



Article

# Methods and Application of Archeological Cloud Platform for Grand Sites Based on Spatio-Temporal Big Data

Yongxing Wu <sup>1,2</sup> , Shaofu Lin <sup>2,3</sup>, Fei Peng <sup>1</sup>  and Qi Li <sup>1,\*</sup>

<sup>1</sup> Institute of Remote Sensing and Geographic Information Systems, Peking University, Beijing 100871, China

<sup>2</sup> Beijing Advanced Innovation Center for Future Internet Technology, Beijing University of Technology, Beijing 100022, China

<sup>3</sup> Institute of Smart City, Beijing University of Technology, Beijing 100022, China

\* Correspondence: liqi@pku.edu.cn

Received: 24 July 2019; Accepted: 24 August 2019; Published: 29 August 2019



**Abstract:** Grand sites are important witnesses of human civilization. The archeology of grand sites has the characteristics of a long period, interdisciplinary study, irreversibility and uncertainties. Because of the lack of effective methods and valid tools, large amounts of archeological data cannot be properly processed in time, which creates many difficulties for the conservation and use of grand sites. This study provides a method of integrating spatio-temporal big data of grand sites, including classification and coding, spatial scales and a spatio-temporal framework, through which the integration of archeological data of multiple sites or different archeological excavations is realized. A system architecture was further proposed for an archeological information cloud platform for grand sites. By providing services such as data, visualization, standardizations, spatial analysis, and application software, the archeological information cloud platform of grand sites can display sites, ruins, and relics in 2D and 3D according to their correlation. It can also display the transformation of space and time around archeological cultures, and restored ruins in a 3D virtual environment. The platform provides increased support to interdisciplinary study and the dissemination of research results. Taking the Origin of Chinese Civilization Project as a case study, it shows that the method for data aggregation and fusion proposed in this study can efficiently integrate multi-source heterogeneous archeological spatio-temporal data of different sites or different periods. The archeological information cloud platform has great significance to the study of the origin of Chinese civilization, dissemination of Chinese civilization, and the public participation in archeology, which would promote the sustainable development of the conservation and use of grand sites.

**Keywords:** grand site; spatio-temporal big data; archeology; cloud computing; Chinese civilization

## 1. Introduction

Grand sites refer to representative sites from various archeological cultures, dynasties or various historical ethnic regimes, which are usually very large in area and have important historical, scientific, artistic, and cultural values [1–3]. Grand sites are important witnesses of the 5000-years of Chinese civilization, which are of great significance to study the origin and development of Chinese civilization. The grand sites in China are characterized by a long history, wide distribution, large quantity, complex types, and a large scale [4]. Grand sites are facing a series of threats such as natural disasters, urban and rural construction, agricultural production, large-scale infrastructure construction, etc. [5], which places the conservation of grand sites under great pressure. The conservation and use of grand sites is generally divided into archeology, conservation planning, exhibition and use, monitoring, and so

on. Archeology is crucial to the conservation of grand sites [1], because the data of archeological investigation and excavations represent the basis for conservation and use.

During archeological investigation and excavation, the geospatial data and descriptive information of particular stratum, ruins and relics will be recorded in detail and sorted. Due to many factors such as strenuous research tasks and the complexity of data processing, less than half of the available data is published within two years after the archeological excavation [6]. With the powerful capabilities of data acquisition, processing, storage, analysis, and visualization, GIS provides effective support to the processing, database building and use of archeological data. In addition, GIS has already been used as an important tool in archeological research, such as settlement analysis and prediction modeling [7,8]. Although GIS has shown obvious advantages in archeological prospection, excavation and other related research regarding data management, spatial analysis, and visualization, it lacks necessary control of data quality and standardization [9–13]. Most of archeological GIS systems are deficient in system interoperability and data sharing, because of stand-alone deployment, the lack of online access, and the lack of measures for data security and disaster recovery [14]. As more and more archeological excavations are undergoing, archeological data are accumulated in large quantities and cannot be processed in time. Therefore, archeological datasets are getting bigger and more complex, which are impossible for traditional GIS to process, store and analyze. A framework of archeological information system should be established from the perspective of archeologists, which is corresponding to what archeologists do in their daily work [15].

Archeological data of grand sites, which have obvious characteristics obtained from spatio-temporal big data, are essentially geo-spatial big data [16]. As key technology of the new generation of GIS, the deep integration of spatio-temporal big data and archeology will have a far-reaching impact on the conservation of grand sites and archeology, specifically manifested in the following four aspects: (a) From the view of archeological concepts and data analysis, the archeological spatio-temporal framework can integrate the structured and unstructured data as a whole, to better visualize, interpret, and understand the past history of human beings. (b) The archeological excavation reports are inevitably led by the subjectivity of archeologists, hence the reports don't fully reflect the real situation of the grand site. The above limitation of archeology and its influences could be reduced to a certain extent by using spatio-temporal big data [17]. (c) With the application of a new generation of information technologies (such as advanced surveying and mapping technologies, big data, Internet of Things, cloud computing) in archeology, the volume and varieties of archeological data has seen a dramatic rise, thus the data processing and mining will be more complicated. Spatio-temporal big data offers a new technology and methodology for data acquisition, processing, storage, organization, analysis, and representation. (d) Archeology is a discipline characterized by interdisciplinary research. The spatio-temporal big data of grand site will offer a new way for archeologists to get in touch with scholars of other disciplines and the public, promoting interdisciplinary research, communication, and collaboration. Spatio-temporal big data can radically transform archeological practice, fostering new research questions, novel data visualization techniques, and new competences. It will also give archeology an enhanced ability to investigate and address those significant questions [18].

Great progress has been made in digital archeology and the digital publication of archeological data [19–22]. Digital photogrammetry is used to carry out survey and mapping of archeological entities in fieldwork [23]. The laser scanning and photogrammetric reconstruction was used to make three-dimensional models of ruins (temples, monuments, etc.), which were published online [20,24,25]. The digitization and digital publication of archeological data have received great attention, while the data aggregation and fusion are very important [26]. Bayesian Neural Network (BNN) was used to establish fusion model of remote sensing data so as to improve the accuracy of archeological information identification [27]. Almost all the archeological data can be digitized, and platforms for spatial analysis and collaborative work were established to promote data sharing and web browsing, using WebGIS, etc. [28]. Projects of data aggregation and cultural heritage management were implemented, which aims to improve the capacities of conservation and promote data sharing [23,26,29,30]. Standardization

is the important task of these projects. It is also necessary to avoid making the data complicated and lengthy [31]. The technologies mentioned above has been proved to be valid and effective in data acquisition in archeology. Meanwhile, they will also generate a large amount of data, whose aggregation, fusion, and representation are very useful to reuse.

In the study of the origin of civilization, it is necessary to collect, process, and integrate archeological data of multiple sites. At the same time, it requires data sharing among scholars. Considering the distinguished features of Chinese grand sites, which often include rammed soil walls, further research on standards and specifications, as well as archeological data aggregation and sharing is required. In this paper, the challenges of archeological data aggregation and sharing was first discussed based on the analysis of the archeological research process, data flow, and the properties of archeological data of grand sites. Then, spatial scales of archeological research, and the classification and coding of archeological data of grand sites in Neolithic Age was proposed, in order to facilitate data acquisition, processing, and representation. The spatio-temporal framework of archeological data of grand sites was also proposed, for the sake of data fusion and spatio-temporal analysis, etc. Finally, the archeological information cloud platform of grand sites based on spatio-temporal big data was designed and built, which was applied to the Origin of Chinese Civilization Project. The methods and platform proposed in this study will promote the aggregation and sharing of archeological data and improve the work efficiency of interdisciplinary research. It will also enhance the scientificity and accuracy of the identification of site value and the interpretation of the past.

## 2. Challenges of Data Acquisition and Data Sharing

Grand sites have a large amount of historical archeological data, which have multi-source, heterogeneous characteristics and obvious spatio-temporal features, and the amount of data is still growing rapidly. Due to the above complexities of grand site archeological data and the lack of unified data processing standards and efficient software tools, it is difficult for grand site data to aggregate, open, and share.

### 2.1. Difficulties of Acquisition and Fusion of Archeological Data

After years or even decades of archeological excavations, many grand sites have accumulated large amount of archeological data. As a result of the limitations of technologies and methods in the past, a large quantity of archeological historical data has not been digitized, and it becomes an urgent task to process those historical data and build a database [32]. Moreover, lots of new archeological data are still being produced rapidly, as a large number of archeological excavations has been carried out every year. If there are no appropriate standards and tools for the acquisition, processing, database building, and use of archeological data, the more archeological excavations carried out, the more numerous and jumbled the data. Then, the data processing and application will face more difficulties.

The difficulties in archeological data aggregation and fusion of grand sites are the following: (a) it is difficult to integrate archeological data generated from different excavations of one grand site (b) it is difficult to integrate archeological data from different grand sites (c) it is difficult to integrate multi-source heterogeneous data at different spatial scales (d) archeological data of grand sites are growing too large, which provides difficulties to data acquisition, processing, and management. More and more technologies are used in archeology, more kinds of data are collected, and the frequency of data acquisition is increasing, therefore the volume of archeological data becomes bigger.

