

# The impact of COVID-19 on the protection of rural traditional village

Li Gang<sup>a,\*</sup>, Wang Fang<sup>b</sup> and Quan Sishi<sup>c</sup>

<sup>a</sup>Wuhan Donghu University, Wuhan, Hubei, China

<sup>b</sup>Wuchang University of Technology, Wuhan, Hubei, China

<sup>c</sup>Siemens Transformer (Wuhan) Co., Ltd, München, Germany

**Abstract.** COVID-19's significant impact on economic and social life has rightfully garnered the attention of citizens and policymakers alike. In response to the pandemic, governments have issued strict guidelines and restrictions to shut down some cities and many rural villages in China. With no cure or vaccine on the horizon, governments are working to mitigate the damage of the lockdowns on rural cultural village. Over the past two decades, rural village has been negatively impacted by terrorism, lack of funding and loss of population. COVID-19 has had similar effect, but in an incredibly short period of time. During the control period of COVID-19, traditional data are widely used in village protection and renewal. Collect and sort out the original data of Huizhou culture to prepare for the subsequent calculation. After the data is ready, the data is processed as the basis of mining its potential application value. In this paper, the key words of Huizhou cultural resources are summarized. The data analysis platform is established. This paper analyzes people's preference for Huizhou cultural resources. To better realize the more effective and far-reaching development and exploitation of Huizhou cultural resources.

**Keywords:** Big data, Huizhou traditional village cultural resources, cultural heritage, mining applications, COVID-19

## 1. Introduction

Because the information about the epidemiological characteristics and transmission mechanism of COVID-19 virus in the early stage of the epidemic has not been completely clear, and there is no rapid diagnosis method. The medical laboratory diagnosis of the virus is mainly through nucleic acid detection, so the whole diagnosis process is relatively complicated, and the diagnosis time of patients is generally long and there is uncertainty. COVID-19's significant impact on economic and social life has rightfully garnered the attention of citizens and policymakers alike. In response to the pandemic, governments have issued strict guidelines and restrictions to shut down theaters, museums, villages, libraries, airports, malls, schools, and restaurants. The damage

of COVID-19 on rural traditional villages is significant. Therefore, the protection of rural traditional villages must be taken seriously. Traditional villages refer to villages with a long history, a relatively complete natural and cultural environment, important historical significance and valuable heritage value. They have important research value for architecture, planning, landscape, humanities and other multi-level disciplines, and are the crystallization of traditional farming culture in China. By the end of 2018, the number of traditional villages in China has reached 6799, among which Huizhou is one of the areas where traditional villages gather the most [1].

To study the development and current situation of the traditional village landscape in Huizhou, we must first understand the natural and human environment in Huizhou. The clan culture has an important influence on the formation and development of the traditional village landscape in Huizhou. The family culture derived from it also has reference significance

---

\*Corresponding author. Li Gang, Wuhan Donghu University, Wuhan 430212, Hubei, China. E-mail: lg325@tom.com..

for the promotion of the traditional village landscape in Huizhou.

Huizhou is located at the junction of Anhui, Jiangxi and Zhejiang provinces, also known as “Xin’an” in ancient times, including Shexian County, Xiuning County, Shexian County, Qimen County, Wuyuan County and Jixi County [2]. Xin’an River winds through [3]. Huizhou area has plain, hill, basin and other landforms, with rolling mountains, rivers and streams, and various natural landscapes.

From the birth of the Fengjian society to the Northern Song Dynasty, the Central Plains has always been the political, economic and cultural center of Chinese civilization, with a relatively dense population [4]. However, due to political turmoil, war and other reasons, there have been several large-scale population migration in the history of Chinese civilization [5]. With the outbreak of “Wuhu Luanhua” at the end of the Western Jin Dynasty, the Central Plains was devastated. Sima Rui, the royal family of Jin, established the Eastern Jin Dynasty in Jianye, and a large number of princes, nobles and rich families crossed the Yangtze River to avoid chaos and protect themselves [6]. At the end of the Northern Song Dynasty, “Jingkang change” took place, and Zhao Gou, the royal family of song, moved to Lin’an to establish the Southern Song Dynasty. According to historical records, “people from all directions gather in two Zhejiang provinces, a hundred times as often.” A large number of people from the Central Plains moved south to Huizhou and Jianghuai areas [7]. The migration of the northern population settled in Huizhou and mingled with the aborigines, resulting in a large number of traditional villages in Huizhou and gradually developed [8–10]. At the same time, several times of “dress south” brought not only the increase of population, but also the spiritual content of the traditional culture and ideology which had been inherited for a long time and accumulated in the Central Plains, which played a guiding and reference role in the development of the clan culture in Huizhou [11].

