

Discussion on the “Digital Recovery” of the Unearthed Texts of the Sutra Cave in the Dunhuang Mogao Grottoes, Lost Overseas for Centuries

Xia Shengping, Zeng Qingying

Tunhuangology Information Center of Dunhuang Academy, Dunhuang Academy, Dunhuang, China

Email address:

dhxiashengp@hotmail.com (Xia Shengping), 1125443239@QQ.com (Xia Shengping), chinyin.tseng@gmail.com (Zeng Qingying)

To cite this article:

Xia Shengping, Zeng Qingying. Discussion on the “Digital Recovery” of the Unearthed Texts of the Sutra Cave in the Dunhuang Mogao Grottoes, Lost Overseas for Centuries. *Mathematics and Computer Science*. Vol. 4, No. 5, 2019, pp. 89-98. doi: 10.11648/j.mcs.20190405.11

Received: August 6, 2019; **Accepted:** October 24, 2019; **Published:** November 8, 2019

Abstract: This paper explores the issue of digital reconstruction of the literature of the Sutra Cave in Dunhuang Mogao Grottoes which had previously been lost overseas for over a century, and the systematic publication of Dunhuang texts scattered around the world on a shared digital resource platform for the academic world, and proposing the “digital recovery” of the unearthed texts of the Sutra Cave in Dunhuang Mogao Grottoes, many of which had been lost overseas for over a hundred years. If Tunhuangology scholars and experts can share these cultural heritage resources and use digital textual research theories and method to carry out Tunhuangology studies, better results can be achieved. This will advance the overall implementation this project, as well as its construction, its overall framework, the adoption of technical routes, and to explore the jointly constructed and shared model of “digitally recovering” the texts of the Sutra Cave in the Dunhuang Mogao Grottoes, so that more experts, scholars, and members of the public can share this precious resource and actively promote the in-depth study of Tunhuangology. It is also hoped that this will achieve the “digital recovery” of the Scriptures of the Sutra Cave in the Dunhuang Mogao Grottoes in a real sense, and provide an effective method for international Tunhuangology scholars, experts, and the public to access its resources quickly and easily.

Keywords: Lost Overseas, Dunhuang Mogao Grottoes, Sutra Cave Texts, Digital Recovery

1. Preface

The Dunhuang Mogao Grottoes are a UNESCO World Cultural Heritage site, and are widely considered one of the greatest treasures of the ancient Silk Road, and a major waystation along it. The grottoes consist of a cave system built between the 4th and 14th century AD, and underwent uninterrupted excavation during dynasties including the Northern Liang, Northern Wei, Western Wei, Northern Zhou, Sui, Tang, the Five Dynasties Period, the Song, the Western Xia, the and Yuan. At present, there are 735 preserved caves in total, 492 of which remain completely intact. There are 492 major caves in the southern section and 243 caves in the northern one (the remaining five have been numbered), forming a networked group of grottoes [1] that are 1680 meters long from north to south, containing 45,000 square meters of preserved murals, more than 2,000 colorful sculptures, 5 wooden structures from the Tang and Song

Dynasties, and more than 50,000 artifacts of literature and art discovered in the Sutra Caves and grotto temples amidst this oasis-like environment of the Gobi Desert [2].

The Dunhuang Mogao Grottoes contain a collection of exquisite murals, colorful sculptures, and architecture. They are the world's largest existing, longest-running, and most abundant corridor of grottoes and storehouse of Buddhist art. It enjoys the reputation being “the Art Museum of the Orient” and is also the longest-running, most influential and well-preserved grotto in China, with the richest cultural significance and an exquisite and superbly rich collection of art.

With its long history, rich heritage, huge information resources, and precious value, the Dunhuang Mogao Grottoes have had a great influence both across China and worldwide, and have become a symbol of the excellent traditional culture and art of the Chinese nation, possessing superbly abundant and unique cultural heritage resources.

At the beginning of the last century, on June 25, 1900, the

Dunhuang Mogao Grottoes were discovered by chance, which had a great influence across China and around the world, becoming one of the largest archaeological discoveries in the early 20th century.

A large number of ancient scripts, enamel paintings, paper paintings, and other works, including tens of thousands of ancient books and some of the world’s earliest printed books, thousands of beautiful enamel paintings, and the world’s earliest paper archives were excavated from the Sutra Cave. It is an underground treasure trove of medieval art. The re-emergence of Sutra Cave, which had been sealed off for nearly a thousand years, has prompted scholars to completely re-evaluate prior understandings of the four civilizations that lay along the Silk Road. Many important cultural relics, paintings, and writings in more than twenty languages have been unearthed in other sites on the Silk Road in China. However, these works are now scattered around the world, making it difficult for scholars to use and share academic resources.

Most of this enormous volume of ancient writings are one-of-a-kind and their sole extant copy, but about 36,000 of these precious cultural relics have been removed outside China’s borders. Collection institutions in the United Kingdom, France, Russia, Japan, India, Germany, Sweden, the United States, Denmark and other countries collectively possess a large number of Dunhuang texts and works. According to currently published statistics, the author has found the following data on unearthed texts in the Sutra Cave: about 57,790 works of Chinese literature; 9,574 pieces of Tibetan literature, and about 588 piece of literature in other languages, [3] and about 1,200 artworks have been unearthed in the Sutra Cave in Dunhuang. These are all inseparable and important parts of the cultural heritage of Dunhuang Mogao Grottoes. Since then, in the world of humanities, an academic boom has emerged following the large-scale excavation of the Sutra Cave and Dunhuang Grottoes, and the international academic study of Tunhuangology has been formed. As a result of this, a large number of research results have been produced in the field of Tunhuangology.

