
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-202304.00636

Research on the Aggregation of Historical Archival Resources Based on Linked Data: Postprint

Authors: Jia Qiong, Wang Ping

Date: 2023-04-01T16:02:49+00:00

Abstract

[Purpose/Significance] Applying linked data technology to aggregate historical archive resources, aiming to explore potential relationships within resource data, discover new knowledge, achieve value augmentation, and thereby optimize knowledge services.

[Method/Process] By reviewing related research on the development, utilization, and aggregation of archive resources, this study constructs a framework for historical archive resource aggregation from the perspectives of objectives, principles, processes, and hierarchical relationships, and employs a “Jing-Shi-Zi-Ji” case study to analyze the resource aggregation process.

[Result/Conclusion] A theoretical guidance framework and retrieval service platform for historical archive resource aggregation from the linked data perspective are constructed, providing an implementation pathway from theory to practice for the development and utilization of historical archive resources.

Full Text

Preamble

Research on the Aggregation of Historical Archive Resources Based on Linked Data

Jia Qiong, Wang Ping

School of Management, Jilin University, Changchun 130022

Abstract:

[Purpose/Significance] The application of linked data technology in historical archive resource aggregation helps uncover potential correlations in resource

data, discover new knowledge, and achieve value-added benefits, thereby optimizing knowledge services. [Method/Process] This paper reviews relevant research on the development, utilization, and aggregation of archival resources, constructs a historical archive resource aggregation framework from the perspectives of objectives, principles, processes, and hierarchical relationships, and employs a “Jing Shi Zi Ji” (Classics, History, Masters, Collections) case study to analyze the resource aggregation process. [Result/Conclusion] The study establishes a theoretical guidance framework and retrieval service platform for historical archive resource aggregation from a linked data perspective, providing an implementation pathway from theory to practice for the development and utilization of historical archive resources.

Keywords: linked data; historical archives; resource aggregation; ontology; metadata

Classification Number: G250

DOI: 10.13266/j.issn.0252-3116.2021.10.011

In the era of big data, network information resources have grown exponentially in scale. Cultural institutions such as libraries, museums, and archives store vast quantities of historical archive resources, yet the digital preservation, development, and utilization of these resources require continuous in-depth advancement to meet users' needs for deep resource access. In recent years, the widespread application and development of linked data have enabled various information resource departments to gradually construct data networks. Linked data employs a triple-based approach to parse digital resources of different forms and collaborates with ontologies for deep-level description, maximizing the revelation of internal relationships among digital resources and thereby achieving interoperability between them. Linked data holds potential for addressing the current challenges of information silos and information overload in the deep organization and development of historical archive resources.

2. Related Literature Research

2.1 Development and Utilization of Historical Archive Resources

Scholars have accumulated substantial beneficial achievements in theoretical methods and strategy construction for the development and utilization of historical archive resources, laying a solid foundation for the micro-level practical application of technology in this domain. L.Y. Chen et al. [1] took intangible cultural heritage as their research object, advocating for the establishment of management information systems aligned with the characteristics of intangible cultural heritage and building interdisciplinary bridges. Niu Li et al. [2] extracted and combined logical frameworks from the general methodological system of digital humanities research and the development and utilization methodology of archival resources in digital environments, forming a methodological pathway system encompassing discovery, reconstruction, and storytelling. Zhao Xueqin et al. [3] constructed a path system for archival information resource de-

velopment based on digital humanities from four perspectives: transforming digital archival information resource integration methods, expanding the scale of archival information resource development, deepening cooperation between archival departments and digital humanities research teams, and broadening the fields involved in archival information resource development. Chen Jiayu et al. [4] constructed a crowdsourced development model for archival stories and proposed development strategies from the resource layer, transformation layer, and dissemination layer. Zhu Lanlan et al. [5] explored diversified development approaches for genealogical archival resources from four theoretical dimensions of digital humanities: deep development of genealogical archival resources based on text mining technology, construction of new genealogical archival databases based on data management technology, reconstruction of family archives and scenarios based on VR/AR technology, and construction of genealogical resource and cultural sharing platforms based on GIS technology. Scholars have also respectively used local characteristic archives [6], Yellow River archival resources [7], oral history archives [8], and other carriers to propose diversified development methods for historical archive resources using digital humanities approaches.