Feudal land ownership and self-sufficient small-scale peasant economy are the main reasons why Huizhou traditional villages are deeply influenced by clan culture. During the northern and Southern Dynasties, the northern part of China was in a state of war for a long time. State taxes are guaranteed. Emperor Xiaowen of the Northern Wei Dynasty issued the land distribution order, which stipulated that all the land owned by the emperor should be distributed to the farmers, and the farmers should pay

taxes to the emperor and undertake certain military forces. Before the middle of Tang Dynasty, equal field system was adopted. In this feudal mode of production, generations of farmers are bound to a fixed small piece of land, farming has become their way of life. In this way, clan members with blood relationship live together and help each other, forming clan culture. Huizhou, with a long history, with the acceleration of urbanization, traditional villages are disappearing [12]. Combined with big data technology, we can develop traditional culture in a new round.

## 2. Big data collection of village information

### 2.1. Overview of big data in traditional villages

The investigation and research of traditional villages in China started late, while big data, as a new concept emerging in recent years, and has few practical case studies in the field of village protection. Although there are some researches based on big data perspective in China, most of them stay at the stage of theoretical research. At the 2014 “national ancient village experience exchange conference and the fourth China ancient village protection and development seminar”, experts at the conference put forward new ideas on the protection and inheritance of ancient villages in the era of big data [13]. This paper discusses how the cloud service platform of ancient villages in the era of big data makes the protection of ancient villages alive, and expounds the significance of the protection of ancient villages; some scholars put forward the role of big data in the protection of ancient villages, and how to face the old villages with big data in the protection and inheritance. Taking the Pearl River Delta as an example, the collection and processing methods of planning data in the process of village planning are analyzed. From the perspective of thinking methods and technology application, big data provides a better resource and data base for village planning.

At present, there are a lot of big data research in the field of urban planning in China, most of which are the application of urban historical blocks. The main reason is that the way of getting big data is more diversified and comprehensive, the city has a high degree of population aggregation, large database and rich database.

The protection and development of ancient villages in China take rural tourism as an opportunity. More and more people begin to realize the impor-

tance and protection significance of ancient villages. Since 2000, urban planning scholars have taken more and more comprehensive factors into account in the study of village protection and development. From economics, architecture, geography and other disciplines, they have conducted a comprehensive and in-depth analysis and Research on the protection methods, protection values and development and utilization of traditional villages

## 2.2. The specific application of network data in the protection and renewal of traditional villages

Through the source of big data, big data can be divided into traditional data and network data. The collection and processing of traditional data mainly includes field survey, interview questionnaire, shooting records and other methods. Finally, collect and implement village topographic map, map and other methods to become spatial planning data. Traditional data is commonly used in village protection and updating. This section mainly introduces the application of network data in the protection and update of traditional villages.

The network data selected by the research institute are mainly from Ctrip, Dianping and Sina Weibo. Specific crawling data can be divided into POI data and text data.

Interest point data is usually converted into interest point data, also known as information point data. Scenic spots, government agencies, companies, shopping malls and restaurants on the electronic map are POI data.

Text data refers to data that cannot participate in the operation, such as English letters, Chinese characters, etc. The text data of this crawl mainly comes from reviews, Ctrip and Sina Weibo.

The captured network data is the text data reflecting the evaluation of a village in Huizhou and the POI data with spatial geographical coordinates.

The POI data collected in this paper is mainly the spatial point data and popular comments of merchants and the login point data of tourists in Sina Weibo. Famous tourist attractions are regarded as. The spatial point of interest data of tourists and enterprises can be used. The spatial distribution and thermal concentration of human social activities are analyzed to provide data basis for planning designers.

