Open Data Services in the Library:
Case Study of the Shanghai Library

Qi Zhang, Shanghai Library, China

　https://orcid.org/0000-0002-4557-9381

Chunjing Zhang, Shanghai Library, China

　https://orcid.org/0000-0002-5654-9786

Zheyu Zhang, Shanghai Library, China

　https://orcid.org/0000-0001-8726-0988

ABSTRACT

The opening and sharing of data are gaining momentum in the era of big data. Libraries have been actively involved in the research and practice of open data. At present, the related research in Chinese libraries is still in progress, which mainly focuses on the introduction of the abroad practices, or on the construction of open data platforms. This paper introduces and analyzes the Shanghai library open data service and provides a useful reference for the open data service of Libraries in China. For the future development of the Shanghai library open data services, this paper puts forward some measures and suggestions that include metadata work, website construction, legal protection, and developer community training.

KEYWORDS

INTRODUCTION

With the advent of the big data era, open data has become a worldwide wave (*Jiang, 2015*). Since the U.S. government's open data portal was launched in May 2009, open data movement has been springing up rapidly around the world (*Li, 2015*). The report, named "Open data: An engine for innovation, growth and transparent governance", was submitted to the European Parliament by the European Commission in November 2010, and set the strategy of coping with big data challenges based on open data" (*Yu, 2014*). In September 2011, Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, Britain, and the United States jointly signed the Open Data Statement at the UN General Conference, and established the Open Government Partnership (OGP). In June 2012, as the first portal in China, the Shanghai municipal government's open data service was open to the public. By August 2017, 53 countries and 165 regions opened their government's open data, 75 countries joined the OGP (*Data. gov, 2017*), and nearly 20 local governments launched their open data platforms in China. Under the global wave of open data, as an indispensable part of the society, libraries are actively involved in open data research and practice, and face the changes and challenges of the times with an open concept and attitude. The open data service of Shanghai library (located at Shanghai in China, for

its location see Figure 1) is one of the cases of open data practices of libraries at home and abroad. On the basis of the detailed introduction of the case, the paper puts forward some suggestions on its development through the comparison and analysis of the open data cases in foreign libraries.

Figure 1. The location of Shanghai Library

Literature Review

For open data, there is no uniform standard definition yet (*Ma &Pu, 2016*). The definitions from Wikipedia (*Wikipedia, 2017*), the British Open Knowledge Foundation (OKF) (*Open data handbook, 2017*), and The World Bank (*World Bank group, 2017*), etc. have been widely used at home and abroad. In addition, the open data principles presented by The Open Data Charter have been more widely recognized. So the paper does not repeat the details of these definitions and principles.

The open data services of foreign libraries are mainly based on bibliographic linked open data, including British, American, German, French, Swedish, Spanish, Hungarian, Finnish, Korean, and other national libraries, and the OCLC, the European digital library, and others. (*Xia, 2015*). At present, many institutions have released large scale related datasets, such as WorldCat.org, Europeana of the European Union Digital Library, the Library of Congress, and the British Library, etc. to arouse the public's attention to the added value of these data. With the increasing number of datasets exposed through the unified interface, the research obstacles of digital humanities and data science will also be reduced, and the public will also benefit from the convenience brought by integration (*Sande & Verborgh & Hochstenbach, 2017*). In 2008, the Swedish National Library published the Swedish Union Catalog LIBRIS, thesaurus and authority files as linked data, and LIBRIS became the first Union Catalog published as linked data as a whole (*Zou& Hu, 2016*), providing a lightweight API based on HTTP to support different data access formats such as MARC-XML, Dublin Core, JSON,

RIS, MODS, RDF, and others (*LIBRIS, 2017*). The Library of Congress has been exposing its famous vocabulary and its Authority Names into linked open data since 2009. It currently provides seventeen downloadable datasets to the public, and the datasets are formatted in RDF/XML, Turtle, or N-Triples (*Marden & Limadeo & Whysel N, 2013*). In 2011, the Biblioth` equenationale de France (BNF) launched its rich linked data resource in order to collect, conserve, and share the documentary heritage of France. The BNF data are presented in RDF or RDF/XML, N3, RDF-NT, and JSON (*Marden & Limadeo & Whysel N, 2013*). In July 2011, the British Library published the British National Bibliography as linked open data, and released a linked open data service on the TSO platform in July 2013 (*Deliot, 2014*). In order to support different target groups, non-library target audiences are subdivided into the general public, academic researchers, the archive community, UK book trade, development community, and schools by the British Library, supporting different standard citation formats, such as CSV, EAD, Endnote, Refworks, RIS, etc. and different standard open data formats such as RDF/XML, Turtle, JSON, MARC-XML, etc. and different access ways, such as download, SPARQL, API and so on (*The British Library, 2017*). Deutsche Digitale Bibliothek (DDB) released an API in November 2013, providing users with open data services based on linked data technology for German culture and technology digital resources, but users need to register an API Key to use the service (*Deutsche digitale bibliothek, 2017*). In April 2014, DDB, with the Service Center for Digitizatin Berlin, the Open Knowledge Foundation Deutschland, and Wikimedia Deutschland hosted the first "Coding da Vinci" hackathon in Berlin to reveal the potential of DDB data and encourage links between programmers, game players, and designers as well as memory organizations such as museums, archives, and libraries. "Coding da Vinci" has been held for four years and was awarded prizes for selected projects spanning from five categories: Most Technical, Most Useful, Best Design, Funniest Hack, and Out of Competition (*Deutsche digitale bibliothek, 2017*). However, the Toronto Public Library adopts a different open data technology route, and provides more than 30 open datasets including catalogue, website search, branch information, cardholders, circulation, children, collection, program, technology, youth, and visits. Most of the datasets are published as CSV, except MARC for catalogue-, JSON for website search-, and KML for geographic information of branch-. Every dataset has metadata, such as introduction, data available, last updated, and frequency of updates (*Toronto Public Library, 2016*). Generally speaking, open data services in foreign libraries started early, and there are plenty of practices. Some of them have become typical cases, which can play an important role in the demonstration and guidance for open data practices in libraries.

At present, the library open data research in China mainly focuses on the introduction of open data practices in foreign libraries, the construction of open research data platform, etc. There is no unified understanding of open data, and the progress of open data practice is slower due to the lack of understanding (*Xiao& Li, 2017*).

