



# Pandemic-Related Media Coverage in Overseas Chinese Media Based on Big Data Analysis: A Case Study of “oushi 1983”

Nan Dai<sup>(✉)</sup>, Yujie Guo, and Miao Wang

Shanghai University of Political Science and Law, No.7989, Waiqingsong Highway, Qingpu District, Shanghai 100024, China  
dainan@shupl.edu.cn

**Abstract.** Under the background of the big data era, analysis and processing technology based on big data has given birth to new methods of humanities research. Taking the reporting practice of overseas Chinese media during the COVID-19 pandemic as the research object and the new media platform “oushi 1983” of Nouvelles D’Europe as the study object, this paper captures the reports published between January 1, 2020, and December 31, 2021, as samples, and conducts big data analysis on the content based on relevant clustering algorithms and text analysis tools. Through iterative computations of the K-means algorithm, 1661 pieces of reports related to the COVID-19 pandemic are clustered and analyzed, so as to make an empirical study on the content theme and emotional bias of the report. By doing so, the reporting pattern and communication function of overseas Chinese media in the international public opinion field is investigated.

**Keywords:** big data analysis · K-means algorithm · overseas Chinese media · pandemic-related media coverage

## 1 Introduction

Since 2020, the COVID-19 pandemic has swept the world. The pandemic has accelerated the evolution of major changes unseen in a century, and also formed a window period for the construction of a new international political and economic order.

The media is the main participant in the construction of international discourse and communication order. In the complex and changeable global public opinion field, the communication value of overseas Chinese media is particularly prominent. Overseas Chinese media mainly refer to mass media outside of mainland China, Hong Kong, Taiwan, and Macau that use Chinese as a way of communication, including newspapers, magazines, online media, radio, television, and various emerging media, majorly founded by settled overseas Chinese [1]. According to incomplete statistics, there are 1019 Chinese media with a certain scale and influence in the world, including 390 newspapers, 221 magazines, 77 television stations, 81 radio stations, and 250 websites, distributed in 61 countries and regions [2]. These overseas Chinese media born and grown abroad are

the mouthpiece of Chinese culture, an important force in today's complex international public opinion environment, as well as an important overseas fulcrum for the spread of China's international political discourse.

## 2 Research Review

### 2.1 Path Clarification of Academic Analysis

As the hometown of overseas Chinese, China has been studying overseas Chinese media since the 1980s. Since the 21st century, the research content and region of overseas Chinese media have been expanding, and the types of media have also extended from the early research of Chinese newspapers and periodicals to Chinese TV, Chinese websites, and new mobile media. In recent years, with the proposal of the "Belt and Road" initiative and the in-depth dissemination of the idea of "telling Chinese stories," the significance of Chinese media to China's diplomacy has become increasingly prominent, and related research has also increased significantly. At the same time, research related to communication effect, communication content, and audience has also increased.

With the improvement of China's international status and influence in recent years, new breakthroughs have been made in the research perspective of Chinese media in the world, for example, *Media and Communication in the Chinese Diaspora: Rethinking Transnationalism* published by the British Routledge Press in 2016, which studies overseas Chinese media from the perspective of transnationalism.

As far as the current research status is concerned, the number of overseas Chinese media research is increasing, but the clarification or current situation descriptions of the development process of Chinese media, as well as macro narratives, account for the majority, with insufficient micro and empirical research. The existing literature focuses on the static analysis of the development pattern, which is still relatively backward compared with the current rapidly changing reality in this field, especially for the latest frontier research and dynamic analysis to be in-depth. Besides, some studies lack theoretical support, specific content research, and objective empirical support [3]. This research is based on big data analysis and processing technology, hoping to make a breakthrough in empirical research in this field through data mining and data analysis.

