnam, and Taiwan on offer - the latest phase of a campaign aimed at rebooting tourism after years of Covid travel restrictions. Open to people living in Hong Kong, HK Express's campaign is giving away 21,626 complimentary tickets to 19 Asian destinations from July 11 to July 24 on a first come, first served

Sentiment analysis

The second highest value is 0.0220, and the date is June 2023. On June 14 2023, Hong Kong authorities planned to introduce 20,000 laborers and workers to deal with a labor shortage in the wake of the pandemic. This reveals that Hong Kong's economy is generally recovering and improving to some extent. For instance:

Hong Kong is set to import around 20,000 workers in a bid to alleviate the labour crunch in the construction, transport and aviation sectors, the government has announced.......The low-skilled labour force fell by around 160,000 people, Secretary for Labour and Welfare Chris Sun said. "Therefore, after the return to normalcy, many industries in Hong Kong are facing the challenge of labour shortages," he said.

Sentiment analysis

The lowest value is -0.0023, and the date is December 2020. On December 24 2020, due to the coronavirus, the number of infections and deaths in many countries rose, such as Brazil. Meanwhile, almost every government and agency invested a lot of funding and resources in developing an effective vaccine. Here is an instance:

Clinical trials of the CoronaVac coronavirus vaccine developed by Chinese laboratory Sinovac have "reached the efficacy threshold" demanded by the World Health Organization, the Brazilian institute charged with its production and distribution said on Wednesday. However, the Butantan Institute didn't publish the results of those trials — the last before authorization. ...... Immunization has been a highly politicized issue in Brazil , where far-right President Jair Bolsonaro has repeatedly said he won't take a vaccine while he's also tried to discredit the CoronaVac jab. The Butantan Institute is supported by Sao Paulo state, whose governor Joao Doria repeatedly clashed with Bolsonaro over the country's coronavirus response and is expected to challenge the incumbent in the 2022 presidential elections. Brazil has suffered the second-largest number of coronavirus deaths in the world after