Since the birth of “Tunhuangology” nearly a century ago, though experts, scholars and institutions have systematically combed over the Dunhuang literature that has been sent overseas, including the artwork, it is difficult for general consumers to review them because the works are scattered in different organizations across the world, and this has brought great confusion to academic scholars about how to make it easier for readers to use and share this precious literature. The helplessness regarding Tunhuangology in China can be seen from Mr. Chen Yanke’s words describing Dunhuang as “the sad history of our nation’s academia” [4]. Many scholars in China and around the world are fascinated by the study of Dunhuang’s cultural relics, and any experts and scholars studying Tunhuangology must mention China’s enormous loss of Dunhuang’s cultural relics. This is a serious point of concern in the academic community. However, to date, there are few database platforms for one-stop and easily retrievable searches of catalogues, documents, images, research catalogue

databases, and full-text databases. There are very few databases in this field that can be shared by the academic community. But what will the future internationalization of Tunhuangology be? Modern digital textual and cultural technologies are quickly evolving. The question is, how to use these digital information technologies to promote the sustainable development of international Tunhuangology? The loss of precious literature and artworks in Dunhuang is not only a major loss to the Tunhuangology community in China, but also a major loss of Chinese history, and a major loss to world cultural heritage.

This paper mainly discusses the digital recovery of the ancient literature of Dunhuang’s Sutra Cave. This generally consists of a compromise method adopted as there is no way for the Dunhuang cultural relics that have been lost overseas to be physically “returned” as a whole. To be exact, the use of a system such as this is necessary to allow the lost ancient works of Dunhuang’s Sutra Cave to be “recovered” in a digital form, or “digitally returned”. By digitally reconstructing and sharing the literature of Dunhuang’s Sutra Cave, Tunhuangology experts, scholars, and members of the public around the world can share this cultural heritage resource and study Tunhuangology with the help of digital research theories and methodologies, so as to achieve better results.

2. The Problems of Digitally Recovering the Unearthed Texts of the Sutra Cave in the Dunhuang Mogao Grottoes

2.1. Recovery of Texts in Dunhuang’s Sutra Cave

How and when will the Dunhuang cultural artifacts lost overseas be returned? This is a hot topic in the Tunhuangology community, and experts and scholars studying Tunhuangology have already discussed this issue. The author agrees with the point of view of Mr. Chai Jianhong, who believes that “First, when we look upon the past, it is necessary to understand history in a realistic way. Second, based on the current situation, everyone must actively promote cultural and academic exchanges. Third, when we look to the future, we should firmly believe that historical problems will be satisfactorily resolved.” [5]. Mr. Chen Yujie’s offers a more detailed point of view in his article *Thinking About the Means of Returning Literature and Cultural Relics Lost Overseas*; that is, though there have been some achievements in the recovery of literature and cultural relics lost overseas, this is a long and complicated process requiring strenuous efforts. With the increasing rate of international exchanges and cooperation, there should be more new option for the recovery of literary and cultural relics lost overseas [6]. The difficulty of recovering Dunhuang’s cultural relics is so great that the issue needs to be viewed objectively and rationally. At present, one set of returned Dunhuang cultural relics have come from the family of the late Japanese calligrapher Aoyama Sanu, which were donated to Dunhuang Academy on October 9, 1997. These Dunhuang cultural relics

were thus returned to their origins, to the site where these items of cultural heritage belong. The catalogue and contents of this donation were published by Mr. Tan Chuanxue. It is mentioned in the published article that Mr. Aoyama Sanu purchased these relics in an ancient-literature bookstore in Tokyo more than a decade ago. He said before he passed away, "In any case, these ancient documents should be returned to China. These artifacts can only have their meaning if they are returned to their original places." [7], Mr. Aoyama Sanu's insights are admirable, and his actions began the first return of Dunhuang cultural relics. Up to now, this is the biggest donation that Dunhuang Academy has ever accepted.

2.2. *Digital Cooperation on the Texts of Dunhuang's Sutra Cave*

The difficulty of recovery or return of physical cultural relics and donation are relatively large. Is it possible to explore other ways? Before this, national institutions collecting the Dunhuang literature were digitizing the literature resources, and the establishment of "The International Dunhuang Project (IDP)" [8] provides a cooperation platform for the Dunhuang literature collection agencies to discuss the protection of collections and cataloging [9]; but its digital progress is very slow, and the organization is set up in the UK. At the beginning of the establishment, its Secretariat is located in the British Library. However, the ownership of the use, management, storage and provision of services in these Dunhuang literature digital resources does not belong to China, and certainly does not belong to Dunhuang. Later, Dunhuang Academy, museums and libraries around the world worked together with Andrew W. Mellon to produce high-quality digital images of the murals and related art and literature of Dunhuang Grottoes in China and incorporate them into an academic electronic archives – "International Dunhuang Archives". Dunhuang Academy and some foreign institutes that collect Dunhuang materials authorized the Mellon Foundation to produce the Mellon International Dunhuang Archive, an academic electronic archive that will provide important information resources for cultural relics protection and other professional scholars. Dunhuang Academy, Mellon Foundation and other participants hope that this pioneering project will promote the work of recording and preserving the art of Dunhuang Grottoes and further promote the development of academic and artistic worlds [10]. At present, we are collecting a large number of ancient literature, enamel paintings, paper paintings and other materials unearthed in Sutra Cave in Dunhuang Mogao Grottoes, which are scattered in libraries and museums around the world. We are using the most advanced digital technology to digitally process these scattered data, and through information integration, knowledge organization to reconnect them with Dunhuang Grottoes digital archives information. Any expert or scholar uses the digital archives of the grottoes to correlate various types of image information and data through related technologies. This is the true purpose of our implementation, and also a very good idea. This is why we want the Dunhuang unearthed literature scattered around the world to be returned

in a digital form. When the international conference was held in Dunhuang in 2015, we formally proposed the "Digital Recovery of Dunhuang Cultural Relics. At the same time, in the "13th Five-Year Plan" of the Ministry of Culture, we proposed "promoting the recovery of overseas Chinese ancient books" [11], which provides strong support for the digital recovery of Dunhuang cultural relics.