POI data is usually represented as point data. Point data generally refers to the data set with geographic coordinate information. The process is as follows:

```

/// <summary>
///
/// </summary>
/// <param name="objStr"></param>
/// <returns></returns>
public static IEnumerable<string> GetArticleKeywords(string objStr)
{
    var idf = new TfIdfExtractor();
    return idf.ExtractTags(objStr, 10, Constants
}

```

Fig. 1. Text segmentation code.

a large number of coordinate information data sets contained in the web page can be collected, public comments and micro blog data can be collected, part of vacancy information and useless information can be eliminated, and the coordinate data sets with latitude as latitude and longitude as LNG can be obtained by writing rules through the locomotive collector or using Python to write connection code.

The text materials collected in this paper are comments of Longchuan village residents and travel notes of Ctrip. By writing the collection rules of locomotive, the text data is obtained and processed by using the function of jiebar loading package in Python platform.

First, load the jiebar language toolkit on the python platform. Manually load the dictionary by loading it in the word segmentation toolkit.

The second step is participle. Participle is to divide a sentence into recognizable phrases. After using readable to write the collected travel notes to the R language platform, we use Jieba participle (Fig. 1) to segment the comments and travel notes, and sort the phrases from high to low.

The last step is to eliminate the related useless words. Words that we, ourselves, and others are more extensive or have no obvious reference can be removed. The frequency of the remaining words is calculated. This is the basis of subsequent analysis.

POI data appears on the spatial graph in the form of points, which is represented by the aggregation of data object points in the space. In this study, POI data were analyzed by thermal analysis and density analysis in spatial analysis.

Thermal analysis is a simple spatial clustering analysis. The specific operation process is to use the powermap package loading function of Excel. The preprocessed POI data with geographic location information is imported into powermap in the form of table, and the spatial aggregation degree of POI point data on the map is represented by color through plug-in electronic map. Divide the latitude and longi-

Table 1  
Statistical data

	Latitude	Longitude
1	30.2626262	118.36669
2	29.9999999	118.56321
3	29.90441	118.5823
4	29.90435	118.6495
5	29.926703	118.6495
6	29.926888	118.287331
7	29.926618	118.289072
8	29.923785	118.285834
9	29.921802	118.286106
10	29.922971	118.286046
11	29.923046	118.285468
12	29.923046	118.286239
13	29.922797	118.28626

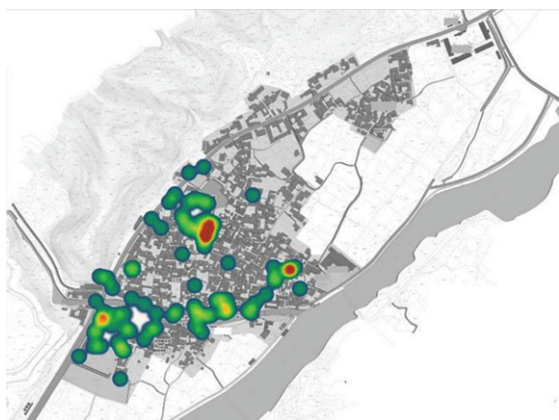


Fig. 1. Thermal agglomeration of POI data in power map.

tude data (Table 1) in POI data into two columns. All data is selected and the plug-in powermap is started. POI data will appear on the map as points. Select “Heatmap” and you can see the aggregation of POI data in space (Fig. 2). In this study, micro blog sign in data can also use thermal analysis method to identify the hot spots of urban population aggregation. The coincidence of several points on the map shows that the population concentration in this area is relatively high, while the population concentration is relatively low.

Density analysis is a common analysis method in GIS space. The principle of point density analysis is to define an area around each point data and add the number of points of interest in the area. Then the data density is obtained by taking the number of points in the range as the numerator and the area in the range as the denominator. In this study, the village area of Longchuan administrative village is the denominator, and the POI data is the numerator. Point density is

analyzed. The spatial aggregation degree of a certain kind of target can be known. Point density is analyzed, and the distribution of commercial and service facilities in urban space can be obtained.