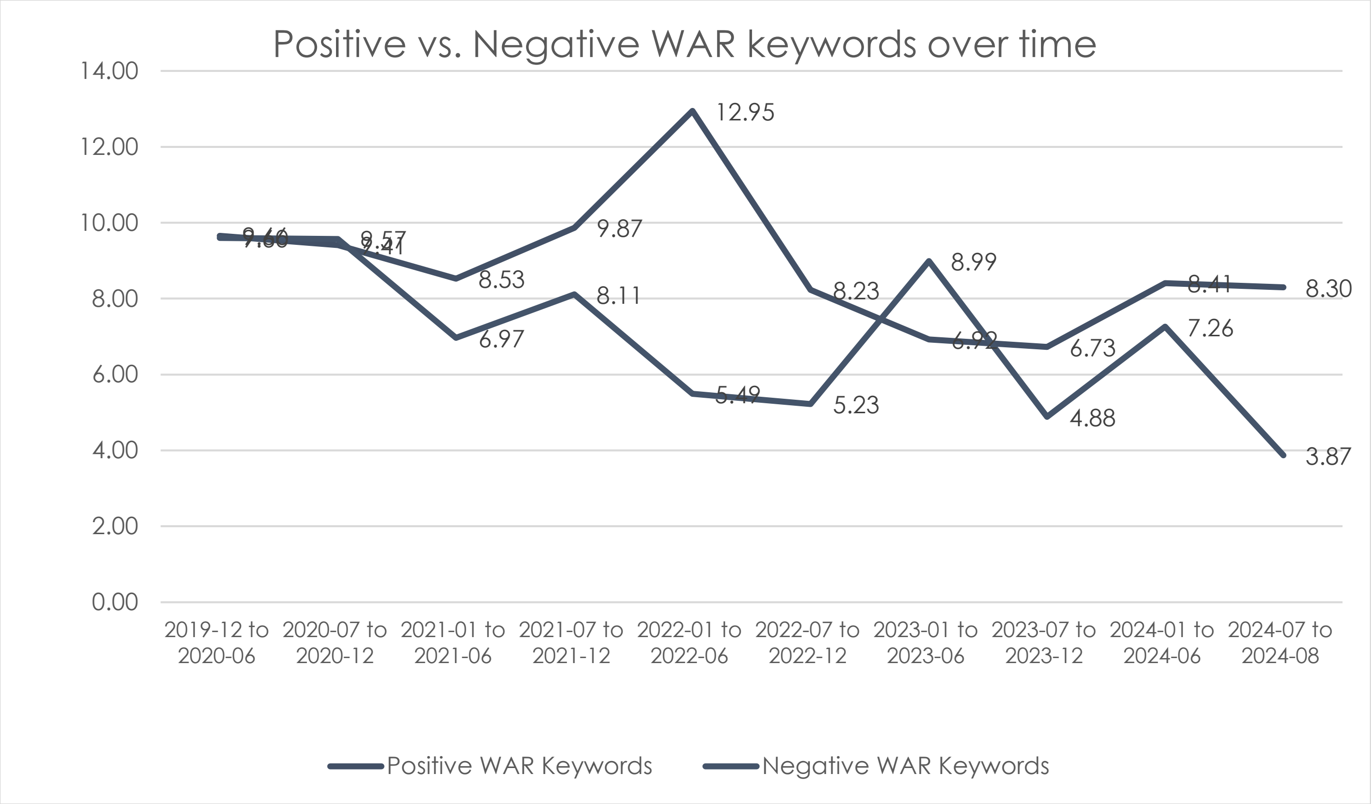
Metaphor analysis

WAR Keywords	**Lemmas of WAR Keywords**					
PROTECTION	*protect*	*protects*	*protected*	*protecting*	*protection*	*protections*
FIGHT	*fight*	*fights*	*fought*			
STRATEGY	*strategy*	*strategies*				
COMBAT	*combat*	*combats*	*combating*			
VICTORY	*victory*	*victories*				
WAR	*war*	*wars*				
THREAT	*threat*	*threats*	*threaten*	*threatens*	*threatened*	*threatening*
VIOLENCE	*violence*	*violent*				
STRUGGLE	*struggle*	*struggles*	*struggling*	*struggled*		
ATTACK	*attack*	*attacked*	*attacking*			

Metaphor analysis

The figure shows the standardized number of positive and negative WAR keywords per 10,000 words during the time period between 2019 to 2024. From January 2021 to June 2022 is the peak period of positive words in the chart.
In other words, it is also the most positive stage for words associated with metaphors.
In the period, the most frequent use of positive words about war metaphors, such as protect, strategy, fight, and so forth, in news texts is aimed at boosting the public's confidence in winning the fight against COVID-19 and overcoming the fear of the virus.

Metaphor analysis



Discussion

Sentiment analysis reflects the emotional disposition towards the novel coronavirus epidemic. The changing trend of the direction. In general, the news have an toward the COVID-19 pandemic. The overall fluctuation is small, and the objectivity of the new is strong.

The overall emotional Polarity in five years from 2019 to 2024 in the corpus was slightly positive (>0). Polynomial regression reveals that the average mood becomes increasingly positive from 2021 to 2024, while it drops softly from 2019 to 2021.
From 2019 to 2021, the general attitude of the news tended to be slightly negative and pessimistic.

Discussion

During this period, many countries were under lockdown, and authorities typically implemented strict quarantine measures. At the same time, there were new deaths and infections every month. These made the public fearful of the unknown. But proactive protective measures and the promotion of vaccines have also given people confidence and hope, so the emotional inclination is slightly pessimistic.

With the gradual reduction of infections and deaths, large-scale public events such as music festivals have resumed, and most citizens have been vaccinated in an orderly manner. News attitudes toward the pandemic have gradually become optimistic.