2.3. *Digitization of the Literature of Dunhuang's Sutra Cave*

At present, Dunhuang Academy is implementing the "Digital Dunhuang Project", and has already provided the world with high-definition digital resources of 30 caves, including the integrated digital resources of the unearthed textual resources of the Dunhuang Sutra Cave, and this is an enormous project. This includes various Dunhuang literature databases that have been under construction, all of which play an important role in different stages of work, but all have their shortcomings. Questions include: how to integrate and utilize these resources; how to use information technology; especially digital information and research technologies to carry out humanities and social science research; whether libraries, museums, archives, information providers, or digital institutions need to think about the above issues; and to explore how Dunhuang texts stored in institutions such as museums in different countries and regions can allow scholars and the public around the world to easily access these resources.

On April 26, 2015, the National Library of France and Dunhuang Academy signed an agreement to give Dunhuang Academy high-definition digital copies of the Dunhuang writings in the NLF's collection, authorizing experts and scholars of Dunhuang Academy to use these items for free for non-commercial purposes. This collaboration is the first time a foreign collection institute has given digital copies of literature of the Dunhuang Sutra Cave to China [12], and this is the first high-definition digital resource of Dunhuang literature obtained by Dunhuang Academy. Since then, the National Library of China has completed the first phase of the "Bibliographic Database for Overseas Ancient Chinese Books" in 2016, which includes all of the high-definition digital resources of the Dunhuang legacy writings presented by the National Library of France to the National Library of China, including more than 2,700 works of Chinese literature from the Dunhuang legacy books, more than 4,000 pieces of Tibetan literature stored in the French National Library, as well as other precious literature in Sanskrit, Uyghur, Khotanese, Sogdian, Hebrew, and other languages, totaling 7,000 items [13]. Later, Northwest University for Nationalities' Institute of Overseas National Literature completed the digital storage of 1,000 volumes of literature from the Dunhuang Sutra Cave currently possessed by France and the U.K, and this built China's first digital library of the Dunhuang Sutra Cave's ancient texts [14].

Since then, these and similar projects have been actively promoted by Dunhuang Academy and the French National Library, which houses Dunhuang literature, and these two institutions have proposed re-digitizing Dunhuang literary

resources. Prior to this, the National Library of France had digitized Dunhuang literature in its collection, but at that time, due to various technical conditions, the accuracy of literature digitization was not sufficient for publication. This time, the two sides would cooperate to jointly digitalize the texts. The results of this digitization will provide a copy of high-definition digital resources to Dunhuang Academy. Using the same authorization system as the previous attempt can be used to make it easy for experts and scholars to use this copy. If this project can be implemented smoothly, it will make a good start for the Dunhuang texts’ “digital recovery”, and for the joint construction and sharing of this digital resource. The Russian State Hermitage Museum has also proposed a bilateral cooperation project to digitalize the literature of Sutra Cave, thereby laying the foundation for the next step.

3. Compilation and Database Construction of the Unearthed Texts of Dunhuang’s Sutra Cave

3.1. The Compilation of the Texts of Dunhuang’s Sutra Cave

According to the current records, it has been just over a hundred years since the discovery of the Sutra Cave. Through the unremitting efforts of several generations of Tunhuangologists, Dunhuang literature has been compiled and published. Historian Chen Yuan compiled his *Records of the Robber of Dunhuang*, which can be said to be the world’s first publicly published catalogue of Dunhuang legacy collections. It is a classified catalogue of Dunhuang writings, and has also made a significant contribution in the history of Tunhuangology and Dunhuang catalogues [15]. Mr. Wang Chongming’s *Index of Dunhuang Legacy Books* was published by the Commercial Press in 1962, and its publication was of epoch-making significance, boasting the most complete and accurate catalogue of Dunhuang legacy writings at that time [16]. Later, scholars such as Zhou Shaoliang, Hao Chunwen, and Liu Jinbao systematically classified the publication of Dunhuang literature [17]. Before and after 2006, Fang Guangchang also made a review and statistical analysis of the literature of Sutra Cave [18]. In 2008, Xia Shengping systematically reviewed and studied the past 20 years of Dunhuang literature, introducing a literature compilation, publication results, subsequent digitalization of the literature and digital platforms, etc. Among them, Shi Pingting and Tai Huili edited the *New Index of Dunhuang Legacy Letters* [19], which served as the latest catalogue of Dunhuang volumes, and corrected the errors of previous catalogues. These are all academic achievements that have made great contributions to Dunhuang literature, and have had a great influence on Tunhuangology. Despite the centuries of research and review of Dunhuang literature, the actual number, type, academic achievements, publications, and databases of the literature of the Sutra Cave can be described as endless. However, even in the modern information age, few of these

databases can be searched online. “The International Dunhuang Project (IDP)” is relatively widely recognized and used. However, it does not have a complete catalogue of all Dunhuang literature that can be retrieved, and its search function has limited convenience. Only the digital catalogue of major national collections and major collection institutions has been digitally completed. There is no complete catalogue of the literature of Sutra Cave. However, the advantage of the IDP is its ability to comprehensively digitize Dunhuang literature achieve certain goals to that end. Its shortcomings are that progress has been very slow, and the retrieval process is inconvenient. Therefore, an initial Dunhuang literature catalogue database has established. After continuous improvement, supplementation, and modification, published cultural relics can be built into a complete catalogue database which compiles all types of cultural relics, including a large number of ancient writings, paintings, drawings, etc.