2.2 Resource Aggregation

Resource aggregation research is relatively abundant in the field of library and information science. T. Heath et al. [9] accelerated the realization of global data integration by constructing Web data and linked data applications. Qiu Junping et al. [10] studied author co-citation within the intelligence domain, demonstrating that co-citation-based collection resource aggregation is an effective method. Wang Hongjun et al. [11] constructed a metadata repository and dynamic link-based resource integration system after metadata collection, integration, and storage, enabling users to retrieve and utilize collection resources through knowledge navigation, information retrieval, and personalized services. You Yi et al. [12] used linked data to achieve semantic linking and resource browsing services for collection resources, and established RDF link creation and maintenance mechanisms in library linked data link management to achieve connections with external data. Qu Hui et al. [13] constructed a semantic co-word analysis model and performed semantic annotation on collection resources based on the results to achieve multi-dimensional aggregation. To a certain extent, archival resource aggregation has also entered the research 视野 of archival science, but the research pace has been relatively slow. Zhang Yunzhong et al. [14] used the “Women at War” case from the NARA digital archival annotation system to construct a digital archival resource aggregation model that helps improve the service benefits of digital archival resources. Shu Zhongmei [15] proposed an archival spatio-temporal ontology model and archival data extraction framework, constructing an archival linked data knowledge graph to expand the development and innovative utilization of archival collection resources and achieve archival resource value-added.

2.3 Research Review

Through a systematic review of domestic and international research status, scholars have formed a relatively systematic research system in terms of methodological pathways and strategy construction for historical archive resource development and utilization, which has greatly promoted development in this field. However, overall, the current research system contains more macro-level theories and fewer technological applications. Linked data-based digital resource aggregation research has been conducted in multiple fields domestically and internationally, attempting to interconnect linked data with intangible cultural heritage, library and information science, government affairs, and other domains to achieve deep resource aggregation. The pace of applying linked data technology to historical archive resource aggregation research in archival science has been relatively slow.

In the digital era of rapidly evolving information technology, the development of historical archive resources faces a series of theoretical and application challenges, including resource dispersion, heterogeneity, and information silos. Linked data provides a feasible research 思路 and solution for historical archive resource aggregation, playing an extremely important role not only in addressing issues such as heterogeneity and information silos among massive dispersed resources but also in promoting the development process of historical archive resources, enriching historical archive resource theories, and broadening the technical routes for historical archive resource aggregation. Based on this, this study takes massive, dispersed, heterogeneous, and value-rich historical archive resources as research objects, introducing linked data methods, technologies, and concepts to achieve deep aggregation of historical archive resources.

3. Historical Archive Resource Aggregation Analysis Framework

Constructing a linked data-based aggregation framework model is an important foundation for achieving historical archive resource aggregation. The following sections explore the objectives, principles, processes, and hierarchical relationships of historical archive resource aggregation to provide a basis for constructing a linked data-based historical archive resource aggregation framework model.

3.1 Resource Aggregation Objectives

The overall objective of historical archive resource aggregation is to focus on user needs, present resource collections to users in the form of knowledge systems, change the traditional presentation of independent knowledge fragments, save users' retrieval time, satisfy users' knowledge needs as much as possible, maximize the 挖掘 of historical archive resource utilization value, and enhance the effectiveness of knowledge services [16]. Specific objectives include:

3.1.1 Resource Organization

Historical archive resources exist in various carriers including paper, photographs, audio recordings, video recordings, and electronic formats, with diverse resource structures and types. Data heterogeneity leads to suboptimal effects in the utilization, development, and sharing of historical archive resources. The ultimate goal of resource aggregation research is to centrally aggregate heterogeneous data information from resource databases to promote information circulation and sharing. This aggregation is not simply integrating dispersed resources but, based on integration, first standardizing resources rationally and then further dividing them into different modules [17]. The final result of historical archive resource aggregation is to break the original model of dispersed, independent, and single resources, reshape the resource structure system, and construct a new, standardized, and systematic resource structure.

