A Study on Comprehensive Industry Data System Based on Computer Exploring Big Data and Resource Integration

He Huang*
Leshan Normal University, Leshan, 614000, China
*Corresponding author: Huanghe@lsnu.edu.cn

Abstract. Sports tourism resources are divided into two types: sports tourism natural resources and sports tourism human resources. The natural resources of sports tourism are divided into landform resources and water resources. The human resources of sports tourism are divided into folk customs and events, venues and buildings. The thesis explores the advantages and disadvantages of Chinese existing sports tourism information system through the combination of big data and sports tourism. At the same time, we analyzed and established the hierarchical structure of the sports tourism information system big data platform and the logical structure of the sports tourism data warehouse, and constructed a sports tourism information data model system based on big data analysis, laying a foundation for further research on the sports tourism information system based on big data.

Keywords: sports tourism; computer big data; comprehensive data; tourism information data model.

1. Introduction
Tourism resources are all kinds of things and factors that can be attractive to tourists in nature and human society, can be developed and utilized for the tourism industry, and can produce economic, social and environmental benefits. Tourism resources include not only physical resources such as tourist destination resources and tourist sources, but also intangible resources such as culture and services. Sports tourism resources are a large category of tourism resources. Sports tourism resources are complementary and mutually beneficial between sports resources and tourism resources. It is a kind of sports tourism activities as a natural factor and social cultural factor that attracts tourists, and stimulates tourists' tourism motivation. To create economic, social, and ecological benefits for the development of the tourism industry as natural and social entities. Sports tourism resources are the conditions for the development of sports tourism and the basis for the development of sports tourism [1]. The standardized classification of sports tourism resources can provide a reliable basis for the further development and utilization of sports tourism resources.

The significance of the emergence of big data technology is not only the huge amount of data processing, but more importantly, the analysis and application of the collected data, the application of big data technology to the sports tourism industry, the construction of a sports tourism information data platform, and the use of data mining technology. The provision of valuable information can promote the rapid development of sports tourism in China.
2. Classification, development and market segmentation of sports tourism resources

2.1. Classification of sports tourism resources

Chinese sports tourism resources are very rich, and the classification methods are also different. Some scholars divide tourism resources into permanent tourism resources and consumable tourism resources according to the way they exist in time and space. According to the different forms of tourism resources, it is divided into natural tourism resources and artificial tourism resources [2]. According to different tourist motives, it can be divided into psychological and spiritual tourism resources and economic tourism resources. No matter how people classify tourism resources, with the development of the tourism industry, the field of tourism resources will continue to expand, and the tourism resource system will be an eternal development concept.

In accordance with international practice, we refer to the viewpoints of many scholars on the classification of tourism resources, and according to the nature of tourism resources, we divide sports tourism resources into two types: sports tourism natural resources and sports tourism human resources. The natural resources of sports tourism refer to the various resources that are endowed by nature, exist and are related to sports activities, and have been developed and utilized to become tourist attractions. The natural resources of sports tourism are divided into landform resources and water resources. The humanistic resources of sports tourism refer to the various resources that the sports elements of human social activities have been developed and utilized to become tourist attractions. Sports tourism human resources are divided into folk customs and events, venues and buildings, etc. (Figure 1).

![Figure 1 Classification of sports tourism resources](image)

2.2. Development and market segmentation of sports tourism resources

"Tourism resource development refers to the development of tourism as a prerequisite, market demand as the guidance, tourism resources as the core, and the focus on exerting, improving and enhancing the attractiveness of tourism resources to tourists, and organizing and planning the development of tourism resources. Economic and technological systems engineering to be utilized." Its essence is to tap the connotation of tourism resources, improve their attractiveness, and turn tourism resources into realistic tourist attractions.