3.2. Constructing the Literature Database of Dunhuang’s Sutra Cave

In the recent past, Ma De and Fang Guangchang also applied for support from the National Social Science Fund of China to create the *Dunhuang Legacy Book Information Resource Library* (2012) to build the new Dunhuang literature database. This library mainly consists of the National Library’s collection of Dunhuang literature catalogues, some digital image resources, and the full texts of the scanned literature. Although a large amount of data has continued to be supplemented and improved in later periods, the advantage is that the original documents of the Dunhuang legacy writings have been recorded and transcribed. However, the project has been closed and its funding was limited. Furthermore, because tasks of post-database maintenance, functional system upgrades, server updates and replacements, data migration, data backup, subsequent data additions, modification, and improvement all require a great deal of manpower and financial resources, there are limitations for subsequent development, and there are difficulties in terms of compatibility with other databases, as well as problems in data exchanges and the sharing of heterogeneous database data. With the joint efforts of the research team, Mr. Ma De’s project has once again been funded, but this is a “Century-Long Project”. The advantage of Ma De and Fang Guangyu’s work is that there are literary records of the Dunhuang Sutra Cave, and the meta-data settings in the database are also reasonable.

However, this database is solely concerned with Dunhuang literature, and cannot meet needs for later database expansion, including all relevant bibliographical data, Dunhuang mural image resources, etc., and cannot solve the problem of “information isolation”. It is also impossible to interface with the “Digital Dunhuang Project”. At present, the implementation of this project can create a shared platform together with the “Tunhuangology Information Resource Library” [20]. The results of this project will be presented in the “Tunhuangology Information Resource Library” as an important pillar of the project to better realize the integration

and data linkages of different types of resources, laying the foundation for future digital resource sharing platforms, achieving the resource and full-text integration of the literature of Sutra Cave, the digital image resources of Dunhuang literature and image resources of Dunhuang grotto murals, as well as “one-stop” search functions expected and hoped for by experts and scholars without the need for cross-platform searching, which expands the significance of conducting research on this topic.

4. Digital Reconstruction of the Unearthed Texts of Dunhuang’s Sutra Cave

Due to historical reasons, Dunhuang artifacts are scattered all over the world, and have been collected by many public and private institutions. Although many documents have been published, these limited resources still cannot meet the needs of scholars. Through cooperation with institutions with a collection of Dunhuang literature, digitalization is carried out together with these institutions, and digital technology is used to digitally reconstruct the literature of the Dunhuang Mogao Grottoes, and discussions have raised the idea of allowing these Dunhuang artifacts, which have been lost for over a century, to be returned in a digital form. To this end, the establishment of a management database of texts from the Dunhuang Sutra Cave, including a catalogue of literature from the Sutra Cave, the catalogue of literary academic research and image resources, and can be used by domestic and foreign experts to retrieve, generalize, classify, sort, and correlate data, so that experts and scholars can realize data association by researching catalogues, literature information, full-text databases, and image resources in a future “Digital Dunhuang Project” to achieve a one-stop search function. This will greatly facilitate scholars’ sharing of the most valuable resources of Sutra Cave in Dunhuang, breaking the barriers of regional and institutional divisions caused by the dispersion of the Sutra Cave’s literary resources.

The “Digital Recovery” of the literature of the Dunhuang Mogao Grottoes’ Sutra Cave allows for the construction of a digital resource sharing platform for Sutra Cave literature, and forms a collection and collation of Dunhuang literature and art including a large number of ancient books from the Dunhuang Mogao Grottoes that have been lost overseas, along with the establishment of an expansible intelligent information system for digital content classification management. Through the development of digital technology, we are actively using the power of science and technology to deeply protect our cultural heritage while deepening the intangible value of data to serve society and the public.

4.1. Digital Reconstruction Requirements for Dunhuang’s Sutra Cave

The development of database management platform can maximize the integration of the original underlying data, and deeply explore the intangible value of the data. The system

organizes the general catalogue of the literature of Sutra Cave along with the academic research catalogue. Through the implementation of this project, it is hoped that a professional and systematic Chinese-made Dunhuang Sutra Cave literature resource management platform will be established. The project can deeply explore the knowledge relevance of related resources, and lay the foundation for the later construction of Dunhuang knowledge maps and the sharing of Dunhuang resources.

For scholars engaged in research on both Dunhuang literature or the Dunhuang Grottoes, this project is highly beneficial. It is also very important for the current connection between the data of the “Digital Dunhuang Project”, which can achieve linkages between the study of Tunhuangology and the cataloguing of Sutra Cave artifacts, the literature catalogue of the Sutra Cave, and the platform of the Tunhuangology Information Resource Library.

The first stage is to complete the overall framework design of the database management platform. On this basis, a central component of the general catalogue database than can retrieve by the literature of Sutra Cave will be established. The system will sort out information and secondary literature research catalogues including the artifact number, the archiving countries, and the place of collection from within the Sutra Cave. This topic has been supported by the Dunhuang Grottoes Protection Research Foundation of China (established in 2018), and the project database management platform has been completed.

In terms of the literature catalogue and research catalogue database of Sutra Cave in Dunhuang, there are studies about its digitization, data survey, data processing, data analysis, the distribution of Dunhuang literature overseas, the number of collections, collection institutions, age, preservation status, and other related information. The platform will be re-integrated according to different resource types to realize the functions of digital content processing, catalogue management, and knowledge association in the digital environment, and finally form a multi-form digital product to provide users with personalized data services.