In September 2013, the University of Chinese Academy of Sciences and the National Science Library held an open information innovation contest on research and education for graduated students of the Chinese Academy of Science. During the 3 months of the contest, 8 databases, such as the union catalog database, Chinese Science Citation Database etc., as well as datasets collected from the Web, ranging from technology, economy, culture, all the way to sociology, and open data processing tool sets (*Chinese Academy of Sciences, 2014*). But it's not an open data service in the real sense, the service was until March 2016, and with the help of achievement of the genealogy digital humanities project, Shanghai Library launched the open data services in the way of an open data application development contest. This project is the first domestic open data service and will have a far-reaching impact in China

RESEARCH METHOD

This paper uses the literature research method, the experience summary method, and comparative analysis method to study the entire cycle of the library open data service. One of the literature research

methods is to search the Chinese journal papers and dissertations on open data service mainly through the Tsinghua Tongfang Journal Full-text Database. The other method is to investigate the practice of open data service of the British Library, the German digital library, the Toronto Public Library, etc. through their websites. In recent years, driven by the Digital Humanities project, Shanghai library has done research on open data service and has practiced open data service. This is a summary of some experiences to be shared.

Through a literature review of open data service projects at home and abroad, the summary of the practice of open data service in Shanghai Library and the comparative analysis of 4 open data services in libraries, this paper puts forward suggestions for improving the practice of Shanghai Library.

DESCRIPTION OF OPEN DATA SERVICE IN SHANGHAI LIBRARY

Background

The Shanghai Library has a rich collection of resources: with about 22,000 types and over 110,000 books of genealogy literature, it has the largest collection of original genealogical documents. It even has a dedicated genealogy reference room. In recent years, more and more readers, and research institutions from home and abroad are showing a growing concern with the blood lineage contained in the Shanghai Library Genealogy Database and multitudinous social, historic and cultural information. In addition to the query of genealogy literature and genealogical information, they also expect Shanghai Library to provide genealogy data services and let the users compile new family trees, renew the old ones with the genealogy data, or make use of the genealogy data on the strength of their own ideas for developing new services. In August 2015, Shanghai library was named the Ministry of Culture Research Center for Public Culture Affairs. Because of a long period of active exploration of the open data service, in March 2016, the Shanghai library decided to run the genealogy open data application development contest and open data service to promote the innovation of public digital cultural services in China. Then the open data service based on celebrity manuscripts and archives was launched in May 2017. The open data application development contest was held for soliciting excellent mobile application prototypes and service ideas. The impact and the effect are better than expected.

Open Data Content and Quantity

The Shanghai library provides open data about genealogy, celebrity manuscripts and archives. It also provides the authority files, such as the China historical chronology, geographical vocabulary, and organization directory on the open data platform (http://data.library.sh.cn/). The genealogy open data is from the linked data-based Genealogy Knowledge Database (http://jp.library.sh.cn/), consisting of 53,987 metadata records of genealogy documents, 608 family names, 1,609 places, and 130,867 people collected from 596 genealogy collection institutions (*Xia, 2016*). The celebrity manuscripts and archives open data is from 240 thousand kinds of library's collections from the early Nineteenth Century to the end of the Twentieth Century, including Sheng Xuanhuai's Archives. The datasets include 108,738 letters, 23,689 telegrams, 17,281 manuscripts, 14,749 decrees, memorials and official documents, 4,169 contracts and articles of association, 18,182 photos and videos, 4,490 signature books, 587 certificates, 20,064 books of account bill, 508 diaries, 4,555 paintings and seal cuttings, 160 objects, 7,067 kinds of paper documents and so on, involving 63,978 celebrity names, and 2,207 place names (*Xia, 2017*).

Open Data Format

All the open data including genealogy open data, celebrity manuscripts and archives open data, and the China historical chronology, geographical vocabulary, and organization directory are RDF format data based on linked data. The Genealogy open data, and celebrity manuscripts and archives open data are generated based on existing metadata. Each work, instance, item, person, organization, and

place which is extracted from metadata, has a HTTP URI. The classes and attributes defined by the genealogy and manuscripts, and archives ontology are used to describe the relationships between these entities and publish genealogy ontology, manuscript and archives ontology as linked open data by RDFs encoding (*Xia, 2016*). The ontology model of genealogy, manuscript, and archives is based on the bibliographic framework (BIBFRAME). The ontology thesaurus is composed of user-defined terms about the genealogy and manuscript, reusable FOAF, Relationship, GeoNames Ontology, Time Ontology, Event and other terms (*Xia, Liu & Zhang, 2014*).

Open Data Service

The Shanghai library open data services follow the CC2.0 protocol. They are licensed under the BY-NC-SA, and provide HTTP URI, Restful API, SPARQL Endpoint for data consumer interface. They all return data in JSON-LD format (*Chen, 2016*) (*Chen, 2017*):

- Using standard API, the RDF data of the resource, its attributes, and values are obtained by accessing the URI of them.
- Through the specific API query interface, the RDF data of the specific types "surname," "celebrity ancestors," "place," "organization," "dynasty," "bibliography," "name," "synonym," "manuscript" and their attributes and values can be obtained, at most 20 pieces a time.
- Obtain RDF data of a resource or a specific type, its attributes and values through the SPARQL query language.

Anyone can use the open data, but they need to register to get API Key first. To help developers to quickly master the API, Shanghai library provides open data interface specification document, makes a detailed description of the access methods, parameters, and returned format, available data scope and data structure, and provides the example code.

Open Data Services Promotion

By drawing lessons from the Shanghai Open Data Apps (SODA) contest, which was organized by Shanghai Municipal Commission of Economy and Informatization and Shanghai Municipal Transportation Commission, the Shanghai library held the open data application development contest to promote the open data to the community along with the announcement of open data service. The contest was composed of six parts: announcement, registration, professional training, submission, assessment, and result and award:

- **Announcement:** The contest announcement was issued in the websites of Shanghai library and all district public libraries and university libraries in Shanghai, and it was also spread across the country online by the Shanghai Library Society. Meanwhile, the roll up banners of announcement were placed in the libraries of Fudan University,Shanghai Jiao Tong University,East China Normal University, and Pudong Library. The announcement includes the schedule and the contest rule and so on.
- **Registration:** Allow individuals or teams to sign up. Each team shall have one team leader and the number of team members are not limited.
- **Professional Training:** Professional training was held after registration in Shanghai Library, and a live webcast was broadcast at the same time. The training courses include open data content, open data access, interface technology, etc.
- **Submission:** Apps for IOS, Android, or micro websites (mobile Web application) can be submitted.
- **Assessment:** The assessment was composed of two phases, document review and oral defense, based on innovativeness, feasibility, technology content, interaction experience, and the degree

of open data utilization. The document review focuses on technical details, which include creativity, the degree of open data utilization, external data relevance, technology advancement, and workload. The oral defense focuses on the overall evaluation work, including theme of innovation, function innovation, interactive experience, and the degree of open data utilization.