3.1.2 Knowledge Services

In historical development and research, how to effectively aggregate historical archive resources is the key problem this paper addresses. In the current environment, experts and scholars have employed numerous ontology construction, linked data, and semantic aggregation methods to analyze historical archive resources, hoping to effectively analyze the distribution patterns of different themes within the same historical archive resources. The aggregation of historical archive resources and their entities provides materials and resources for in-depth exploration of historical themes and research structures. With technical support, such as using scientific mapping and visualization techniques in the study of historical themes and research structures, experts and scholars can quickly clarify the structure of historical research [18].

3.1.3 Scientific Communication

Scientific communication is a relatively large-scale and common social phenomenon in academia, with scientific cooperation being its most typical manifestation. Traditional methods of obtaining historical archive resources through individual collection have various unavoidable problems. Compared with traditional data acquisition methods and research approaches for historical archive resources, the academic correlation network between historical archive resources provides a good research platform for resource network research, helping researchers conduct academic exchanges more effectively.

3.1.4 Interdisciplinary Integration

In historical research thematic structures, using interdisciplinary and cross-domain approaches has been a consistently hot topic of focus. Typically, experts and scholars engaged in historical archive research delineate resources in a certain historical field based on specific resource classification criteria, then use quantitative analysis methods to analyze the correlations or mutual citation situations between different types of entities contained in the resources, such as historical figures or historical materials, thereby revealing the status of interdisciplinary analysis and communication.

3.2 Resource Aggregation Principles

Referring to relevant standards and specifications such as archival metadata and DC metadata, this paper establishes metadata standards and specifications for historical archive resources to meet the needs of defining, repairing, organizing, interpreting, managing, preserving, and maintaining resource entities and digital objects, sustaining the lifecycle activities of historical archive resources. The aggregation process should adhere to the following principles: (1) Historical archive resources to be aggregated should be heterogeneous, independently existing, and distributed; (2) The first step of resource aggregation is collection, and the second step is aggregation—that is, first organizing existing historical archive resources into a whole, then sequencing the resources to achieve aggregation; (3) The aggregated system should provide a searchable platform for historical archive researchers and related scholars; (4) The ultimate purpose of aggregation is to improve resource utilization rates and facilitate the conduct and deepening of historical archive-related research. Therefore, the aggregation of historical archive resources is not simply aggregating various types and sources of heterogeneous historical archive resources into the same Web network environment but, more importantly, achieving deep aggregation of historical archive resource services and resources.

3.3 Resource Aggregation Process

Through aggregation, historical archive resources can be divided into various entities, their detailed attributes and relationships analyzed, resource catalogs organized, and catalog association and retrieval functions enabled—for example, clearly displaying all resources related to a certain dynasty, historical figure, or document, and associating all content expressions and carrier manifestations covered by a certain historical archive resource. Linked data-based resource semantic aggregation can reveal the structure of surface-level content and deep-level knowledge units in historical archive resources, with particular focus on establishing interconnections between resource entities to build a knowledge network.

The historical archive resource aggregation process encompasses stages of collection, resource description, semantic aggregation, and visualization presentation, as shown in Figure 1 [Figure 1: see original paper].

3.3.1 Resource Collection Stage

Historical archive resource data includes different types from different systems, exhibiting heterogeneous characteristics in storage distribution. When collecting historical archive resources purposefully, the characteristics and reliability of sources should be clarified, and historical archive resources must be generated from authentic and valid environments.

3.3.2 Resource Description Stage

Resource description is the prerequisite and foundation for resource collection, requiring compliance with unified metadata standards and frameworks. To facil-

itate computer recognition and processing, this paper reuses relatively mature metadata standard systems to uniformly describe data resources from different sources, structures, and types.