The literature of Dunhuang’s Sutra Cave that has been lost overseas is currently located in different museums, libraries, public, and private institutions in different countries. Other artifacts are collected in major institutions in China, and can be distinguished in terms of database meta-data settings. This paper mainly discusses the systematic collection, review, sorting, cataloging, and classification of Dunhuang literature from overseas, and reviews the catalogues of academic research results of literature of Sutra Cave in Dunhuang together with the existing digital resources of Dunhuang Academy’s Tunhuangology Information Resource Library. By using the intelligent information linkages within the system and the database management platform of the second phase of the Tunhuangology Information Resource Library, which is currently under development, the data processing, data indexing, data integration, and database association of the texts of Dunhuang’s Sutra Cave and its literature research catalogue have finally been completed.

4.2. Digital Reconstruction Design Concepts of Dunhuang’s Sutra Cave Texts

The entire system of the management platform has been designed by layer. Each layer communicates with every other through loose coupling. From bottom to top, the system consists of a standard protocol layer, a data layer, an interface layer, an application support layer, a business application layer, and an access layer. The security of the system requires a control layer for secure access. The system integrates the catalogue of literature resources of Sutra Cave in various countries, and combines the digital image database of the

Sutra Cave with the academic resources of Tunhuangology. Through exploring the internal connection among the global literature catalogue of the Sutra Cave, the digital image database of the Sutra Cave and the various resources of the Tunhuangology Resource Network, the system realizes the following flowchart: the Sutra Cave literature catalogues → texts’ digital image resources → the literature research catalogue → full-text databases, forming the knowledge sequence of the Sutra Cave, eliminating the issue of information isolation, and finally providing services for researchers of Dunhuang’s Sutra Cave.

The detailed structure is shown below:

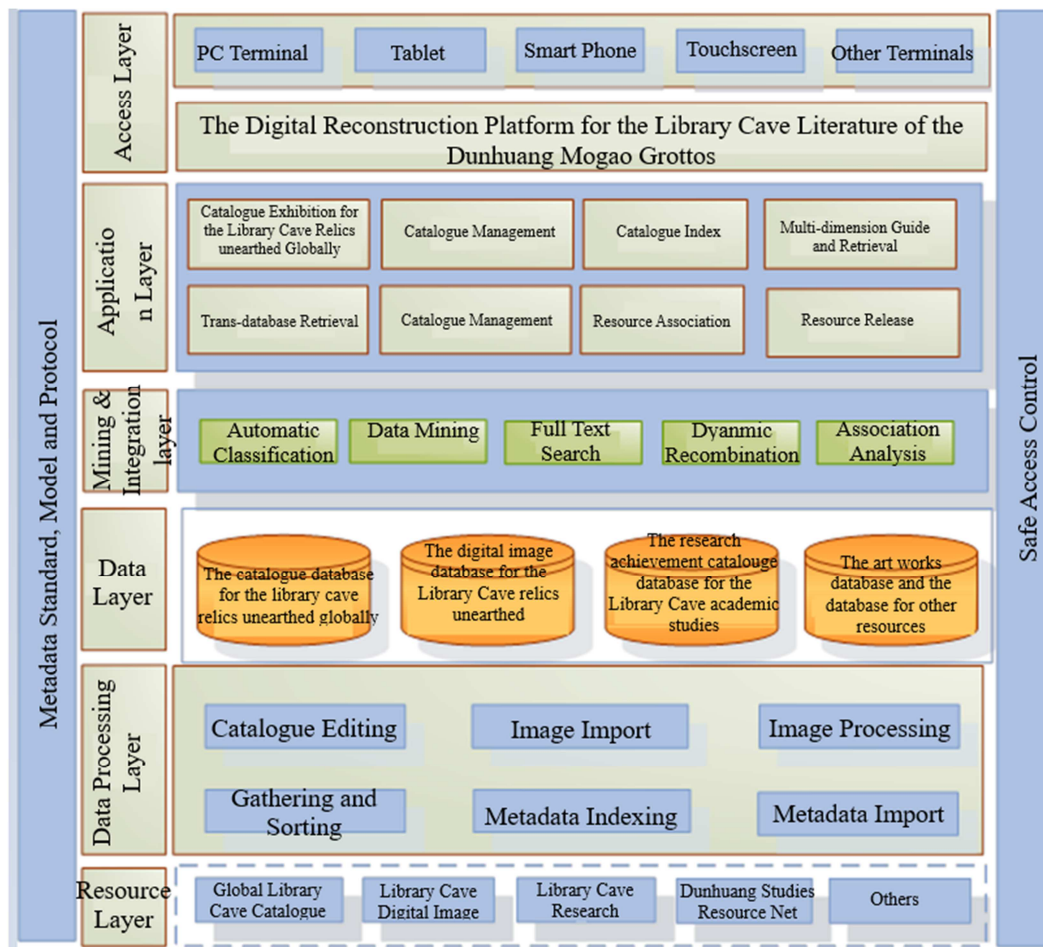


Figure 1. Design of Digital Reconstruction Platform Uneareded Texts of the Dunhuang Mogao Grottoes’ Sutra Cave.

To this end, the platform mainly serves to help Dunhuang experts and scholars across China and around the world to systematically search the collections and catalogues of various artifact collections. Through the implementation of the project, the texts of Dunhuang’s Sutra Cave can be systematically reviewed, and a searched database can be established to comprehensively sort out the textual catalogue of Sutra Cave, and set up multiple meta-data.