- **Result and Award:** The contest set up one first prize, two second prizes, and five third prizes. The reward for first prize is 10 thousand yuan, each second prize is 5 thousand yuan, and each third prize is 2 thousand yuan. The Popularity Award was added in 2017. Via the WeChat account, the public can vote for eight different favorite works once a day within 10 days. The number of votes of works in the top 8 wins the Popularity Award, and the highest number of votes of a work wins the best Popularity Award. There are no money rewards, and the result of the contest was issued on the day of the oral defense.

Open Data Contest Propaganda

In addition to building a contest website for announcement, schedules, rules, and training materials, results of the award and awarded works, Shanghai library also held an exhibition, and a sharing session to promote open data service and an open data contest. The contest exhibition in the hall of Shanghai library introduced the contest situation, motivation of winners, work conception, work features, contest harvest of winners, etc. The sharing session mainly introduced the awarded works and their technology development experience by the winners in the way of live exchange between all the participants. The sharing session was held during the exhibition.

Open Data Services Effect

Shanghai library promotes the open data by means of contest, and obtains obvious results.

- The open data API access has increased.

During the 2016 contest, the three-month access to the open data API received a total view of 55,571 times. Visits to the Genealogy Knowledge Database Service Platform and Shanghai library Genealogy Reading Room also rose sharply. During the 2017 contest, the three-month access received a total view of 245,678 times, rose 342 percent, and the enthusiasm does not vanish with the end of events.

- The public has an application requirement for library's open data.

The two contests actively attracted the participation of 131 teams and 336 people in total. There were 71 teams and 195 people signed up for the 2017 contest, and there was an 18 percent growth of teams and a 38 percent growth of people on a year-to-year basis (*Figure 2*). Participants were from Shanghai, Beijing, Guangdong, Jiangsu, Zhejiang, Hubei, Anhui, Heilongjiang, Jilin, Sichuan, Taiwan, and other places. The teams from Shanghai accounted for about three-fourths. Participants include the students and social workers that were distributed among all ages (*Figure 3*). About half of the participants were college students with diverse disciplines, which exceeded our expectations. Because at the very beginning, we expected company developers to expand our new services based on their experiences. Astonishingly, it was the college students who brought fresh ideas to the contest.

More than half of the participants are aged from 21 to 30. There was a substantial growth of participants under the age of 20 in 2017, with an increase from 8% to 18%. Participants under the age of 30 accounted for about 75% of the participants (*Figure 4*). This reveals that even at a traditional culture event, young people are also full of enthusiasm. This may be due to their intrinsic love for digital technology, mobile applications, and dynamic interaction. This inspires us, and in future

Figure 2. The comparison of participant number
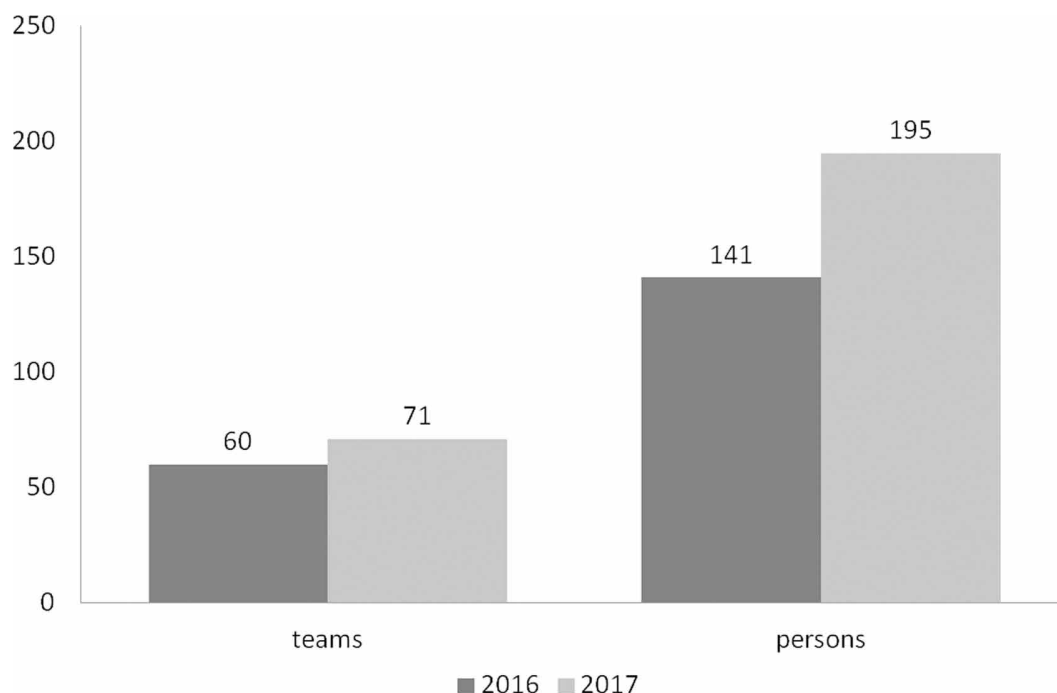
events, we should be geared towards a broader population, no matter what themes they are, historical or contemporary.

- The potential values of open library data can be inspired by mining deeply, new innovative applications and values can be generated by integrating with the external data.

According to the type of works submitted for two years (*Figure 5*), with the development of social networks, people are more inclined to submit mobile web applications. These applications are very lightweight, and can be spread very quickly through WeChat and other social platforms.

