Research on the Digital Humanities Practices in Chinese Libraries: A Case Study of Shanghai Libraries

WANG SHEN, JIUYU CHEN, JIA GUO, CHUANG HONG, AND JUN DENG

ABSTRACT
This paper examines the practices of digital humanists in Chinese libraries, with a particular focus on those that have good application results in digital humanities (DH). First, based on research into human databases created by Chinese libraries and analysis of information collected from 152 libraries, we selected three libraries for a case study whose research findings of DH projects have been put into practical application. Next, we collected relevant materials by means of interviews, literature research, network research, and other methods. Finally, the paper identifies some important issues, such as motivations, content, and cooperation, that need to be addressed when analyzing these programs by a case study. Some achievements have been made because of efforts by a number of Chinese libraries to develop DH, but accompanying these achievements are challenges. The paper puts forward the core problems for DH in Chinese public libraries and tries to give a numerical solution scheme that will help to promote the development of DH.

INTRODUCTION
In recent years, libraries have been unfolding support for digital humanities (DH) (Posner 2013). The library and information science (LIS) community has made efforts to provide new facilities and develop new services to meet humanities scholars’ changing research behaviors and needs by employing digital tools and methods (Zhang, Liu, and Mathews 2015). American university libraries have carried out DH research and practice. Subsequently, Chinese libraries have also carried out DH research. “DH” implies humanities-based research, teaching, and intellectual engagement conducted with digital technologies and resources. The use of these technologies may be prosaic or transformative (Zorich 2008, 4). It has
strong practicality, and its research results are often reflected in the form of DH projects, which can effectively reveal the development of the field from the perspective of DH project practice. China has 3,153 public libraries at the national, provincial, municipal, and county levels and 830 university libraries (Chinese Library Association and National Library of China 2018, 67). Hundreds of libraries were researched, and we identified and selected the subjects of a case study. The results highlight that the DH projects of Shanghai Library, East China Normal University Library (ECNUL), and Shanghai Jiao Tong University Library (SJTUL) have entered the application stage. We selected the libraries mentioned above as the case study participants. We show the construction process, difficulties, and solutions of DH in Chinese libraries and provide examples of the experience for the development of DH in libraries in China and the world.

**Background**

DH research has entered its third stage (Berry 2011). Third-wave DH points the way in which digital technology highlights the anomalies generated in a humanities research project and leads to the questioning of the assumptions implicit in such research. China started late on DH research, and there is still a gap between its research results and those reached through DH research in America.

We retrieved a total of 136 articles on DH in Chinese libraries. We abandoned irrelevant articles, short reviews, and introductions to meetings. Finally, we selected a total of 110 articles. After analysis, the research topics were mainly divided into four classes: recent DH research at home, recent DH research abroad, the theory and methods of research for the DH, and the practices of DH projects in Chinese libraries. Among these topics, the research literature of DH theory and methods is the most common, indicating that DH research in Chinese libraries is still concerned with the theoretical research and simulation phase. There is more research on DH in foreign countries than on the practices of the DH projects. The literature review on foreign DH focuses on the status quo of DH research, construction and practice, and so forth, as shown in Table 1 (which can be accessed at https://doi.org/10.5281/zenodo.3676322). There are few reviews of China’s DH research, as shown in Table 2 (which can be accessed at https://doi.org/10.5281/zenodo.3676322). Chinese libraries pay attention to the theory of DH and the construction and organization methods of various information resources such as ancient books, Buddhist cultural heritage, and cultural relics as shown in Table 3 (which can be accessed at https://doi.org/10.5281/zenodo.3676322).

The DH Project of the Ancient Books supported by the Guangxi University Library and the DH Project of Huizhou Documents supported by the Sun Yat-sen University Library are still under construction. The China Local History Document Database (CLHDD) supported by SJTUL, the
China Genealogy Knowledge Service Platform (CGKSP), the celebrity manuscript library of Sheng Xuanhuai, and the inquiry platform supported by Shanghai Library are now in the application phase. The first phase of the Heterogeneous Fangzhi Integration Platform (HFIP) of the ECNUL is completed, and the second phase of construction is under way, as shown in Table 4 (which can be accessed at https://doi.org/10.5281/zenodo.3676322).

According to our findings, a total of sixty-one Chinese libraries published papers in the field of DH. The top three libraries in the number of published papers are the Shanghai Library, the ECNUL, and the SJTUL. Each of the other libraries that published papers in the field of DH published three or fewer papers. Results of the literature survey on DH studies indicate that the Shanghai Library, the ECNUL, and the SJTUL are in leading positions both in the field of DH research and in the field of project practice in Chinese libraries, the latter of which is the focus of this paper.