### 2.2 Path Design of the Research

In order to analyze the pandemic reporting and communications of overseas Chinese media through Internet means, the paper takes COVID-related reports from the Wechat public account "oushi 1983" of *Nouvelles D'Europe* as the study object, captures reports published between January 1, 2020, to December 31, 2021, in a targeted manner through relevant clustering algorithms and text analysis tools, and obtains 1661 reported data related to COVID-19 as study samples after data denoising, clustering, and merging. The research classifies and organizes the text data according to the keyword tags of news reports, combined with mathematical statistical methods and content analysis methods, to figure out the rules of the text, refine the characteristics of the report content, and interpret the perspective and tendency of "oushi 1983".

### 3 Research Method

Clustering analysis is an active research section in the field of data mining in the era of big data. As a quantitative method, it is often used in data mining and data comparison to extract similar features, thus providing a basis for clustering and classification. The paper adopts the K-means clustering method based on division, which is an unsupervised learning algorithm proposed by J. B. MacQueen [4] in the 1970s, which is also one of the most widely used clustering algorithms, with advantages of simplicity, high efficiency, and good result distribution. In this research, the K-Means algorithm is used to cluster the data of 1661 COVID-19 pandemic-related reports after data cleaning.

The clustering process is based on the Euclidean distance principle. The data samples are first divided into K clusters, and the centers of K initial clusters are determined. Assuming that X is the sample space and x is a single sample, K samples are randomly selected from the sample space as the initial vector, the set composed of cluster centers is D, and the distance from the remaining samples to the centers of K clusters is calculated respectively. The formula is as follows:

$$\text{dist}(x, \mu_i) = \sqrt{\sum_{k=1}^K |x - \mu_i|^2}, \mu_i \in D \tag{1}$$

Determine which cluster each sample belongs to according to the distance size, update the mean vector according to the result, and repeat the process until it is optimal. A smaller value of E indicates a smaller sample distance within the cluster and a higher similarity.

$$E = \sum_{i=1}^K \sum_{x \in X} |x - \mu_i|^2 \tag{2}$$

By exploring the functional relationship between features, dimension reduction is realized, data samples are expanded in space in the form of points, and better clusters are obtained through iterative calculation.

### 4 Analysis of Research Samples

#### 4.1 Content Theme

After two rounds of cluster analysis based on word frequency and reporting themes, six media coverage themes related to the COVID-19 pandemic are formed, namely, real-time pandemic tracking in Europe, anti-pandemic measures by the European government, pandemic in France, anti-pandemic story in China, mutual assistance between China and Europe, and impact of the pandemic.

As can be seen from Table 1, in the six types of reports about the pandemic, the three types with the highest proportion are “real-time pandemic tracking in Europe,” “anti-pandemic measures by the European government,” and “pandemic in France.” These are also the information most closely related to European Chinese society, the most direct service object of “oushi 1983”. The remaining three categories are “anti-pandemic story in China,” “mutual assistance between China and Europe,” and “impact of the pandemic,” accounting for 18%.

**Table 1.** Proportion of Pandemic-Related Media Coverage Themes in “oushi 1983” [created by the authors]

Media Coverage Theme	Quantity	Proportion
Real-time Pandemic Tracking in Europe	705	42%
Anti-Pandemic Measures by European Government	288	17%
Pandemic in France	373	23%
Anti-Pandemic Story in China	45	3%
Mutual Assistance between China and Europe	22	1%
Impact of the Pandemic	228	14%

Data Statistics Period: January 1, 2020, to December 31, 2021.

## 4.2 Emotional Bias

In terms of the guidance of media reports on the emotional bias of the public, its orientation affects public opinion to a certain extent. Based on the difference between the report object and the report content, the research divides the 1661 samples related to the COVID-19 pandemic from the “oushi 1983” into two categories, namely, China’s pandemic-related media coverage and Europe’s pandemic-related media coverage. These two types of research samples are divided according to the emotions displayed in the reports, and the emotion bias is calculated with jieba, a Chinese word segmentation tool jieba, and Hownet, an emotion dictionary (Table 2).

