



# Governance Pathways for Public Digital Cultural Data in a Multi-Source Heterogeneous Context

Yuxiang Luo

Hainan Technology and Business College, Hainan, China

tg667788@xzcstudio.com

**Abstract.** After analyzing the current dilemmas faced by public digital culture in China, we find that there are still problems such as low level of digitalization of resources, obvious single-dimensional characteristics of cultural services, and insufficient data collaboration. On this basis, we will also explore an effective path for the governance of public digital cultural data in China from several aspects, including the guidance of national policies, the richness of information resources, the diversity of data processing methods, and the use of modern information technology to expand its service functions.

**Keywords:** Multi-source heterogeneous; Public digital culture; Data governance

## 1 Introduction

With the rapid development of information technology and the explosive growth of digital resources, public digital cultural data, as an important part of modern society, is gradually highlighting its important value. Public digital cultural data not only contains rich historical and cultural information, but also reflects the dynamics of social development and changes in people's lives. It is an important force to promote social progress and cultural inheritance. However, under the background of multi-source heterogeneous public digital cultural data governance, it faces many challenges and problems. How to effectively integrate, manage and utilize these data has become an important issue that needs to be urgently solved in the digital culture field.

The multi-source heterogeneous background refers to the diversity of data sources and the difference in data structure. In the digital age, public digital cultural data comes from different institutions, platforms and individuals, including libraries, museums, archives, the Internet, etc. There are significant differences in format, standards, quality, etc. This multi-source heterogeneous characteristic makes data governance more complex and difficult, requiring cross-field and cross-departmental collaborative cooperation and innovative practices. Promoting the prosperous development of digital culture and providing strong support for social progress and cultural inheritance. Therefore, studying the governance path of public digital cultural data under the background of multi-source heterogeneity is of great theoretical significance and practical value.

## 2 Approaches on Public Digital Cultural Data Governance in A Multi-Source Heterogeneous Context

Currently, China has entered a critical period of accelerating the construction of a modern public cultural service system. Focus is placed on the "14th Five-Year Plan" to continuously promote the integration of the public cultural service system and advance its high-level development.[1] Therefore, in the next phase of development, how to effectively integrate public digital cultural resources becomes a crucial issue that needs urgent resolution. This project proposes an overall architecture comprising three levels: "Data-Platform-Service" from a resource integration perspective (as shown in Figure 1).

Based on this framework, a method for integrating public cultural information is proposed, which includes data integration, platform integration, and service integration. Information integration is a prerequisite to achieving this goal. With consideration of the current resource supply condition, the objective is to achieve overall planning and coordinated construction without duplicative efforts.[2]

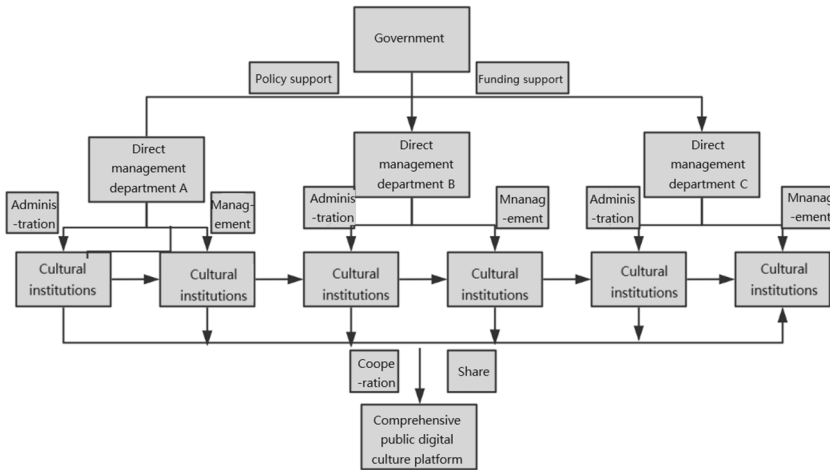


Fig. 1. Public digital cultural comprehensive service platform

It is essential to expedite the establishment of accompanying standards and systems to ensure the continuous growth of innovative resources. These measures play a positive role in enhancing the efficiency of public cultural services, optimizing resource structure, activating existing resource stock, and promoting the national modern public cultural service system.

