# Research on Protection and Development Planning of Traditional Villages Based on Tourism Big Data Analysis

**Junkai Wang**
*Zhengzhou University (New Campus), Zhengzhou, 450001, Henan, China*

***Abstract****: Traditional village refers to a village which was formed earlier, has rich traditional resources, has certain historical, cultural, scientific, artistic, social and economic values, and should be protected. Combined with the related technologies of artificial intelligence and tourism big data, this paper discusses the construction of intelligent tourism analysis model. Dynamic streaming data processing in the model collects and updates data from social networks in real time. Then, combined with the machine learning algorithm in the field of artificial intelligence, a novel user-centered recommendation strategy is proposed to achieve the purpose of recommending suitable tourist attractions for users. Using users' browsing traces, this paper analyzes users' sources, browsing interests and preferences, and provides reliable data basis for traditional village protection and development planning.*

***Keywords****: Tourism big data, Traditional villages, Protection, Development planning*

## Ⅰ. Introduction

Urbanization and new rural construction have brought about social progress and economic development, but these traditional villages with a long history have been annihilated in the torrent of history. We don't know anything about the culture contained in these disappearing villages. While accelerating the socio-economic development, we didn't protect our traditional civilization, history and culture. Nowadays, how to effectively protect these precious historical heritages and how to deal with the relationship between the protection and development of traditional villages have become the core of our thinking.

Nowadays, with the development of social information technology, people's social activities in neighborhoods, villages and cities will generate a large number of data, which include information on tourism, economy, transportation, food and culture. The amount of data is growing at a fast speed, and it is time-sensitive and dynamic. These data not only arouse the academic circles' thinking and application of data, but also increase the cognition and concern of countries for data. In smart tourism services, it technologies such as Internet of things [1], artificial intelligence, big data, wireless sensor networks and short-range communication have been used to provide tourists with accurate travel information and rich travel experience [2-3]. In order to further improve the planning level of traditional village protection and development, combining artificial intelligence and big data technology, this paper discusses the construction of smart tourism big data analysis model and designs efficient recommendation strategies.

## Ⅱ. Traditional Villages and Tourism Development

Ancient villages used to be used to refer to some villages with relevant historical heritage. These villages still have rich traditional cultural forms and cultural accumulation after the changes of the times, and the historical context is relatively intact. Historic village is a legal concept, which was passed by legislation in China and given legal protection. Traditional villages are derived from ancient villages, but they have more emphasis in their connotation-cultural value.

Traditional villages include three cultural connotations [4]:

(1)The existing traditional architecture is complete.

(2)Village location and pattern keep traditional characteristics.

(3)Intangible cultural heritage live transmission.

Tourism development has both opportunities and challenges for traditional villages. In recent years, people's living standards have been continuously improved, and the continuous development of tourism in various places has enabled us to enter the era of national tourism. However, the reinforced concrete in cities has restrained our longing for nature, and the countryside has become the place we yearn for. Traditional villages have attracted a large number of tourists because of their unique landscapes and profound cultural heritage. Tourism development has also become the main way to protect the living state of traditional villages and a new driving force for traditional development.

The village itself is an important tourism resource, so the protection and inheritance of the village pattern, historical buildings and customs and culture is the prerequisite for realizing the sustainable development of traditional village tourism. Tourism development without protection will only lead to over-exploitation and over-commercialization of traditional villages, and tourism development based on protection will be sustainable and operable.

## Ⅲ. Tourism Big Data Analysis of Traditional Village Protection

### *A. Smart Tourism Big Data Analysis Model*

Wisdom is the premise of intelligent tourism, which cannot be separated from high-tech support. By using big data, artificial intelligence and other technologies and hardware, the real-time and dynamic collection of tourism industry information and tourist behavior information is realized, the personalized needs of users are integrated and analyzed, and various tourism schemes suitable for tourists are designed [5]. From the collection of tourism data to the analysis of tourists' needs, to data integration, data cleaning, data analysis, deep learning and design of appropriate tourism schemes, the whole process involves big data, artificial intelligence and other related technologies.

The purpose of tourism is to provide tourists with a more comfortable tourism experience, more satisfying their own tourism products and more convenient supporting facilities. In order to make tourism more intelligent and meet the individual needs of tourists, based on big data, artificial intelligence and other technologies, a smart tourism big data analysis model based on artificial intelligence as shown in Figure 1 is constructed.

