

Research Article

Library Resource Construction and Interconnection Model Based on Linked Data

Xinyuan Fan 

Library, Sichuan University, Chengdu 610065, Sichuan, China

Correspondence should be addressed to Xinyuan Fan; xy-fan@scu.edu.cn

Received 7 May 2022; Revised 29 June 2022; Accepted 12 July 2022; Published 4 August 2022

Academic Editor: Imran Shafique Ansari

Copyright © 2022 Xinyuan Fan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Heterogeneous and self-distributed massive data has always been the core issue of knowledge openness and knowledge sharing. Because the library has a wide range of data sources, various types of resources, and many operating platforms, a large amount of heterogeneous data is stored in it, which greatly affects the sharing of resources in the library. Linked Data, as one of the implementation methods of Semantic Web technology, can promote the interconnection and interoperability of data in different fields and promote the opening and sharing of knowledge. Therefore, this paper proposes the construction of library resources based on Linked Data, aiming to study how to optimize the association rules of Linked Data in the context of the Semantic Web and promote the continuous realization of open sharing of information resources. Experiments show that, under the action of the optimized association rules, the support degree between data can reach 79%, and the data association degree can be increased to 84%.

1. Introduction

Libraries have shouldered the burden of inheriting culture and history since ancient times. In the Internet environment, although digital technology has been extremely developed, people still retain many information resources with unique historical significance in the library. However, the standards and storage protocols used to preserve resources in libraries are not the same. To a certain extent, this leads to massive resources being isolated and discrete from each other. In order to connect the isolated data with each other, the library field has adopted many methods, such as data fusion and resource integration. Based on this, the library community began to seek breakthroughs in Linked Data. However, the current Linked Data is often aimed at the interconnection of the same document resources and cannot realize the association of different document resources that are related to each other.

Linked Data, as a way of linking and publishing data, realizes the interconnection of different literature resources. Therefore, the library resource construction based on Linked Data proposed in this paper will continuously promote the interlibrary resource interoperability and realize the real data interoperability. At the same time, the research of the

article will further exert the advantages of Linked Data, so that knowledge discovery and knowledge sharing will continue to become the mainstream of society. And the development of Linked Data has also promoted the development of the World Wide Web to a certain extent.

The innovation of this paper is as follows:

- (1) The article proposes a new library resource construction plan and library resource interconnection model and links the Linked Data technology with the library, which is bound to promote the open sharing of library resources.
- (2) The article combines theory with practice, which not only describes the results of Linked Data optimization from the theoretical level, but also develops and enriches Linked Data theory from the actual situation of library lending.

2. Related Work

Many scholars have provided a lot of references on the research on Linked Data technology, library resource construction, and library resource interconnection mode.

Building on existing infrastructure, Hu et al. built a web-based system for real-time tracking of red tides caused by toxic dinoflagellates. This red tide threatens human and environmental health in the eastern Gulf of Mexico every year. The system integrates different data through a customized web interface and can also automatically detect the correlation between different data [1].

Vander Sande et al. identified the shortcomings of commonly used Linked Data publishing methods and took into account the current lack of a large amount of Linked Data exposed by LAM. They proposed a new way to publish Linked Data. They also demonstrated this by combining queries with other relational data sources to archive the DBpedia version [2].

Nuzzolese et al. proposed a new approach to Linked Data exploration. The method uses encyclopedic knowledge schemas as relevant criteria for selecting, organizing, and visualizing knowledge. The pattern was discovered by mining Wikipedia's link structure and assessed through user-based research. Based on this, they built a tool called Aemoo. It supports knowledge exploration driven by encyclopedic knowledge schemas and integrates data from heterogeneous sources, namely, static and dynamic knowledge, as well as textual and Linked Data [3].

Radulovic et al. proposed a quality model for Linked Data. This provides unique terminology and reference for Linked Data quality specification and assessment. The quality model described above specifies a set of quality characteristics associated with Linked Data and a formula for metric calculation. In addition, they have extended the W3C Data Quality Vocabulary, so that it can be used to capture quality information and quality characteristic information for specific Linked Data [4].

Hartmann et al. proposed a method that can improve the reliability of data association. They noted that the progress robustness of the data is measured by a metric. This criterion acts at the beginning of data generation and tends to deviate from the initial correlation where the data is generated. This has serious implications for the correlation between data [5].