Word frequency analysis refers to the frequency of keyword phrases in crawling text data to perceive the public’s cognition of the region. By using word segmentation technology to eliminate high-frequency meaningless phrases, the word frequency with comment tendency (good scenery, good water quality, etc.) is calculated. The text data used in this paper are Ctrip’s travel notes and comments on the public comment network. Through the analysis of high-frequency vocabulary and low-frequency vocabulary of public cognition, we can know the satisfaction and cognition degree of the masses to the village.

Co-occurrence analysis is to select representative words on R platform, set phrase vector, and analyze the most closely related words. Different from word frequency analysis, co-occurrence analysis clusters some high-frequency words to observe the relevance of scenic representative words, so as to understand the image of the village in the public.

This chapter mainly introduces the role of big data in village protection and update, focusing on the role of network data in traditional village protection and update: providing a variety of data; assisting public participation; assisting social research. Then it introduces the POI data and text data that need to be collected in this study, and describes the specific process of data collection, processing and analysis in detail, so as to prepare for the next step of analysis.

### 3. The way to improve the development and utilization level of Huizhou traditional village cultural resources in the era of big data

Historical and cultural cities have been strengthened. The cultural resources excavation and cultural ecology of historical and cultural blocks and ethnic towns are protected as a whole. Excellent traditional culture is inherited and carried forward. The development of local characteristic culture is promoted. The cultural memory of the city is preserved. Whether from the perspective of protecting the existing cultural heritage of traditional villages, or from the perspective of mining the contemporary value of traditional cultural heritage. In the context of big data, digital technology support and platform development to break through the development and utilization

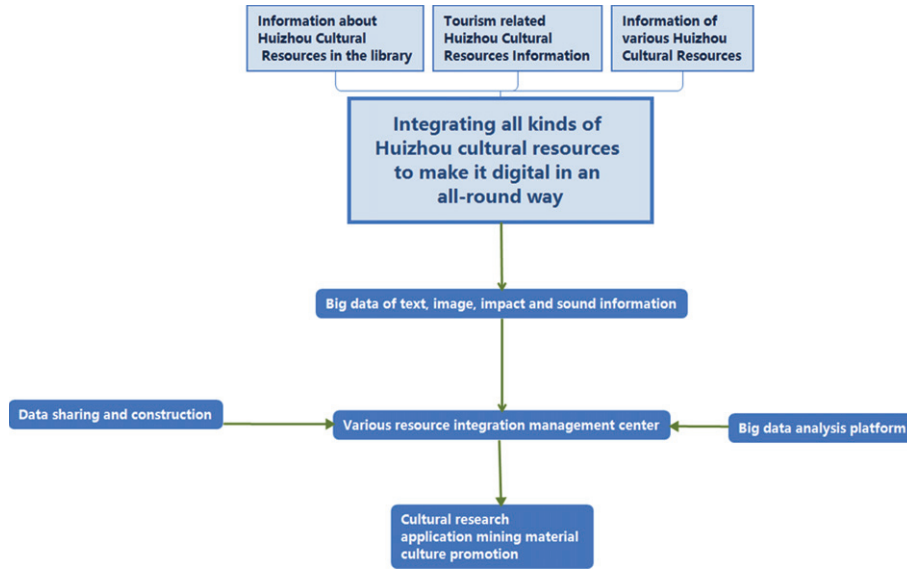


Fig. 2. Big data applied to Huizhou cultural resources framework.

level of traditional village cultural resources should be fully relied on [13]. Huizhou traditional villages are the precious heritage of Huizhou history and culture. In the era of big data, the concept of openness, equality, cooperation and sharing is adhered to. The development and utilization level of Huizhou traditional village cultural resources can be comprehensively improved from the following aspects:

### 3.1. Using big data thinking to open digital protection and inheritance work

The traditional village culture of Huizhou is rich in human concept and harmonious culture, which has good practical value. In the era of big data, through the Internet+technology platform, a social public cultural service cloud platform with the theme of Huizhou traditional village culture, which covers all aspects and is accessed by multiple terminals, is built. All the cultural resources of traditional villages in Huizhou are put into the digital network in various ways, placed on the cloud platform, open to the society, and let the traditional village culture come out from the historical depth and regional scope. Let the public really know and understand, and can enjoy the service of cultural resources of Huizhou traditional villages at any time, appreciate the traditional culture of Huizhou, participate in the interaction of traditional culture of Huizhou villages, so as to truly protect and inherit the traditional villages of Huizhou. Through the cloud

platform, we actively explore the “smart” service transformation path to realize the cultural resources of traditional villages in Huizhou, to provide digital cultural services for the public, and to provide content resources services for more cultural and art creators.