### 2.2. Difficulties of Archeological Data Sharing

Archeology is a discipline that studies the past such as ancient societies, human beings, relations between human and nature, etc., based on the objects left over from ancient human activities [33]. Because of the irreversibility of archeological excavations, the characteristics of interdisciplinary study and the uncertainties of cognition of grand sites [34], archeological data are the basis for different archeologists and researchers of other disciplines to study grand sites. For some grand sites, archeological data are even

the only way to perceive and understand them. Technically and methodologically speaking, there are the following difficulties in archeological data sharing of grand sites: (a) archeological excavation data resources have not been digitized in time; (b) lack of unified data standards and specifications hampers data integration from different sites; (c) an unified platform for data sharing is still in absence.

### 3. Methods for Aggregation and Fusion of Grand Site Archeological Data

The aggregation and fusion methods of spatio-temporal big data of grand sites include the spatio-temporal framework of grand site archeology, the spatial scales of grand site archeological research, the classification and coding of grand site archeological data, and the data acquisition, processing, and database building of grand sites.

#### 3.1. Spatial Scales of Archeological Research on Grand Sites

Archeological research is always related with spatial scales, which are the important characteristics of grand site spatial-temporal data and plays important role in archeological data fusion. The spatial scales of archeological data were studied, to select that at which scale archeological data should be aggregated and fused. Different researchers adopted different spatial scales to facilitate their research purposes [35–37], for instance, Chinese Neolithic culture can be divided into three contact zones or nine major regions according to the geographical environment and archeological findings [38,39].

As the main parts of grand sites, sites, ruins, and relics are the main archeological research objects. The archeological research of grand sites can be summarized as 5W1H (who, when, where, what, why, and how) [6]. According to the archeological research objects, research content and spatial distribution of grand sites, this study divided the archeological research of grand sites into seven spatial scales: relic, ruin, site, site groups, watershed (cultural area or region), nation, and globe. Archeological researchers can obtain different information and knowledge from different spatial scales, and obtain all-round three-dimensional information of research objects through multi-scale data integration. Table 1 shows the spatial scale division of archeological research on grand sites and the corresponding contents of archeological research.

**Table 1.** Spatial Scales of Archeological Study and Corresponding Content.

Serial Number	Scale	Corresponding Research Content
1	Relic	Relic type, manufacture/generation, use, material, process, time, implied meaning, etc.
2	Ruin	The function, layout, construction, abandonment process and reasons of the relics, including culture, life and social significance, etc.
3	Site	The formation, development, abandonment process and reasons of the site, the functional zoning of the site, the cultural, social and living conditions of the site in ancient time. the stratum structure of the site, the contribution of the site to the origin and development of civilization, etc.
4	Site groups	The spatial and temporal relationship between the sites, the communication between the sites, and the distribution of functions (such as central settlement and general settlement), etc.

Table 1. Cont.

Serial Number	Scale	Corresponding Research Content
5	Cultural area /region/Watershed	The origin, formation and development of civilization in the region, the distribution characteristics of sites in the region, the prediction of regional sites, etc.
6	Nation	The division of national cultural area, the comparison of civilizations in different watersheds, the study on the origin of Chinese civilization, etc.
7	Globe	A Comparative Study between Chinese Civilization and Other Civilizations

### 3.2. Spatial-Temporal Framework of Archeological Data of Grand Sites

The archeological spatio-temporal framework is the basis for discussing the pedigree of archeological culture. The division of archeological culture and the determination of its location and periods are the basic research of archeology. The scientificity and reliability of other archeological studies have much to do with the validity of the archeological spatio-temporal framework [40]. If the archeological spatio-temporal framework is wrong, it is almost impossible to get correct results of other archeological studies such as value evaluation of grand sites, relationships among different sites, etc. Table 2 shows the spatio-temporal framework of archeological culture established from the perspective of archeologists.

**Table 2.** Spatio-temporal Framework of Archeological Culture in Mountainous Areas of Western Liaoning Province during the Summer solstice and Warring States Periods [40].

Era	West (West of Nulu'er Tiger Mountain)	East (East of Nulu'er Tiger Mountain)
Late Warring States Period		Yan culture
Early and Middle Warring States Period	“Water Spring Ruins” “Jinggouzi Ruins” “The Ruins of Wudaohezi”	Late “Linghe Ruins”
From Western Zhou Dynasty to Spring and Autumn	Xiajiadian upper culture	Early “Linghe Ruins”
Late Shang Dynasty	“Wei yingzi type”	Wei yingzi type
Early Shang Dynasty		Late Xia Jia Dian Lower Culture
Xia dynasty		The Early Stage of Xiajiadian Lower Culture

Archeological data of grand sites, which are only valuable in specific time and space, have very distinct spatio-temporal properties. If the spatial or temporal information is lost, the value of archeological data will be greatly reduced [16]. The time in archeological research is divided into relative time (such as archeological culture or site staging) and absolute time (such as C14 dating and AD dating). Archeological space can be divided into horizontal space and vertical space, and also can be divided into modern space and ancient space. The horizontal space has both the current administrative divisions and the space in the sense of archeology. The vertical space is mainly stratum, which also has a sense of time. Generally, the lower the stratum, the earlier the time. The watershed, administrative division and grids of exploration are modern spaces, which mainly serve as spatial indexes. Ancient spaces include cultural area, site boundary, relics and so on, which are the focus of archeological research. In order to realize the archeological unification of the physical space, the information space and the cognitive space, it is necessary to establish an integrated spatio-temporal framework from the perspective of big data. The framework would not only conform to the archeological spatio-temporal framework from the perspective of archeologists, but also satisfies the processing, database building,

and visual representation of archeological big data. Table 3 is the spatio-temporal framework of archeological spatio-temporal big data that is established from the perspective of big data.

**Table 3.** Spatio-temporal Framework of Archeological Data of Grand Sites.

Time Expression	Spatial Scales
Archeological culture staging	Administrative division
Site staging	Geographical area/basin
Dating	Cultural area
Years ago	Site scope
The Year of Cadres and Branches	Site division
Imperial calendar	Site functional area
Dating of dynasties	ruin (Grid of excavation)
C14 dating	Relic

In the spatio-temporal framework of Table 3, relations of different times, spatial associated relations and spatio-temporal associated relations were established. A four-level modern spatial index sequence of “watershed, province (municipality directly under the central government), city and county” and a five-level archeological spatial sequence of “archeological culture, site boundary, site functional area, ruin and relic” are established. Other modern spatial indexes also include site zoning, grids of excavation, etc. The three spatial sequences above are independent, and also can be used in an integrated way. They provide macro to micro forms of spatial logic, such as “nation, province (municipality directly under the central government), city, county, site, site division/functional area, ruin, relic”.

The archeological spatio-temporal framework from the perspective of big data is not only the basis for data acquisition, data processing, data storage and database building, but is also the basis for data visualization. It can help to trace back the archeological excavation process of grand sites and restore the grand sites in the information space with the spatio-temporal concept used by archeologists.

### 3.3. Classification and Coding of Archeological Data

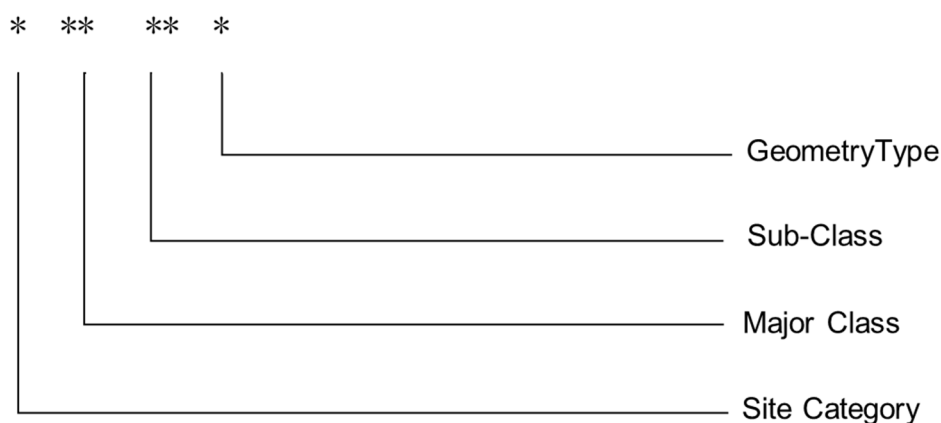
Data classification and coding is mainly used for data acquisition, data storage, data management, data retrieval and exchange, which is one of the key issues to realize information exchange, integration and sharing within and between systems. In the process of data acquisition, coding can be used as the identification of thematic data types, and can also be used to check the accuracy and integrity of data, to modify or reorganize the data layer [41]. Archeological data classification and coding will classify the data according to its spatio-temporal characteristics, attributes, and contents, and then formulate rules to code the classified data. According to the data classification and coding, the features, categories, correlation, and basic attribution of the data are determined.

In this paper, the linear classification method was used to classify and code the archeological data of Neolithic sites from 3500 BC to 1500 BC. According to the whole process of archeological research and the conservation of sites, the data are divided into seven major classes (Site location, auxiliary positioning, site functional areas, ruins, relics, literature, images and video). Each class can be divided into several sub-classes. Table 4 is the description of classification.