Chen et al. introduced a data interface conversion method, which is used to convert specific semantic structure data into the data format that people need through a special conversion method. The data conversion method can convert all semantic structures on the current semantic web. But the converted data format supports only a few formats. But its conversion speed is fast, and the error rate is low [6].

Zhao et al. found an optimal planning model for data association when they researched three renewable energy sources: wind, photovoltaic, and concentrating solar energy. This mode minimizes the discreteness of the data. Using this model, they designed different interconnection models of renewable energy for different regions. The data capacity of its renewable energy can be planned by this model [7].

3. Linked Data and Library Resource Construction

3.1. The Semantic Web. Every kind of information and resource has its own semantics. Relying on the recognition of

semantics, people can easily find and distinguish various data [8]. But for computers, the semantics of information resources make little difference. Therefore, in order to better communicate and cooperate between humans and machines, the concept of the Semantic Web has been proposed. The Semantic Web is an intelligent web. It makes the entire Internet a universal information exchange medium by adding semantics to documents on the World Wide Web.

3.1.1. Semantic Web Technology. Because computers cannot identify the semantics of different information, people specifically add computer-understandable semantic identifiers to resources on the Internet. And this technology that allows computers to understand the semantics of information is called Semantic Web technology [9]. Semantic Web technology turns the Internet into a data network with the same standards. In this network, both humans and computers can easily read and publish information. The hierarchical structure of the Semantic Web is shown in Figure 1. The Semantic Web technology has undergone many developments, and its protocols and standards have also undergone several revisions. However, no matter how many times it is changed, technologies such as RDF and SPARQL are always the foundation of Semantic Web technology.

3.1.2. RDF. RDF stands for Resource Description Framework. It is a standard language proposed by W3C to describe the metadata of all resources on the Internet. RDF provides a unified description standard for resources on the Internet and also facilitates the construction of specific Internet semantic associations [8]. With a unified resource description framework, people have realized the integration of network information resources and also standardized the data resources on the network that are not good or bad. RDF realizes the expression of information with the least constraints, and the data information it expresses is independent of different applications and is beneficial to the data interconnection between different applications. RDF adopts a simple description method; that is, based on XML technology, it uses triples composed of three elements of subject, predicate, and object to represent the semantic information of resources. This semistructured XML description of (S, P, O) triples is also very convenient for storage and use. A set of RDF data can form an RDF directed graph. Every RDF triple exists in this form. Figure 2 is a diagram showing an example of the structure of an RDF triple.

3.1.3. SPARQL. RDF provides a new way of expressing information resources on the Internet, but this standard language cannot be recognized based on traditional query languages [10]. Therefore, in order to adapt to the development of the RDF language, people developed an RDF query language: SPARQL. The query language can query any information resource represented by RDF, and its syntax is very simple. The syntax of SPARQL is rewritten by people referring to relational database SQL language, so the

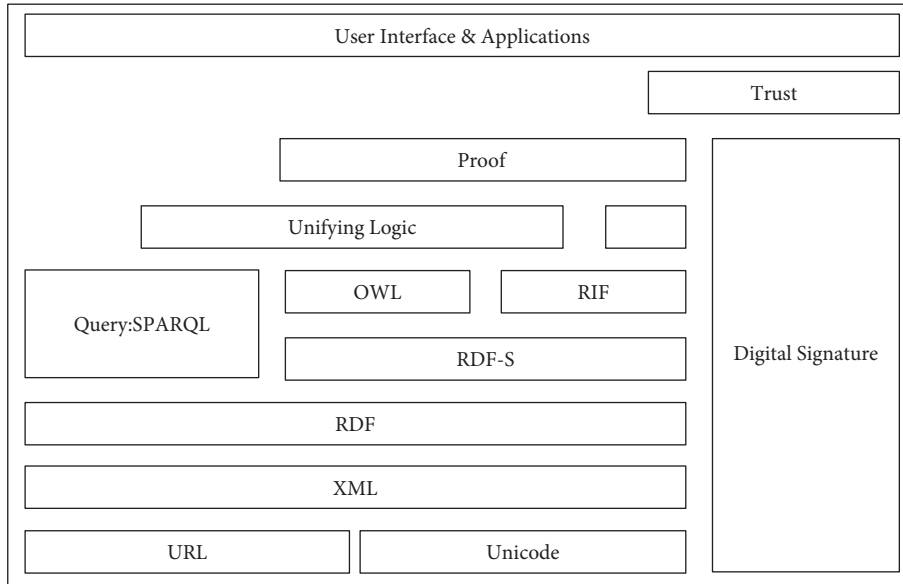


FIGURE 1: Hierarchical architecture of the semantic web.