**METHODS AND CASE STUDY FRAMEWORK**

**Methods**

In order to explore the challenges faced in the construction of DH in Chinese libraries, we used quantitative and qualitative research methods to conduct in-depth research on the DH construction of Chinese libraries. Based on our literature search, we found that both public libraries and university libraries have conducted research on DH. So, at the preliminary investigation stage, the locally developed human database of fifty-two public libraries and the top hundred university libraries were surveyed through online and telephone research. Working from the definition of DH and DH Center (Zorich 2008), the following case selection criteria were developed:

- The library builds digital collections as scholarly or teaching resources
- It creates tools for authoring (i.e., creating multimedia products and applications with minimal technical knowledge or training)
- It offers DH training (workshops, courses, academic degree programs, postgraduate and faculty training, fellowships, and internships) or reports, or lectures, programs, conferences, or seminars on DH topics for general or academic audiences
- The amount of data exceeds 10,000 records

**Case Study Framework**

Aiming to find out about the practical use of DH in Chinese libraries, we interviewed the director of library Jilin Province Library and other DH specialists and did some research by reading literature to construct the
case study framework before making case analysis, which consists of these six aspects: driving factors, project content, implementation process, achievements and application, construction mode, and problems. Driving factors play an irreplaceable role in the development of DH in China. Project content, specific implementation process, and construction mode are essential components. The achievements and application component are a reflection of the achievements of DH construction in China. Problems are challenges for China's DH construction. Accordingly, focusing on the specific performance of the six aspects, this paper sums up the difficulties Chinese libraries encountered during the course of DH construction and puts forward corresponding solutions. A semistructured interview guide was designed based on this research framework (the interview guide can be seen in Appendix A, which can be accessed at https://doi.org/10.5281/zenodo.3677841).

**Data Collection**

*Locally Developed Humanities Databases in Libraries*

As can be seen from the definition of data humanities, “DH is a diverse and still emerging field that encompasses the practice of humanities research in and through information technology, and the approaches or exploration of how the humanities may evolve through their engagement with technology, media, and computational methods” (Alliance of DH Organizations n.d.); DH needs to provide scientific research support services. The main function of the public library is public service, not only research service. Therefore, we have investigated only the locally developed humanities databases of public libraries at the national, provincial, and municipal levels. DH means massive data and providing scientific research tools and training services, which requires capital investment and technical input. We selected the libraries of the top hundred universities in China as survey subjects because of their high level of funding and support. Due to a large number of locally developed humanities databases, we chose a database that had more than a thousand records and was still in use at the time of the study. The initial data collection time for this study was November 2018; after that, we added more data in May and June 2019.

Case study subjects were determined based on the literature, Internet, and phone survey results and case selection criteria. The three DH subjects we chose are at Shanghai Library, SJTUL, and ECNUL, as shown in Table 5 (which can be accessed at https://doi.org/10.5281/zenodo.3676322). The CGKSP at Shanghai Library has 54,000 records and provides tools for humanities research. Shanghai Library did a four-day training for DH. The HFIP at ECNUL has 67,753 records and provides visualization and annotation tools. The CLHDD at SJTUL has 35,000 records and provides
statistical tools. They participated in and organized conferences on DH. The three libraries are located in Shanghai, probably because Shanghai is an international economic and technological innovation center.

**Cases Data Collection**

After determining that these three libraries would compose the case study, according to the requirements of case data collection, we used a variety of methods to collect data. The data include a field visit to each library, the DH-related report meeting, the interview with the person in charge of the DH project, and the literature survey on the digital human-system (details can be seen in Table 6, which can be accessed at https://doi.org/10.5281/zenodo.3676322). Cuijuan Xia is a senior engineer of the Shanghai Library System and Network Center and the director of the CGKSP DH project. Xin Li is a research librarian of ECNUL and the director of the HFIP DH project. Jie Zhang is a librarian of SJTUL and the main participant of CLHDD DH project.

**Data Analysis**

**Analysis of Locally Developed Humanities Databases in Public Libraries**

We investigated 148 locally developed humanities databases of one national, thirty-one provincial, and twenty municipal public libraries. The National Library of China has the largest number of locally developed humanities databases. The number of locally developed humanities databases in provincial libraries is uneven, ranging from zero to seventeen. The top eight libraries accounted for 63 percent of the total (see Figure 1), indicating that locally developed humanities databases are relatively concentrated. Of the 148 databases, 53 have more than 10,000 records. This shows that China, as an ancient country with a historical civilization, has a large number of cultural resources. The main data type is images: 113 databases contain picture resources in .pdf and .jpg formats. Only four databases have full-text textualizing data, and some databases—such as the celebrity manuscript library of Sheng Xuanhuai supported by the Shanghai Library—have not completed the textualization of all resources. Among the fifty-two public libraries, only the Shanghai Library has carried out DH projects. Other libraries did not involve DH and did not provide scientific research tools and DH training (see Appendix B, which can be accessed at https://doi.org/10.5281/zenodo.3676322).

**Analysis of Locally Developed Humanities Databases in University Libraries**

Compared with public libraries, there are fewer locally developed humanities databases in university libraries. There are 161 locally developed humanities databases in one hundred university libraries and 54 databases with data records greater than 10,000. The range in the number of locally
developed humanities resource databases among university libraries, ranging from zero to seven, is smaller than the range among public libraries. The data types are also mostly images. The index type is 31, of which 17 databases have more than 10,000 records, indicating that there are many resources to be developed in university libraries. Four databases contain full-text textualization data resources, one in the ECNUL and the other three in the SJTUL, indicating that these two university libraries are more in-depth in revealing the content of resources. (see Appendix C, which can be accessed at https://doi.org/10.5281/zenodo.3676322).

In summary, our research found that Chinese libraries contain a wealth of humanities resources, but those resources are lacking in development. Most resources are only partially digitized. Only three libraries, the Shanghai Library, the ECNUL, and the SJTUL, have reached the application stage and are able to provide scientific research tools and DH training.