**Table 4.** Classification of Archeological Data of Grand Sites.

Coding	Major Class	Sub-Class
1	Site location	Site, site boundary
2	Auxiliary positioning	Site zoning, site grid, origin of coordinates, exploration methods, etc.
3	Site functional area	Tomb area, residential area, palace area, handicraft workshop area, etc.
4	Ruins	Palaces, houses, tombs, pits, wells, roads, etc.
5	Relics	Pottery, jade, gold ware, animal bones, etc.
6	Literature	Archeological diaries, journal articles, newspaper articles, archeological excavation reports, degree papers, etc.
7	Images and videos	Plans, sections, photos, videos, etc.

The coding of the established classification system is to adopt six-digit coding, and the coding and classification are in one-to-one correspondence, as shown in Figure 1.

**Figure 1.** Code Rules for Site Data (The sign “\*” represents a digit.).

The coding rules for site data are as follows:

- The first digit from the left is the category code, with “1” representing the ancient Neolithic site. Numbers 2–9 are reserved for subsequent research on data classification and coding of other types of sites.
- The second and third digits from the left represent major categories.
- The fourth and fifth digits from the left represent the sub-classes.
- The last digit represents the geometric type of the class. Number 0, 1, 2, 3, and 4 represent the non-geometric type (or the geometric type does not need to be considered), point type, polyline type, polygon type and volume type, respectively.
- For categories without the next level classification, the coding bits corresponding to the lower level classification are filled with “0”, but the geometric type of the last element is filled according to the actual element type, not necessarily “0”.

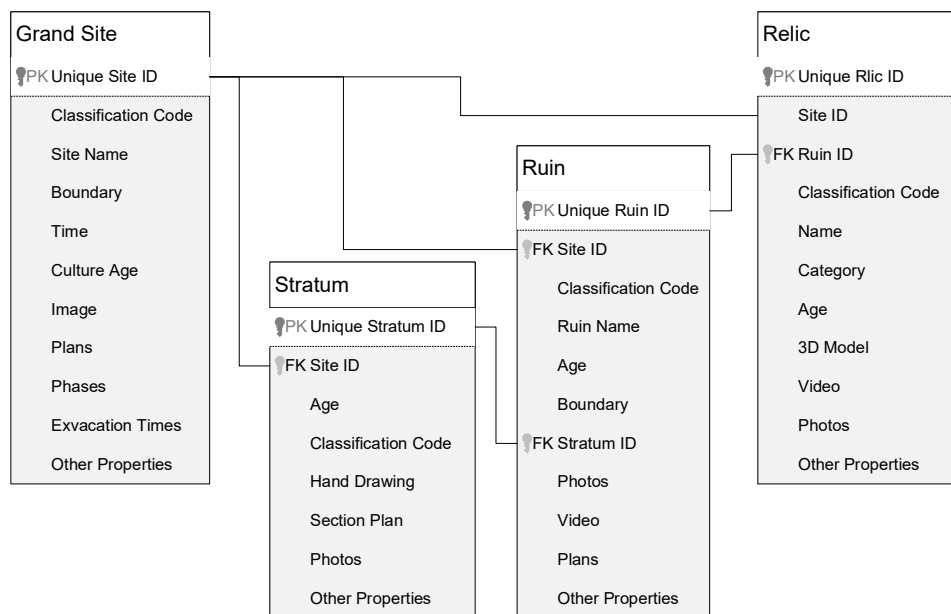
Examples of classification and coding of grand site data in the scale of 1:10~1:100 are shown in Table 5.

**Table 5.** Examples and descriptions of classification and coding of Neolithic site data with scale of 1:10~1:100.

Major Class	Sub-Class	Sub-Class Coding	Explanation
Relics	Pottery	105010	“1” Neolithic site, “05” means relic, “01” means pottery, “0” means no geometric type
	Jade article	105020	
	Human ruins	105040	
Ruins	Palace	104013	“1” Neolithic site, “04” means ruin, “01” palace, “3” polygon
	Well	104033	“1” Neolithic site, “04” indicates ruin, “03” well, “3” polygon
	Pit	104043	

3.4. Processes of Archeological Data Acquisition, Processing and Fusion

Original archeological records and collected archeological data have various formats (forms, written records, maps, photos, video, etc.). Organization and processing of archeological data is an important step to realize the conversion from data to information and its continuous use. Figure 2 shows the spatial association among sites, ruins and relics and the basic information of which to be recorded. Based on the aforementioned spatio-temporal framework of archeological data, classification and coding, spatial scales of archeological research and the spatial association among archeological data shown in Figure 2, this study proposed the process of acquisition, processing and fusion of archeological spatio-temporal big data of grand sites (Figure 3).



**Figure 2.** Information recorded and relations of sites, stratums and ruins and relics.



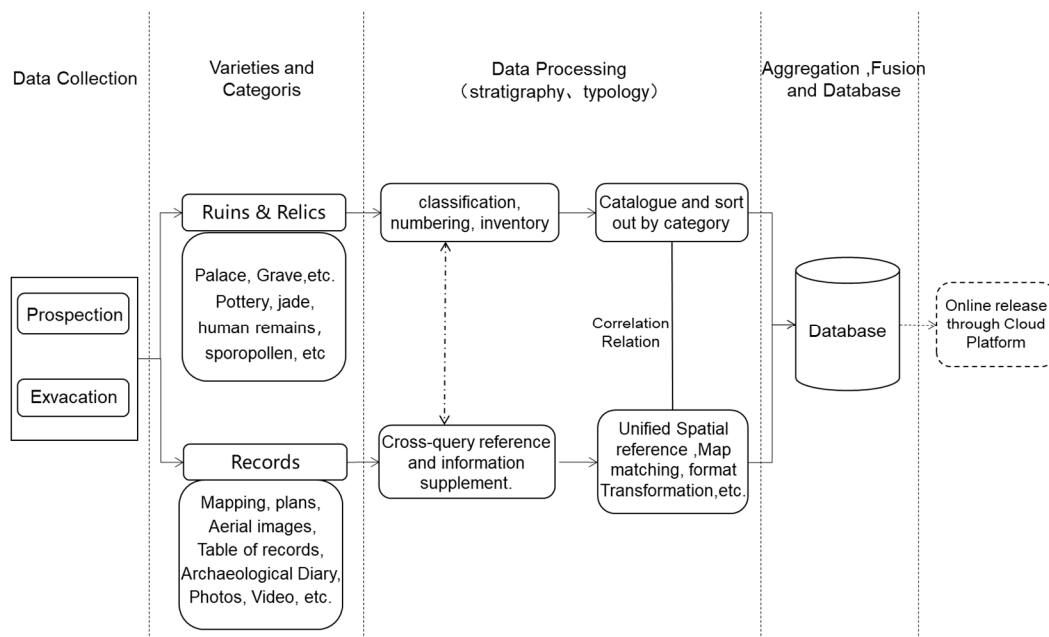


Figure 3. Process of archeological data acquisition and fusion of grand sites.

#### 4. Construction of Archeological Information Cloud Platform for Grand Sites

##### 4.1. Archeological Spatio-Temporal Data of Grand Sites

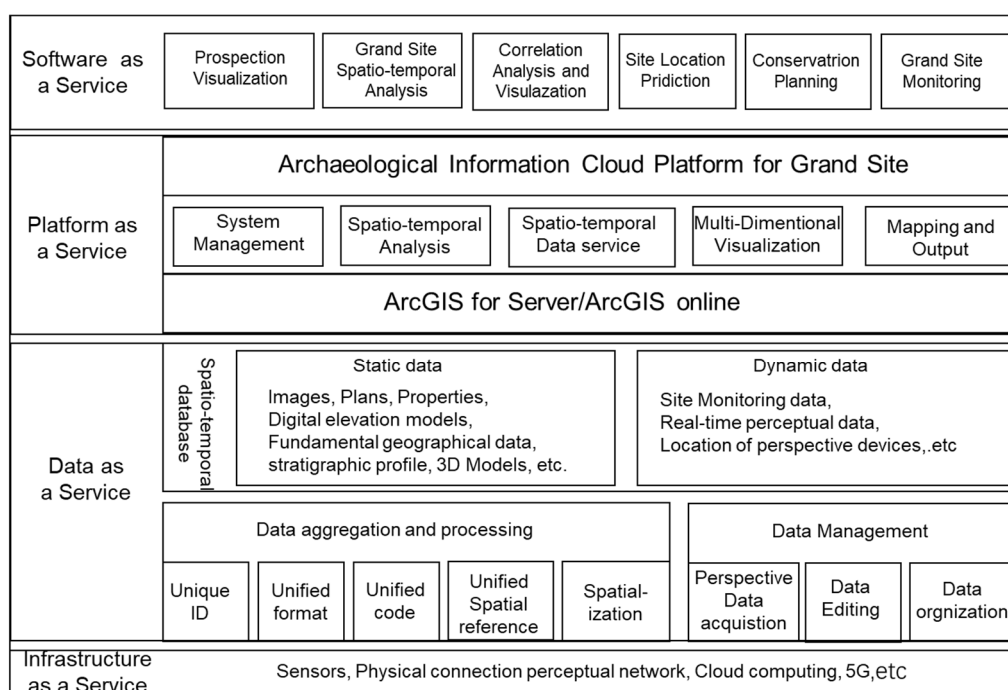
The archeological spatio-temporal database of grand sites mainly includes fundamental geographic data covering the site and nearby, data generated by archeological excavation and investigation, environmental information data, three-dimensional model data, etc.