References

Amara, A., Hadj Taieb, M. A., & Ben Aouicha, M. (2021). Multilingual topic modeling for tracking COVID-19 trends based on Facebook data analysis. Springer Science+Business Media, LLC, part of Springer Nature. https://doi.org/10.1007/s42979-020-00352-5

Benrouba, F., & Boudour, R. (2023). Emotional sentiment analysis of social media content for mental health safety. Social Network Analysis and Mining, 13(1), 1-21. https://doi.org/10.1007/s13278-022-01000-9

Bhardwaj, M., Mishra, P., Badhani, S., & Muttoo, S. K. (2024). Sentiment analysis and topic modeling of COVID-19 tweets of India. International Journal of System Assurance Engineering and Management, 15(5), 1756-1776. https://doi.org/10.1007/s13278-022-01000-9

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3(Jan), 993–1022. https://dl.acm.org/doi/10.5555/944919.944937

Boon-Itt, S., & Skunkan, Y. (2020). Public perception of the COVID-19 pandemic on Twitter: Sentiment analysis and topic modeling study. JMIR Public Health and Surveillance, 6(4), e21978. https://doi.org/10.2196/21978

Charteris-Black, J. (2009). Metaphor and political communication. In A. Musolff & J. Zinken (Eds.), Metaphor and discourse (pp. 97–115). Palgrave Macmillan. https://doi.org/10.1007/978-1-349-36585-5_6

Damevski, K., Chen, H., Shepherd, D. C., Kraft, N. A., & Pollock, L. L. (2018). Predicting future developer behavior in the IDE using topic models. IEEE Transactions on Software Engineering, 44, 1100–1111. https://doi.org/10.1109/TSE.2017.2771683

De Smedt, T., & Daelemans, W. (2012). Pattern for Python. Journal of Machine Learning Research, 13, 2063–2067. https://doi.org/10.5555/2188385.2343710

Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. https://arxiv.org/abs/1810.04805

Follett, L., Geletta, S., & Laugerman, M. (2019). Quantifying risk associated with clinical trial termination: A text mining approach. Information Processing & Management, 56, 516–525. https://doi.org/10.1016/j.ipm.2019.03.006

Ge, J., Alonso Vazquez, M., & Gretzel, U. (2018). Sentiment analysis: A review. In M. Sigala & U. Gretzel (Eds.), Advances in social media for travel, tourism and hospitality: New perspectives, practice and cases (pp. 243-261). New York, NY: Routledge.

Heintz I et al (2013) Automatic extraction of linguistic metaphor with lda topic modeling Inproceedings of the First Workshop on Metaphor in NLP

Jacob Devlin, M., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. Retrieved from https://arxiv.org/abs/1810.04805

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. Multimedia Tools and Applications, 78, 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia tools and applications, 78, 15169-15211.

Jurafsky, D., & Martin, J. H. (2009). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (2nd ed.). Prentice Hall, Pearson Education International.

Kaur, C., & Sharma, A. (2020). Social issues sentiment analysis using python. In Proceedings of the 5th International Conference on Computing, Communication and Security (ICCCS) (pp. 1-6). IEEE. DOI: 10.1109/ICCCS49374.2020.00040

Kaur, C., & Sharma, A. (2020). Twitter sentiment analysis on coronavirus using textblob. EasyChair2516-2314.

Khan, W., Ghazanfar, M. A., Azam, M. A., Karami, A., Alyoubi, K. H., & Alfakeeh, A. S. (2022). Stock market prediction using machine learning classifiers and social media, news. Journal of Ambient Intelligence and Humanized Computing, 1-24. https://doi.org/10.1007/s12652-022-03866-z

Liu, B. (2010). Sentiment analysis and subjectivity. In Handbook of natural language processing (Vol. 2, pp. 627-666). https://doi.org/10.1201/9781420085768-c23

Liu, L., Tang, L., Dong, W., Yao, S., & Zhou, W. (2016). An overview of topic modeling and its current applications in bioinformatics. SpringerPlus, 5, 1608. https://doi.org/10.1186/s40064-016-3258-5

Maaten, L. v. d., & Newell, G. (2020). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2008.07999. Retrieved from https://arxiv.org/abs/2008.07999

Machová, K., Mikula, M., Gao, X., & Mach, M. (2020). Lexicon-based sentiment analysis using the particle swarm optimization. Electronics, 9(8), 1317. https://doi.org/10.3390/electronics9081317

McFarland DA et al (2013) Differentiating language usage through topic models. Poetics 41(6):607–625

Mathur, A., Kubde, P., & Vaidya, S. (2020). Emotional analysis using Twitter data during pandemic situation: COVID-19. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 845–848). IEEE. https://doi.org/10.1109/ICCES48766.2020.9138079