The information of Huizhou cultural resources can be obtained from different ways. Figure 2 shows the basic framework of applying big data to Huizhou cultural resources. The information of Huizhou cultural resources can be collected from resource sharing, and use big data analysis to better mine and apply the information of Huizhou cultural resources. At the same time, mining and summarizing all the relevant information of Huizhou cultural resources, transmitting it to the data management center, integrating the data resources, then reasonably distributing the digital resources according to the characteristics of Huizhou cultural resources, and finally applying the data of Huizhou cultural resources to the research of culture and art field, promotion, inheritance and other applications according to different needs.

In the construction of Huizhou cultural resources data management platform, first of all, it is necessary to summarize all kinds of Huizhou cultural resources data and resources, including books, collections, intangible cultural heritage materials and various kinds of artificial recommendations and other information, then through the digital transformation platform, it is transformed into data, and unified scheduling through the data management platform. Intangible cultural resource management platform

includes data co construction, sharing, data representation and modeling, data mining, data analysis, data collection, data sorting, data classification and cloud platform and other related functions and sub modules. After a series of data collation, using the designed mobile Internet software, the platform information can be an important data for cultural research according to the relevant content of customers' mobile search. It can also be promoted as graphics, images, news, etc. to increase people's sense of national identity, enhance cultural self-confidence, so that everyone can become a contributor to cultural heritage. Figure 3 shows the basic framework of intangible cultural resources data management platform

Using big data technology for application and decision-making, data analysis is the top priority. It can better understand people's demand for Huizhou cultural resources, so as to better apply and inherit valuable document resources. Data analysis platform can be roughly divided into data integration, data storage, data calculation, data analysis and result display and interpretation.

Data integration is an important part of data development, which can collect many resources to make it data, including relational database data extraction, real-time data collection, file data collection, database real-time replication, etc. The collected data should be stored to make it easier to extract. Data storage includes data mart and data warehouse. After data storage, data can be retrieved at any time for calculation. Calculation data is the key of the whole big data application, which includes batch calculation, flow calculation, content calculation and query calculation. These calculations can reasonably push Huizhou cultural resource information according to user demand or market demand. To promote the rapid development of intangible cultural undertakings, data mining is particularly important. After data calculation, reasonable analysis is required, that is, descriptive analysis, predictive analysis and in-depth data analysis are carried out through data mining algorithms. These effective data calculations are the premise of displaying results, decision support and other mining applications. Finally, it is necessary to display the results of computation and mining in mobile or multimedia. The forms of display are various, which can be summarized as visual display and result interpretation. Data analysis and its platform construction are the difficulties and key points of the whole system, and the most important part of mining intangible cultural heritage resources. Figure 4

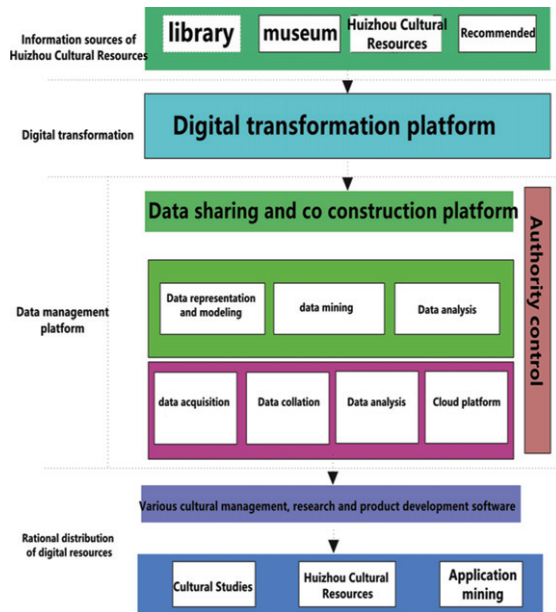


Fig. 3. Basic framework of data management platform.