- Fundamental geographic database.** The fundamental geographic data include four categories—digital elevation model, topographic map, satellite images and aerial images. The scales of vector data covering the grand site and nearby is 1:250,000 and 1:10,000, as well as 1:500 of core area of grand site. The raster data include ETM with resolution of 30 m, SPOT images with resolution of 2.5 m, QuickBird images with resolution of 0.6 m, and aerial images with 0.1 m, which covered different periods.
- Archeological database.** Under a unified spatio-temporal framework, a spatio-temporal database of grand sites is created, which is mainly about data of sites, ruins, relics and spatial associations among them. Specifically, the database includes site boundaries, site functional areas, site plans, site description information, data of ruins and relics, and so on.
- Environmental information database.** The environmental information database includes the environmental data of the site and nearby areas, such as the water system, residential area, vegetation and landform, etc.
- Three-dimensional model library.** Based on the data of archeological excavation (sites, ruins and relics), three-dimensional models are constructed under the guidance of archeologists. The archeologists, who carried out the excavations, have a good understanding of archeological excavation data and literature data. After years of studies, they know what the ruins were, how they were built, how they were used, and so on. With three-dimensional models and the interpretation of the sites given by archeologists, the past of the site can be restored in the virtual environment.

##### 4.2. Architecture of Archeological Information Cloud Platform for Grand Sites

By using centralized management and intelligent scheduling of computing resources, cloud computing conveniently and dynamically provides users on-demand with services such as computing,

storage, application software, and data through the network [42]. Based on the above characteristics and advantages of cloud computing, the archeological information cloud platform for grand sites is conducive to the formulation and implementation of unified archeological data standards and specifications. The platform can provide better system security and data security strategies, more convenient software services and data services, and technically provide online data access for multidisciplinary researchers and the public. Besides, the archeological information cloud platform allows archeological institutions to focus on research and reduce investment in hardware facilities, application software, data and platforms, etc. The platform can improve the overall information level of the archeology industry. Figure 4 is the logical architecture diagram of the archeological information cloud platform for grand sites. The architecture is divided into four layers—infrastructure layer, data layer, platform layer, and application layer.



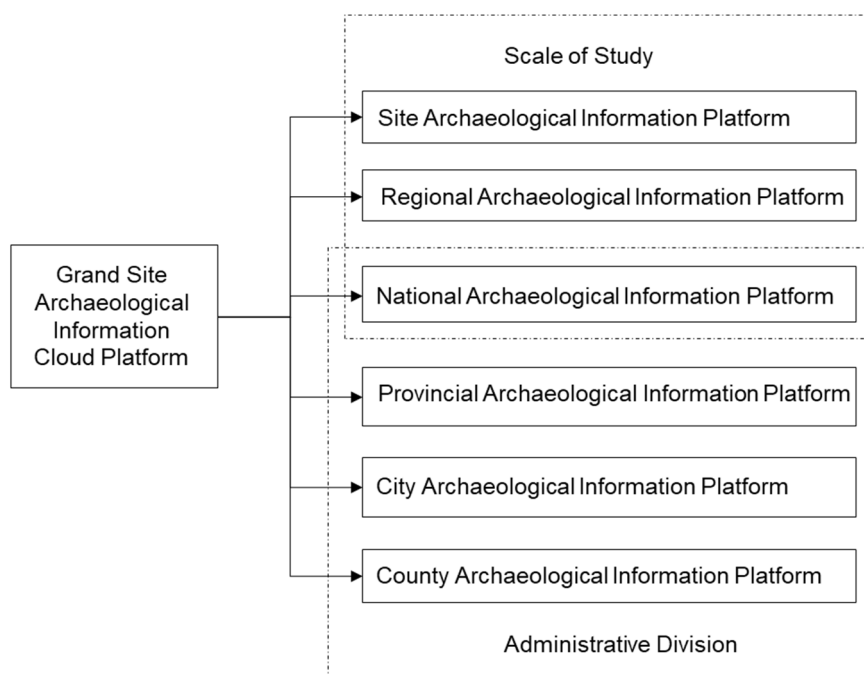
**Figure 4.** System Architecture of Archeological Information Cloud Platform for Grand Sites.

According to the content, process and complexity of archeological research, the archeological information cloud platform for grand sites must meet the following three needs. First, it can conduct data acquisition, editing, management, inquiry, statistics, mapping, and output of archeological data. Second, it has the ability to carry out spatial analysis on archeological data to transform data into information, and then information into knowledge. It provides quantitative analysis and auxiliary decision support for archeological research. Third, grand sites can be displayed dynamically and multi-dimensionally. Through multi-scale dynamic display of the site, ruins and relic data, the spatial form and settlement environment of the grand site can be visually displayed, and the key ruins can be restored in three dimensions through modeling.

#### 4.3. Archeological Information Cloud Platform Deployed on Demand

The archeological information cloud platform for grand sites can be logically divided into different archeological information platforms on demand. Based on the various types of cultural heritage administrative departments, the archeological information cloud platform can be logically divided into different platforms through the control of data authority and function authority. The platform can be divided into the national archeological information platform for grand sites, the provincial (municipality directly under the central government) archeological information platform for grand sites,

and the municipal and county archeological information platform for grand sites, etc. The platforms meet the management requirements of the administrative departments of cultural heritage at all levels (Figure 5).



**Figure 5.** Archeological information cloud platform logically divided on demand.

Many sites in a certain region need to be researched for some kinds of archeological studies, such as the studies of the origin of civilization in the Yellow River basin, the origin of civilization in the middle and lower reaches of the Yangtze River basin, and the formation and development of Erlitou culture, etc. Therefore, an archeological information platform is needed to aggregate data of different grand sites in the same platform, so as to support the interdisciplinary study. For this kind of archeological study, the above-mentioned archeological information cloud platform can play a very important role. It can be logically divided into different regional archeological information system on demand, for example, the Yangtze River basin archeological information platform. There is no need to build a new archeological information system from scratch, which saves lots of time and investment funding, and fully reflects the flexibility of the cloud platform.

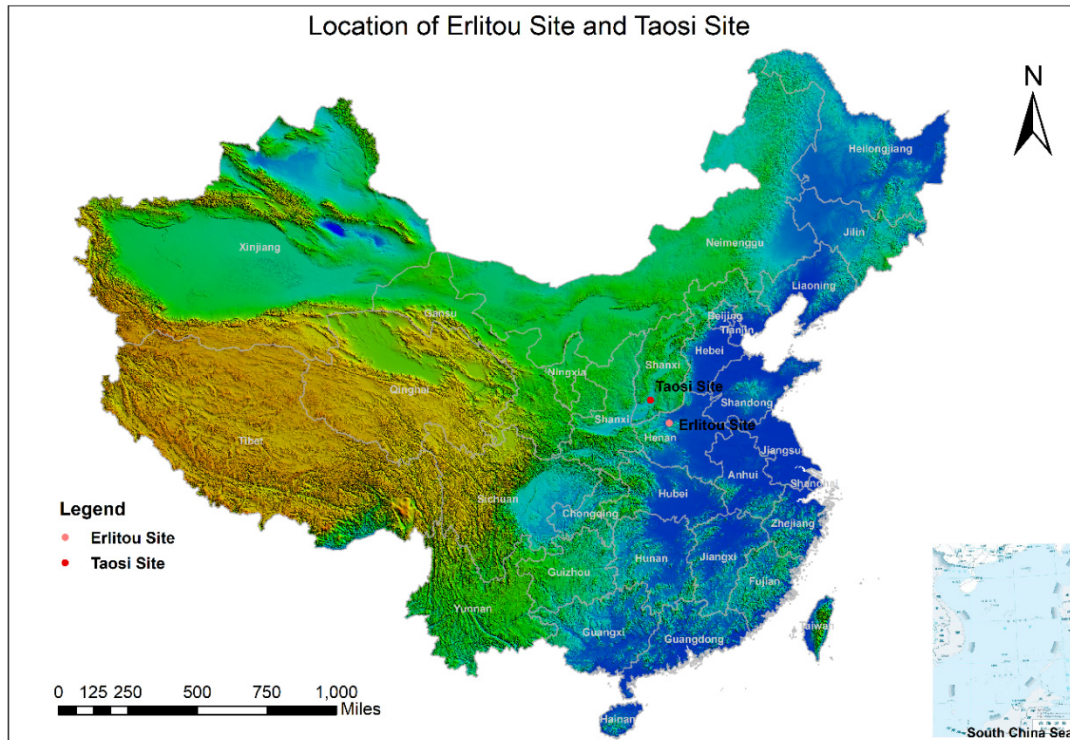
## 5. Case Study

“The Origin of Chinese Civilization Project” is a major scientific research project in the field of Chinese history and culture. It studies the origin, formation, and development of Chinese civilization and explores the background, causes, nature, and characteristics of the origin of Chinese civilization. The project takes archeology as the core, and carries out research on the origin of Chinese civilization from different angles and levels, as well as all aspects through a multidisciplinary approach.