3.3.3 Semantic Aggregation Stage

The core of the entire process is semantic aggregation. When aggregating historical archive resource semantics, one can: (1) Use prior semantics for semantic classification aggregation, such as aggregating resources with the same semantic features based on semantic similarity of themes or content; (2) Conduct resource aggregation based on the organization of ontology concepts and conceptual relationships, such as revealing semantic relationships between metadata elements like personal names, institutional names, and titles in resource instances and performing semantic linking and mapping to achieve resource aggregation.

3.3.4 Visualization Stage

Visualization is the result presented in visual form. On the one hand, it can intuitively display complex links and associations in historical archive resources; on the other hand, among massive data and complex associations, users can utilize their intuitive cognitive and comprehension abilities to discover potential knowledge. Setting appropriate presentation forms and visual distances according to the importance of historical archive resources, in line with the visual thinking and habits of general users, is the desired outcome of visualization. The visual presentation of aggregated historical archive resources facilitates the joint mobilization of users' senses and perception, thereby assisting users in knowledge discovery.

4. Linked Data-Based Historical Archive Resource Aggregation

To enable historical archive resources to be semantic and capable of widespread linking with other resources, this paper utilizes linked data technology for their reorganization. Therefore, the content construction of historical archive resources is the focus of this section's research. The following sections propose linked data-based historical archive resource aggregation methods from the perspectives of hierarchical relationships and frameworks for historical archive resource aggregation.

4.1 Hierarchical Division of Historical Archive Resource Aggregation

This paper explores historical archive resource aggregation from three levels: ontology-based, internal, and external resource aggregation, as shown in Figure 2 [Figure 2: see original paper].

4.1.1 Ontology-Based Resource Aggregation

Historical archive resources consist of archival resources from different sources and types. To facilitate unified resource description, unified metadata standard formats need to be established to achieve interoperability between meta-

data. Referring to the OAI metadata harvesting protocol [19] and other standards, heterogeneous metadata (MARC, EAD, LIDO, TEI, etc.) from different sources, types, and formats are mapped to metadata ontologies to obtain datasets conforming to description specifications. XSTL is used to transform datasets through XML conversion to generate metadata conforming to RDF/XML description specifications, achieving semantic aggregation of heterogeneous metadata.

4.1.2 Internal Aggregation Based on Linked Data

According to linked data publishing principles, historical archive resources are named using URIs and their resource semantics are described. Using semantic similarity algorithms provided by OCLC, the similarity between historical archive resources represented in RDF format is calculated to 挖掘 relevant resources with various implicit relationships. Comprehensive utilization of widely applied vocabularies and mappings such as SKOS, VIAF, GEO, FRBR, and FOAF [20] transforms these implicit semantic relationships into explicit RDF semantic relationships. RDF can link heterogeneous resources to form a knowledge network, thereby completing intra-domain resource aggregation.

4.1.3 External Aggregation Based on Linked Data

The publication of historical archive resources requires linking with internal and external datasets from multiple data channels. URIs are used to identify and locate RDF-described resources and RDF links to related associated resources, and the semantic annotation tool Annocultortool [21] is used to annotate the linked data identified by URIs to obtain Web reference resolution HTTP URIs. Using HTTP URIs as bridges, cross-domain aggregation of historical archive resources with external related open linked datasets is achieved through services provided by the LOD Cloud (Linked Open Data Cloud).

4.2 Structural Framework for Historical Archive Resource Aggregation

The linked data-based historical archive resource aggregation framework is divided into three layers: data layer, publication aggregation layer, and service application layer, as shown in Figure 3 [Figure 3: see original paper].

4.2.1 Semantic Model Construction

For the content description of historical archive resources, different metadata standards exist in different historical periods. To enable non-uniform historical archive resource description information to have consistent management standards and models, this paper references METS and RDF and other relevant provisions to construct an open model that aggregates multiple metadata standards [22]. The model primarily includes core resource classes: data collections describing different data sources, historical archive resource objects contained in data collections, and digital manifestation forms of resource objects. To enable consistent semantic description of historical archive resources of different types, sources, and formats, the ORE standard is adopted, using the ore:Aggregation

class to associate all historical archive resource objects and datasets, forming a data collection network. Additionally, to strengthen associations in descriptions, this model creates metadata attributes for finding related historical archive resources, including name, identifier, version, definition, etc., and uses unified description standards to associate resource objects for better aggregation.