There are generally two types of sports tourism market. First, according to the nature of sports tourism resources, it can be divided into: popular sports tourism resources, professional sports tourism resources, ethnic sports tourism resources, etc. Among them, popular sports tourism resources include leisure sports tourism: fishing, horse riding, diving, golfing and so on. Fitness sports tourism: climbing, dancing, basketball, bowling, table tennis, tennis, badminton, bodybuilding, fitness activities, etc. Sports spectator tour: to watch sports events of various types and scales, etc. Professional sports tourism resources include sports event tourism: professional athletes participating in various sports competitions and so on. Competitive sports tourism: all kinds of athletes or tourists with stunts participate in the tourism activities caused by sports events, such as gliding, sailing, surfing, skiing, etc. Exciting sports tourism resources include adventure sports tourism: bungee jumping, rafting, rock climbing and other adventure sports trips. Ethnic sports tourism
resources include various folk festivals, folk customs and other sports tourism activities. It should be
noted that there is no absolute boundary between these categories, and sometimes they can overlap.

The second is to classify according to the source market of sports tourism: according to the
different ages of tourists, such as middle-aged and elderly people, young people, teenagers, etc.,
different projects are developed to meet the needs of different groups of people [3]. According to the
different economic strength and consumption power of consumers, they are divided into luxury type,
standard type and economic type. The luxury type has relatively high requirements for the
development of sports tourism resources, and generally pursues sports tourism activities that are
comfortable, high-end services, strong consumption power and high-level. The standard type focuses
on simplicity and practicality in terms of equipment and equipment, does not care whether it is a
famous brand, does not require high service and environment, and does not pursue luxury
consumption. Economical equipment and equipment are simple, pursuing economic benefits.
According to the different nature of tourists' participation, it is divided into viewing type and
participation type. The viewing type is mainly sightseeing, such as watching large-scale sports events,
natural scenery and cultural landscapes. Participation types are mainly sports activities that some
sports tourists can experience firsthand. According to the different genders of sports tourists, it can be
divided into male-suitable projects and female-suitable projects.

3. Construction of Sports Tourism Information System

3.1. The overall structure of the sports tourism information system

When constructing the sports tourism information system, fully consider the shortcomings of the
existing system, and build a new type of sports tourism information system platform based on the
application of big data technology. The overall structure of the sports tourism information system is
shown in Figure 2 below:

![Figure 2](image-url)

**Figure 2** The overall structure of the sports tourism information system

3.2. Hierarchical structure of big data platform of sports tourism information system

The construction of the big data platform of the sports tourism information system integrates big data
technology, and the underlying data source is supported by the so-based platform. Data includes both
traditional structured data, semi-structured data and unstructured data [4]. However, according to the
different data types of the sports tourism information system, HADOOP and PMP engines are used to
calculate and process the sports tourism information system data and forwarded to the upper
information data processing unit. The processed sports tourism information data is provided by the
resource scheduling and service interface to the high-level, and a unified model can be obtained and
handed over to the terminal application. Figure 3 shows the hierarchical structure of the sports tourism
information big data platform.
3.3. **Logical structure of sports tourism information data warehouse**

The logical structure of the sports tourism information data warehouse is a key to data processing. It has a great impact on the final use effect of the entire system and the overall data usage efficiency [5]. The following aspects have an important impact on the construction of the sports tourism data warehouse. Figure 4 shows the logical structure diagram of the sports tourism information data warehouse.

![Figure 4: Logical structure diagram of the sports tourism information data warehouse](image-url)
3.3.1. Storage of massive data. The sports tourism information system platform based on big data technology is a large-scale data management information system, which will inevitably involve the storage of massive data about sports tourism. When designing the actual system, it is necessary to consider the adaptability of the hardware equipment and the structure of the tourism data storage, and make necessary processing and design. This includes the storage structure and specific algorithms of the sports tourism data, the disk array of the database, etc. Also consider the issue of cleaning up and maintaining the database.

3.3.2. ETL technology. ETL technology is the main technology for the realization of the data warehouse mechanism. It is a technology for extracting, converting, cleaning and loading the data resources to be integrated. The extraction process is a selection process. By selecting relevant data from sports tourism big data sources, it is necessary to use corresponding information extraction methods to finally realize the standardization and standardization of sports tourism information data resources. The data conversion process is the process of unifying different sports tourism information formats. Data cleaning technology is the extraction process of sports tourism information data to remove illegal data information [6]. Data loading is to load the cleaning result data into the data warehouse according to certain rules, and finally form the physical and logical storage structure of the sports tourism data warehouse.