### 5.1. Multi-Source Heterogeneous Data Fusion of Different Sites

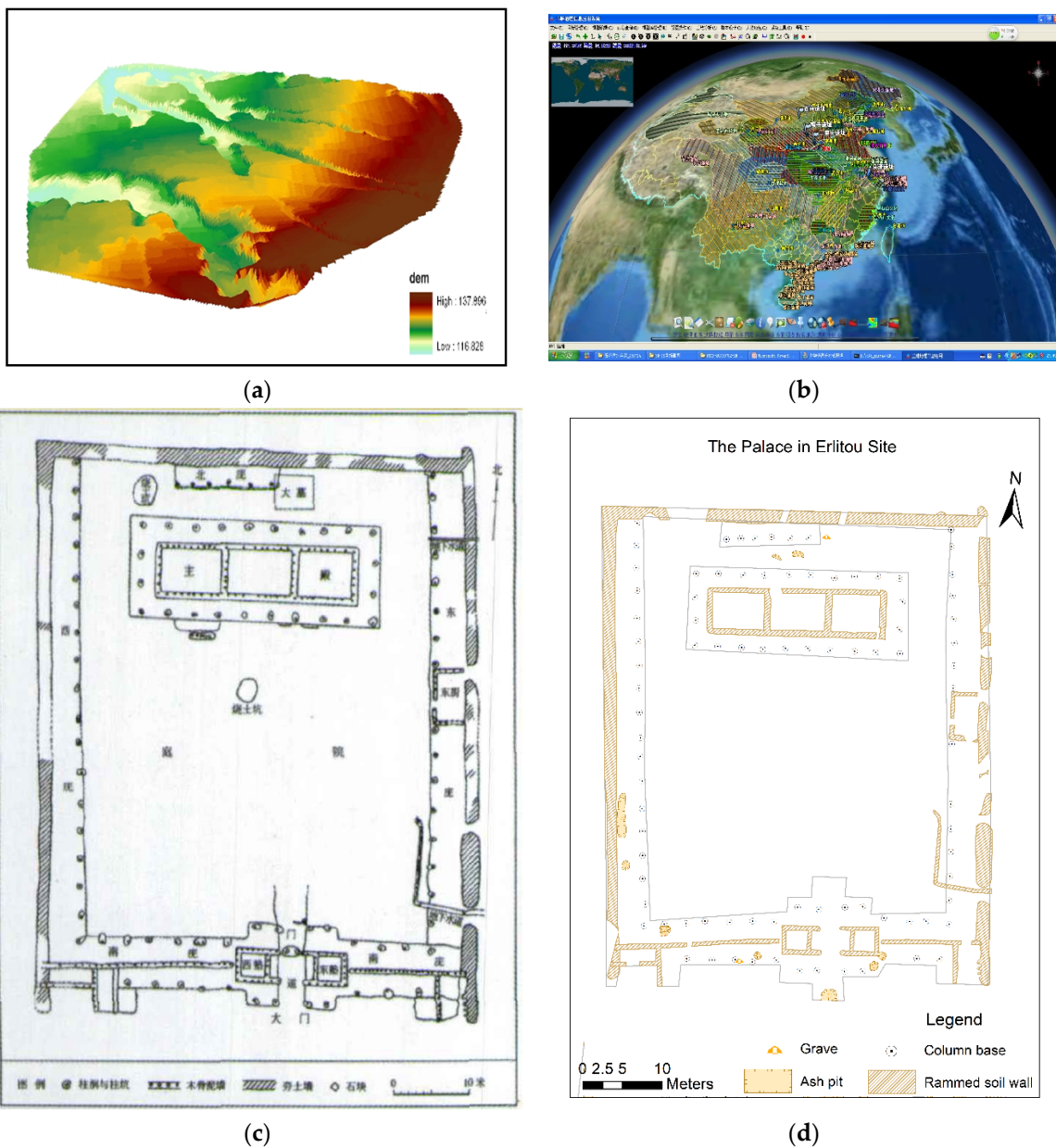
Erlitou Site, divided into four stages, covers an area of about 3 million square meters, and its duration is from 1750 BC to 1500 BC [43]. The site is located in Luoyang Basin of Yiluo River basin, Erlitou Village, Yanshi County, Henan Province. Since its first discovery in 1959, the Erlitou site has gone through dozens of archeological excavations [44]. Taosi Site covers an area of more than 3 million square meters, and it is located in Taosi Township, Xiangfen County, Shanxi Province. Through C14 dating, it is concluded that the age of Taosi site is from 2500 BC to 1900 BC [45]. It is an important

large-scale city site in Linfen Basin in the middle reaches of the Yellow River and the site is divided into early, middle, and late stages [46]. The first archeological excavation of Taosi site was carried out in 1978, and dozens of archeological excavations have been carried out so far. Figure 6 shows the geographical location of Erlitou site and Taosi site.



**Figure 6.** Location of the Erlitou Site and Taosi Site in China.

According to methods of archeological spatial-temporal data aggregation and fusion highlighted above, such as archeological data spatio-temporal framework, classification and coding, spatial scales and so on, archeological data of Erlitou and Taosi sites was digitized and fused, following by database building. Data sources include satellite images, aerial images of unmanned aerial vehicles, 1:500 DEM (digital elevation model), CAD format site plans, and so on. Lots of data were collected from archeological excavation reports of the two sites, such as ruin plans, ruin properties, stratigraphic profiles, relic properties, and relic photos, etc. All the data of sites, ruins, and relics were given time information. Erlitou is divided into stages one to four, and Taosi is divided into early, middle, and late stages. Ruins and relics were classified into different tables according to the classification and coding proposed in Section 3.3. Meanwhile, the same spatial reference was used to establish the spatial affiliation among sites, ruins, and relics. In data fusion representation, data with different precision or scale are displayed according to different spatial scales. Figure 7 shows the different kinds of data collected and processed in this study. Figure 7c is the raster plan of the Palace, which is the most important ruin of the Erlitou Site. Figure 7c was digitized and processed according to the methods mentioned above, and the result was vector data with properties, which was represented as Figure 7d.



**Figure 7.** Display for the integrated data of Erlitou Site and Taosi Site. (a) 1:500 digital elevation model of Taosi Site. (b) All the data displayed in same Globe. (c) Palace plans in archaeological report [47]. (d) Palace plans after digitized and classified.

### 5.2. Multi-Scale Correlation Display

According to the system architecture in Section 4.2, based on ArcGIS Server and ArcGIS online, an archeological information cloud platform of grand site was developed. The default page of the platform is shown in Figure 8.

The main functions of the platform include browsing multi-dimensional spatial information of grand sites, query statistics, archeological spatial analysis, spatio-temporal analysis, archeological mapping, etc. It can provide researchers with support services such as correlation analysis and display of sites, ruins and relics, data mining of archeological excavation data in different periods, analysis of man-land relationship, restoration of sites and comparative analysis of the conservation of different sites in the same period, etc. Figure 9 is a correlation display of archeological data at different spatial scales (sites, functional areas of sites (zones), ruins and relics).



Figure 8. The Default page of Archeological Information Cloud Platform for Grand Site.

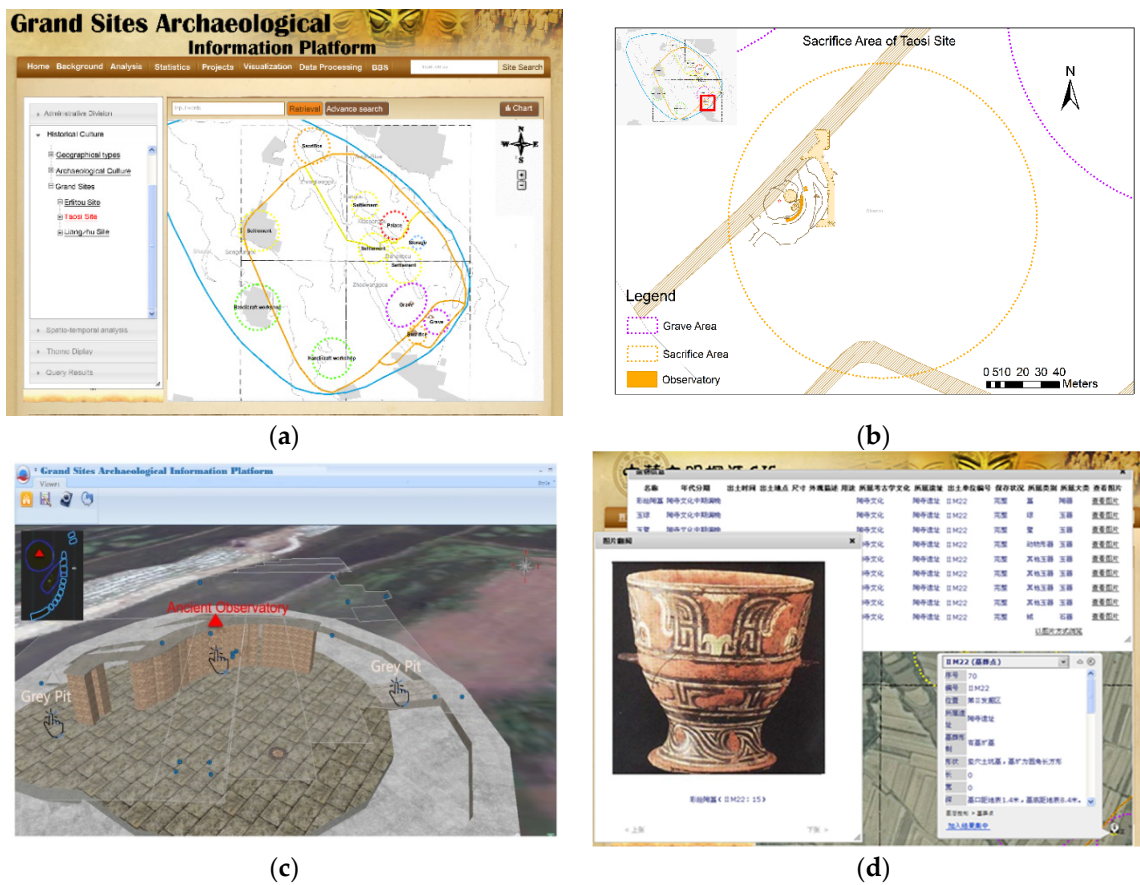


Figure 9. Archeological Information Display of Grand sites at Different Spatial Scales. (a) Plans of Taosi Site (Site Scale). (b) Sacrifice Area of Taosi Site (Functional Area Scale). (c) The observatory in the sacrifice area (ruin scale). (d) All relics discovered in in the ruin IIM22 (relic scale).