**Case Data Analysis**

Content analysis of the data was done using NVivo11 pro computer-assisted qualitative data analysis software.

To ensure the accuracy and effectiveness of the research, we used interview data, literature, and DH platforms, repeatedly analyzed and verified
the information, and finally completed the revision and demonstration of the research results.

**Case Study Results**

*The Driving Factors of DH*

**Project Factors.** The main source of funding for DH construction comes from projects at the national, municipal, and university levels. The main source of funding for the Shanghai Library’s CGKSP is a special funding project for Information Development of the Shanghai Economics and Information Technology Commission. It is also funded by the National Social Science Fund of China. The HFIP of the ECNUL relies on the National Social Science Fund of China and the university-level Humanities Platform Project. The main fund of the CLHDD of SJTUL comes from the Cao Shuji National Social Science Fund of China of the History Department of Shanghai Jiao Tong University, the Shanghai Social Science Planning Major Project, and the Shanghai municipal education commission key projects of humanities and social science innovation, relying on the 985 Phase III Special Construction Project of the university and CALIS Phase III Special Database Construction Project of the Ministry of Education.

**Resource Factor.** The three libraries have large numbers of humanities resources. Shanghai Library has a large collection of resources, including celebrity manuscripts and Chinese ancient books. Shanghai Library is the institution with the largest number of original genealogies at home and abroad (Xia et al. 2016, 29). ECNUL is rich in Fangzhi resources, with a total of nearly fifty thousand kinds of books, including nearly twenty thousand e-books. It also integrates other local resources of the National Consortium of Normal University Library in China (NCNULC). Shanghai Jiao Tong University Local Document Center and its History Department transfer a large number of historical documents to the SJTUL. These resources involve all aspects of society—economy, life, work, etc.—which are the most widely documented groups that reflect the traditional society in the Southeast (Tang and Zhang 2016, 219).

*The Demand for Resource Integration.** Chinese libraries hope to complete the organization, integration, association, and integration of single type resources to multiple types of resources and then to multisource heterogeneous resources through digital humanistic methods. During our interview with Cuijuan Xia, she said, “After the digital library has developed to a certain extent, a metadata plan has been developed, and there is more and more data, and it is necessary to construct an overall solution to complete the data fusion.” The integrating of the unique resources of different libraries can more effectively utilize scattered, isolated, and closed special resources (Li, Zhang, and Wang 2017, 49). “As one of the important in-
stitutions for collecting local historical documents, the library undertakes the task of digitizing the information, and organizing the information will help to promote the progress of academic research in the digital age,” Jie Zhang said. “Data-driven humanities research” is a new opportunity to enter the next generation of libraries. Providing faceted visual browsing, semantic search, and even knowledge mining services for humanities research scholars helps to break the isolation of various resource databases in the library, promoting data opening and knowledge flow (Xia et al. 2016).

Service Requirements. Meeting the needs of scientific research services is an important function of DH platforms. CGKSP had a large number of users of scientific research. A large number of historical documents can attract researchers’ interest. By integrating special collections at the library, HFIP creates large structured data sets and improves the authority of human research supported by resources. The main purpose of the Chinese local historical document database is to serve scientific research activities and help to discover new research contents (Li, Zhang, and Wang 2017, 49). As a public library, the Shanghai Library is also driven by the demand to provide services for the public. “Research results are to serve the public. . . . A public library like ours cannot just focus on data, and it must serve the readers,” director Xia said. New technology is supposed to be employed to provide online services for more users (Xia et al. 2016).

Project Content
• Shanghai Library–CGKSP. The genealogy records a large number of figures, economics, folklore, population, and other materials, and plays a powerful role in promoting the study of humanities and social sciences such as history, sociology, and folklore. It can provide commonsense services and intelligent root-seeking services for the general public plus knowledge discovery and knowledge mining services for humanities researchers, as well as bibliographic control and knowledge increment functions for the researcher. The resources include more than 54,000 genealogical catalogs containing 608 surnames included in the China National Genealogy.
• ECNU-L-HFIP. Local chronicles, known as Fangzhi in Chinese, involve records of different administrative levels, including national records, district records, county records, township records, and prefecture records. Different categories of content are also incorporated in such records, such as the content concerning irrigation, education, and so on. Based on the platform, an environment covering learning, researching, communicating, and teaching is hopefully to be available for users.
• SJTUL-CLHDD. The historical documents in the CLHDD were collected between 2006 and 2013 and amount to 350,000 documents and
1.5 million pages. The local literature includes the following thirteen categories: contracts, account books, taxes, administration, litigation, letters, daily and reference books, family etiquette, religion, genealogy, drama, educational examinations, and medicine. The database has specific metadata and provides cross-navigation, data statistics, and other functions that not only help researchers find their own literature, but are also likely to help researchers discover new research topics.