The 1st-prize and 2nd-prize winners of 2016 have high novelty and application value. The 1st-prize was an Android App for non-Chinese native speakers who learn about Chinese surnames and cultures, and learn Chinese characters. Combining the genealogy, surname culture, and Chinese Character culture cleverly, the 1st-prize winner realizes the link between Shanghai Library Genealogy Knowledge Database and other famous data pools, such as DBpedia, Wikitionary and GeoNames, and presents the potential semantic values of Shanghai Library Genealogy Open Data in the World Linked Data Cloud. By using the concept of Crowd sourcing, one of the 2nd-prize winners manages to build an online coordination and edition platform, which not only mines the SNS social attributes of genealogy, but also connects it with the official history and local chronicles, hence inspiring individual enthusiasm for the crowd creation of historical figures and events. Taking spatio-temporal information as the point cut, the other 2nd-prize winner mines deeply the genealogy data on the strength of the strong map service function of GIS, and offers a variety of visual presentation approaches. Taking Hu's genealogical data as the example, it restores and visually presents their family migration path. Some ideas and technologies have been adopted by the second stage development of the Shanghai Library Genealogy Knowledge Database Service Platform. The works of 2017 have higher completeness, and more works are combined with external data to enable users to be acquainted with more details of history and humanities (*Figure 6*).

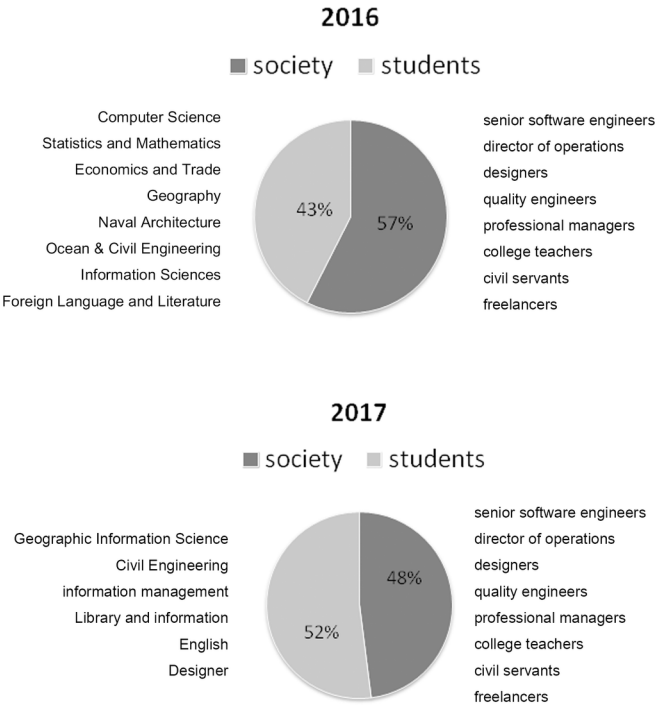**Figure 3. The comparison of occupational distribution**

2016

society   students

Computer Science
Statistics and Mathematics
Economics and Trade
Geography
Naval Architecture
Ocean & Civil Engineering
Information Sciences
Foreign Language and Literature

43%   57%

senior software engineers
director of operations
designers
quality engineers
professional managers
college teachers
civil servants
freelancers

2017

society   students

Geographic Information Science
Civil Engineering
information management
Library and information
English
Designer

48%   52%

senior software engineers
director of operations
designers
quality engineers
professional managers
college teachers
civil servants
freelancers

Figure 4. The comparison of age distribution

2016

4%   8%
18%
70%

■ 11-20 years old
■ 21-30 years old
■ 31-40 years old
■ 40 years old

2017

8%   18%
20%
54%

■ 11-20 years old
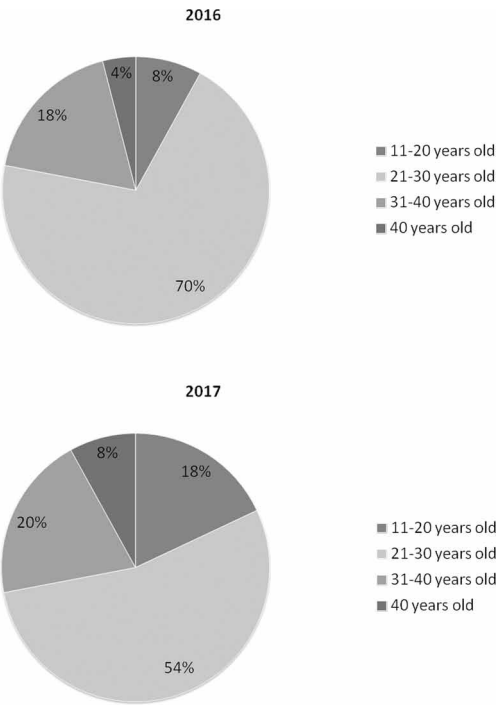■ 21-30 years old
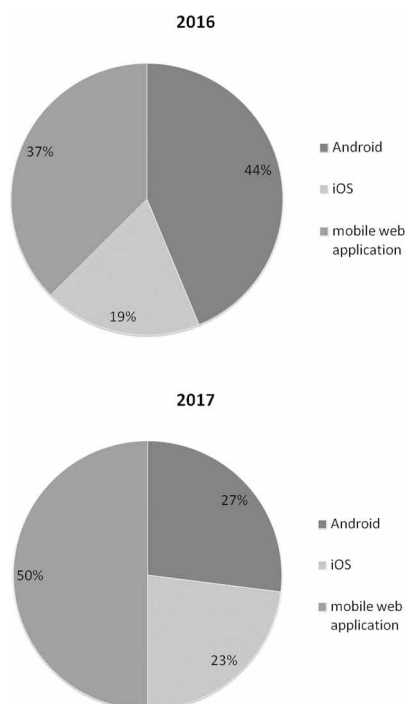■ 31-40 years old
■ 40 years old

Figure 5. The comparison of submission

**COMPARISON AND ANALYSIS OF THE FIVE CASES OF OPEN
DATA SERVICE IN LIBRARIES AT HOME AND ABROAD**

Although nearly twenty local governments in China have launched the open data platform one after
another since 2012, (*Sohu, 2017*), it's still in the preliminary exploration stage, and the open data
concept has not been popularized in the society (*Gao, 2015*). The Shanghai library open data service
project is the first case in Chinese libraries. Compared to non-library fields of open data services
in China, the project has a better understanding of the concept and principles of open data, and a
tendency of structured and linked data in the Internet age. The paper chooses five open data service
cases of the library for comparative analysis (*Table 1*).