Figure 9a is the plans of the Taosi site, which shows the overall information of the Taosi site at the site spatial scale, including the site boundary, site functional areas (palace area, sacrifice area, burial area, residential area, and handicraft area), etc. Figure 9b is a plan view of the sacrificial function area of the Taosi site, showing the plan layout of each ruin in the sacrificial area. Figure 9c is a three-dimensional restoration of the ruin called observatory in the sacrificial area, and Figure 9d is a display of all the relics discovered in the ruin coded IIM22. The display includes a list of relics information, photos of relics, three-dimensional models of relics, etc.

Through the multi-scale correlation display, it is easy to understand the spatial association and subordinate relationship among sites, ruins and relics, and to know which ruins and relics were found in the grand site, as well as the layout of the grand site. Meanwhile, this kind of display also provides a good way for researchers and the public to understand the process of archeological excavation of grand sites and what has been discovered. It will encourage more people to pay attention to grand sites and to participate in the conservation of grand sites [48]. The multi-scale correlation display helps to restore sites and ruins in virtual environment, to evaluate the value of grand sites, and to draw up more scientific conservation plans, which is a great help for dealing with all kinds of threats faced by grand sites.

### 5.3. Spatio-Temporal Evolution Analysis Based on Archeological Cloud Platform

Yangshao culture, Longshan culture, and Hongshan culture are important parts of the broader Chinese culture and history. The spatio-temporal analysis of archeological cultures can dynamically show the spatial distribution and influence of archeological culture in a specific period of time. It is helpful for researchers to intuitively understand the transformation process of archeological cultures and to analyze the origin and development law of Chinese civilization. Figure 10 shows the spatio-temporal evolution of archeological cultures in China from 7000 BC to 2000 BC (The time step interval is 1000 years) [37].

As can be seen from Figure 10, the Chinese civilization originated from the Yellow River basin, the Yangtze River basin and the Songhua River basin, and from 5000 BC to 2000 BC, the Chinese civilization showed a multi-point concurrent trend. Based on spatio-temporal data, the platform can simply, conveniently, and intuitively show the spatio-temporal evolution of archeological cultures and the spatial relationship of different archeological cultures (distance, partial overlap, etc.).

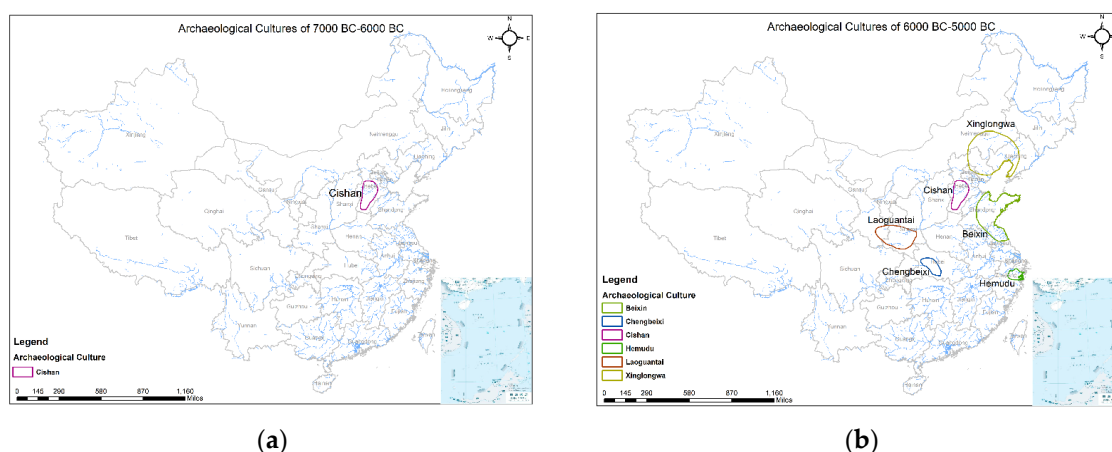
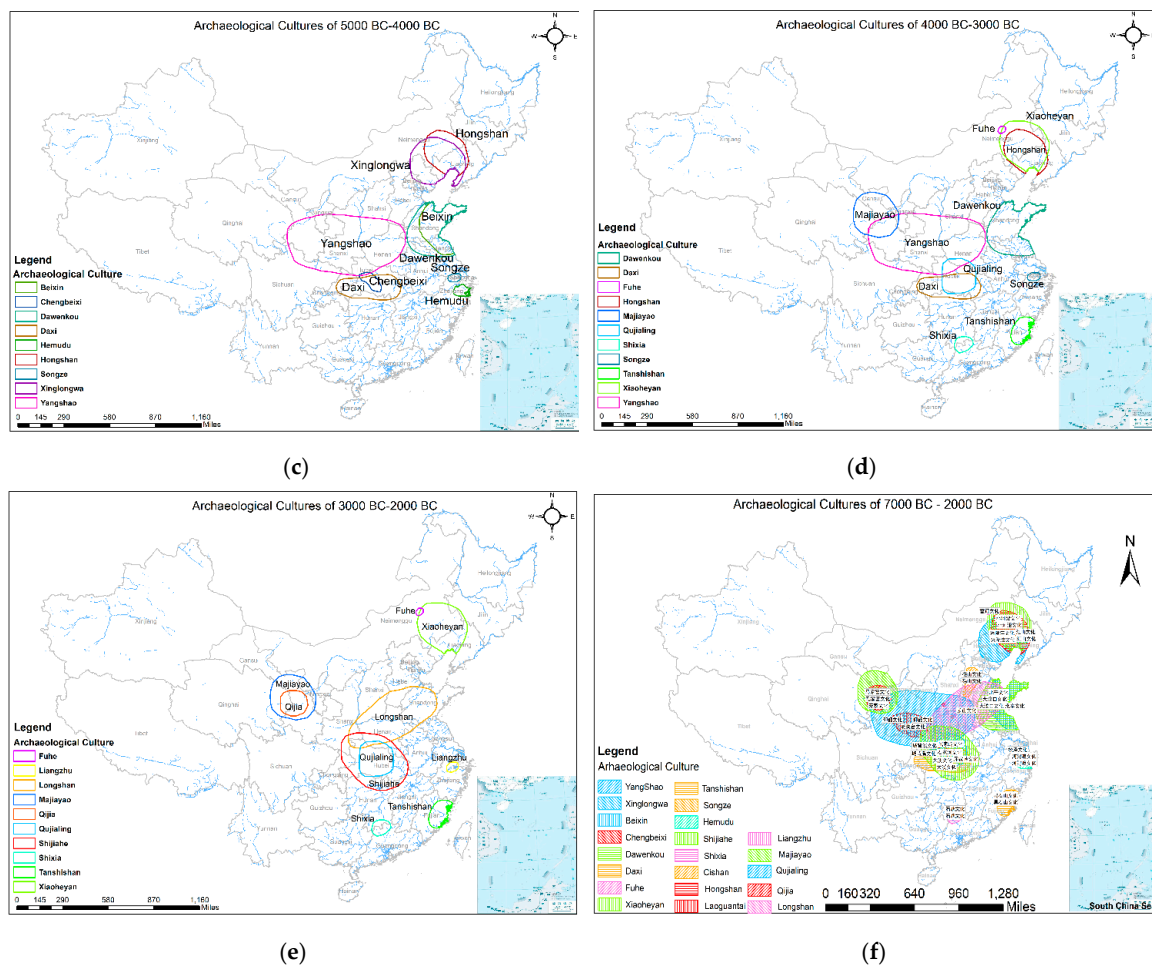


Figure 10. Cont.



**Figure 10.** Spatio-temporal Evolution of archaeological cultures from 7000 BC to 2000 BC. (a) Archaeological Cultures from 7000 BC–6000 BC. (b) Archaeological Cultures from 6000 BC–5000 BC. (c) Archeological Cultures from 5000 BC–4000 BC. (d) Archeological Cultures from 4000 BC–3000 BC. (e) Archeological Cultures from 3000 BC–2000 BC. (f) Archeological Cultures from 7000 BC–2000 BC.