The Construction Process, Achievements, and Application

Through the survey of the DH projects of the three libraries, we found that the project construction process mainly follows the requirements analysis: design, data processing, platform system implementation achievements, and applications. The construction processes of each library are slightly different based on the implementation content. For example, the documents collected by SJTUL have not yet been digitized, and resource sorting has been added before the demand analysis. In the requirement analysis phase, the three DH projects focus on user requirements. In addition, the CGKSP proposes genealogy bibliographic control requirements for data reuse and sharing (Xia et al. 2016, 29). The local historical literature database proposes a literature digitization requirement. The design phase consists of three main parts: data model design, system design, and functional design. The three DH projects adopt different processing methods according to their resources. The three DH projects in the implementation phase of the platform system are implemented in cooperation with outsourcing companies. The following is a detailed introduction to the construction procedure.

Resource Sorting. The literature collected by the SJTUL is sourced from field surveys or market acquisitions and requires insecticide dust removal, coding, inventory packing, leveling, and repairing. These tasks increase construction cost, time, and labor.

Demand Analysis. Demand analysis is an important task for the proper construction of DH projects. The users of the CGKSP include the general public and scientific researchers. The library hopes to provide general users with the root-seeking ancestor requirements with limited known information, provide scientific users with knowledge discovery of specific research topics, and provide an open platform for user-generated content (Xia et al. 2016, 30). The genealogy bibliographic control requirements for data reuse and sharing include the establishment of a union catalog of global genealogy and the control of genealogical data concepts based on the World Wide Web, which supports the sustainable development of bibliographic control (Xia et al. 2016, 29). The users of the HFIP include normal alliance organizations and scientific researchers. The main requirement is to integrate the heterogeneous resources of the institutional
pavilion to provide a research environment for scientific research users (Lu and Li 2018, 38). The main requirement of the CLHDD is to develop more analytical tools that are useful for researchers (Zhao 2016). Since the literature has not yet been digitized, the need for effective digitization has also been proposed (Zhao 2016).

**Design.** The humanities resources in CGKSP and HFIP are completely digitized and partially datalized. In order to make better use of the existing metadata, ontology technology is used to design the data model. The reason is that ontology can integrate existing heterogeneous metadata. Since the original data of CLHDD have not been digitized and cataloged, a metadata scheme with more content characteristics has been designed. CGKSP was developed to meet the needs of the public users and scientific researchers in terms of functional design. They pay particular attention to the user-friendliness of the interface. HFIP and CLHDD aim to build a scientific research environment. They will add more scientific tools.

- **Shanghai Library–CGKSP.** The data model of the CGKSP is based on the genealogy knowledge ontology. The genealogy ontology is based on BIBFRAME. The core model Work-Instance-Item of the associated bibliographic data model has been adopted. They also reuse the attributes in FOAF, the terms of Geonames to describe the residence area of genealogy, the terms of Schema.org, W3COrganization to describe the collection institutions involved in genealogical resources, the W3CTimeOntology to describe time information, and the unique attributes such as “pedigree name” and “dynasty” (Xia et al. 2016, 31). The entity description information is organized based on the Resource Description Framework (RDF) statements and encoded in a standard serialization format. The function is designed to meet the needs of the general public to browse and find the roots (Xia et al. 2016, 34). For scientific researchers, the platform designs advanced search based on concept matching and exploring and discovering based on time-space association. It also provides a platform for user communication and interaction (Xia et al. 2016, 35).

- **ECNU–HFIP.** The data model design is mainly based on identifying the metadata provided by the affiliate institution. A vocabulary suitable for the specific application scenario is selected to establish shared core metadata ontology to promote semantic integration and interoperability between heterogeneous metadata. By extending the core ontology, a specialized metadata ontology is built to implement a semantic description of specific resources (Li, Zhang, and Wang 2017, 51). The HFIP provides functions including unified search, analysis tools, GIS tools, a crowdsourcing platform, API interface, full-text browsing, and so forth. They hope to develop more analytical tools that help researchers.

- **SJTUL–CLHDD.** The data model design of the database mainly includes
Resource classification and metadata design. A total of thirteen primary classifications have been designed for resource classification (Zhao and Tang 2014). The design of the metadata scheme is based on the Dublin Core set, including metadata items such as the subject, title, time, and so on. It also sets corresponding metadata schemes for the special document, such as targets and amount elements (Zhao 2016). System functions include resource push, statistics, faceted data display, and cross-search (Zhao 2016).

Data Processing. Since collection metadata is already available, they mainly use technical means to complete data cleaning and conversion in the CGKSP and HFIP projects. Because the original data of CLHDD have not been digitized and cataloged, all literature resources had to be manually coded. This caused difficulties in project implementation. Due to a large number of users, Shanghai Library’s crowdsourcing platform is developing well for full-text textualizing. Although HFIP has built a crowdsourcing platform for full-text textualizing, it is currently not functioning due to limited users.

- **Shanghai Library–CGKSP.** The RDF data of the CGKSP is based on existing metadata. These data are stored in RDB (relational database) format and are needed to convert the RDB format to RDF data in Turtle format (also known as RDB2RDF). They mainly use two open-source and automatic conversion tools: DB2Triples and Open Refine (Xia et al. 2016). They are designing a crowdsourcing platform for full-text textualizing.

- **ECNUL-HFIP.** The HFIP cleans, transforms, merges, and fuses the obtained data according to the core metadata ontology and the metadata provided by the data provider and stores the processed data in the core database (Lu and Li 2018). The ECNUL also cooperates with the data company Chinese Online to conduct text processing of the full text of the chronicles (Li, interview).