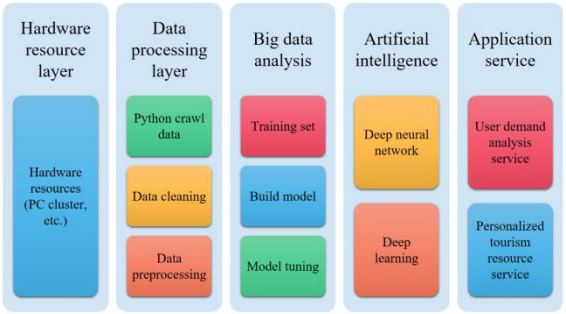


*Fig.1 Big Data Analysis Model of Smart Tourism Based on Artificial Intelligence*

### *B. Data Model*

The model contains two main entities: user $U$ and tourist attraction $S$. Tourist attractions can be specific archaeological sites, sculptures or pictures displayed in museums, historical buildings and famous squares in downtown areas, etc. In addition, a tourist attraction can be associated with a specific PoI, which is defined by a set of geographical coordinates.

For the tourist attraction $s_i \in S$, its annotation can be expressed by tuples:

$$\xi_{s_i} = \left(a_1, a_2, \cdots a_N; b_1, b_2, \cdots b_M\right) (1)$$

In which $a_i, b_i$ is ontology attribute and metadata respectively. One or more attachments of a tourist attraction (e.g., text related to posts or multimedia data of published images) can be associated with a given tourist attraction.

Regarding users, two different types of information are processed: static information is recorded in the user's configuration file, while dynamic information is recorded in the user's data log. For user $u_i$ , its user profile can be represented by tuples:

$$\ell_{u_i} = \left(x_1, x_2, \cdots, x_N; y_1, y_2, \cdots, y_M\right) (2)$$

In which $x_i, y_i$ represents the ontology attribute of the user and the metadata of the configuration file respectively.

Given a tourist attraction set $S$ and a user set $U$ , the user data log can be represented by a tuple set:

$$\psi = \left(t, u, s, \overline{a}_1, \cdots \overline{a}_N\right) (3)$$

In which $t, u, s, \overline{a}_i$ respectively represent time stamps, users, tourist attractions and attributes used to describe the user's behavior towards specific objects.

The above data model constitutes a knowledge base, in which users' information can be analyzed to infer what tourist attractions a specific user may be interested in. The knowledge base can be represented by graph $\left(V, E\right)$, where $V$ is a node set, and nodes can be users, tourist attractions, labels, etc. $E$ is a collection of edges, which represent the relationship between nodes.

Because modern computers can't intelligently recognize human's natural language, and can't directly process unstructured data such as text, the preprocessed text data needs to be transformed into a structured form, so that the computer can “read” and then recognize and process it. This transformation process is the formal representation of text. Words, phrases, phrases, etc. are all commonly used text forms, and common text form representation models include Boolean model, probability model and vector space model [6]. Among them, the text representation of vector space model is better, and it is also a frequently used text representation method.

The vector space model can be described as: given set $T\left\{t_1, t_2, \cdots, t_n\right\}$ is $m$ features appearing in the text, let $w_i$ represent the weight of the $i$ th feature in the text $D$ , and $D$ can be expressed as:

$$D = \left\{t_1, w_1; \cdots, t_2, w_2; \cdots; t_n, w_n\right\} (4)$$

Among them, the weight can be calculated by using Boolean weight method, word frequency weight method or $TF - IDF$ weight method.

***C. Feature Weighting***

After transforming the tourist review text into vector space model, we can get a sparse matrix, which is usually a high-dimensional space. In the process of machine learning, the excessively high feature space dimension may cause “dimension disaster” [7-8]. It is necessary to weight the text features, which plays a very important role in the classification results. The process of feature weighting is the process of giving different weights according to the

contribution of each feature to the classification results. Commonly used feature weighting methods are Boolean weighting, word frequency weighting and $TF-IDF$ weighting.

$TF-IDF$ is a widely used feature weight calculation method. Its main idea is that the most effective feature words for document classification should be those words that appear frequently in one document but rarely in other documents. The calculation formula can be expressed as follows:

$$W(i,j)=tf(i,j)\times\log\left(\frac{N}{n_i}\right) \quad (5)$$

Where, $tf(i,j)$ is the number of times the feature item $t_i$ appears in the text $d_j$, $n_i$ is the number of texts containing $t_i$ features, and $N$ is the total number of documents.

### *D. Experiment and Result Analysis*

The corpus used in the experiment is the commonly used travel emotional vocabulary of online reviews, and 1500 reviews from 5 scenic spots are selected as corpus, and the emotional tendency of these reviews is manually marked.

The experimental environment is visual studio 2008 and SQL server 2005, and the LIBSVM toolbox is used to realize the SVM machine learning model [9]. In this experiment, we used the evaluation indexes of accuracy rate, recall rate and F-means to test the experimental effect. The experimental results are shown in Figure 2.