## 6. Discussion

With the application of technologies in the archeology and conservation of grand sites, such as digital photogrammetry, three-dimensional laser scanning, GIS, remote sensing and so on, the question of how to integrate the data of multiple sites and multiple spatial scales to promote archeological data sharing still needs further research. The methods proposed in this paper are mainly to promote the standardization of grand site archeological data acquisition, processing, database building and use, which will facilitate the aggregation and sharing of archeological data. In the situation of digital archeology, it is mostly technicians from other professional fields, not archeologists, who collect data in the fieldwork and later process data in-lab, making use of digital technologies. Therefore, it is particularly important to establish standard specifications from the perspective of archeology and technology [18].

Classification and coding provide one of the most important standards for archeological data acquisition and processing, and are the basis of data sharing and data quality inspection [41]. Through classification and coding, it is easy to know which class the data should be classified into. This will simplify archeological data acquisition and processing, and promote standardization. It is fundamental to direct data acquisition, processing and warehousing in the field of archeological excavation, reducing intermediate links and improving data processing efficiency. Technologies such as digital photogrammetry and laser scanning are used to carry out data acquisition in the fieldwork, but these raw data require a lot



of subsequent indoor processing [23–25]. Information must be extracted from the raw data according to archeological requirements. Under these circumstances, the extracted information will be classified and stored into different feature classes according to the classification and coding, which potentially improves the data quality and promotes data reuse. Spatial scale is the basis for archeological data acquisition and data representation, which indicates what data should be collected and facilitates the determination of data scale and resolution under different archeological research scales. Correspondingly, during visualization, related data will be extracted from the database according to the displayed spatial scale. As shown in Figure 9, the data, data scale, and data resolution of each scale are different. For example, for the same tomb, when the plane of the grave area is displayed, the tomb is represented as a point. When the layout of the tomb is studied, it is represented as a polygon. To sum up, the classification and coding and spatial scales play a vital role in archeological data acquisition, processing, aggregation, and reuse.

Archeological data have significant spatio-temporal characteristics. The spatio-temporal framework of archeological data is the foundation of archeological data aggregation and data organization, which is also the basis of spatio-temporal analysis of the formation of grand sites and the evolution of civilization. As shown in Figure 10, with a unified space-time benchmark, the evolution of archeological culture in China from 7000 BC–2000 BC is displayed, which intuitively tells the public and scholars the evolution process of archeological cultures in China. The excavation time, data collection time and corresponding archeological cultural period of the archeological entities were recorded during data acquisition and processing. Using these times and the spatial association among sites, ruins, relics and strata, we can trace the archeological excavation process back in the information space, although archeological excavation is irreversible in field.

Archeological information cloud platform is flexible, dynamic and on-demand regarding resource allocation, with better system security and data security strategies. With a large number of historic sites, China has five levels of cultural heritage administration departments from top to bottom. There is at least one archeological research institution in every province. If all these administration departments and institutions build information platforms individually, it will require a huge investment and a heavy workload. In a sense, it is also a huge waste. Through the control of data authority and function authority, a new archeological information platform can be logically and quickly generated to meet different management and research needs. Thus, the investment on informatization would be greatly reduced. More importantly, an archeological information cloud platform is conducive to the implementation of unified data standards and data sharing, providing a more convenient way of completing data online processing and online publishing, greatly shortening the time for the publication of archeological data, and providing archeological research results. The safety and convenience of the platform would be attractive for scholars to use the platform. If the platform is used by more and more scholars to gather data, study and communicate, it will become a public platform for researchers, data owners, the public, and administrators to participate in the conservation of grand sites. If this occurs, then the platform will promote archeological data sharing and interdisciplinary research.

The methods and platform in this paper are very valid and efficient in the Erlitou site and Taosi site. There are hundreds of thousands of sites in China, and the types and periods of these sites are different. The paper only proposed the classification and coding of grand sites from 3500 BC–1500 BC in the Neolithic Age, and its validity for other sites needs to be verified. The feasibility and necessity of the archeological information cloud platform was investigated, and a platform was designed and developed for Exploring the Origin of Chinese Civilization Project. As for how the platform can be popularized and applied to all archeological institutions and administration departments in the country, the current study does not demonstrate who should be in charge of the establishment and operation of the platform and the development of the corresponding management mechanism. These areas will be studied in future work.

## 7. Conclusions

The paper has given an account of the challenges of archeological data acquisition and aggregation. The dissertation has investigated the spatial scales of archeology, and seven spatial scales and the corresponding data are presented. The classification and coding of archeological data of grand sites from 3500 BC–1500 BC was also proposed, so as to simplify the data processing, improve data quality, and promote data sharing. From the perspective of archeologists and big data, a spatial-temporal framework of archeological data was established. Corresponding to this framework, an archeological cloud platform of grand sites was developed. Taking the Erlitou site and Taosi site as examples, the methods proposed in this study is verified to be valid. The origin of civilization is a common research topic for human beings. As a result of the nature of interdisciplinary collaboration, the research on the origin of civilization needs archeological data of many grand sites and the results of archeological studies. Due to the archeological features of the grand sites mentioned earlier in this paper, archeological data are of great importance to archeological research. The purpose of the current study was to provide a standard specification and information platform for data acquisition, processing, aggregation and fusion for archeological data of different grand sites or different periods, and finally promote the sharing of archeological data, which gives full play to greater value and produces more benefits [48].

This study has shown that the classification and coding, spatial scales and spatial-temporal framework can effectively collect, process, aggregate, fuse, and organize multi-source heterogeneous data of different grand sites or different periods. Basic information, such as data categories and geometry types, can be obtained through the code, which facilitates data retrieval and quality inspection in the information system. The platform offers multi-scale and multi-dimensional representation of archeological data of grand sites, according to the customs of archeologists. The correlation and spatial relationship of sites, ruins and relics can also be displayed intuitively. With the advancement of data opening and sharing, an archeological “data ocean” will eventually be formed [16]. The archeological information cloud platform can provide services to archeological institutions, such as data acquisition, processing, storage, representation and spatial analysis, etc. Besides, it also offers lots of support to various types of interdisciplinary research, such as data, collaborative work, visualization, analysis tools, and so on. The third general investigation of immovable cultural heritages has shown that there were 193,282 sites with varying types, sizes, and historical periods in China. In addition to Erlitou and Taosi, there are more than 30 other grand sites studied in project of the Origin of Chinese Civilization, including the Liangzhu site. With the application of internet plus, big data, cloud computing, Internet of Things, and other technologies in the archeology and conservation of grand sites, the varieties and quantity of archeological data will continue to grow rapidly. The results of the current study are of great significance to the conservation of grand sites, especially for those under threat.

Almost every profound transformation in archeology is closely related to the penetration of natural science into archeology [49]. The application of spatio-temporal big data and cloud computing in archeology will provide a new paradigm for archeology [18,50]. Data opening and sharing is a new trend, which increasingly affects archeology as well. Archeologists should embrace cloud computing and big data with a more positive attitude. Together with administration agencies and experts from other disciplines, archeologists can actively explore the corresponding management mechanism, data privacy, data intellectual property, data ethics, a data resource catalog, and other standards and norms for archeological data opening and sharing. In this case, the change will promote orderly and safe sharing of archeological data, and the value of archeological data will increase. Furthermore, it will promote the sustainable development of conservation and use of grand sites through sharing.

**Author Contributions:** All authors would like to describe the contributions in details as follows: Conceptualization, Y.W., F.P., and Q.L.; Methodology, Y.W., S.L. and F.P.; Software, Y.W.; Validation, Y.W. and F.P.; Formal Analysis, Y.W., S.L.; Data Curation, F.P. and Y.W.; Writing—Original Draft Preparation, F.P.; Writing—Review and Editing, F.P., S.L., Q.L. and Y.W.; Visualization, Y.W.; Supervision, Q.L., S.L. and Y.W.

**Funding:** This research was funded by the project of “Research on the Application of GIS and VR Technology in the Exploration of the Origin of Chinese Civilization”, National Science and Technology Support; (Project Number: 2013BAK08B07).