- **SJTUL-CLHDD.** In the digitization stage of the literature, since most of the local historical documents are Chinese handwritten materials, the traditional Chinese characters are difficult to identify, and the main method is manual coding (Tang and Zhang 2016).

System Platform Implementation. None of the three libraries have the capacity to implement the system platform. They choose to cooperate with software companies. The libraries lack skills in software development in DH projects. They are better at demand analysis and design in the early stage of the project.

- **Shanghai Library–CGKSP.** CGKSP implementation mainly employs software companies. During the implementation process, the staff of the
Shanghai Library is responsible for the development of implementation documents. Then, the library staff and software developers discuss until an agreement is reached. The staff of the software development company are trained by the DH team of the Shanghai Library (Xia, interview).

- **ECNU-HFIP.** The construction and implementation of the HFIP are mainly carried out by cooperative software companies. The research tools provided include open-source data analysis software such as Baidu (Li, interview).

- **SJTUL-CLHDD.** SJTUL and a data service company work together to develop the database. The data service company has experience in data management and data platform development, and cooperation is smooth (Zhang, interview).

**Achievements and Application.** Shanghai Library has certain advantages in terms of the number of users. One of their innovative uses of open data has been to host an application competition. The competition promotes the use of genealogical data. HFIP is mainly applied in the NCNULC. CLHDD has been published by Shanghai Jiao Tong University Press. However, only a few research results are currently based on DH platforms.

**Shanghai Library–CGKSP.** The CGKSP provides search functions based on title, surname, place of residence, and so on. The advanced search function allows multiple combinations and visualization of search results. Users can use the tool to search by time-space and map. The platform supports crowdsourcing editing, data publishing as linked data, and data reusing and sharing. It also provides APIs and open datasets. As a pioneering library for the open data of linked data technology in China, the CGKSP has received more attention. There are hundreds of thousands of users participating in the platform, and some of them are humanistic researchers. In addition, there are hundreds of town libraries that use the platform. Based on this platform, the Shanghai Library has hosted an open data application competition. The competition has been successfully held two times since 2016. About two hundred people from more than one hundred teams signed up for the competition (Zhang and Xia 2018). During the competition, the number of visitors to the genealogy knowledge base data service platform and the number of readers in the genealogy reading room increased significantly (Zhang et al. 2017).

**ECNU-HFIP.** At present, the platform has completed the first phase of data collection based on the NCNULC. It can retrieve 65,826 pieces of information and realize the visualized regional browsing function and tag cloud technology by introducing GIS technology and word segmentation technology. The means of revealing include title/responsible search, tag cloud, and others, increasing the retrieval accuracy of resources (Cheng and Zhang 2018). The platform publishes Fangzhi records in the form of linked data and provides training and visual assistant tools.
SJTUL-CLHDD. The China local historical literature library platform has been published by Shanghai Jiao Tong University Press. The functions of the China local historical literature library platform include resource push, statistics, faceted data display and cross-search, navigation, and others. Many universities including Harvard, Stanford, and Columbia have purchased the database. Some research results based on the platform have been published (Zhao 2015).

DH Project Construction Mode
There are eight members of the DH project team at Shanghai Library, five members in ECNUL, and sixteen members in SJTUL. Due to limited team members, the three libraries cooperate with different persons to build the DH projects.

Resource Sorting—Cooperation with Historical Experts. The SJTUL cooperates with the Department of History and the Local History and Literature Research Center in processing local historical documents. The history department sends professionals to participate in on-site guidance and carry out the protective restoration of the literature (Zhao and Tang 2014).

Demand Analysis and Design Phase—Collaboration with Humanities Experts. In the demand analysis and genealogy ontology design stage of the project, the Shanghai Library takes the local team as the core and collaboration with humanities experts. They design the system’s functional requirements according to the scientific research content of the experts and focus on analyzing the data that experts use for scientific research (Xia, interview). In the initial stage of the DH project, the ECNUL did sufficient research to fully understand the research needs of humanities researchers, thus ensuring that the platform can support humanities research. The funds of CLHDD are mainly from the history department of Shanghai Jiao Tong University. History department researchers lead the design of the metadata plan. The main purpose is to better meet the research needs of humanities and realize the display of composite context (Zhang, interview).

Data Processing Stage—Collaboration with the Outsourcing Company, Crowdsourcing, and Graduate Students. The genealogy metadata should be mapped one by one to the framework of genealogy ontology. The Shanghai Library cooperates with the development company. For the textualization of the full-text content, the Shanghai Library adopted a crowdsourcing approach (Xia, interview). The resources of the HFIP are mainly provided by the NCNULC. The textualization of the full text of the Fangzhi is in cooperation with the data service company (Li, interview). The local historical literature is almost all handwritten material, and it is difficult to identify traditional characters in historical document processing. Graduate students from East China Normal University were introduced to digitize these documents (Zhang, interview).
Platform Development and Implementation Phase—Collaboration with the Outsourcing Company. The specific implementation of the CGKSP mainly adopts the form of on-site development of outsourcing software companies. The ECNUL cooperates with the data service company to implement the HFIP. The China local historical literature database was developed by professional publishing houses and was published and promoted online.