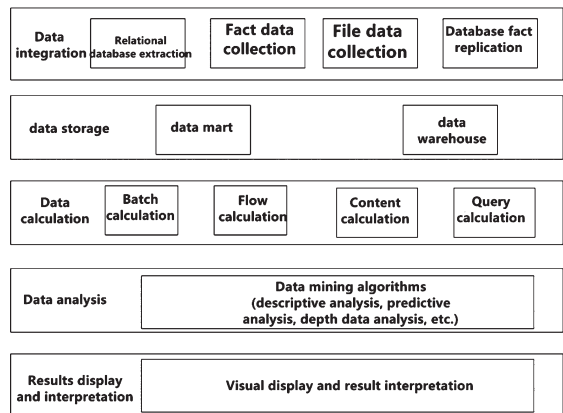


Fig. 4. Basic framework of data analysis platform.

is the basic frame diagram of data analysis platform construction.

Using big data technology to mine Huizhou cultural resources, first of all, through the collection and management of the original data to prepare for the subsequent calculation, after the data preparation, the data is processed to make it data-based, mining its potential applications. Through the operation mode of platform design, the data of Huizhou cultural resources are interpreted to form systematic knowledge. Through the original knowledge and characteristics, it can be better used. Figure 5 shows the specific implementation process of Huizhou cul-



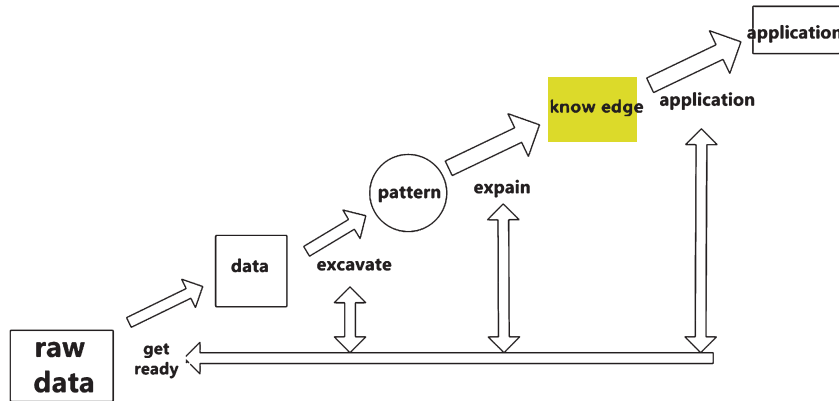


Fig. 5. The implementation process of cultural resources mining in Huizhou.

tural resources mining, in which the links of mining, interpretation, application and preparation are interconnected. Therefore, some data in the preparation stage can be directly applied to other links as required. If some data of other links are not satisfied with the processing, they can be reprocessed and applied again. The integrated process can make the data more valuable and more effective.

It is a systematic project to apply big data technology to the utilization of cultural resources in Huizhou. It is a multidisciplinary research and innovation in culture, art, informatics, communication and other fields. Using data management and analysis platform to protect, inherit and utilize cultural resources, new ideas and means are developed in the development of information technology and cultural communication innovation. According to the design concept of this paper, we can find out more applications of Huizhou cultural resources, and provide more effective technical support and decision support for the research of “digital cultural integration”.

#### 4. Experiment and results

The methods, results and efficiency of data analysis are very important, which are related to the accuracy of the system and the effectiveness of the application value.

In the process of data collection of Huizhou cultural resources, the sources of information can be divided into many different ways, including information collected and disseminated in traditional ways, such as intangible cultural heritage resources or areas where inheritors are located, on-site archives

such as cultural centers, libraries or museums, information released by traditional media such as television broadcasting, newspapers and magazines. There are also information that is transmitted and preserved by new media, such as web pages, electronic databases, Huizhou cultural resources website, micro-blog, WeChat official account, forum, social platform, etc. These information have different characteristics of isomorphism and isomerism, as well as different attributes of structured, unstructured or semi-structured, which need to be processed by different data processing methods. In addition, there are many kinds of cultural resources in Huizhou, each of which has its own unique characteristics and key words. In addition to the relevant information found through the above ways, we should also find historical books or related collection materials. Collect the historical background, development genes, national characteristics, etc. related to intangible cultural heritage resources, and collect more keywords, including dynasties, years, relevant figures, roles, classifications and other surrounding words, so as to improve the integrity of data and ensure the effectiveness and greater efficiency of subsequent data processing and analysis.