3.4. Sports tourism information data model
In the sports tourism information data model, in order to describe an unstructured data, various attributes of the data object are classified into five attribute categories: basic attributes, environmental attributes, behavior attributes, characteristic attributes, Content attributes. Each attribute class includes multiple different attributes. The specific values of these attributes are different and can have different data types. As shown in Figure 5:
4. Structural analysis of influencing factors in the sports tourism system

4.1. Systematic thinking of the factors affecting sports tourism
The WSR system methodology is a system theory with Eastern philosophical speculation. The basic core of its philosophy and philosophy is to consider not only the object's aspect of the object, but also how these objects can be better used in the aspect of things when dealing with complex problems. At the same time, understanding problems, handling problems and implementing management decisions are all inseparable from the human aspect. Take WSR as a system to study complex issues systematically, completely and hierarchically [7]. It contains the basic idea: the determinants of anything are not single, they are all in an interactive system of purpose, value, instrumentality, etc., and the influencing factors on things are always in the surrounding environmental system.

Sports tourism contains many components, which are concentrated in the composition and relationship of people, things and things. It is the interaction between people (tourists) and people (staff), and people and things. It is people's sports tourism needs and tourism. The process of combining the district environment to achieve the expected goal is the result of coordination among the various elements in the tourist district. In the final analysis, sports tourism is realized by tourists. The occurrence of tourism behavior depends largely on the cognitive level, values and economic conditions of tourists. It is caused by the integration of people, finances, materials, and information. Then, investigating the influencing factors of the sports tourism system will inevitably start from two levels inside and outside the system.

4.2. The structure division of influencing factors of the sports tourism system under system thinking
As a whole, the influencing factors of the sports tourism system are composed of internal and external layers. The outer layer is the environmental impact layer, which mainly provides conditions for the occurrence of sports tourism activities and restricts the scale and benefits of internal operations. The
inner layer is the operational influence layer. It is the foundation for the normal operation of the sports tourism system. It is also the essential embodiment of the realization of the value of tourists' sports tourism. At the same time, it is the foundation for the existence and development of enterprises [8]. It is the internal motivation of sports tourism activities and is at the core. Both are different levels of the influencing factors of the sports tourism system, but they act on tourism companies and tourists at the same time, and they are interrelated, interdependent, and mutually restrictive. On the one hand, external environmental factors provide power for the demand market of the sports tourism industry, which is the prerequisite and foundation for the smooth operation of the industry. On the other hand, internal factors "re-creatively" open up the sports tourism market by acting on people's sports tourism practice, and promote the improvement and effect of external environmental factors. External environmental factors will affect internal factors to varying degrees at the same time, such as economic factors, personal value orientation, rules and regulations, etc., while acting externally, they will directly or indirectly affect the components and systems of the system in different ways. The overall structure of the internal influencing factors is shown in Figure 6.

![Figure 6](image)

**Figure 6** Structure diagram of influencing factors of sports tourism system

5. Effectiveness test of sports tourism system based on big data processing algorithm

5.1. Design of the algorithm for the recommendation of sports tourist attractions

This article defines that after tag extraction of scenic spots, the more common tags between two scenic spots, the greater their correlation. Suppose that the tag set extracted for scenic spot A is denoted as $L_A = \{l_{a1}, l_{a2}, ..., l_{an}\}$, and the tag set extracted for scenic spot B is denoted as $L_B = \{l_{b1}, l_{b2}, ..., l_{bn}\}$. $L_C$ represents their common tag set, then $L_C = \{l_{c1}, l_{c2}, ..., l_{cn}\}$. Use the text relevance calculation method to calculate the relevance between scenic spots. This article uses the commonly used cosine relevance method to calculate. Suppose the degree of association between scenic spot A and scenic spot B is $\text{sim}(A, B) \in [0, 1]$. Where $l_{ak}$ and $l_{bk}$ represent the number of common tags between scenic spot A and scenic spot B, and their correlation degree is:

$$\text{sim}(A, B) = \frac{\sum_{k=0}^{p} (l_{ak} \times l_{bk})}{\sqrt{\sum_{k=0}^{m} l_{ak}^2 \times \sum_{k=0}^{n} l_{bk}^2}}$$  \hspace{1cm} (1)

5.2. System test

During the experiment, we extracted tags from the encyclopedia of scenic spots, counted the number of each tag vocabulary, and used the cosine correlation formula to calculate the correlation between
scenic spots. Therefore, by calculating the degree of relevance, the degree of similarity between scenic spots is 0.147. Calculating all 3,205 scenic spots, forming about 514 million pairs of correlation values, using the big data results to analyze, Table 1 lists some of the experimental results.