Miller, M., Banerjee, T., Muppalla, R., Romine, W., & Sheth, P. A. (2017). What are people tweeting about Zika? An exploratory study concerning its symptoms, treatment, transmission, and prevention. JMIR Public Health and Surveillance, 3(2), e38. https://publichealth.jmir.org/2017/2/e38/

Musolff, A., & Zinken, J. (Eds.). (2009). Metaphor and political discourse. Palgrave Macmillan. https://doi.org/10.1057/9780230224156

Radim, R., & Sojka, P. (2010). Gensim: Topic modelling for humans. In Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks. ELRA, Valletta.

Rana, A., Ayoub, T., Ghilzai, S. A., & Shehzad, W. (2024). Metaphorical framing of the COVID-19 pandemic in Pakistan: A corpus-driven critical analysis of war metaphors in news media. PLoS ONE, 19(10). https://doi.org/10.1371/journal.pone.0297115

Rahul, K., Jindal, B. R., Singh, K., & Meel, P. (2021). Analysing public sentiments regarding COVID-19 vaccine on Twitter. In Proceedings of the 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 488-493). IEEE.

Saberi, B., & Saad, S. (2017). Sentiment analysis or opinion mining: A review. International Journal of Advanced Science and Engineering Information Technology, 7(5), 1660-1666. https://doi.org/10.18517/ijaseit.7.5.2153

Semino, E., Demjén, Z., Demmen, J., Koller, V., Payne, S., Hardie, A., et al. (2017). The online use of violence and journey metaphors by patients with cancer, as compared with health professionals: A mixed methods study. BMJ Supportive & Palliative Care, 7(1), 60–66. https://doi.org/10.1136/bmjspcare-2014-000785

Sengupta, M., El Baff, R., Alshomary, M., & Wachsmuth, H. (2024, June). Analyzing the use of metaphors in news editorials for political framing. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (pp. 3621-3631).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Journal of Machine Learning Research, 18(1), 1-61. Retrieved from http://jmlr.org/papers/v18/1706.03762.html

Wicke, P., & Bolognesi, M. M. (2021). Covid-19 discourse on Twitter: How the topics, sentiments, subjectivity, and figurative frames changed over time. Frontiers in Communication, 6, 651997. https://doi.org/10.3389/fcomm.2021.651997

Wicke, P., & Bolognesi, M. M. (2020). Framing COVID-19: How we conceptualize and discuss the pandemic on Twitter. PLoS ONE, 15(9), e0240010. https://doi.org/10.1371/journal.pone.0240010

Yin, H., Song, X., Yang, S., & Li, J. (2022). Sentiment analysis and topic modeling for COVID-19 vaccine discussions. World Wide Web, 25(3), 1067-1083. https://doi.org/10.1007/s11280-021-00897-0

Zeng, H., Chen, J., Jiang, M., Gu, J., Peng, B., Ahrens, K., & Huang, C-R. (2024). Legitimation Strategies in Corpora of Government News on COVID-19: A Topic Modeling Approach. Submitted.

Zhang, L. (2022). COVID-19 sensing: Negative sentiment analysis on social media in China via BERT model. Journal of Social Media Studies, 3(2), 45-59.

Zhu, N., Zou, P., Li, W., & Cheng, M. (2012). Sentiment analysis: A literature review. In Proceedings of the 2012 IEEE International Symposium on Management of Technology (pp. 572-576). IEEE.

References

Ahrens, K., & Jiang, M. (2020). Source domain verification using corpus-based tools. Metaphor and Symbol, 35(1), 43-55.

Ahrens, K., & Zeng, W. H. (2022). Referential and evaluative strategies of conceptual metaphor use in government discourse. Journal of Pragmatics, 188, 83-96.

Al Husain, M. A. (2020). The metaphorical modeling of COVID-19 pandemic crisis in Infosphere online news articles. Journal of Education College Wasit University, 1(40), 655-672.

Amaireh, H. A. (2022). Covid-19 is War, Water & a Person: Metaphorical Language of the Coronavirus Disease in" the Jordan Times" Newspaper. Theory and Practice in Language Studies, 12(7), 1286-1293.