The Dilemma of DH in Chinese Libraries

Through literature research, data analysis, and case studies of DH in Chinese libraries, the results show that the development of DH at Chinese libraries is still in its infancy and there are few DH projects reaching the application level. In order to understand the reasons, we analyze in-depth interview data and point out the typical problems encountered in the construction of DH projects in Chinese libraries and some problems that have not yet been solved.

Resource Data Processing Problem

Most libraries provide a wealth of resources for research and learning. Chinese libraries are not lacking in special resources, but these resources are not fully utilized. Most resources are only partially datalized. The documents of HFIP of the ECNUL are currently only partially textualized (Li, interview). The CLHDD of SJTUL still has a batch of documents not entered into the database (Zhang, interview). Text recognition of ancient books is very difficult, and researchers have to manually identify the text. Faced with a large number of paper resources, it is hoped that a large number of people will do data processing and then purchase according to the topic (Li, interview).

Funding Problem

The problem of resource data processing is accompanied by a shortage of funds. The foundation of DH is massive data, which come from the digitization and textualization of paper resources. The digitization and textualization of massive resources require a lot of money. This is also why DH can take the lead in economically developed cities. The funding for DH projects in the three libraries comes from outside; there is no federal money budgeted for the DH projects within the libraries. The genealogy resources of the Shanghai Library have been digitized. CGKSP is completed with only a small amount of funds. The resources of other unfinished databases rely on the Shanghai Library crowdsourcing platform. But funding is still a big problem (Xia, interview). The textualization of the full text of the library of the ECNUL is also subject to funding restrictions (Li, interview). The local literature library of the SJTUL is funded by a collaborative humanities research team. The university has limited funds for supporting the construction of DH in the library. The digitized
and cataloged materials are only part of the paper collection. Due to the lack of follow-up financial support, only a small amount of scanning is now being done (Zhang, interview).

Technical Problems of Text Recognition and Indexing
If technical tools can be used to realize text recognition and indexing of the literature, resource processing and funding problems will be greatly improved. The literature survey of Chinese libraries shows that they have a weak ability to develop technical tools and can rely only on other research teams. Due to the particularity of Chinese characters, text recognition and indexing of ancient books are difficult. Handwritten documents are more difficult to do text recognition and indexing with because of the large differences in writing. The three libraries use manual methods for textualization and indexing of the full text, which greatly limits the digitization process of resources and further restricts the development of DH projects in China.

DH Project Promotion and Application
Judging from the results of the literature survey, there are a few humanities achievements completed by applying three DH platforms. How to promote the application of DH projects is an important issue. It is necessary to investigate the scientific users of the platform and scholars who research in this subject area.

Enlightenment on DH
Datalization and Textualization of Library Resources
The digitization and textualization of resources are the base and prerequisite of the success of DH projects. They could provide an ideal starting point for DH construction. Currently, there are still some resources awaiting digitization (directory index only) or partial datalized (only simple metadata) in Chinese libraries. If these resources were fully textualized and datalized, DH projects would be rapidly developed.

Exploration of Multiresource Fund Access
The source of funds for the DH project of the Chinese Library is the National Project, the Government Informationization Project, and scientific projects of other humanities research teams. DH construction demands sufficient funds. If the core resources are completely digitized and textualized and the team members master the core technology, the demand for funds is not very large. However, if the core resources are not completed, then a lot of money is needed to complete the infrastructure work of the project. In this case, it is necessary to broaden the channels of funding. For example, the three libraries in the case study applied for vari-
ous forms of funding support from the National Social Science Fund, the National Natural Science Fund, the provincial government’s Informationization Project, and other research teams.

**Strengthening of Technical Cooperation**

Facing huge quantities of information that requires digitization and textualization, DH scholars are in urgent need of new tools and platforms to organize, index, recognize characters, search, excavate, and analyze. Shanghai Library and ECNUL have used ontology technology and RDF format to build a data model. The Shanghai Library uses the OpenLink Virtuoso object-relational database to store data and GIS technology to achieve time and space browsing and display a knowledge map of associated data. It provides an API call interface. The ECNUL uses tools such as GIS technology and word cloud to visualize search results. It also provides visualization tools such as Rising Sun charts, relationship charts, tree charts, and migration charts. The CLHDD provides search statistics function. These three DH projects use technology that focuses on the organization of resources and visualization, but deeper processing technologies like automatic identification of resources, text segmentation, entity identification, and data mining are less applied. It is recommended to work with experts from the DH Research Center and natural language processing experts for further improvements.

**Innovative DH Promotion and Application Model**

The application of DH is limited by service objectives. The Shanghai Library has set up a good role model to break through this problem using a contest to encourage the public to make use of DH, and they succeeded. The contest not only boosted the profile of the CGKSP but also developed other information services. The HFIP of the ECNUL has been promoted and used mainly within the National Normal College Library Alliance because it is still new. The CLHDD of the SJTUL is mainly promoted by the publishing company. Humanities scholars and libraries have a different understanding of humanities resources. Postgraduate training can be used to drive scholars to use the DH platform. DH can continue to develop healthily only if it better meets the needs of users in many aspects and extends applications to new users.