### Table 1 Some experimental results of scenic spot relevance

| A   | B   | Correlation | A   | B   | Correlation | A   | B   | Correlation |
|-----|-----|-------------|-----|-----|-------------|-----|-----|-------------|
| 251 | 450 | 0.996       | 199 | 446 | 0.424       | 342 | 429 | 0.36        |
| 34  | 188 | 0.967       | 243 | 267 | 0.423       | 7   | 243 | 0.36        |
| 14  | 154 | 0.816       | 174 | 298 | 0.423       | 2   | 441 | 0.359       |
| 130 | 324 | 0.816       | 410 | 414 | 0.423       | 174 | 219 | 0.359       |
| 22  | 177 | 0.8         | 49  | 51  | 0.422       | 414 | 449 | 0.359       |
| 265 | 440 | 0.719       | 219 | 298 | 0.421       | 199 | 345 | 0.358       |
| 241 | 316 | 0.647       | 219 | 237 | 0.417       | 199 | 369 | 0.358       |
| 9   | 164 | 0.609       | 199 | 418 | 0.416       | 94  | 156 | 0.358       |
| 355 | 451 | 0.572       | 291 | 420 | 0.416       | 218 | 369 | 0.358       |
| 414 | 423 | 0.572       | 174 | 237 | 0.414       | 216 | 320 | 0.358       |
| 118 | 166 | 0.515       | 165 | 405 | 0.413       | 141 | 199 | 0.358       |

6. Conclusion

The development of sports tourism resources must have a complete support system, and this support system must be highly comprehensive and synergistic. This requires not only the active planning and operation of sports tourism managers, but also the support and participation of the whole society, especially the intervention of economists, tourism scientists, sports scientists and government management agencies. Based on the application of big data technology, this article builds a new sports tourism information system platform and builds a system evaluation model to verify the effectiveness of the system proposed in this article.

Acknowledgments

Project Source: Leshan Normal University New Arts Research and Reform Practice Project.
Project Name: Research on the Integrated Development of "Sports + Tourism" (Number: JG2021 XW-12).

References

[1] Kurtzman, J. & Zauhar, J. Sports tourism consumer motivation. Journal of Sport Tourism, vol. 10, pp. 21-31, January 2005.
[2] Chiu, L. K. Ramely, A. & Wafi, A. Make green growth a priority: issues and challenges in organising green sports tourism events. Malaysian Journal of Sustainable Environment, vol. 7, pp. 53-71, January 2020.
[3] Tu, Y. F. & Hwang, G. J. Trends and research issues of mobile learning studies in hospitality, leisure, sport and tourism education: A review of academic publications from 2002 to 2017. Interactive Learning Environments, vol. 28, pp. 385-403, April 2020.
[4] Langenbach, M. & Tuppen, J. The concept of localised outdoor sports tourist systems: its application to Ardèche in south-east France. Journal of Sport & Tourism, vol. 21, pp. 263-286 April 2017.
[5] van der Kruk, E. & Reijne, M. M. Accuracy of human motion capture systems for sport applications; state-of-the-art review. European journal of sport science, vol. 18, pp. 806-819, June 2018.
[6] Weed, M. Sports tourism theory and method—Concepts, issues and epistemologies. European sport management quarterly, vol. 5, pp. 229-242, March 2005.
[7] Liang, S. Schuckert, M. Law, R. & Masiero, L. The relevance of mobile tourism and...
information technology: an analysis of recent trends and future research directions. Journal of Travel & Tourism Marketing, vol. 34, pp. 732-748, June 2017.

[8] Roxas, F. M. Y. Rivera, J. P. R. & Gutierrez, E. L. M. Framework for creating sustainable tourism using systems thinking. Current Issues in Tourism, vol. 23, pp. 280-296, March 2020.