Busso, L., & Tordini, O. (2022). How do media talk about the Covid-19 pandemic? Metaphorical thematic clustering in Italian online newspapers. arXiv preprint arXiv:2204.02106.

Charteris-Black, J. (2011). Politicians and rhetoric: The persuasive power of metaphor. Springer.

Charteris-Black, J. (2004). Corpus approaches to critical metaphor analysis. Basingstoke: Palgrave Macmillan.

Charteris-Black, J. (2009). Metaphor and gender in British parliamentary debates (pp. 139-165). Palgrave Macmillan UK.

Charteris-Black, J. (2021). Metaphors of coronavirus: Invisible enemy or zombie apocalypse?. Springer Nature.

Chen, J. Z., Ahrens, K., & Tay, D. (2024). 'Luxurious' metaphors in luxury hotel websites in Singapore and Hong Kong: A mixed-methods study. Applied Corpus Linguistics, 4(2), 100090.

Chen, T., Dai, M., Xia, S., & Zhou, Y. (2022). Do messages matter? Investigating the combined effects of framing, outcome uncertainty, and number format on COVID-19 vaccination attitudes and intention. Health Communication, 37(8), 944-951.

Choi, E. P. H., Hui, B. P. H., & Wan, E. Y. F. (2020). Depression and anxiety in Hong Kong during COVID-19. International journal of environmental research and public health, 17(10), 3740.

Colak, F. U. (2023). COVID-19 as a metaphor: understanding COVID-19 through social media users. Disaster Medicine and Public Health Preparedness, 17, e159.

Fellbaum, C. (2010). WordNet. In Theory and applications of ontology: computer applications (pp. 231-243). Dordrecht: Springer Netherlands.

Flusberg, S. J., Mackey, A., & Semino, E. (2024). Seatbelts and raincoats, or banks and castles: Investigating the impact of vaccine metaphors. Plos one, 19(1), e0294739.

Fu, F. (2024). Analyzing metaphor patterns in COVID-19 news pictures: A critical study in China. Plos one, 19(2), e0297336.

Guliashvili, N. (2024). "What's COVID-19 Vaccine Like?"–From A Cognitive Frame to A Similative Proposition. ESI Preprints, 20(2).

Harnovitasari, T. D. (2017). An analysis of Metaphor and simile and jane eyre novel by charlotte bronte (Doctoral dissertation, IAIN Ponorogo).

Jung, J., Horta, H., & Postiglione, G. A. (2021). Living in uncertainty: The COVID-19 pandemic and higher education in Hong Kong. Studies in Higher Education, 46(1), 107-120.

Khaliq, M., Joseph, R., & Rai, S. (2021). Covid is war and vaccine is weapon? COVID-19 metaphors in India. In Proceedings of the 18th International Conference on Natural Language Processing (ICON) (pp. 431-438).

Koller, V., Hardie, A., Rayson, P., & Semino, E. (2008). Using a semantic annotation tool for the analysis of metaphor in discourse. Metaphorik. de, 15(1), 141-160.

Kozlova, T. (2021). Cognitive metaphors of covid-19 pandemic in business news. In SHS web of conferences (Vol. 100, p. 02004). EDP Sciences.

Krennmayr, T. (2011). Metaphor in newspapers.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), Metaphor and thought (2nd ed.) (pp.202–250). Cambridge, UK: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980, 2003). Metaphors we live by. University of Chicago press.

Li, E. C. Y., & Lee, P. H. (2023). The geopolitics of disease prevention: Military analogies against COVID-19 in Hong Kong, Taiwan, and beyond. Media, Culture & Society, 01634437231216444.

Liu, Y. (2023). (Re) framing to (re) evaluate: metaphors in cross-national COVID-19 news translation.

Liu, Y., & Tay, D. (2023). Modelability of WAR metaphors across time in cross-national COVID-19 news translation: An insight into ideology manipulation. Lingua, 286, 103490.

Mujagić, M. (2018). Dangerous waters metaphor in news discourse on refugee crisis. Metaphorik. de, 28, 99-131.

Thank You