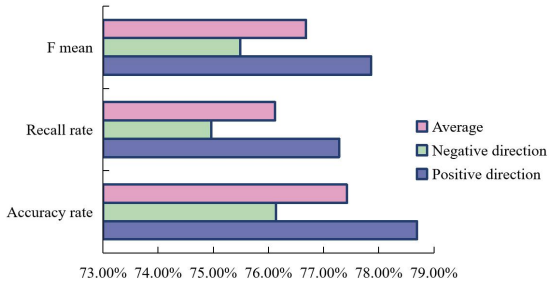


*Fig.2 Experimental Result*

The research shows that the correct rate of emotion classification experiment based on SVM model can be as high as 80% [10]. From the experimental results, the method designed in this paper takes the tourism emotion dictionary as the feature, uses $TF-IDF$ to calculate the feature weight, and uses SVM machine learning model to classify the emotion tendency. Although the correct rate does not reach a very high value, the method is simple, efficient and easy to realize, and the values of the three indexes are still satisfactory.

## 4TRADITIONAL VILLAGE PROTECTION AND DEVELOPMENT PLANNING

### *E. Traditional Village Database Construction*

Traditional village database system is implemented by three layers: data layer, service layer and former client display layer (Figure 3).

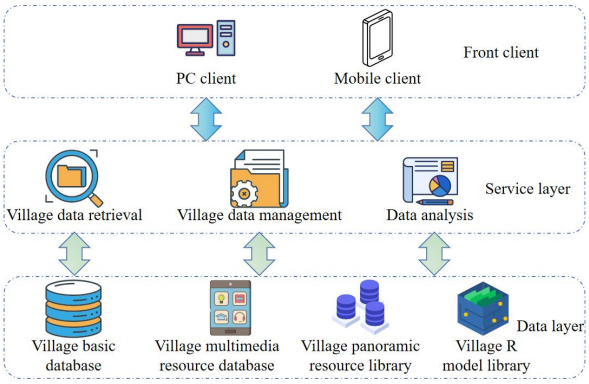


*Fig.3 Technical Architecture of Traditional Village Database System*

(1)Data layer

The data layer provides data organization and storage functions, and realizes the construction of digital resources in traditional villages, such as non-legacy buildings and ancient buildings, involving construction subjects, protection and utilization systems, information resource organization, retrieval and long-term preservation, and provides data service functions for the service layer.

According to different data types, the data layer provides four sub-database systems, namely, village basic database, village multimedia resource library, village panoramic resource library and village VR model library. The village basic database provides village basic data, such as village introduction, village location, village site selection and other data storage and organization, and most of the data exist in structured and semi-structured forms.

(2)Service layer

Service layer provides data management, data retrieval and big data analysis. The village data management service provides the functions of warehousing, modifying and deleting village data. The data analysis service provides data analysis function, and according to the user's visit records, it can complete the hot spot calibration of system visit, the analysis of users' concerns about village data and the hot spot analysis of traditional village visit, which is convenient for tourism management departments and tourism enterprises to customize, accurately publicize and market tourism products.

(3)Former client

The former client interface is a user-oriented resource outlet and an interface for database to realize retrieval service. According to the resource structure design of the data layer, the functions of searching within columns and across columns are realized by setting the first-level directory and the corresponding second-level and third-level directories in the former client user interface. In the current era dominated by mobile Internet, the database system must realize the mobile former client besides the traditional PC former client.

***F. Renewal and Utilization of Traditional Buildings***

Traditional villages contain a large number of traditional dwellings, and even some villages are declared with traditional architectural communities as objects of cultural preservation, which shows the significance of traditional architecture. Traditional buildings can be renewed and utilized, and existing buildings can be fully utilized to reduce the impact of new buildings on village texture and style.

Renovation of buildings is mainly carried out from three aspects: structure, internal function and external style. Traditional dwellings take a long time to be built, so it is necessary to properly reinforce and replace their structures to increase their safety, mainly focusing on the load-bearing components such as pillars and vertebrae, and the necessary ones can be replaced by local government organizations. The internal function mainly depends on the villagers' subsequent use requirements. Villagers can choose the mode according to their own situation, and each mode is given a transformation template, which selects a typical residential building in the village for demonstration. The exterior style part is mainly based on architectural guidelines, and does not conflict with the traditional village style when updating. The general principle is to emphasize naturalness, not to seek for good scale, not to seek for high standards, but also to pay attention to human touch and simplicity, and not to pursue novel materials.

On the basis of GIS, a comprehensive survey of various cultural resources is carried out and the survey results are input into the computer, forming an exclusive database of traditional villages. A service system including query and retrieval, graphic and image display, statistical evaluation and 3D visualization is finally built from different databases. Users can query various details of traditional villages through the Internet.

Under the digital background, the network technology is very developed, and the public open platform can be the website of government official information disclosure, the APP containing scenic spots and multifunctional services, or even the 3D map reflected on the computer or mobile phone by using remote sensing image technology, and the virtual traditional village built on the Internet and APP by using GPS and 3S technology. Netizens and tourists can click on different scenic spots to enter the simulation interface and experience 3D effects, just like being in a traditional village. Visitors and users can learn about the culture, activities, business and other information of traditional villages in real time, while managers can publicize the information of traditional villages in real time through APP. Users can also learn about the buildings to be repaired in time through their feedback.