Platform integration is key. Although initially, China prioritized projects like the National Cultural Information Resource Sharing Project, Digital Library Promotion Project, and Public Reading Room Construction Project in the public digital cultural services, and the National Public Culture Cloud was officially launched on November 29, 2017, as a major platform, the platforms are not fully interconnected yet, which

limits their ability to provide accurate services. The needs of multilingual users, particularly from minority groups, have not been sufficiently addressed, necessitating the construction of a mobile, personalized, and multilingual comprehensive service platform[3].

Service integration is the goal. The essence of service lies in people-oriented approaches, aggregating content and uncovering knowledge at the semantic level of massive, heterogeneous resources, thereby achieving deeper integration and utilization. This provides the public with higher quality, diverse, and convenient services. To enhance the efficiency of public cultural services, collaboration across various sectors is necessary to promote coordinated, balanced development among urban and rural areas, regions, ethnic groups, and the nation. The paper discusses the integration of resources in public digital cultural services, hoping to provide references for future promotion and improvement of public cultural digitalization.

### **3 The Plight Faced By Public Digital Culture in China**

#### **3.1 Low Level of Resource Digitization and Obvious Single-Dimensional Characteristics of Cultural Services**

Among the plight faced by public digital culture in China, the low level of resource digitization and the obvious single-dimensional characteristics of cultural services are two key issues. The low level of resource digitization is a common problem. Many cultural heritage resources, including historical documents, ancient buildings, artworks, etc., have not been fully digitized and sorted out. This not only brings challenges to the preservation and inheritance of resources, but also limits the breadth and depth of public digital cultural services. The lack of digital resources means that it is impossible to fully utilize modern technologies to spread and promote, and it is impossible to meet the growing digital cultural needs of the public. The prominent single-dimensional characteristics of cultural services are also an urgent problem to be solved. At present, many public digital cultural services are mainly for academic research and teaching, with relatively single application scenarios and lack of richness and diversity[4]. Such service models often ignore the cultural needs of the general public, and cannot make full use of the advantages of digital technology to enhance the interactivity and experience of cultural services. In the digital age, the public's demand for cultural services has changed dramatically. They are more eager to participate in the process of cultural creation and sharing, rather than just being recipients of cultural services.

#### **3.2 Insufficient Data Collaboration and Access Barriers Exist**

Insufficient data collaboration and access barriers are another two important plights faced by public digital culture in China. Insufficient data collaboration is mainly reflected in the relatively low degree of data sharing and collaborative work among different cultural institutions, different regions and different industries. This is mainly affected by the traditional fragmented management system and the traditional concept of "large and complete, small and complete" resource construction[5]. Due to the lack of

unified data standards and sharing mechanisms, each institution works in its own way, resulting in the inability to effectively integrate and share data resources, forming information islands and causing waste and duplication of resources. The access barrier refers to the various obstacles faced by the public in accessing public digital cultural resources. These obstacles can come from the technical level, such as incompatible formats of digital resources, unstable access platforms, etc.; they may also come from the institutional level, such as copyright protection, privacy protection and other policies that restrict access and use of resources. These access barriers not only affect the efficiency and usability of the public to obtain digital cultural resources, but also limit the development and innovation of public digital cultural services[6].

### **3.3 Metadata Description Standards are Multi-Source Heterogeneous, Data Between Systems is Incompatible**

The specifications of metadata are multi-source heterogeneous, and the data between systems is incompatible. The reason lies in the lack of a unified data standard and specification. The metadata description standards and formats used by various cultural institutions in the digitization process are also inconsistent, resulting in inefficient communication and sharing of data between different systems. The multi-source heterogeneous metadata description standards bring the following main problems:

(1) Difficulty in data sharing: Due to the inconsistent metadata description standards adopted by each system, data conversion and mapping are required during sharing, increasing the difficulty and cost of data sharing.

(2) Ambiguity in data understanding: Different metadata description standards may cause ambiguity in the interpretation and meaning of the same data item in different systems, thus affecting the quality and utilization effect of the data.

(3) Data island phenomenon: Due to the lack of unified data standards and specifications, data between systems cannot be effectively integrated and shared, resulting in data islands, limiting the value and application range of the data[7].