**Acknowledgments:** This study was funded by the Ministry of Science and Technology of the People’s Republic of China Research on the Application of GIS and VR Technology in the Exploration of the Origin of Chinese Civilization Program (2013BAK08B07). We thank all members of the project.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Zhang, Z.P. Problems of the Conservation of Grand sites in China. *Archaeology* **2008**, *1*, 18–26.
- Chen, T.B. Analysis of the Conservation Planning and Technological Innovation of Chinese Great Sites. *Southeast Cult.* **2009**, *2*, 23–28.
- Lu, L.; Lu, R.Y.; Zhou, Q.H. Research critique and prospect of major heritage site issue in China. *J. Xi’an Univ. Archit. Technol. (Nat. Sci. Ed.)* **2012**, *4*, 517–522.
- Shan, J.X. Protection for the Grand Site of Relics and Relative Policies. *Archicreation* **2009**, *6*, 24–25.
- Lu, J.S. The Big Cultural Sites in China: Protection Cases, Existing Problems and Suggestions. *Fudan J. (Soc. Sci. Ed.)* **2005**, *6*, 130–136.
- Renfrew, C.; Bahn, P. *Archaeology: Theories, Methods and Practice*; Shanghai Ancient Books Publishing House: Shanghai, China, 2015.
- Danese, M.; Masini, N.; Biscione, M.; Lasaponara, R. Predictive modeling for preventive Archaeology: Overview and case study. *Cent. Eur. J. Geosci.* **2014**, *6*, 42–55. [[CrossRef](#)]
- Danese, M.; Masini, N.; Biscione, M.; Lasaponara, R. Gis and Archaeology: A Spatial Predictive Model for Neolithic Sites of the Tavoliere (Apulia). In Proceedings of the First International Conference on Remote Sensing and Geoinformation of the Environment (RSCY2013), Paphos, Cyprus, 8–10 April 2013; Volume 8795, p. 879501.
- Gao, F.; Wang, X.; Wu, L.; He, H. The application of 3S Technology in the Exploration of Site. *Sci. Conserv. Archaeol.* **2009**, *1*, 9–14.
- Liu, J. Theory and Practice of Digital Archaeology. *Relics South* **2007**, *1*, 1–8.
- Yang, R.; Lu, P.; Wu, H. The Analysis of Spatial Centralized Character of Settlements at Peiligang Culture Period in Henan Province. *Areal Res. Dev.* **2012**, *31*, 150–155.
- Mccoy, M.D.; Ladefoged, T.N. New Developments in the Use of Spatial Technology in Archaeology. *J. Archaeol. Res.* **2009**, *17*, 263–295. [[CrossRef](#)]
- Arias, V.M. Application of Gis and Spatial Data Modeling to Archaeology A Case Study in the American Southwest. Ph.D. Thesis, University of New Mexico, Albuquerque, NM, USA, 2013.
- Ioannides, M.; Magnenat-Thalmann, N.; Fink, E.; Žarnić, R.; Yen, A.-Y.; Quak, E. Bridging Archaeology and GIS: Influencing Factors for a 4D Archaeological GIS. In *Digital Heritage; Progress in Cultural Heritage: Documentation, Preservation, and Protection*; Springer International Publishing: Cham, Switzerland, 2014; pp. 186–195.
- Lock, G.; Pouncett, J. Spatial thinking in archaeology: Is GIS the answer? *J. Archaeol. Sci.* **2017**, *84*, 129–135. [[CrossRef](#)]
- Mccoy, M.D. Geospatial Big Data and archaeology: Prospects and problems too great to ignore. *J. Archaeol. Sci.* **2017**, *84*, 74–94. [[CrossRef](#)]
- Zhang, Z.P. On the limitations of archaeology. *Palace Mus. J.* **1999**, *2*, 67–69.
- Gattiglia, G. Think big about data: Archaeology and the Big Data challenge. *Archäol. Inf.* **2015**, *38*, 1–12.
- Lasaponara, R.; Masini, N. Living in the Golden Age of Digital Archaeology. In Proceedings of the International Conference on Computational Science and Its Applications, Beijing, China, 4–7 July 2016; Lecture Notes in Computer Science. Gervasi, O., Murgante, B., Misra, S., Eds.; pp. 597–610.
- Neamtu, C.; Comes, R.; Popescu, D. Methodology to create digital and virtual 3d artefacts in archaeology. *J. Anc. Hist. Archaeol.* **2016**, *3*, 65–74. [[CrossRef](#)]
- Comes, R.; Tompa, V.; Bodi, Ş.; Neamtu, C.; Török, K.F. From theory to practice: digital reconstruction and virtual reality in archaeology. *J. Anc. Hist. Archaeol.* **2018**, *4*, 80–88.

22. Opitz, R. Publishing Archaeological Excavations at the Digital Turn. *J. Field Archaeol.* **2018**, *43*, 68–82. [[CrossRef](#)]
23. Previtali, M.; Valente, R. Archaeological documentation and data sharing: Digital surveying and open data approach applied to archaeological fieldworks. *Virtual Archaeol. Rev.* **2019**, *10*, 17–27. [[CrossRef](#)]
24. Scianna, A.; La Guardia, M. 3D virtual ch interactive information systems for a smart web browsing experience for desktop pcs and mobile devices. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *42*, 1053–1059. [[CrossRef](#)]
25. Valente, R. Digital workflow to improve osteoarchaeological documentation. *Digit. Appl. Archaeol. Cult. Herit.* **2019**, *13*, e00097. [[CrossRef](#)]
26. Kansa, E.C.; Kansa, S.W.; Wells, J.J.; Yerka, S.J.; Myers, K.N.; DeMuth, R.C.; Bissett, T.G.; Anderson, D.G. The Digital Index of North American Archaeology: Networking government data to navigate an uncertain future for the past. *Antiquity* **2018**, *92*, 490–506. [[CrossRef](#)]
27. Agapiou, A.; Sarris, A. Beyond GIS Layering: Challenging the (Re) use and Fusion of Archaeological Prospection Data Based on Bayesian Neural Networks (BNN). *Remote Sens.* **2018**, *10*, 1762. [[CrossRef](#)]
28. Galeazzi, F.; Richards-Rissetto, H. Editorial Introduction: Web-based Archaeology and Collaborative Research. *J. Field Archaeol.* **2018**, *43*, S1–S8. [[CrossRef](#)]
29. Marwick, B.; Birch, S.E.P. A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing. *Adv. Archaeol. Pract.* **2018**, *6*, 125–143. [[CrossRef](#)]
30. McKeague, P.; Veer, R.V.; Huvila, I.; Moreau, A.; Verhagen, P.; Bernard, L.; Cooper, A.; Green, C.; Van Manen, N. Mapping Our Heritage: Towards a Sustainable Future for Digital Spatial Information and Technologies in European Archaeological Heritage Management. *J. Comput. Appl. Archaeol.* **2019**, *2*, 89–104. [[CrossRef](#)]
31. Cooper, A.; Green, C. Embracing the Complexities of ‘Big Data’ in Archaeology: The Case of the English Landscape and Identities Project. *J. Archaeol. Method Theory* **2016**, *23*, 271–304. [[CrossRef](#)]
32. Meyer, É.; Grussenmeyer, P.; Perrin, J.-P.; Durand, A.; Drap, P. A web information system for the management and the dissemination of Cultural Heritage data. *J. Cult. Herit.* **2007**, *8*, 396–411. [[CrossRef](#)]
33. Xia, N. What is Archaeology? *Archaeology* **1984**, *10*, 931–935.
34. Xu, P.F. Archaeological Research on Chinese Civilization Formation. *Jilin Univ. J. Soc. Sci. Ed.* **2005**, *1*, 15–21.
35. Yan, W.M. Origin of Chinese Civilization: Journal of Chifeng University. In Proceedings of the Special Series of the 5th Hongshan Culture Summit Forum, Inner Mongolia, China, 10 August 2010.
36. Su, B.Q. A New Investigation of the Origin of Chinese Civilization. *Soochow Acad.* **2013**, *4*, 5–11.
37. Yan, W.M. Unity and Diversity of Prehistoric Culture in China. *Cult. Relics* **1987**, *3*, 38–50.
38. Tong, Z.C. Multicenter Development Theory and Unbalanced Development Theory of Chinese Neolithic Culture—On the Law of Chinese Neolithic Culture Development and the Origin of Chinese Civilization. *Cult. Relics* **1986**, *2*, 16–30.
39. Zhang, G.Z. *Archaeology: Rethinking Some Basic Concepts and Theories*, 1st ed.; Joint Publishing: Hong Kong, China, 2013.
40. Zhao, B.F. Researches on space-time framework and pedigree relations of Archeology Culture of Xia Dynasty to Warring States of the mountainous area of west. *Res. China's Front. Archaeol* **2006**, *4*, 32–69.
41. He, J.B.; Li, X.T. Research on the Ontology of Geographic Information Classification and Coding. *Geogr. Territ. Res.* **2002**, *18*, 1–7.
42. Mell, P.; Grance, T. *The NIST Definition of Cloud Computing*; NIST Special Publication: Gaithersburg, MD, USA, 2011; Special Publication 800-145.
43. Zhang, X.L.; Qiu, S.H.; Cai, L.Z.; Bo, G.C.; Wang, J.X.; Zhong, J. Establishment and perfection of archaeological chronological sequence of Xinzhai-Erlitou-Erligang culture. *Archaeology* **2007**, *8*, 74–89.
44. Xu, H.; Zhao, H.T. Review on the Cultural Stages of Erlitou Site—Centering on Tombs Excavated with Bronze and Jade Rites. *Relics South* **2010**, *3*, 44–52.
45. Zhao, C.Y.; Yuan, J.; He, N. Analysis of Strontium Isotope ratio of Enamel of Animals unearthed from Taosi site in Xiangfen County, Shanxi Province. *Quat. Sci* **2011**, *1*, 22–28.
46. Gao, J.T. Preliminary Investigation on Settlement Pattern of Taosi Site. *Cult. Relics Cent. China* **2007**, *3*, 13–20.
47. Xu, H. Erlitou: Most of China. *Chin. Cult. Herit.* **2009**, *1*, 50–67.
48. Mccoy, M. The Race to Document Archaeological Sites Ahead of Rising Sea Levels: Recent Applications of Geospatial Technologies in the Archaeology of Polynesia. *Sustainability* **2018**, *10*, 185. [[CrossRef](#)]

49. Li, Q. Overview of Scientific and Technological Archaeology. *Wuhan Univ. J. (Humanit. Ed.)* **2007**, *60*, 676–681.
50. Martinho, T. Researching Culture through Big Data: Computational Engineering and the Human and Social Sciences. *Soc. Sci.* **2018**, *7*, 264. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).