Selecting the High Value and Good Quality Data

The primary issue with open data services is opening up what type of data, and the open data should
have a certain data size, but also without intellectual property issues, to be of potential value. From the
table, we can see that these five cases of open data are in line with the above requirements, but also have
their own characteristics. When starting the open data service in 2016, the Shanghai Library selected
high-quality and high-value genealogy data as open data mainly based on the aspects of resource
characteristics, reader requirements, and resource data quality. Genealogy is the most characteristic of
resources in the Shanghai Library. It is famous for the largest number of original Chinese genealogies
in the world, which not only has great influence but also has drawn wide attention from research
institutes both at home and abroad. Readers who search their genealogies online and offline have
been pouring in and some readers have expressed the hope that they can use their genealogy data to
renew their genealogy background. Secondly, genealogy is a kind of historical map of the descendants
and deeds of the family recorded in a special form containing surnames, family migration, lineage
catalogue, deeds, customs, relationships and so on (*Wang, 2007*). Moreover, When Shanghai Library

Figure 6. The characteristics of the award-winning works
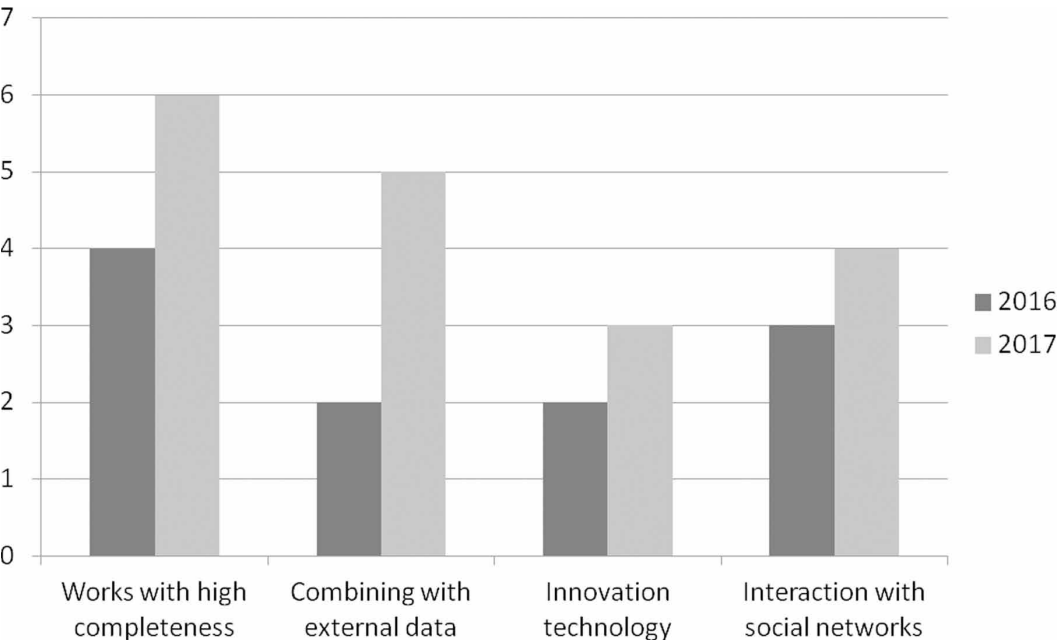
led the compilation of thematic union catalogue "The General Catalogue of Chinese Genealogy" in 2000, they formulated the genealogy description metadata specification and cataloging rules (*Zhou & Gu, 2005*). Based on the union catalogue, the data concerning surnames, people, places, time, and institutions were further regulated and supplemented in 2013 for the genealogical digital humanities project. Based on the genealogical metadata scheme, we designed a BIBFRAME Pedigree ontology, a publishing family tree ontology with related data technology. The genealogy ontology based on BIBFRAME was designed and published with linked data technology. The genealogy ontology gives semantics to the data, linked data techniques, and encodes data in a standard format that allows the machine to understand the semantics and handle the relationships between the data. Celebrity manuscripts and archives span nearly 200 years from the beginning of Nineteenth Century to the end of Twentieth Century. It is an important research resource about modern history, cultural, economic, and social problems, and continues the historical and cultural theme, BIBFRAME ontology format and the linked data publishing technology. Therefore, the potential value of the data is higher and the quality is better.

The Higher Degree of Data Openness

Machine Readability is the most basic principle of open data. All five cases in the table provide a machine-readable data format. Four of them use the W3C's 5-Star open standard recommended by Tim Berners Lee.

For open data, its openness has two dimensions: the legal openness and the technical openness (*World Bank group, 2017*). At present, the main types of open data license agreements established by international organizations include the Creative Commons license agreement, the open data sharing license agreement, the open government license agreement, etc. (*Yang, Xia &Xu, 2016*). Three cases in the table used the Creative Commons license agreement, and one adopted the Open Government License Agreement. The Shanghai Library adopts CC-BY-NC-SA in the public domain. Although CC-BY-NC-SA is not the ideal type of open data license agreement (*Huang & Li, 2016*), compared

Table 1. Comparative table of five library's open data services

Institute	The British Library	Deutsche Digitale Bibliothek	The Toronto Public Library	The Library of Congress	The Shanghai Library
Release date	2011.7	2013.11	Not found	2009	2016.3
Datasets	Metadata about published books, serials	The cultural and Scientific Heritage of Germany in digital form	more than 30 open datasets about Catalogue, Event, and Feed Data from the website, branch information etc.	19 main datasets, including LCSH, LCNAF, LCC, etc. preservation dataset, small datasets	Metadata about genealogy, celebrity manuscripts and archives
Quantity	3.9 million records	23,721,478 objects, including 9,341,578 with digitized media	2.4 million catalogue records, etc.	500 million triples	290 thousand records
Update frequency	monthly	monthly	annually	Not found	Not found
Linked open data	YES	YES	NO	YES	YES
Delivery Format	CSV, EAD, Endnote, Refworks, RIS & BibTex, RDF/XML, Turtle,JSON, MARC21/ MARCXML	DC, DenkXWeb, indexMeta, ESE, EAD, METS/MODS, MARCXML, LIDO	XML,MARC/ BZ2,JSON,KML,CSV	RDF/XML (MADS, SKOS) N-Triples (MADS and SKOS) JSON (MADS/RDF and SKOS/RDF) MADS - RDF/XML MADS - N-Triples MADS/RDF - JSON SKOS - RDF/XML SKOS - N-Triples SKOS - JSON MADS/XML MARC/XML	RDF JSON-LD
Delivery Method	Download SPARQL/Linked Data API	API	Download	Search Download Command-line RESTful API	SPARQL/Linked Data API
Data Access Interface	SPARQL Editor SPARQL Endpoint	HTTP URI, RESTful API	NO	HTTP URI(Humans),RESTful API(machines)	HTTP URI RESTful API,SPARQL Endpoint
API key	NO	YES	NO	No	YES
License	CC0 1.0	CC0 1.0 CC3.0 CC-BY,CC-BY-ND,CC-BY-SA,CC0,Public Domain,dl-de/by for different data in Coding da Vinci	OPEN GOVERNMENT LICENCE – Toronto Public Library	Not found	CC2.0 CC-BY-NC-SA
Usage guide for license	6 aspects	4 aspects	6 aspects	Not found	Not Found
Website/ Platform	http://bnb.data.bl.uk/	https://api.deutsche-digitale-bibliothek.de/.	https://opendata.tplcs.ca/	http://id.loc.gov/	http://data.library.sh.cn/
Website/ Platform column	Home Getting Started Documentation FAQs Contact us	Home, About us, Help, Discover	Introduction, Library data, library Open data policy, Contact us	Linked Data Service, About, Main Dataset Descriptions, Preservation Dataset Descriptions, Small Datasets Descriptions, Search, Download, Technical Center, Contact Us, Privacy Policy	Introduction, Lists of Authority and ontology, Usage guide of API
Administrative Metadata	Description Update frequency Dataset Date Size(KB)	Description Update quantity License Data type Created Last Updated	Description Date Size format Data available Last Updated Update frequency	Description URI Last Updated Data Formats	Not found
Activity	Not found	Coding da Vinci	Not found	Legislative Data Challenge, Congressional Data Challenge	Open data application development contest
Number of Activity	Not found	2014 to 2017 4 times	Not found	2013,2017 twice	2016 to 2017 twice