## **4 China's Public Digital Cultural Data Governance Path**

Public service and cultural inheritance. The governance of digital public information requires the coordination and cooperation of all sectors of society, and also involves interests in many aspects. Therefore, it is necessary to guide, regulate and coordinate the process of data governance through national policies and regulations. In public cultural services, the government bears the main responsibilities and plays a leading role. Only under the authoritative leadership of the government can all public cultural institutions actively participate, so that public digital cultural data can really be connected and effectively managed. Ultimately, under the guidance of various public cultural competent departments, a shared platform is jointly established to achieve the social value of public digital culture. In the process of public digital cultural governance, the government needs to unify coordination, unified dispatching, unified standards, and finally unified implementation, and establish a "one-stop" comprehensive public digital

cultural service platform with regional characteristics and comprehensive characteristics, as shown in Figure 1:

#### **4.1 Continuously Seek New Resource Providers to Expand the Scope of Public Digital Cultural Service Resources**

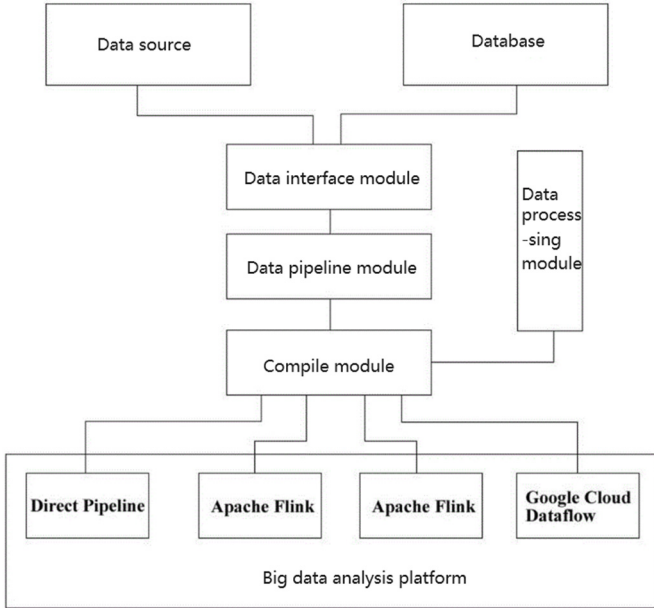
In discussing the governance path of China's public digital cultural data, finding new organizations to provide resources and expanding public resources for digital cultural services is undoubtedly an important strategy. This approach aims to ensure the extensive coverage and in-depth satisfaction of public digital cultural services, and continuously promote the sharing and innovative application of cultural data.

The government can promote cooperation between different departments to integrate digital cultural resources scattered in the fields of culture, education, science and technology to form a unified service platform. Encourage the private sector, especially cultural enterprises and technology companies, to participate in the construction of public digital cultural service resources. They can provide advanced technical support and rich content resources. By participating in international cultural cooperation projects, high-quality cultural resources from abroad can be introduced, while promoting the overseas dissemination of China's excellent culture.

Evaluate the newly added resource providers to ensure that the quality and service level of the resources they provide meet the requirements of public digital cultural services. Sign cooperation agreements with new resource providers to clarify the rights and obligations of both parties and ensure the legal use and sharing of resources. Upgrade and improve the existing public digital cultural service platforms to accommodate more types and larger scale data resources[8].

#### **4.2 Adopt Diversified Data Governance Methods and Store Heterogeneous Digital Cultural Information Resources**

When discussing the path of public digital cultural data governance in China, adopting diversified data governance methods and storing heterogeneous digital cultural information resources is crucial. This can not only ensure the comprehensiveness and diversity of digital cultural information, but also improve the efficiency and flexibility of data governance. As shown in Figure 2:



**Fig. 2.** Diversified data governance methods

Diversified data governance methods refer to adopting different governance strategies and management methods according to different types, sources, formats of digital cultural information resources. This can maximize the preservation of the original appearance and value of the data, while facilitating subsequent data analysis, utilization and dissemination. Specifically, the following data governance methods can be adopted:

1. Metadata governance: Unified description and classification of digital cultural information resources through metadata (data describing other data) to facilitate subsequent retrieval and utilization.

2. Data cleaning and integration: Improve data quality and consistency by cleaning and integrating raw data, eliminating redundant, erroneous, invalid data. As shown in Table 1:

**Table 1.** Data management and preprocessing strategies

Keywords	Key Points
Data management and preprocessing strategies Subject: Data cleaning and Standardisation	<ol style="list-style-type: none"> <li>1. Identify and handle missing values, outliers and duplicate records to ensure data integrity and consistency.</li> <li>2. Standardize data formats, data types and data scales to eliminate barriers caused by heterogeneity.</li> <li>3. Apply data validation rules and constraints to ensure data accuracy and credibility.</li> </ol>