**Conclusion**

DH has brought new opportunities and challenges to humanities research. The role of libraries in DH is a theme of Chinese library research. In recent years, Chinese libraries have begun to build some DH projects, but overall both theoretical study and practical applications in this area are still in the incipient stage. We used a comprehensive approach involving a review of the literature, statistical analysis, and case studies to
study the current situation and difficulties of DH construction in Chinese libraries. We want the world to understand the development status of DH in Chinese libraries and to promote the development of DH in Chinese libraries. In terms of corpus construction, although the Shanghai Library has published a standardized database of information such as names and places, it lacks images and text recognition tools and indexing corpora and needs to be strengthened further.

DH is a rapidly developing research field, and its research content, methods, and tools are also evolving. Chinese libraries need to constantly explore application strategies of the semantic web, natural language processing, data visualization, and other technologies in the DH projects to strengthen the development of DH tools. Due to time constraints, a limitation of the study is that some databases are being upgraded and the information collected may not be comprehensive. Another limitation is that there is not research on every Chinese university and public library, and there may be libraries with better constructions that are not included in the study.

ACKNOWLEDGMENTS
The authors gratefully acknowledge the financial support of the National Natural Science Foundation of China: Research on Multi-source Heterogeneous Online Product Review Data Fusion and Knowledge Discovery Based on Graph Model (grant number 71974075) and the National Social Science Fund of China: Research on Knowledge Aggregation and Knowledge Discovery of Historical Archives Resources from the Perspective of DH (19BTQ102).

REFERENCES
Alliance of DH Organizations. n.d. “DHQ on Digital Humanities.” http://www.digitalhumanities.org/dhq/about/about.html
Berry, David M. 2011. “The Computational Turn: Thinking about the DH.” Culture Machine 6 (3): 5–10. https://doi.org/10.3917/presa.185.0209.
Chen, Tao, Cuijuan Xia, Wei Liu, and Lei Zhang. 2015. “Research and Implementation of Visualization Technology for Linked Data.” Library and Information Service 59 (17): 113–19. https://doi.org/10.13266/j.issn.0252-3116.2015.17.017. (陈涛, 夏翠娟, 刘炜, 张磊. 2015. “关联数据的可视化技术研究与实现.” 图书情报工作59 (17): 113–19. https://doi.org/10.13266/j.issn.0252-3116.2015.17.017.)
Cheng, Jing, and Yi Zhang. 2018. “A GIS-based Visualization Design for the Integration of Heterogeneous Library Resources.” Library Forum 38 (10): 47–54. https://doi.org/10.3969/j.issn.1002-1167.2018.10.008. (程静, 张毅. 2018. “基于GIS的图书馆异构资源整合可视化设计.” 图书馆论坛38 (10): 47–54. https://doi.org/10.3969/j.issn.1002-1167.2018.10.008.)
Chinese Library Association and National Library of China. 2018. China Library Yearbook: 2017. Edited by Han Yongjin, 67. Beijing: National Library of China Publishing House. (中国图书馆学会, 国家图书馆. 中国图书馆年鉴: 2017. 北京: 国家图书馆出版社: 67.)
Li, Fang, Jin Chen, and Xin Wang. 2015. “Planning and Practice for the Historical Documents Digitalization in Recent Library Holding of Shanghai Jiao Tong University.” Journal of Academic Libraries 33 (2): 77–83. https://doi.org/10.3969/j.issn.1002-1027.2015.02.012. (李芳, 陈进, 王昕. 2015. “上海交通大学新藏地方历史文献的数字化建设规划与实践.” 大学图书馆学报 33 (2): 77–83. https://doi.org/10.3969/j.issn.1002-1027.2015.02.012.)
Li, Xin, Yaxiu Yu, and Jing Cheng. 2019. “Data Optimization Based on Crowdsourcing—A Case Study on the Construction of Digital Local Chronicles.” *Library Forum* 39 (2): 73–79, 118. https://doi.org/10.3969/j.issn.1005-6041.2019.02.010. (李欣, 于亚秀, 程静. 2019. “基于众包的数据优化—以数字方志特藏资源建设为例.” 图书馆论坛 39 (2): 73–79, 118. https://doi.org/10.3969/j.issn.1005-6041.2019.02.010.)

Li, Xin, Yi Zhang, and Zhili Wang. 2017. “DH Research Demand of Library’s Heterogeneous Special Resource Integration.” *Digital Library Forum* (11): 48–53. https://doi.org/10.3772/j.cnki.jlis.160014. (李欣, 张毅, 汪志莉. 2017. “图书馆异构特藏资源整合的数字人文研究需求.” 数字图书馆论坛 (11): 48–53. https://doi.org/10.3772/j.cnki.jlis.160014.2017.11.008.)

Liu, Wei, and Ying Ye. 2017. “Exploring Technical System and Theoretical Structure of Digital Humanities.” *Chinese Library Journal* 43 (5): 32–41. https://doi.org/10.13530/j.cnki.jljis.170002. (刘炜, 叶鹰. 2017. “数字人文的技术体系与理论结构探讨.” 中国图书馆学报 43 (5): 32–41. https://doi.org/10.13530/j.cnki.jljis.)

Liu, Dan, and Xin Li. 2018. “A Digital Humanities Platform Based on the Integration of Heterogeneous Resources.” *Library Forum* 38 (10): 38–46. (鲁丹, 李欣. 2018. “整合异构特藏资源构建数字人文系统.” 图书馆论坛 38 (10): 38–46.)