with the domestic library service sharing license, its legal openness is relatively in a higher level. To the technical level of openness, Machine Readable is the most basic attribute of open data, which can be unstructured data, semi-structured data, and structured data (*Zheng, 2015*). All five cases in the table provide a machine-readable data format. Four of them used the W3C's 5-Star open standard recommended by Tim Berners Lee. The RDF data abstraction model and its various serialization formats such as RDF/XML, Turtle, JSON-LD of Shanghai Library's open data reached the highest star rating in the five-star rating system on the format of open data and the use of semantic web standards, were in line with the development direction of open data structurization and association advocated by the W3C. So having technical openness is held in higher degree over other aspects.

The Formation of "Five in One" of Open Data Service and Promotion Mode

Opening data is not an end in itself, however, making data useful is the purpose (*Zheng, 2015*). Open data does not directly produce value, and it is only when open data is taken advantage of by society that it really makes sense (*Zheng, 2015*). To promote the reuse of open data, the "open government data index " of the Organization for Economic Cooperation and Development (OECD) lists the open data application development contests (such as for mobile clients) as one of the indicators of reuse (*Zheng & Liu, 2016*).

The "Shanghai big data development implementation opinions" puts forward to support Shanghai open data innovation application competition and other data competition activities (*Shanghai.gov.cn, 2016*). Adopting the experience of home and abroad, the Shanghai library holds open data contests, launches open data services, and collects excellent works and ideas from the whole country. Hence, they can increase public awareness of open data and its significance for helping libraries to open up the potential value of data.

Some auxiliary activities, e.g. the technical training for participants held soon after the registration, and the one-day sharing session during the one-week exhibition after the contest, succeeded in sharing brilliant ideas of awarded works. In this way, we integrated contest, training, exhibition, and a sharing session into vital parts of open data services, which constitutes a Five-in-One paradigm for open data service and promotion.

From the comparison with four other libraries, the Shanghai library open data services still has the following deficiencies:

- Open data formats are not diversified.
- The website is simpler and can't provide more details to data users.
- The lack of administrative metadata.
- The usage guide for license is not clear enough.

SUGGESTION

The open data service is a nice try that the Shanghai library studies to absorb domestic and abroad open data research and Practices. The open data service is also a thought and exploration that the Shanghai Library transforms and innovates in a public library in the era of big data. It provides a useful reference for the open data service of Libraries in China. Combined with the status of open data at home and abroad, and the implementation of this case, the author believes that the Shanghai library open data services in the future should be strengthened in the following areas.

Strengthen the Metadata Work

The open data is not only just a single data, it is more of a certain kind of data or data combinations, so it is often called the open dataset. Open data is not only the opening of data, but also the opening of the process, that is, including a series of data operations such as data source, acquisition, and processing

(*Ma &Pu, 2016*). Metadata is a kind of description data, and it is very important to strengthen the open data management, improve the open data sets discoverability, help users understand, and select the data (*Yu, Zhai & Lin, 2016*). The United States clearly specified the metadata content in its open data policy memorandum, a G7 commitment to participate in a metadata GitHub mapping test (*i199IT, 2014*), the "open data certificate" of the British Open Data Research Institute (ODI) (*Open data certificate, 2017*), and "2017 Chinese local government data open platform report" all claim that providing metadata is one of the evaluation indicators. In January 2014, W3C launched the data open metadata recommendation standard DCAT (Data Catalog Vocabulary) (*Maali & Erickson, 2014*) etc. Metadata is an important component of open data, and has the same important status as open data. Therefore, the Shanghai library open data services need to strengthen the basic work metadata, but also to pay attention to the semantic and machine-readability of metadata, and refer to the domestic and foreign existing metadata specification and application examples in order to make metadata schema and publish the schema on its open data platform.

Strengthen the Construction of Open Data Website

The establishment of a unified government open data portal is a common practice among governments and regions about the open data service, which is no exception in foreign libraries (*Wu, 2015*). There are usually two ways which are "proprietary" and "unified embedded." The former is a particular open data platform that all open data gather in, the latter is a column of a website (2017). Although the approach is different, a common feature is publishing open data centrally and providing open data consumption interfaces or downloads centrally. In addition, open data service is not just a simple open data portal, but is more of a dynamic system that provides the open data and the reuse of open data, and inspires the innovation among different stakeholders (*Zheng& Liu, 2016*). The data guide, data acquisition, data applications, and interactive communication (*Zheng & Gao, 2015*) is therefore the basic functions of the open data portal. At present, the Shanghai library has a professional open data platform, but it is more of a dynamic for all the open data. The Shanghai library open data services collect the existing open data and publish them on the open data platform uniformly, establish the competition channel that links the open data application development contest website in order to show the open data application case, strengthen the interaction and communication, and provide convenience for the user acquisition, learning and innovation applications.