This also serves as a catalogue of academic research results related to Dunhuang literature, including information management, access, sorting, classification, association, retrieval, and other cataloguing and archival functions. After the completion of the project, the main results are presented in

the form of a database platform. It can be used to realize the “one-stop” search of information on the platform of the “Tunhuangology Information Resource Library”, which will greatly facilitate the use of Dunhuang’s Sutra Cave texts by scholars around the world. This valuable resource adopts a jointly constructed and shared model, actively promoting the internationalization of Tunhuangology in the digital information science and research environment, which is the real purpose of this “digital recovery”.

4.3. Designing the Metadata of the Texts of Dunhuang’s Sutra Cave

The meta-data standards of the digital reconstruction

project of the literature of Dunhuang's Sutra Cave are based on advanced Chinese and international Conceptual Reference Models (CRM), with experts such as Liu Hua, Zhao Yajie, Wang Ping, Han Chunping having done research [21] in this field according to the needs of the construction of the project resource library. Metadata standards have been formulated that meet the project requirements, and these assist in the editing of the standardized cataloging of digital resources of Sutra Cave. The project has higher standards for meta-data design, with strong scalability, and related elements can gradually be added. Considering the combination of the future system and the Dunhuang digital resource library, the system needs to be combined with the digital meta-data standard of Dunhuang to plan while designing the meta-data. It includes: detailed records of the collection country, collection location, collection number, name, inscription, age, etc., and the establishment of the project *Literature of the Sutra Cave in Dunhuang and the Literature Research Catalogue Database*; currently, the first phrase is mainly based on Chinese literature. A systematic review of the literature catalogues of academic research results of the literature of Sutra Cave in Dunhuang, statistics of collection institutions, and databases of management, classification, and association. The goal of these is mainly to establish the association and integration of the underlying data of a collection catalogue database, a researching catalogue database, and existing "information resource pools". To establish the literature database of Dunhuang's Sutra Cave, digital technology and Internet technology can be used to integrate the research results of Tunhuangology studies, the literature of Sutra Cave, relevant ancient books, and the full-text data of Tunhuangology into the "Tunhuangology Information Resource Library". The *Literature and Academic Research Catalogue Database of Dunhuang's Sutra Cave* is a special feature within the "Tunhuangology Information Resource Library".

4.4. From Digitalization to Knowledge Association of the Texts in Dunhuang's Sutra Cave

Through a multi-dimensional relationship, the literature resources of the Dunhuang Mogao Grottoes have been interrelated to form a resource-related network. The system supports multi-dimensional data mining and the establishment of various catalogues of different resources based on content mining, content, and attribute association, including static and dynamic associations. Static association consists of statically writing the association relationship within the database, while the dynamic one only extracts the association features. When the resource association is performed, a query is dynamically calculated according to the set parameters.

The system supports the unified mining of various resources. Based on content mining, it can automatically establish the relationship between various resources. The system can automatically mine various types of resources managed in the content resource library, and establish associations between content similarities. The system administrator can perform secondary editing and management on the automatic mining association relationship. First, the

mining module reads the content resources and related configuration information, then extracts the content features and establishes an associative relationship and saves the extracted association relationship to the database. The administrator then performs manual secondary modifications or management and maintenance activities. The display platform of Dunhuang Sutra Cave resources integrates a variety of resources including the "Global Literature Catalogue of Dunhuang's Sutra Cave", "The Academic Research Results Catalogue of the Literature of Dunhuang's Sutra Cave", "Digital Imaging of Dunhuang's Sutra Cave", the "Tunhuangology Resource Network", etc. Based on the B/S System structure, the system applies application system structures such as Web technology, database technology, XML language development technology, object-oriented technology, etc. to realize the seamless integration of digital resources and user terminal services. Digital resources are produced once, after which they can distributed and utilized in multiple terminals and channels.

5. Technical Methods Adopted by the Digital Management Platform of the Sutra Cave Texts

The development of the database management platform has been mainly based on the core technology of CNKI, which has stronger technical advantages both in terms of database architecture and platform function design. The previously developed "Tunhuangology Information Resource Library" has been in operation for five years, and the second phase is now in the design and transformation stage and is just in time with the project, allowing the more successful achievement of the integration and association of different resource types and heterogeneous databases.

5.1. Iterative Development Under the Premise of Top-Level Design

The development of the entire platform follows the development process of customized developmental software system standards, and focuses on top-level and detailed design processes. With the support of the project consulting team, (requirements analysis, summary design and detailed design of the system should be completed firstly) preliminary tasks including the requirements analysis, summary design, and detailed system design have been completed. An iterative development model has been adopted under the guidance of these design documents. After repeated iterations, the development of the prototype system has ultimately been able to meet the users' actual business needs.

5.2. Adopting an SOA-Based Architecture

Based on the advanced concept of SOA, a flexible and scalable system architecture has been achieved that combines openness and scalability. Service-Oriented Architecture (SOA) is the leading application architecture, boasting strong

integration capabilities. It provides basic IT application functions through the concept of business services. These services can be freely arranged, interconnected, integrated, and adjusted at any time to flexibly match both new and future needs. For example, in the SOA architecture, business functions such as meta-data generation, data exchange, external retrieval, and user authority authentication are encapsulated and published as services through standard interfaces, and deployed on system data and functional integration platforms as a service. Any application that wants to access another application can determine the attributes and calling formats of the accessed service through service discovery and representation systems, thus achieving cooperation between standardized applications, satisfying the loose coupling principle between the application systems, completely avoiding any impact on the other party's applications due to the use of unilateral systems and adjustments within the program.

5.3. Unified User Management & Rights Management

In this system, not only has single sign-on been enabled for each web portal, but also for each mobile terminal, involving usernames/passwords, digital certificates, and support plug-ins. It has a centralized permission control function, provides a role-based permission control mode, supports permission control of dynamic resources, and provides a secure secondary login function for certain sensitive information after the initial login has been completed. Unified user management provides precise privilege partitioning for specific users, effectively allowing high-volume user information maintenance and user rights management. The system supports user information management throughout the product life-cycle, including user information registration, user information modification, user rights assignment, and user deletion.