4.2.2 Linked Data Publication and Aggregation

After model establishment, the first step is local resource data linking, followed by establishing linked datasets for historical archive resource semantic services through vocabulary mapping, entity recognition, and data aggregation operations, and finally defining vocabulary terms and ontology-described data information concepts, terms, and logical structures. The aggregation achieved through establishing relationships between datasets in this paper enables resource sharing between datasets. This process can be described as: (1) Finding different vocabulary descriptions for the same concept or relationship through vocabulary mapping; (2) Establishing associations between the same entity data in different datasets through entity recognition; (3) Using the aggregation process to verify, filter, and merge conflicting attributes of the same entity object through manual intervention.

4.2.3 Aggregation Services

At this level, not only can historical archive resource data be semantically linked, but resource access and retrieval services can also be provided for users. The core value of historical archive resource aggregation services is to improve data quality and provide structured experience and information with unique insights. Using Web Service models with semantic support for knowledge representation, description, and reorganization of historical archive resources enables the reuse and sharing of historical archive resource digital data, enhancing their utilization value.

5. Application Example Analysis of Historical Archive Resource Aggregation

To enhance the reliability and usability of the historical archive resource aggregation framework, this paper uses a “Jing Shi Zi Ji” case study to validate the framework, employing Protégé and D2R for ontology construction and dataset publication respectively.

5.1 Metadata Scheme

Historical archive resource types are complex. Describing digital resources requires not only revealing the objects themselves but also interconnecting historical archives, making description more difficult than for other single information resources. The core metadata elements for historical archive resources in this paper reference FRBR, CDLS, DC, CDWA, and other standards, combined with the attributes of historical archive resources and archival data, with horizontal

extension and vertical expansion, divided into 4 categories and 21 elements, as shown in Table 1 .

Table 1. Core Metadata for Historical Archive Resources

Category	Elements
Historical Archive Description Information	Name, Era, Category, Identifier, Subject, Source, Creation, Specification, Location
Digital Object Management Information	Creator, Time, Copyright
Digital Object Description Information	Filename, Size, Format, Type
Related Resources	Related Persons, Related Knowledge, Related Events, Related Digital Resources

5.2 Ontology Construction

This paper constructs a Jing Shi Zi Ji domain ontology based on the seven-step method [23] and Protégé technology, with the specific implementation process shown in Figure 4 [Figure 4: see original paper].

(1) Clarify the domain and scope of the ontology. The constructed ontology targets the Jing Shi Zi Ji categories within historical archive resources.

(2) Reuse existing ontologies to improve work efficiency and obtain more interactions. Through literature review and investigation, existing Jing Shi Zi Ji metadata mostly follows the *Cataloging Rules for Ancient Books* revised in 2006, with unified cataloging and classification that fully reveals main information [24].

(3) Enumerate terms and concepts in the Jing Shi Zi Ji domain, obtaining as many relevant concepts and terms as possible to lay the foundation for defining hierarchical relationships later.

(4) Define classes and class hierarchical structures, selecting relevant vocabulary from the Jing Shi Zi Ji domain terms and analyzing logical relationships between terms to construct hierarchical structures.

(5) Define properties and property constraints as a conceptual framework for hierarchical relationships. To clarify the relationship between properties and classes, property descriptions need to be added. In Protégé, types such as Class, String, and Number can be used. Partial data and object properties of Jing Shi Zi Ji are shown in Figure 5 [Figure 5: see original paper].

To enable clearer and more complete expression of ontology semantic relationships, after defining conceptual semantic relationships, the ontology is visualized, as shown in Figure 6 [Figure 6: see original paper]. Compared with the

Jing Shi Zi Ji class hierarchy, ontology visualization can reflect richer and more specific connection descriptions.