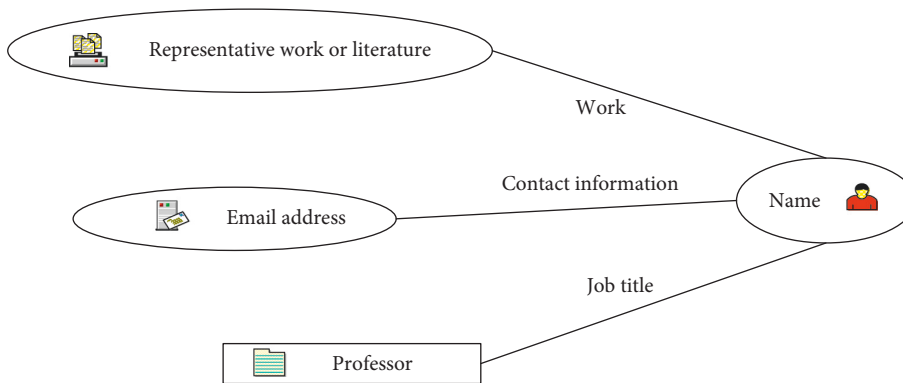


FIGURE 2: RDF triple example diagram.

language structure of the two is very close. As a data query access language, SPARQL is suitable for both local and remote queries. It enables easy remote access to Linked Data Networks and federated retrieval of different types of RDF resources.

3.2. Linked Data Technology. Linked Data is one of the implementations of Semantic Web technology. It can continuously promote the interconnection of data in different fields and promote the opening and sharing of knowledge. It is a data processing technology that applies the RDF format model [11]. Linked Data adds machine-understandable semantic description information to resources on the World Wide Web and realizes the sharing and association of Web resources. This standardizes the entire Internet resources and initially builds a semantic data network that is strictly described, tightly connected, and capable of automatic evolution. In this semantic network, every user can quickly and accurately find and utilize information resources on the Internet. The relationship

between Linked Data and the Semantic Web is shown in Figure 3.

The advent of Linked Data presents a huge opportunity for libraries. Linked Data allows libraries to provide services according to a normative standard schema, thereby truly integrating themselves into the entire world of data and information [12]. In recent years, the application of library Linked Data has made great progress, but generally speaking, it is still in its infancy. The application of Linked Data in various fields is being explored and discovered by experts and scholars, and new results will be continuously achieved.

3.2.1. Features and Principles of Linked Data. The characteristics of Linked Data are mainly reflected in that it establishes association relationships for different data and uses this relationship to provide object discovery and identification services [13]. Therefore, the construction and interconnection of Linked Data are the key to the release and application of Linked Data.

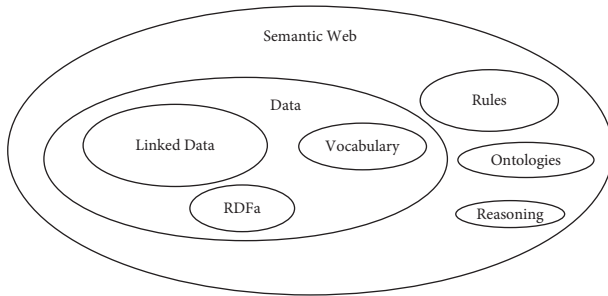


FIGURE 3: The relationship between linked data and the semantic web.

Linked Data building interconnection is to first build separate datasets and then build links between the separate published linked datasets. Linked datasets after interconnection will achieve greater practical value. The publishing process of Linked Data refers to adding meaningful and related information to the Internet, so that Linked Data can be more easily retrieved and used by people. The key point of Linked Data is to make the relationship between data units have certain semantics. The W3C organization takes into account both the realization of the Linked Data function and the simplicity and practicality and finally formulated four Linked Data release guidelines:

- (1) It uses URI as the identifying name of any resource object.
- (2) Through HTTP and URI, people can locate objects specifically, and objects can refer to each other.
- (3) When the URL of an object is queried, it provides information in RDF form or SPARQL standard.
- (4) It provides relevant links as much as possible to point to other URLs in order to discover more objects [14].