Remember that the collection of all Huizhou cultural resources we want to represent is  $D = \{d_1, d_2, d_3, \dots, d_N\}$ , and the set of words in all Huizhou cultural resources (also known as the dictionary) is  $T = \{t_1, t_2, t_3, \dots, t_N\}$ .

Item  $j$  of Huizhou cultural resources is expressed as  $d_j = \{w_{1j}, w_{2j}, w_{3j}, \dots, w_{nj}\}$ , where  $w_{1j}$  represents the weight of the first word  $t_1$  in Huizhou Cultural Resources  $j$ , and the larger the value is, the more important it is; other vectors are interpreted

similarly. In order to represent the  $j$  item of Huizhou Cultural Resources  $d_j$ , the key is to calculate the value of each component in the  $d_j$ . If the word  $t_1$  appears in item  $j$  of Huizhou cultural resources, the value of  $w_{1j}$  is 1; if  $t_1$  does not appear in item  $j$  of Huizhou cultural resources, the value of  $w_{1j}$  is 0, and the number of times that the word  $t_1$  appears in item  $j$  of Huizhou cultural resources can also be selected as  $w_{1j}$ :

$$TF - IDF(t_k, d_j) = TF(t_k, d_j) \cdot \log \frac{N}{n_k} \quad (1)$$

Among them,  $TF(t_k, d_j)$  is the number of occurrences of the  $k$  word in Huizhou Cultural Resources  $j$ , and  $n_k$  is the number of articles including the  $k$  word in all Huizhou cultural resources. Finally, the weight of the  $k$  word in the item Huizhou Cultural Resources  $j$  is obtained by the following formula:

$$w_{k,j} = \frac{TF - IDF(t_k, d_j)}{\sqrt{\sum_{s=1}^{|T|} TF - IDF(t_k, d_j)^2}} \quad (2)$$

Let  $N(u)$  be the set of cultural resources that user  $u$  likes, and  $N(v)$  be the set of resources that user  $v$  likes. To calculate the similarity between  $u$  and  $v$ , use Jaccard formula:

$$w_{k,j} = \frac{TF - IDF(t_k, d_j)}{\sqrt{\sum_{s=1}^{|T|} TF - IDF(t_k, d_j)^2}} \quad (3)$$

Cosine similarity:

$$w_{uv} = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|} \quad (4)$$

Suppose there are four users: A, B, C, D; there are five items: a,b,c,d, e. The relationship between users and items (users like items) is as follows:

$$w_{uv} = \frac{|N(u) \cap N(v)|}{\sqrt{|N(u)| \times |N(v)|}} \quad (5)$$

Data processing based on data collection, data analysis and other technologies is the key technology research of big data technology applied in Huizhou cultural resources mining.

In this paper, data simulation is used to carry out the experiment of cultural resources mining in Huizhou. Through data collection, analysis, modeling, calculation and other processes, the final conclusion is reached.

First of all, the key words should be collected. This paper takes the clan, Huizhou opera and silk as examples. Through all kinds of books, film and television works, museums, web information and cultural resources directory, crawling and collecting

Table 2  
Big data analysis results

Probability term	Keywords					
	History	Character	Story	Effect	Works	Total
Patriarchal clan	0.36	0.40	0.21	0.23	0.58	0.20
Anhui Opera	0.34	0.50	0.46	0.32	0.46	0.32
Silk	0.30	0.33	0.32	0.56	0.68	0.48

Table 3  
Traditional data analysis results

Probability term	Keywords					
	History	Character	Story	Effect	Works	Total
Patriarchal clan	0.26	0.35	0.20	0.19	0.35	0.23
Anhui Opera	0.24	0.45	0.35	0.22	0.30	0.44
Silk	0.26	0.34	0.29	0.45	0.54	0.49