(6) Ontology instantiation. Jing Shi Zi Ji ontology instances are established after classes and their properties are clarified.

(7) Ontology evaluation. Verify whether the ontology conforms to logic, including the rationality of classes, properties, and instances.

5.3 Linked Data Publication

Most current portal websites and information systems store data in relational databases, which should therefore be converted to RDF format. With the development and application of linked data, scholars have developed tools such as D2R, SquirrelRDF, and Virtuoso to convert database data into linked datasets. Using D2R as an example, its architecture is shown in Figure 7 [Figure 7: see original paper].

D2R primarily provides RDF data query interfaces, which can be called by RDF storage, HTML browsers, and SPARQL clients. D2R mainly includes the D2R server, d2rq engine, and d2rq mapping language. Its primary function is to use customizable d2rq mapping to convert data from databases into virtual RDF format data and then convert it into actual RDF files.

D2R's built-in startup service program can be used directly. Starting `d2r.server.bat` requires only the command “generate-mapping -u root -p 123123 -o db_{relic}.ttl jdbc:mysql://db_{relic}?useSSL=false” to execute the generated mapping file, then running D2R displays “[Server started at http://localhost:2020/]” to access the D2R service page in a browser.

5.4 Resource Aggregation

Since D2R's method for publishing linked datasets can only be retrieved through SPARQL, which is not suitable for ordinary retrieval users, this paper adopts MYSQL+ECLIPSE development technology to design a Jing Shi Zi Ji digital service platform for historical archive resources to aggregate resources.

This platform provides users with retrieval methods such as theme, era, and location. For example, when searching for “Zuo Zhuan” (Commentary of Zuo), retrieval can be performed using related keywords such as “Zuo Zhuan,” “Zuo Qiuming,” and “Spring and Autumn Zuo's Commentary.” In the search results, relevant information including name, images, theme, and figures is displayed. After double-clicking the name, the platform provides research results from related scholars, primarily research papers. Due to potential internal relationships among massive historical archive resources, users can click to query related events and figures, displaying retrieval resources for related events and figures, as shown in Figure 8 [Figure 8: see original paper].

In the digital humanities perspective, developing, 挖掘, and applying traditional

cultural resources has become a research hotspot. To discover implicit data relationships, potential knowledge, and value-added benefits in historical archive digital resources, humanities computing tools are needed. Linked data assigns unique URIs to resource entities and uses RDF to link digital resources, thereby forming a data network. This paper constructs a historical archive resource aggregation framework from the levels of metadata description, metadata-ontology mapping, RDF data publication, and semantic aggregation and access, providing theoretical guidance for historical archive data resource aggregation. The historical archive resource aggregation framework provides an implementation path system for the “Jing Shi Zi Ji” case study, and the final established historical archive data resource aggregation service platform can meet users’ needs for querying and retrieving historical archives.

References

- [1] CHEN L Y, WANG B Z, CHEN B. Research on digital museum for intangible cultural heritage[C]//IEEE international symposium on IT medicine and education. Jinan, 2009: 734-737.
- [2] NIU L, GAO C X, ZHANG Y F, et al. Discovery, reconstruction and storytelling: paths and methods of archival research from the perspective of digital humanities[J]. Journal of Library Science in China, 2021, 47(1): 88-107.
- [3] ZHAO X Q, DANG Z, LI T E. Problems and countermeasures in the development of archival information resources from the perspective of digital humanities[J]. Beijing Archives, 2021(1): 18-22.
- [4] CHEN J Y, TANG L L, WANG X Y. Research on crowdsourced development models for archival stories in the we-media era[J]. Zhejiang Archives, 2021(1): 28-30.
- [5] ZHU L L, BO T Y. Diversified development of genealogical archival resources from the perspective of digital humanities[J]. Zhejiang Archives, 2019(10): 31-33.
- [6] LU D D, NIE Y X. Development paths for local characteristic archival resources from the perspective of digital humanities[J]. Shanxi Archives, 2020(4): 29-33.
- [7] LI Z F, YU Z S. Research on the development and utilization of Yellow River archival resources from the perspective of digital humanities[J]. Shanxi Archives, 2021(1): 73-79.
- [8] GAO S, WANG X N. Research on the development and utilization of oral history archival resources from the perspective of digital humanities[J/OL]. Shanxi Archives. [2021-04-09]. <https://kns.cnki.net/kcms/detail/14.1162.G2.20210112.1728.002.html>.
- [9] HEATH T, BIZER C. Linked data: evolving the Web into a global data space[J]. Synthesis lectures on the semantic Web: theory and technology, 2011, 1(1): 1-136.
- [10] QIU J P, ZHOU Y. Analysis of deep aggregation models and services for collection resources based on author co-citation: taking ontology research in library and information science in CSSCI as an example[J]. Library and Information Service, 2014(7): 19-24.