5.4. Componentization & Component Development Technology

The major means of database software development has evolved from a focus on the base source code, object code, and class library (object-oriented technology), to today's component-based development technology. A component is a software module with a specific function that can do almost any task. The novel concept of components, with their high degree of reusability, offer an entirely new way of thinking about software design. The component integrates the idea of hardware-based chip-centric processing into the analysis, design, and construction of software, and componentization allows these elements to be organized together to form different components. As components build up, developing software using these components is as easy as using building blocks. Component technology is by far the best and fastest-growing technological method for reusing software. It solves issues of reusability, poor adaptability, and long temporal cycles of software development.

5.5. XML-Based Data Support and Web Service Technologies

XML (Extensible Markup Language) is a simplified SGML (Standard Generalized Markup Language) developed by ISO (International Organization for Standardization). It is a meta-language used to define other markup languages. It emphasizes the description of information structures and content rather than the appearance of information. It is a language independent of format, platform, and application. XML can define a self-describing document structure, and has the support of additional standards such as RDF, XQL, and DOM. Through XML, the interoperability between heterogeneous databases can be realized and the independence of the original database can be maintained.

5.6. Using Computer Cluster Technology to Achieve Dynamic Automatic Load Balance

A computer cluster is a parallel or distributed system of computers connected to each other. From the outside, they are essentially just a system that provides a unified service to the outside world. The database platform project of cataloguing Sutra Cave texts has a broad variety of contents and services, so it has high requirements for system performance. Based on the above factors, the system performs computer clustering according to its service function, which can effectively satisfy needs for a reliable, scalable, stable, and operationally efficient system.

6. Conclusion

6.1. Detailed Measures for the Digitalization of the Texts of Dunhuang's Sutra Cave

Regarding the specific implementation plan for the “digital recovery” of the texts of the Dunhuang Mogao Grottoes' Sutra Cave, the first step is to establish a catalogue database of literature of Dunhuang's Sutra Cave and build a database sharing platform for domestic and foreign experts to search, sort, summarize, and classify resources to lay the foundation for resource sharing in the future “Digital Dunhuang”. This project will develop and explore technology for the platform, and explore feasible implementation routes for the industry. Secondly, this project represents the integration of the Tunhuangology Catalogue Database and the literature database of Sutra Cave, which will realize a one-stop search system for research catalogues, literature information archives, full-text databases of research literature, and retrieval of image resources between the two. A third goal of the system is to comprehensively and systematically sort and publish the status of Dunhuang literature collection, their preservation status, statistical information, etc., providing a systematic searchable database for the academic community to play its due role, rather than having these databases lie dormant in museums. As Dr. Wu Xin from the Metropolitan Museum of the United States said, “The emergence of the Dunhuang texts has changed the situation of Chinese historical research, and

scholars can restore the original appearance of these works through these texts”, [22] thus exploring the construction and sharing model of “digitally recovering” the literature of Dunhuang’s Sutra Cave that has been previously lost overseas, allowing experts, scholars, and members of the public from around the world to share these precious resources and actively promote the in-depth study of Tunhuangology.

6.2. The Significance of the Digital Recovery of the Dunhuang’s Sutra Cave Unearthed Texts

The significance of this project includes the comprehensive and systematic publication of Dunhuang literature from all over the world, revealing the relationship between the heritage value of Dunhuang literature and the Dunhuang Grottoes; allowing the Tunhuangology community to fully understand the relationship between the value of these precious cultural heritages and the related literature; further exploring how to better protect and utilize this precious resource and how to carry out international collaboration. This provides a shared platform for the academic community to use digital resources more widely, and realizes a true “digital recovery” of the lost texts of the Sutra Cave in Dunhuang’s Mogao Grottoes. It provides an effective way for international Dunhuang experts, scholars, and public to access resources quickly and easily.

In the digital age, the use of multimedia digital technology can help protect, utilize, and pass on this precious cultural heritage, and the digitalization of cultural heritage is an inevitable development trend in the information age. With the help of digital networking technology, information such as the fine murals of the Dunhuang Grottoes and the literature of Sutra Cave can be virtually reproduced or reconstructed, and researchers can easily access, compare, and collate resources without temporal or spatial constraints, and this will greatly promote the study of Tunhuangology and the promotion of Dunhuang culture [23]. It will also enhance the preventive protection of cultural heritage and promote the study of Tunhuangology, and play an important role in promoting the management of the world cultural heritage of the Mogao Grottoes and the international Tunhuangology community. The “Digital Dunhuang Project” implemented by Dunhuang Academy will integrate all the resources connected to the cultural heritage site of the Dunhuang Mogao Grottoes. The first step is to let the cultural relics of Dunhuang’s Sutra Cave “digitally return” to their source of origin, the Dunhuang Mogao Grottoes, which gives the site the respect it deserves. Secondly, through the joint efforts of international organizations and Tunhuangology community, we will truly realize the “recovery” of physical cultural relics, and will be the most exciting moment in the international Tunhuangology community. This will allow the splendid cultural heritage of Dunhuang to be shared by all.

Today, we are implementing the “Digital Dunhuang Project”, which will integrate the unearthed literature of Sutra Cave in Dunhuang, including the “digital recovery” or “digital return” of literature from the Sutra Cave that has been lost overseas. We will use the most advanced digital and information technologies to implement the digital protection and dissemination of cultural heritage, which will undoubtedly bring our long history and cultural heritage to life

more than ever before. Our existing reserves of precious cultural heritage will also help to spread this deep, exquisite, historical, and brilliant culture to the rest of the world, giving the world a greater respect for China.