### *G. Village Development Planning*

The continuation of the life of traditional buildings in villages is the best means to protect the historical context. On the premise of maintaining the original layout and form, we should try our best to give new life to traditional buildings in villages by reusing as much as possible, that is, in terms of functions. Planning and development should be organically combined with conservation and inheritance, and development land should be reserved for the long-term development of villages.

Protect all the streets and lanes in the core protected area and dredge them in the local silted areas, so as to make them an effective and continuous road system. All streets and lanes are not allowed to be rebuilt arbitrarily for commercial and other purposes. For newly built buildings, the specific data of color, height, etc. must be controlled to keep the original street style. Keep the twists and turns of streets and roadways, protect the nodes such as ancient trees and squares in streets and roadways, and keep their overall landscape features.

Combined with the renovation of traditional village buildings and the adjustment of land use, the planning will increase the land for social living facilities and greening for residents. Mainly divided into public facilities, living facilities and production facilities:(1) public facilities: keep the function of the cultural activity room as a cultural exhibition; The damaged buildings and idle land on both sides of the main roadway in the village will be transformed into public green spaces or squares; Improve the sanitation facilities, infrastructure and fire control facilities in the village. (2) Living facilities: Restore the original public stadium in the east of the village as a square for the villagers to rest. (3) Production facilities: Maintain existing irrigation facilities. Maintain existing agricultural roads. Add high-level pool and feed water pump to meet the production and living needs of villagers.

## Ⅳ. Conclusion

The protection and renewal of traditional villages should take protection as the primary premise. From the perspective of tourism big data analysis, it can be concluded that the public cognition is insufficient in understanding the material form and other historical buildings of villages, and lacks the overall macro impression; The cognition of

village intangible culture is not comprehensive, and only stays in the traditional village food and tourism culture. Tourism development must be based on protection. The principle of “moderation” in commercial development is one of the most basic protection principles for traditional villages. Commercial development cannot assimilate traditional villages and avoid disorderly development. It is the best development to combine development with protection. The future research work is mainly to strengthen the cluster analysis of tourism network evaluation words, find out the common points in tourists' evaluation, provide data support for tourism enterprises to improve their products and services, and realize “science and technology help tourism”.

**References**

[1] Wang Nan, Zeng Rong. Characteristics of Protection and Development Planning of Traditional Village Relocation for Poverty Alleviation--Case of Traditional Village Protection and Development Planning for Goga Village in Tadian Town, Eshan County[J]. Yunnan architecture, 2019, 000(003):146-151.
[2] Sun Li, Sun Meiling. Study on the Protection and Development of Traditional Village Landscape--Taking Sundu Village of Luoyang City as an Example[J]. Heritage and conservation research., 2019, 004(005):22-26.
[3] ZHONG, Guoqing, LIANG, et al. The Construction Planning of Traditional Village Homestay Based on Minimal Intervention:A Case Study of Shizhai Village,Sihui,Zhaoqing,Guangdong[J]. Journal of Landscape Research, 2020, v.12(03):111-117.
[4] Zhang Hai, Lu Song. Knowledge Structure of Tourism Research on Domestic Traditional Village[J]. Journal of west anhui university., 2019, 035(002):106-111.
[5] DING, Keren. Landscape Protection Mode of Traditional Villages in Ningbo[J]. Journal of Landscape Research, 2020, v.12(03):73-77.
[6] Koerniawaty F T, Ardika I W, Budiarsa M, et al. The Indigenous' Participation in Preserving Cultural Heritage of Bena Traditional Village as a Tourist Attraction in Bajawa, Ngada Regency, East Nusa Tenggara[J]. Management: English version., 2019, 007(006):P.577-581.
[7] Kuilboer J P. Big data and innovation in tourism, travel, and hospitality: managerial approaches, techniques, and applications.[J]. Computing reviews, 2020, 61(5):175-176.
[8] Alaei A R, Becken S, Stantic B. Sentiment Analysis in Tourism: Capitalizing on Big Data[J]. Journal of travel research, 2019, 58(2):175-191.
[9] Wang C, Guo N, Lian M, et al. Differences in Regional Media Responses to China's Holistic Tourism: Big Data Analysis based on Newspaper Text[J]. IEEE Access, 2020, PP(99):1-1.
[10] Zhao Hongyan, Li Qiaoxing, Wu Junfang.. Extension Innovation Schemes and their Evaluation to Integrate the Tourism with Other Ones under Big Health in Guizhou[J]. Journal of Guangdong University of Technology., 2019, 036(003):1-7,15.