- [11] WANG H J, LIU M, LI Y, et al. Integration and implementation of library resource integration system based on metadata repository and dynamic linking[J]. Chinese Journal of Medical Library and Information Science, 2011(10): 65-67, 70.
- [12] YOU Y, CHENG Q. Discussion on collection resource aggregation model based on linked data[J]. Information Theory and Practice, 2013(1): 109-114.
- [13] QU H, QIU J P. Research on collection resource aggregation based on semantic co-word analysis[J]. Information Science, 2016, 34(2): 15-20.
- [14] ZHANG Y Z, FENG S S. Research on digital archival annotation system resource aggregation based on topic maps[J]. Library and Information Service, 2018, 62(14): 116-124.
- [15] SHU Z M. Research on archival knowledge graph construction under the background of digital humanities[J]. Shanxi Archives, 2020(2): 53-60.
- [16] DONG K. Research on multi-level deep aggregation of digital literature resources[D]. Wuhan: Wuhan University, 2014.
- [17] YANG X M. Research on dynamic semantic aggregation of ubiquitous learning resources[J]. e-Education Research, 2014, 35(2): 68-73.
- [18] ZHANG M. Social network analysis based on keywords of ethnic minority archives[J]. Heilongjiang Archives, 2019(2): 23-24.
- [19] DING Q J, SU J, ZENG J X. Research on national metadata database and its collaborative construction framework[J]. Information Studies: Theory & Application, 2020, 43(10): 88-92, 80.
- [20] ZHANG P T. Analysis of the linking of bibliographic data in the British Library[J]. Journal of the National Library of China, 2015, 24(4): 103-113.
- [21] WANG P, HUANG X P. Research on semantic fusion methods for digital cultural resources based on linked open data: a case study of Europeana[J]. Library and Information Service, 2016, 60(12): 29-37.
- [22] European data model primer[EB/OL]. [2021-04-09]. https://pro.europeana.eu/files/Europeana_{Profession
- [23] LI D P. Research on the construction of discipline knowledge map based on ontology[D]. Wuhan: Central China Normal University, 2011.
- [24] QU Y W. Organization of ancient book digital resources based on linked data[J]. Encyclopedia Knowledge, 2020(27): 5-7.

Author Contributions

Jia Qiong: Responsible for paper writing and revision.

Wang Ping: Responsible for research 思路 framework design.

Jia Qiong, Wang Ping

School of Management, Jilin University, Changchun 130022

Abstract: [Purpose/Significance] The application of linked data technology in the aggregation of historical archive resources can help explore potential correlations in resource data, discover new knowledge, and realize value increment, thereby optimizing knowledge services. [Method/Process] This paper sorted out

related researches on the exploitation and utilization of archival resources and aggregation, constructed a framework for historical archive resource aggregation from the aspects of objectives, principles, processes, and hierarchical relationships, and analyzed the resource aggregation process using the “Jing Shi Zi Ji” case study. [Result/Conclusion] The study constructed a theoretical guiding framework and retrieval service platform for historical archive resource aggregation from the perspective of linked data, providing an implementation path from theory to practice for the development and utilization of historical archive resources.

Keywords: linked data; historical archive; resource aggregation; ontology; metadata

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.