These four principles involve the definition of resources, resource identification, resource description, and linking respectively. One aspect of this criterion is to locate Web resources using URIs that can be referenced by the HTTP protocol, so that after the resource provider publishes and deploys, users or applications can obtain and use these associated data and achieve interconnection through the HTTP protocol. On the other hand, it uses RDF and other methods to describe various heterogeneous data resource entities in the Web that we think are meaningful, points to related resources by means of RDF links, etc. and reveals the semantic relationship between resources. It finally converts unstructured data on the network and structured data using different standards into structured data that follows the unified Linked Data standard.

3.2.2. Key Technologies. We now know that Linked Data is a technology built on top of Web technologies. The Web technology mainly involves three aspects: HTTP, URL, and HTML.

- (1) HTTP, which is a Web rule that defines in detail the mutual communication between a browser and a

World Wide Web server, is a data transfer protocol for delivering documents over the Internet to the World Wide Web. HTTP is a passport for the circulation of information resources throughout the network.

- (2) URL, which is an identifier for locating data information resources on the World Wide Web, is generally composed of three parts: the naming mechanism for accessing the data resource, the host name for storing the data resource, and the name of the data resource itself [15].
- (3) HTML is a markup language that can be used to describe the current web document [16]. The reason why HTML is called Hypertext Markup Language is because its text documents contain so-called hyperlink points. The so-called hyperlink is actually a pointer in the form of a URL. By activating the pointer, the user can make the browser in use conveniently obtain a new web page address, thereby obtaining new information.

There is a common interaction among HTTP, URL, and HTML to jointly generate services. If the World Wide Web is likened to a giant intertwined web, then HTML is the intersection of this web. The URL is the identifier of the location of each intersection, which is used to mark these intersections and is a unique identifier for each intersection. And HTTP is every intertwined thread on this big web. It interweaves the intersections of the entire network in series to form a large whole [17]. The goal of Linked Data is to build an interconnected structured and data-based giant semantic web. The Semantic Web uses the RDF form to describe the network resources of data information, so the Semantic Web using Linked Data generally requires a data model in the form of RDF.

3.2.3. Data Association Rules. Before talking about association rules, we first need to understand how resources are composed [18]. Resource integration is the most common way of composing data resources, and association rules are how to discover and find integrated resources. In the current association rules, people generally use the keyword matching method; that is, the related resources are defined according to the corresponding keywords.

But this method ignores the discovery of other related resources, so its association rules are not perfect. As an information service provider, the library has introduced new technologies and new ideas in recent years and is also researching and developing new resource integration methods and new association rules according to its own characteristics. Based on this, resource integration and association rules based on Linked Data have become the first choice for libraries. Figure 4 is a structural diagram of resource integration based on Linked Data. This figure mainly reflects the effective integration of network information resources by converting the current document Web into the data Web and establishing associations between different data through URLs.

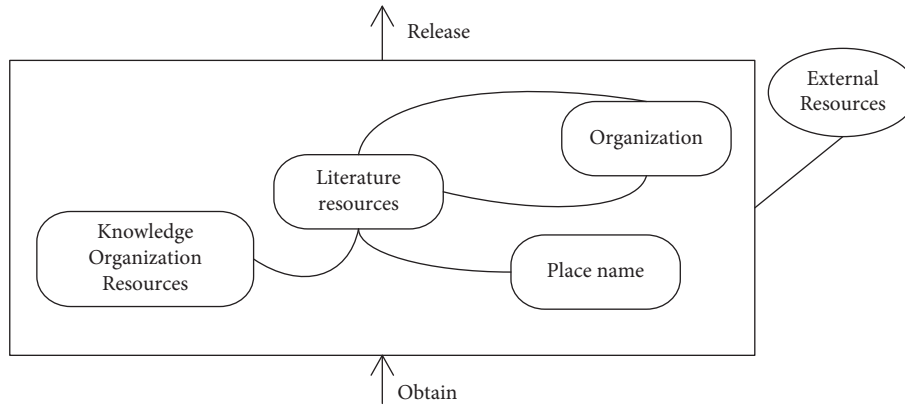


FIGURE 4: Resource integration based on Linked Data.