Strengthen the Legal Guarantee of Open Data

Matching the corresponding open license protocol for a data set is an important step to open data, and it is the main way to guarantee the openness of data from the legal perspective. Data open authorization is typically included in Web site statements, an exemption clause, or service agreements on an open data platform. The Toronto public library open data policy on its website announces the open data policies in detail. They consist of policy objectives, basic principles, policy statement, accountability, data set definition, and accessory. The accessory "Open Government License-Toronto Public Library", informs the user that data can be free to use, what rights are included in free use, under what circumstances can the data be used freely, etc. (*Toronto public library, 2017*). The British Library provides link to the Creative Commons CC0 1.0 Universal Public Domain Dedication in its linked bibliographic data open web site. In the full terms and conditions of its web site, the document is highlighted as a user guide, not the legal contract, and described in detail about data usage, data support, using and sharing, disclaimer, etc. (*The British Library, 2017*). In addition, the non-discrimination and open authorization of open data breaks the restriction of sharing conditions and specific sharing parties in traditional data sharing (*Gao, 2015*). Therefore, in the Shanghai library open data services, to refine the CC-BY-NC-SA content helps users to fully understand the agreement, but also maximize the law of openness based on defining the open data boundary with open data services deeply.

Attention to Data Developers' Community

The generation, opening, use, effect, and feedback of the open data is a dynamic process of circulation (*Zheng, 2015*), which involves multiple stakeholders, various factors of interaction, and the cause and effect loop. It is a complex and dynamic ecosystem. In this ecosystem, when data is open, the way and how well the data users use will ultimately determine the effectiveness of open data (*Zheng, 2015*). From a technical perspective, the user is divided into two groups, those with a non-technical background and those with a technical background. These two groups have significant differences in data use and data utilization ability. The main difference between them is that only the users with a technical background are likely to become data developers. Data developers can be divided into teams and individuals. The teams are usually teams of research and development companies, or groups of free combination. Both teams and individuals are the basic users of open data and play an important role in the open data ecosystem. Therefore, in the open data service of Shanghai library, we should pay attention to nurturing, mentoring, and promoting the data developer community and its growth. In particular, the R & D Company is regarded as a partner to maintain the interaction with developers. When necessary, to assist developers in transforming one-off activities into long-term projects (*i199IT, 2014*), to stimulate data developers enthusiasm for the use of data innovation and power, and help to form a positive cycle of an open data ecosystem. Hence, it makes contributions to the formation of a positive cycle of open data ecosystems.

CONCLUSION

The life of data lies in sharing (*Beijing Youth Daily, 2017*), the more sharing, the stronger the vitality, and the premise of sharing is opening. Data is the soil for the smart society and datamation of all businesses is the foundation of intelligence (*Xu, 2016*). Therefore, the open data is not only the wave of the world, but also the essential requirement of the development of the big data era. Although open data is confronted with many challenges such as data quality, data management and data ecology, opening up is inevitable. The cases of Shanghai library show that open data services help libraries with playing a positive role to libraries in the release of data values, improvement of their services, and promoting the exchange of knowledge and enhancement of public data quality. The library should open data and this will initiate a brilliant future for libraries and their users. Therefore, enabling libraries to establish a favorable space in open data.

REFERENCES

i199IT. (2014). *Open data certificate (2017). What can we learn from the G8 "Open Data Charter"?* Retrieved August 23, 2017 from http://www.199it.com/archives/270072.html

Beijing Youth Daily. (2017, June 4). In big data era, cooperation and sharing can win more. *Beijing Youth Daily*, p. A2.

Cabinet Office. (2017). *G8 Open Data Charter and Technical Annex*. Retrieved June 25, 2017 from https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex

Chen, S. S. (2016). *Genealogy Open Data API training*. Retrieved July 20, 2017 from http://pcrc.library.sh.cn/zt/opendata/doc/家谱开放数据接口培训.pptx

Chen, S. S. (2017). *Manuscript Open Data API Instructions*. Retrieved July 20, 2017 from http://pcrc.library.sh.cn/zt/opendata/2017/doc/6.手稿开放数据接口说明.pdf

Chinese Academy of Sciences. (2014). *Research and education open innovation application contest completed.* Retrieved July 18, 2017 from http://www.cas.cn/jypx/gzdt/201401/t20140120_4025735.shtml

Data.gov. (2017). *Dataset*. Retrieved June 10, 2017 from https://catalog.data.gov/dataset

Deliot, C. (2014*). Publishing the British National Bibliography as Linked Open Data*. Retrieved July 10, 2017 from http://www.bl.uk/bibliographic/pdfs/publishing_bnb_as_lod.pdf

Deutsche digitalebibliothek. (2017). *"Coding da Vinci" – Deutsche DigitaleBibliothekorganises Germany's first culture hackathon*. Retrieved July 12, 2017 from https://www.deutsche-digitale-bibliothek.de/content/ueber-uns/aktuelles/coding-da-vinci-deutsche-digitale-bibliothek-veranstaltet-ersten-deutschen-kultur-hackathon

Deutsche digitalebibliothek. (2017). *Frequently Asked Questions*. Retrieved July 12, 2017 from https://www.deutsche-digitale-bibliothek.de/content/faq#188

Gao, F. (2015). Open Data: Concepts, Status and Opportunities. *Big Data Research*, *1*(2), 9–18.

Huang, R. H., & Li, N. (2016). A Study of License Types of Open Data. *The Library*, (8), 16–21.

Jiang, G. X. (2015). *The global open data movement is booming*. Retrieved June 10, 2017 from http://www.cssn.cn/dzyx/dzyx_xyzs/201503/t20150330_1566110.shtml

Li, Y. (2015). *The four characteristics of the national government's open data*. Retrieved June 10, 2017 from http://www.360doc.com/content/15/0429/16/25235726_466847374.shtml

LIBRIS. (2017). *Technical and format information*. Retrieved July 02, 2017 from http://librishelp.libris.kb.se/help/tech_eng.jsp?open=tech

Ma, H. Q., & Pu, P. (2016). Research on the Connotation and Theoretical Foundation of Open Data *Library*. *Theory into Practice*, (11), 48–54.

Maali, F., & Erickson, J. (2014). *Data Catalog Vocabulary (DCAT)*. Retrieved September 10, 2017 from http://www.w3.org/TR/vocab-dcat/

Marden, J., Limadeo, C., & Whysel, N. (2013). Linked open data for cultural heritage: evolution of an information technology. In *ACM International Conference on Design of Communication*. ACM.