Acknowledgements

This article is one of the phased achievements of “Tunhuangology Academic Data Collection and Research” (17ZDA213), a major project of the National Social Science Fund of China, and is one of the research results of “The Unearthed Texts of the Sutra Cave in Dunhuang and Database Project of Literature Research Catalogue”, a project of China Dunhuang Grottoes Conservation Research Foundation.

References

- [1] Yunhe, Pan and Jinshi Fan. (2003). Chapter 1. The Reality and Virtuality of Dunhuang. Hangzhou: Zhejiang University Press. p23.
- [2] Jinshi, Fan. (2006). Management and Monitoring of World Cultural Heritage Based on World Cultural Heritage Value: A Case Study of Dunhuang Mogao Grottoes. *Dunhuang Studies*. 6 (J), p1-5.
- [3] Shengping, Xia. (2008). Collection, Arrangement and Publication of Dunhuang Literature in the Past Twenty Years. *Dunhuang Studies*. 5 (J), p51-58.
- [4] Yanke, Chen (1930). Chen Yuan’s Record of the Robbery of Dunhuang, *Research Journal of Historical Language*. (1).
- [5] Jianhong, Chai. (2000). The Past and Future of Tunhuangology. In: Jiayu, Song and Zhong, Liu *British Collection of Dunhuang Chinese and Tibetan Literature Studies-Commemorating the 100th Anniversary of Dunhuang Literature Discovery*. Beijing: China Science Press.
- [6] Yujie, Chen. (2018). Thoughts on Means of Returning Archives Lost Overseas. *Archives*. 4 (J), p21-24.
- [7] Chanxue, Tan. (1999). Dunhuang Literature and Three Proofs Donated by Aoyama Sanu. *Dunhuang Studies*. 2 (J), P47-54.
- [8] International Dunhuang Project. Available: <http://idp.nlc.gov.cn/>. Last accessed 15th Jan 2011.
- [9] Translated by Hong, Wei and Shitian Lin. (2005). Digital Dunhuang, Zebei and Xuelin-Commemorating the 10th Anniversary of the International Dunhuang Project (IDP). *Journal of The National Library of China*. 52 (2), p34-38.
- [10] Jinshi, Fan. (2009). Application of Digital Technology in the Protection and Display of Dunhuang Grottoes. *Dunhuang Studies*. 6 (J), p1-3.
- [11] National Ancient Books Protection Work Plan during the 13th Five-Year Plan Period. Available: <https://mp.weixin.qq.com/s>.
- [12] Xiantang, Zhang and Shuyan, Gu. (2015). Dunhuang Academy and the National Library of France Accelerating the Digitization of Dunhuang Posthumous Writings. Available: <http://www.chinanews.com/cul/2015/04-27/7236984.shtml>. Last accessed 27th Apr 2015.

- [13] Yongjun, Xiao. (2018). The French Collection of Dunhuang Legacy Text Resources Can Be Read Online. Available: https://guoxue.ifeng.com/a/20180309/56599936_0.shtml. Last accessed 9th Mar 2018.
- [14] Yi, Zhu. (2018). Digital Recovery Realized for Overseas Dunhuang Literature. Available: http://www.cssn.cn/zx/bwyc/201809/t20180921_4567026.shtml. Last accessed 21th Sep 2018.
- [15] Huawen, Bai. (2004). Brief Introduction to the Establishment and Development of Dunhuang Catalogue Work in China. *Journal of Dunhuang and Turfan Studies*. 7 (J), p163.
- [16] Shaoliang, Zhou. (2000). Chinese Scholars' Contribution to Dunhuang Literature Cataloging. In: Jiayu, Song and Zhong, Liu *British Collection of Dunhuang Chinese and Tibetan Literature-Commemorating the 100th Anniversary of Dunhuang Literature Discovery*. Beijing: China Science Press.
- [17] Chunwen, Hao. (1998). Retrospective and Prospects of Dunhuang Literature and History Research. *Historical Research*. 1 (J). Jinbao, Liu. (2009). Significant Progress in the Compilation of Dunhuang Literature in the Past 30 Years. *Social Science Front*. 9 (J), P122-126.
- [18] Guangchang, Fang. (2006). Dunhuang Included in the Literary Catalog of Tibetan Buddhism. *Dunhuang Studies*. 3 (J), p86-99.
- [19] Pinting, Shi and Huili, Tai (2000). *New edition of the Dunhuang Legacy Writings*. Beijing: Zhonghua Book Company.
- [20] Official Website of the Tunhuangology Information Resource Library. Available: <http://dh.dha.ac.cn/>.
- [21] Hua, Liu; Yajie, Zhao; Daqing, Gao and Shucheng, Zhao. (2003). Design and Structure of Meta-data Standards for Tunhuangology Digital Library Testament. *Journal of Shanghai Jiaotong University*. S1. Ping, Wang; Jiezheng, Pi; Chunyun, Hao; Weiguo, Lu and Xiaopeng, Yuan. (2006). Construction Plan Design for Digital Dunhuang. *New Technology of Library and Information Services*. 11. Hong'en, Li and Shiqing, Wang. (2007). Meta-data Design Principles of Tunhuangology Digital Library. *Dunhuang Studies*. 3.
- [22] Wei, Cai. (2007). The Background of Dunhuang Cultural Relics in Russia. *Life Weekly*. 11 (J), P60.
- [23] Shengping, Xia. (2011). Digital Age: The Heritage Protection of World Cultural Heritage Site, Dunhuang Mogao Grottoes and the Organization and Management of Information Resources. *Digital Library Forum*. 9 (J), p49-59.