3.3. *Interconnection of Library Resources.* With the development of computer and network technology, the structure and scope of library collections have undergone tremendous changes. It is no longer limited to traditional formats such as paper documents and is more derived into digital resources [19]. Network and digital technology not only bring opportunities for library development, but also bring new challenges to library management and services. The application of digital technology makes the resources of the library richer, more diverse, and wider. However, different digital resources are often in the hands of many libraries, widely distributed in different databases, and the data storage formats and organizational forms are also different, so the problem of heterogeneous distribution of data arises. Different databases also contain overlapping and repeated content, and the degree of knowledge correlation is restricted. Massive, distributed, and heterogeneous data is not conducive to users' search and use. In order to improve the accuracy and comprehensiveness of the information users need to obtain, it is necessary for the resources in the library to be interconnected. And with the improvement of user needs, the deep integration of resources is an inevitable requirement for the development of resource interconnection construction to a certain stage. The flow chart of library Linked Data dynamic service is shown in Figure 5. The figure reflects the four parts of the library dynamic service composition platform: data layer, management layer, business layer, and user interface.

Linked Data can be linked to other resources, and it supports structured data, which can be well used for retrieval and browsing of semantically related data. Therefore, linking library resources to achieve association based on a single data set is the greatest application value of Linked Data. For the library, the construction of data association is the premise of realizing the browsing and expansion of the library's associated data. Therefore, the method of building the association of library Linked Data is extremely important, and it is also an important part of the integrated management of library Linked Data. The association construction method of library Linked Data is mainly divided into two categories: unique identification method and RDF association method according to the universality of its application.

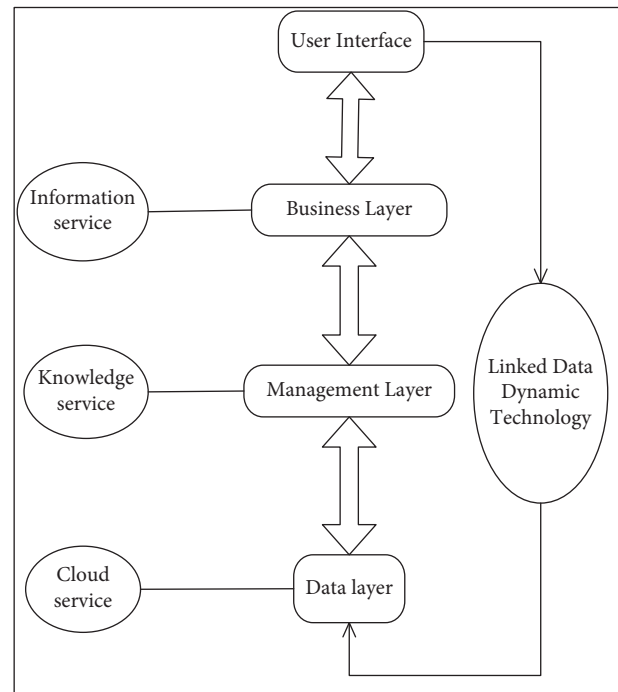


FIGURE 5: Flowchart of linked data dynamic service.

3.3.1. *Status Quo of Library Resource Interconnection.* In the Internet environment, the library is both a publisher and a receiver of information. But there is a big difference between the types of Linked Data in libraries, and between them. The basic types of library Linked Data are shown in Table 1. Moreover, in the World Wide Web environment, libraries are publishing more and more data, and the relationships between different data sets are becoming more and more complex. At this time, as long as the state of any data set changes, it may lead to the disconnection of the related data link. Therefore, the interconnection of library Linked Data is the basic premise of Linked Data management [20]. With the connection of library resources, network resources and real resources can be merged, and new connotations are continuously given to resources.

TABLE 1: Basic types of library Linked Data.