OGP. (2017). *About OGP*. Retrieved June 22, 2017 from https://www.opengovpartnership.org/about/about-ogp

Open Data Certificate. (2017). *Open Data Certificate-FAQ Page*. Retrieved August 23, 2017 from https://certificates.theodi.org/en/faq

Open Data Charter. (2017). *Principles*. Retrieved June 28, 2017 from https://opendatacharter.net/principles/

Open Data Handbook. (2017). *What is Open Data?* Retrieved June 21, 2017 from http://opendatahandbook.org/guide/en/what-is-open-data

Opengovdata.org. (2017). *The Annotated 8 Principles of Open Government Data*. Retrieved June 23, 2017 from https://opengovdata.org/

Sande, M. V., Verborgh, R., & Hochstenbach, P. (2017). Toward sustainable publishing and querying of distributed Linked Data archives. *The Journal of Documentation*, *2017*(1).

Shanghai.gov.cn. (2016). *Opinions on the implementation of big data development in Shanghai*. Retrieved August 18, 2017 from http://www.shanghai.gov.cn/nw2/nw2314/nw2319/nw2404/nw41165/nw41166/u26aw50223.html

Sohu. (2017). *2017 China local government open data platform report*. Retrieved June 22, 2017 from http://www.sohu.com/a/145532032_468714

The British Library. (2017). *British Library Catalogue Datasets in RDF and CSV*. Retrieved September 12, 2017 from http://www.bl.uk/bibliographic/pdfs/british_library_catalogue_dataset_tc.pdf

The British Library. (2017). *Sharing British Library Open Metadata with Non-Library Communities*. Retrieved July 10, 2017 from http://www.bl.uk/bibliographic/pdfs/sharing-bl-open-metadata-non-library-communities.pdf

Toronto Public Library. (2017a). *Toronto Public Library Open Data*. Retrieved July 15, 2017 from https://opendata.tplcs.ca/

Toronto Public Library. (2017b). *Toronto Public Library Open Data*. Retrieved September 10, 2017 from http://www.torontopubliclibrary.ca/terms-of-use/library-policies/open-data.jsp

Wang, H. M. (2007). The root seeking feelings of Chinese people from the genealogy. *Library Journal*, *26*(02), 85–88.

Wikipedia. (2017). *Open data*. Retrieved June 21, 2017 from https://en.wikipedia.org/wiki/Open_data

World Bank Group. (2017a). *Open Data Defined*. Retrieved June 23, 2017 from http://opendatatoolkit.worldbank.org/en/essentials.html

World Bank Group. (2017b). *Open Data Essentials*. Retrieved August 10, 2017 from http://opendatatoolkit.worldbank.org/en/essentials.html

Wu, M. (2012). Open Data in the United Kingdom and the United States Government Application and Implications. *Library and Information*, *2012*(1), 127-130.

Xia, C. J. (2015). From Portals to Platforms: The Implementation Technologies for Libraries. *Library Journal*, *34*(9), 19–22.

Xia, C. J. (2016). *Genealogy Open Data-Content and structure*. Retrieved July 20, 2017 from http://pcrc.library.sh.cn/zt/opendata/doc/开放数据竞赛培训.pptx

Xia, C. J. (2016). *Instructions*. Retrieved July 20, 2017 from http://data.library.sh.cn/tools/rdb2rdf/manual.pdf

Xia, C. J. (2017). *Celebrity manuscripts and Archives Linked Open Data-Content and structure*. Retrieved July 20, 2017 from http://pcrc.library.sh.cn/zt/opendata/2017/doc/3.名人手稿及档案关联开放数据——内容及结构.pdf

Xia, C.J., Liu, W., & Zhang, L. (2014). A Genealogical Ontology in the Form of BIBFRAME Model. *Library Tribune*, *2014*(11), 5-19.

Xiao, M., & Li, G. (2017). The Open Data Service of Library in the Environment of Big Data. *Journal of the Sichuan Society for Library Science*, *2017*(1), 23-26.

Xu, Z. P. (2016). *Data is Everything*. Retrieved September 12, 2017 from http://www.ccmw.net/article/106004.html

Yang, M., Xia, C.J., & Xu, H.B. (2016). A Study of the License Agreement of Open Data and Its Application in Domestic Libraries. *Library Tribune*, *2016*(6), 91-98.

Yu, M. Y., Zhai, J., & Lin, Y. (2016). Open Data Core Metadata Research in Chinese Local Governments. *Journal of Intelligence*, *35*(12), 98–104.

Yu, Y. (2014). *EU big data development strategy*. Retrieved June 10, 2017 from http://www.mofcom.gov.cn/article/i/jyjl/m/201412/20141200826137.shtml

Zheng, L. (2015). *The current situation, problems and difficulties of open data in China local government*. Retrieved August 15, 2017 from http://www.199it.com/archives/395885.html

Zheng, L. (2015). Study on Open Government Data: Definitions, Factors and Interactions. Chinese Public Administration, 2015(11).

Zheng, L., & Gao, F. (2015). The Study of Open Government Data Platform: Frame, Status and Suggestion. *E-Government*, *2015*(7), 8-16.

Zheng, Y.P., & Liu, M,C. (2016). The Status and Problem of Open Data Evaluation:Based on the Contrast and Analysis of Foreign Open Data Evaluation *E-Government*, *2016*(8), 84-93.

Zhou, Q.F., & Gu, Y. (2005). Advance of "Catalogue of Chinese Family Tree." *New Century Library*, *2005*(1), 11-13.

Zou, M. C., & Hu, Y. (2016). Research on Bibliographic Data Association of National Libraries in Europe and America *Library*. *Theory into Practice*, (11), 61–66.

Qi Zhang is a research librarian at the Shanghai Library in China, and graduated with her Bachelor's degree and Master's degree in Library Science Department from East China Normal University. Her research interests include open data, virtual reference service and library network service. She has published many papers in Chinese library professional journals.

Chunjing Zhang is now a research associate at the Shanghai Library in China. She graduated with her Bachelor's degree and Master's degree in the department of information science from East China Normal University in 2001 and 2004. Her research interests include digital library and open data. She has published several professional papers in Chinese research journals.

Zheyu Zhang is an assistant engineer at the Shanghai Library in China. She graduated with her Bachelor's degree and Master's degree in the department of computer science from Shanghai Ocean University in 2012 and 2015. Open data and public digital cultural service are her research interests. She has published 2 professional papers in Chinese research journals.