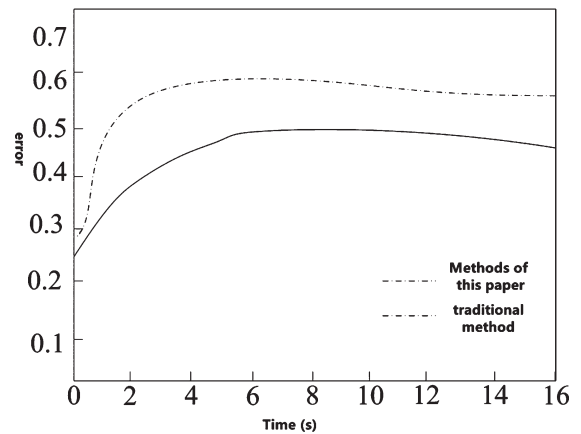


Fig. 6. Analysis results.

their keywords, using the existing digital conversion technology, data conversion and storage. In the data simulation, 100 volunteers were randomly selected to participate in the experiment. The data analysis formula proposed in this paper was used for calculation and comparison with the traditional data analysis. Finally, the results of Tables 2 and 3 are obtained by calculation.

According to Tables 2 and 3, the analysis results are obtained by binary semantic fitting, as shown in Fig. 6.

The data analysis method proposed in this paper is used and the frequency of searching specific works in clan culture is 0.58. The frequency of searching characters is 0.50, and the frequency of searching specific works in silk is 0.68. Compared with traditional data processing methods, the error is large and the accuracy is low.



## 5. Conclusion

Huizhou traditional villages are the embodiment of traditional culture and the sustenance of nostalgia. During the period of COVID-19, the epidemic prevention work in rural areas was carried out in a jealous way. From the perspective of big data, actively promote the construction of traditional village level big data platform. The content innovation of cultural resources has been deepened. In the protection phase of COVID-19, data resources are fully utilized. The formation and development of the new industrial chain is promoted, and the development path of Huizhou traditional villages is explored with the concept of open sharing. The continuation of contemporary emotion is the best way for the public to “look at mountains, water and nostalgia”, and it is also the permanent connection of traditional culture.

## References

- [1] M. Chen, S. Mao and Y. Liu, Big data: a survey, *Mobile Networks & Applications* **19**(2) (2014), 171–209.
- [2] Lynch, Clifford, Big data: How do your data grow? *Nature* **455**(7209) (2008), 28–29.
- [3] I.A.T. Hashem, I. Yaqoob, N.B. Anuar et al. The rise of ‘big data’ on cloud computing: review and open research issues, *Information Systems* **47** (2015), 98–115.
- [4] H.V. Jagadish, J. Gehrke, A. Labrinidis et al. Big data and its technical challenges, *Communications of the Acm* **57**(7) (2014), 86,88–94.
- [5] Madden, Sam, From databases to big data, *IEEE Internet Computing* **16**(3) (2012), 4–6.
- [6] A. Jacobs, The pathologies of big data, *Communications of the ACM* **52**(8) (2009), 36,38–44.
- [7] A. Aisling O’Driscoll, B. Jurate Daugelaite, D. Roy and B. Sleator, ‘Big data’, Hadoop and cloud computing in genomics, *Journal of Biomedical Informatics* **46**(5) (2013), 774–781.
- [8] Wang, Lei, Zhan, Jianfeng, Luo, Chunjie et al., Big-DataBench: a big data benchmark suite from internet services, *IEEE* (2014), 488–499.
- [9] N.V. Chawla and D.A. Davis, Bringing big data to personalized healthcare: a patient-centered framework, *Journal of General Internal Medicine* **28**(suppl 3)(S3) (2013), S660–S665.
- [10] M. Zaharia, R.S. Xin, P. Wendell, et al. Apache spark: a unified engine for big data processing, *Communications of the Acm* **59**(11) (2016), 56–65.
- [11] V. Cevher, S. Becker and M. Schmidt, Convex optimization for big data: scalable, randomized, and parallel algorithms for big data analytics. *Signal processing magazine, IEEE* **31**(5) (2014), 32–43.
- [12] Gang-Hoon Kim, Silvana Trimi and Ji-Hyong Chung, Big-data applications in the government sector, *Communications of the Acm* **57**(3) (2014), 78–85.
- [13] L. Einav and J. Levin, Economics in the age of big data, *Science* **346**(6210) (2014), 715–715.