Library linked data type	Resource publisher	Resource affiliates
Linked Data in the library	Data service provider in the library	Other resources in this library
Interlibrary linked data	In-library and other library data publishers	Related resources from other libraries
Off-library linked data	External data publisher	Resources on the internet other than libraries

3.3.2. The Role of Linked Data in the Construction of Library Resources. Linked Data, as the best practice of the Semantic Web, has a huge space for resource integration and sharing. Linked Data aims to get rid of the lack of semantics of existing network information. By publishing and linking structured data, it can achieve semantic association between scattered and heterogeneous data islands, thus promoting the evolution of traditional file network to shared data network. Linked Data, as the main means of promoting the realization of the data network, naturally has the advantage of integrating resources. Linked Data can integrate resources into a truly seamless and infinitely open whole and can also enhance the semantic correlation between resources by combining with ontology technology. With such technical support, the interconnection of library resources can be put on the agenda. And, in this process, through the resource description of the associated data, users will be able to find the required resources more smoothly. Then, based on this library, refined push and information resource management can also be achieved. After repeating the above two steps, the library can publish Linked Data information independently and then continue to build into an intelligent and convenient library, truly realizing the integration of library and resources. The construction process of library resources based on Linked Data is shown in Figure 6.

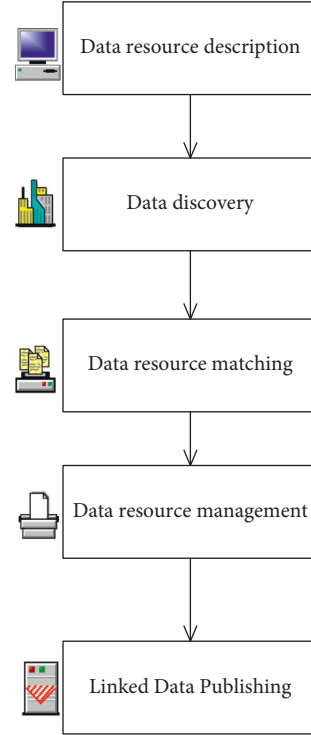


FIGURE 6: Library resource construction process based on Linked Data.

4. Library Resource Construction and Interconnection Model Based on Linked Data Optimization

We learned above that Linked Data is a behavioral scheme in the Semantic Web that aims to connect both data and reality, which is feasible and realistic. However, the structured format provided by Linked Data is relatively rigid and cannot adapt to the ever-changing reality. Therefore, to establish the correlation between data, it is necessary to mine its association rules first. At present, the traditional association rules cannot realize the effective sharing of data, so optimizing the association rules has a great effect on the construction and sharing of library resources.

In traditional association rules, S is the degree of association, D is the data set in rule mining, I is the set of items in the data set, and its function expression is shown in the following formula:

$$S_{\alpha} = \sum_{i=1}^n \frac{D(n)|I_i - I_n|}{\sum_{i=1}^n D} \quad (1)$$

In the above formula, all subsets of I are called item sets of D . It defines K as a set in I , and then, K is also an itemset in D , which is defined as follows:

$$K = \{I_1, I_2, I_3 \dots I_n\} \quad (2)$$

Each item set has a unique resource identifier X ; when this identifier is consistent with the regular characters in the data set, we call the item set support data set, which is defined as

$$P = \sum_{i=1}^I K_X \frac{\text{count}(D)}{\|\sum_{i=1}^n I\|} \quad (3)$$

$$\text{count}(D) = \|\sqrt{k_n}\| \times \sum_{i=m}^I K_{\alpha}$$

$\text{count}(D)$ is the number of occurrences of regular characters in the dataset, and P is the support of the item set K to the association rules in the dataset. The larger the value is, the higher the correlation between the item set and the data set is, and the smaller the value is, the lower the correlation is.

However, the degree of association only describes the number of occurrences of the itemset in the entire data set, and it cannot describe the situation of the itemset itself. Therefore, in order to study the credibility of the item set itself, we introduce the concept of confidence, which is defined as follows:

$$C = \frac{\sum_{i=m}^L I \times K\alpha}{\sum_{i=n}^I \tan D_\theta}. \quad (4)$$

Among them, C represents the size of the confidence, and the larger the value is, the less credible the item set is. In order to further study the relationship between support and confidence in the dataset, we limit it.

$$\lim_{I \rightarrow 0} \frac{\sqrt{S_{n+1}}}{\sqrt{S_n}} \geq S_{\min}, \quad (5)$$

$$\lim_{I \rightarrow \infty} \int_0^\infty (C_n \times C_{n+1}) d_n \geq C_{\min}, \quad (6)$$

S_{\min} represents the minimum value of support, and C_{\min} represents the minimum value of confidence. If there is an item set I_0 that satisfies both equations (5) and (6) in the limit state, the item set is called a strong association rule. In this case, datasets with strong association rules are often of interest to readers and users.