3. Data visualization: Visually present digital cultural information resources in forms like charts, images, etc. to facilitate user intuitive understanding and utilization. Algorithm improvement can be made through formulas (1) and (2), and tfrsp function in Matlab can be used to obtain two-dimensional spectrum and time-frequency domain graphs for data visualization[9].

$$X(k) = \sum_{j=1}^N x(j)\omega_N^{(j-1)(k-1)} \text{ Formula (1)}$$

where,  $\omega_N = e^{(-2\pi i)/N}$

$$S_z(t, f) = |\text{STFT}_z(t, f)|^2 = \left| \int_{-\infty}^{+\infty} z(t')\eta^*(t' - t)e^{-j2\pi ft'} dt' \right|^2 \text{ Formula (2)}$$

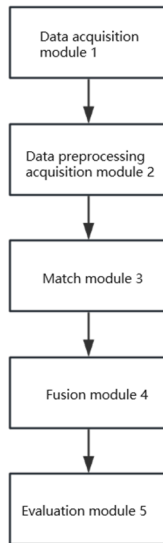


Fig. 3. Distributed storage system

4. Data mining and analysis: This project plans to use data mining and analysis methods to extract useful information and rules from massive data to assist scientific decision making and research.

(1) Heterogeneous digital cultural information resources refer to digital cultural resources with diverse data types, complex structures, and different formats. In order to effectively store these resources, appropriate storage strategies and technical means need to be adopted. Specifically, the following aspects can be considered:

(2) Build a distributed storage system: By building a distributed storage system, disperse and store the digital cultural information resources across multiple nodes, improving data security and reliability. As shown in Figure 3:

(3) Adopt cloud computing technology: Use the elastic scalability and on-demand payment features of cloud computing technology to provide flexible support for the storage of digital cultural information resources.

(4) Implement data encryption and security measures: Encrypt and implement security measures for digital cultural information resources to prevent data leakage and illegal access[10].

## 5 Conclusion

The era of big data has brought us many new things and also brought new changes to our lives. But no matter how much data there is, if it cannot be used for the society and play its role, it would just be a mirage. We must make effective use of the new value of big data for public digital culture, promote the overall deployment of national cultural governance, solve the problems of unbalanced regional cultural development and unequal access to public culture, while also contributing to social progress. At the same time, appropriate intervention measures need to be taken for data governance activities, such as risk control, regulatory compliance testing, and ensuring data privacy, so as to achieve "data-driven smarter governance".

## Acknowledgment

Project supported by the Education Department of Hainan Province, project number: Hnjg2023-201

Project supported by the Education Department of Hainan Province, Project number: Hnky2023-76.

## References

1. Jin Yu. Research on the Realization Path of Digital Cultural Governance of Public Libraries in the Era of Big Data [J]. Journal of Traditional Chinese Medicine Library and Information Science, 2016, 40(3):27-30.
2. Han Bing. Research on the Realization Path of Digital Cultural Governance of Public Libraries in the Era of Big Data [J]. Computer Enthusiasts (Popular Edition) (Electronic Version), 2020(3):941-942.
3. Zhou Jianxin. The inherent mechanism, dilemma review and realization path of national cultural digitization [J]. Journal of Yantai University (Philosophy and Social Sciences Edition), 2024, 37(1):68-77.
4. Wei Jingzhu, Wang Zheng. Research on collaborative governance of public cultural data: Connotation, category and theoretical framework [J]. Library and Information Knowledge, 2022, 39(6):67-77,157.
5. Chang Y C, Dai H J, Wu J C Y, et al. TEMPTING system: A hybrid method of rule and machine learning for temporal relation extraction in patient discharge summaries[J]. Journal of biomedical informatics, 2013, 46: S54-S62.
6. Dredze M, Kulesza A, Crammer K. Multi-domain learning by confidence-weighted parameter combination [J]. Machine Learning, 2010,79(1-2): 123-149.
7. Yang Yan, Tian Jiming. Research on the construction of digital culture governance system based on the integration of science and technology and humanities [J]. Modern Intelligence, 2020,40 (10): 43-51.



8. Zheng Jianming. Innovation and Thinking of digital Culture Governance Path under the Environment of Big Data [J]. JJournal, 2016, (06): 1-5.
9. Zhang Weidong. Research on the digital governance path under the small data architecture [J]. Work in university Library, 2016,36 (06): 46-48 + 56.
10. Liu Wei, Zhang Qi, Zhang Zhe Yu. Research on Public Cultural Services for Big Data Innovation [J]. Construction of the Library, 2016, (03): 4-8 + 12.

**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.