Statistics show that the proportion of positive reports on the pandemic in China is higher. The main positive reports include praising a series of timely and strict pandemic prevention and control measures implemented by the Chinese government, praising medical workers who are working hard against the pandemic, etc. In addition, “oushi 1983” also reports China’s assistance to European countries in terms of sending medical experts abroad, curing critically ill patients, and sharing experiences in fighting the pandemic. In terms of neutral-emotional reports, the main reports focus on providing an objective and clear interpretation of Chinese policies to the European public, and providing valuable news and information services to overseas Chinese.

As can be seen from Table 3, the reports on the COVID-19 pandemic in Europe by the “oushi 1983” show more negative emotion, covering ineffective government pandemic prevention measures, uncontrolled deterioration of the pandemic, shortage of medical

**Table 2.** Emotional Bias of “oushi 1983” to China’s Pandemic-Related Media Coverage [created by the authors]

Media’s Emotional Bias		
Positive	Neutral	Negative
53%	45%	2%

Data Statistics Period: January 1, 2020, to December 31, 2021.

**Table 3.** Emotional Bias of “oushi 1983” to Europe’s Pandemic-Related Media Coverage [created by the authors]

Media’s Emotional Bias		
Positive	Neutral	Negative
15%	48%	37%

Data Statistics Period: January 1, 2020, to December 31, 2021.

supplies and beds, rising mortality rate, rising unemployment rate, family crisis, student dropout, and so on. The reports with neutral emotions are mostly about pandemic data updates, pandemic-related knowledge dissemination, and information on government pandemic prevention deployment and decision-making.

## 5 Conclusion

Based on data analysis and empirical research, we can see that in the COVID-related reports, overseas Chinese media, due to their unique location advantages, present the global pandemic report in Chinese with their local practice on the one hand and construct a Chinese narrative on the global fight against the pandemic in China’s external communication on the other hand. Overseas Chinese media play its unique communication function in the international public opinion pattern.

- (1) Facing audiences in different countries and regions, overseas Chinese media collect information, track and analyze the pandemic trends on time, present comprehensive and real local pandemic reports, and exert the information dissemination function in public health emergencies.
- (2) As an important bridge connecting home and abroad, overseas Chinese media actively report China’s effective anti-pandemic measures and charitable donations, resolutely oppose the reports stigmatizing China, and play the integration function in international public opinion.
- (3) Overseas Chinese media play the “composite” [5] function in media image construction and build the international image of China as a “responsible major country” with the domestic media.
- (4) The anti-pandemic reports of overseas Chinese media have built an anti-pandemic media front with full news and information, widespread scope, strong public opinion guidance, and good communication effect, built a bridge for the exchange of anti-pandemic information at home and abroad, and played the role of constructing discourse power in international communication capacity.

As brand-new digital research data, the size of the big data sample is huge and rich with a large time span. Based on statistics on the relevant data of the pandemic report from January 1, 2020, to December 31, 2021, the paper mines and summarizes the report samples through the iterative calculation of the K-means algorithm, and depicts the research object in a relatively comprehensive, complete and objective manner, thus

breaking through the limitations of previous research, as well as enhancing the empirical and scientific nature of related research.

## References

1. Qiu L. Dialectical Analysis of the Role of Overseas Chinese Media in China's External Communication[J].International Communications, 2015(10).
2. Lv Z Y. The More Complex the International Public Opinion Environment, the More it is Necessary to Play the Role of Chinese media: the Practice and Exploration of China News Service in Building a World Chinese Information Center [J].International Communications, 2015(10).
3. Ming D.A Critical Review of China's Research on Overseas Chinese Media (1979 — 2015) — An Analysis on the China National Knowledge Infrastructure Database[J].Journal of Overseas Chinese History Studies.2017(02).
4. J.B.Mac Queen. Some Methods for Classification and Analysis of Multivariate Observations[M].Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,no.1.1967.pp.281–297.
5. Jie H. Comparison of Overseas Chinese media and Chinese Domestic Media in Building National Image [J]. The Press, 2016(03).

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