However, under such association rules, the selected datasets often cannot be adaptively updated and adjusted, and such rules cannot promote resource sharing well. Because, under this rule, the data resources that users can finally see are discrete and isolated, therefore, based on the original association rules, we reoptimized its recommendation mechanism and association mode. First of all, we define a discrete set A for the isolated and discrete cases of the data. It can preprocess all candidate data, and its definition is shown in the following formula:

$$A = \left\{ \frac{\sqrt{D_i} - \sqrt{I_n}}{\log(\text{peer}(I))} \leq S, \quad \text{peer}(I) \subseteq X_i \right\}. \quad (7)$$

$\text{peer}(I)$ is an attribute value in the set, which specifies that the data attribute in the set can only appear once.

$$\text{peer}(d) = \frac{I_n}{I_i} \sum_i X_i, \quad (8)$$

$$X(i) = X^2(n) + \sum_{I \subset X} y \frac{1}{\text{count}(I)}.$$

After the assignment of the weight of $X(i)$, the discrete data will be organized together by the new weight to form an item set. But this item set B is different from the previous item set because each data in this item set has an independent weight.

Different weights of data will be given when the data is called, and the weight will directly affect the number of occurrences of the data and ultimately determine the relevance of the data to other data. The weight assignment function is as follows:

$$W = b^{(X_i/I_i)} * \text{count}_i(n) + \alpha^{i-1} * \text{count}_{i-1}(n-1) + 1,$$

$$b_I = X_n * I_{b1} + X_{n-1} * I_{b2} + y, \quad (9)$$

$$y = b^{i+1} * I_i^n + b^2 * n + b,$$

b is the observation parameter under different exponential decay, and α^i defines the initial weight. y is the weight decay factor. When there is a new data increment or the data changes, it will lead to a new observation parameter, namely, b_I . It eventually affects the weights to change, generating new weights W .

In this way, the data can be dynamically adjusted and transformed, and the degree of association between the data sets will also be continuously adjusted as the data changes.

In order to further reduce the isolation and discreteness of the data, we need to introduce an observation function T , which is defined as follows:

$$T = \sum_{i=1}^{\infty} \frac{\sqrt{\mu}}{\sigma}, \quad (10)$$

σ and μ represent the standard deviation of the data sample and the mean of the data isolation, respectively. If the final value of T is greater than the determined minimum value we set in advance, then we say that the data sample is highly isolated. Alternatively, when the value satisfies a normal distribution, and the value is greater than a preset data confidence level, we also call it data isolation.

Data samples that are isolated from each other can be clustered by the Means algorithm. Its function is expressed as

$$d = \sqrt{\sum_{i=0}^n (a_i - b_i)^2}, \quad (11)$$

$$a = \frac{y_b - y_a}{\sqrt{\text{Var}^d}}, \quad (12)$$

It is easy to prove from formulas (11) and (12) that

$$y^d = \frac{\text{Var}^d}{\text{count}(d)} - \left(\frac{L^d}{\text{count}(d)} \right)^2. \quad (13)$$

It can be shown that the more the frequent items contained in the discrete data, the farther the distance between it and the assigned initial cluster, that is, the weaker the similarity between the data. Then, we proceed to calculate its mean for each cluster. The calculation method is as follows:

$$\tilde{Y} = \frac{\sum_{L \subset X} (L_i - L_{i-1})^2}{\text{Var}_m}, \quad (14)$$

$$L_i = \ln n \otimes \left(\sum_{i=1}^{\infty} (\ln x - Y_n) \right).$$

\tilde{Y} and L are the mean results of the clusters and the cluster centers between the clusters. After repeating this step, the mean will begin to converge and approach a constant indefinitely. Usually, we use the error value to define

$$Q = \sum_{i=1}^m \sum_{d \in I} |p - m_i|^2, \quad (15)$$

m is the final number of clusters obtained, and Q is the square of the sum of errors over all the data in the dataset. p is a random point in the data space, representing a randomly assigned data object. The smaller the error, the higher the final aggregation degree of the data. The larger the error, the higher the isolation of the data, and the weaker the correlation. In order to realize the interconnection of data and make the data more relevant, the next article will start from the actual library data and continuously verify the data interconnection mode optimized based on Linked Data.