Poole, Alex H., and Deborah A. Garwood. 2017. “The Conceptual Ecology of DH.” *Journal of Documentation* 73 (1): 91–122. https://doi.org/10.1108/JD-05-2016-0065. (波尔, Alex H., and Deborah A. 加伍德. 2017. “概念生态学：数字人文.” 文献与信息学 73 (1): 91–122.)

Posner, Miriam. 2013. “No Half Measures: Overcoming Common Challenges to Doing DH.” *Library Administration & Management* 53 (1): 43–52. https://doi.org/10.1080/01930826.2013.756694. (波斯纳, 米丽安. 2013. “不做半途而废的事：克服数字人文的共同挑战.” 图书馆管理 53 (1): 43–52.)

Tang, Meng, and Jie Zhang. 2016. “Study on the Process of Local History Document Sorting from the Perspective of Source—Taking the Library of Shanghai Jiaotong University as an Example.” *Cultural Journal* 12: 219–22. (汤萌, 张洁. 2016. “来源视角下的地方历史文献整合流程研究—以上海交通大学图书馆为例.” 文化学刊 12: 219–22.)

Xia, Cuijuan, Wei Liu, Lei Zhang, and Wenjing Zhu Wenjin. 2014. “A Genealogical Ontology in the Form of BIBFRAME Model.” *Library Forum* 34 (11): 5–19. https://doi.org/10.3969/j.issn.1002-1167.2014.11.002. (夏翠娟, 刘薇, 张磊, 朱雯晶. 2014. “基于书目框架 (BIBFRAME) 的家谱本体设计.” 图书馆论坛 34 (11): 5–19.)

Xia, Cuijuan, Wei Liu, Tao Chen, and Lei Zhang. 2016. “A Genealogy Data Service Platform Implemented with Linked Data Technology.” *Chinese Library Journal* 42 (3): 27–38. https://doi.org/10.13530/j.cnki.jljis.160014. (夏翠娟, 刘薇, 陈涛, 张磊. 2016. “家谱关联数据服务平台的开发实践.” 中国图书馆学报 42 (3): 27–38. https://doi.org/10.13530/j.cnki.jljis.160014.)

Xia, Cuijuan, and Lei Zhang. 2016. “The Application of Linked Data in Shanghai Library’s Service of Genealogy Digital Humanities.” *Library Magazine* 35 (10): 26–34. https://doi.org/10.13663/j.cnki.lj.2016.10.004. (夏翠娟, 张磊. 2016. “关联数据在家谱数字人文服务中的应用.” 图书馆杂志 35 (10): 26–34. https://doi.org/10.13663/j.cnki.lj.2016.10.004.)

Zhang, Lei, and Cuijuan Xia. 2018. “A Study on Open Data Service of Library for Digital Humanities: Shanghai Library’s Open Data Contest as an Example.” *Library Magazine* 37 (3): 33–38, 48. https://doi.org/10.13663/j.cnki.lj.2018.03.006. (张磊, 夏翠娟. 2018. “面向数字人文的图书馆开放数据服务研究—以上海图书馆开放数据应用开发竞赛为例.” 图书馆杂志 37 (3): 33–38, 48. https://doi.org/10.13663/j.cnki.lj.2018.03.006.)

Zhang, Ying, Shu Liu, and Emilee Mathews. 2015. “Convergence of Digital Humanities and Digital Libraries.” *Library Management* 36 (4–5): 362–77. doi:10.1108/LM-09-2014-0116. (张迎, 刘书, 马熙思. 2015. “数字人文与数字图书馆的融合.” 图书管理 36 (4–5): 362–77. doi:10.1108/LM-09-2014-0116.)

Zhao, Siyuan. 2015. “Land Market, Credit Mechanism and Relationship Network in Huizhou Rural in the 19th Century.” *Modern Chinese History Studies* (4): 82–97. (赵思渊. 2015. “19世纪徽州乡村的土地市场、信用机制与关系网络.” 近代史研究 (4): 82–97.)
Wang Shen is an associate professor of the Information Management Department, School of Management, Jilin University. Her research interests include library services and technology, information retrieval, user information behavior, and social media. She has published a monograph and a number of research papers in journals and conference proceedings. She can be reached at shenwang@jlu.edu.cn.

Jiuyu Chen is a postgraduate student of the Information Management Department, School of Management, Jilin University. Her research interests focus on digital libraries and DH. She can be reached at 361266083@qq.com.

Jia Guo (corresponding author) is a lecturer in the Financial Management Department, School of Management, Jilin University. His research interests include knowledge management, user information behavior, and financial information management. He has published a monograph and a number of research papers in journals and conference proceedings. He can be reached at GuoJia@jlu.edu.cn.

Chuang Hong is a doctoral student in the School of Management, Jilin University. His research areas include DH, knowledge management, and library science. He can be contacted at exybean@126.com.

Jun Deng, PhD, is the deputy director of the Information Management Department, School of Management, Jilin University, a position she has held since 2015. For fifteen years, she has led and participated in more than ten important research projects, focusing on digital information resources management, development and utilization, library intelligence, and records management. In addition to her project-management work, she was a postdoc and visiting scholar at the University of Buffalo in 2010 and 2018. Her current areas of research are DH, especially knowledge aggregation and knowledge discovery. She can be reached at dengjun9722@163.com.