To this end, this paper collects the corresponding codes of some bibliographies in the library and the datasets to which they belong. It aims to study the relationship between different bibliographies and provide a reference for the recommendation mechanism and interconnection mode of library data in the following. The corresponding code of the bibliography is shown in Table 2.

Table 2 shows that the codes corresponding to different borrowing items of the library are unique. Based on this, we found that the corresponding book code of "Semantic Web Technology Research" is Tp312/12, and the corresponding book code of "Linked Data Method" is Tp312/137.13 by checking the relevant coding records of the library. The corresponding book code of "Programming Language Design Case Course" is Tp312/342, and the corresponding book code of "RDF Data Standard Research" is Tp312/486. "Semantic Web and Linked Data" corresponds to Tp312/122, "C Language Programming" corresponds to Tp312/623, and "Java Programming Fundamentals" corresponds to Tp319/145.

For convenience and brevity of presentation, we only label the relevant bibliographies in the table and then compose them into a dataset. Dataset $D = \{A. \text{Semantic Web Technology Research}, B. \text{Linked Data Methods}, C. \text{Programming Language Design Case Tutorial}\}$, and then we explore association rules for this dataset. Figure 7 shows the relevance degree of related bibliography and the exploration of association rules.

Figure 7(a) shows that when the relevance degree of the bibliography is detected, the original association rule can only detect the bibliography that is closest to the keyword. In this case, the correlation between the bibliographies is relatively high, which can reach 50%. Specifically, when checking the bibliography related to the programming language design case tutorial, under the original association rules, only one C language programming was included in the retrieval result, but the JAVA program foundation related to it was not detected. Figure 7(b) shows that when the RDF data standard research is used for correlation research, the detection correlation results are relatively low, only about 10%.

In order to further study the defects of the original association rules, we subdivided the original data set and divided the following transaction set table, as shown in Table 3. Then, based on the transaction table, we further calculated the confidence between different categories. The results are shown in Table 4.

Table 3 shows that different types of bibliographies can actually correspond to different data sets. For example, RDF data standard research can belong to both the data set of Linked Data and the data category of semantic web

technology research. The purpose of dividing specific bibliographies into transaction sets is to further subdivide their associations, which can provide data support for our subsequent research.

Table 4 shows that after different transaction sets are combined, it produces different candidate datasets. It undergoes techniques such as connection and screening, and the final candidate dataset will be less and less. But in this process, the confidence of association rules will change accordingly. It takes Semantic Web and Linked Data as an example, which belong to two datasets A and B together. When we match the bibliography with the association rules, the first thing to be found is the Semantic Web technology research and the associated data method, and the confidence level at this time is 0.6. However, after connecting the data, we will find that the RDF data standard research actually belongs to this category, and the confidence level becomes 1 at this time.

To investigate whether this method is feasible, we compared it with several common association rule algorithms on the market. Figure 8 shows the data association degree of different association rules under 8 consecutive experiments.

Figure 8(a) shows that, at the beginning of the experiment, our method performs comparable to the other two algorithms in terms of data correlation. In the first 4 experiments, the data correlation of our method can reach a minimum of 60%, and it has achieved initial results. Figure 8(b) shows that our method maintains relatively good correlation in the latter few experiments. Compared with the other two algorithms, it has a relatively large advantage, and the data correlation degree is basically stable at 80%.

Figure 9(a) shows that, compared with the traditional data association rules, our optimized association rules have higher data support, and the highest can reach 3.4. Figure 9(b) shows that, after optimization, our association rules not only improve in support, but also have certain advantages in confidence. This shows that the data correlation based on our association rules can basically reach more than 75%.

But 75% data correlation does not solve the real-life dynamic data transformation problem. Based on this, we add a weight adjustment strategy to the optimized association rules. Table 5 shows the weighted association rules.

Table 5 shows that, compared with the original data confidence, the weighted confidence can more accurately reflect the relationship between the data. It uses the Semantic Web and Linked Data as examples. When we match the bibliography with the weighted association rules, the researches on Semantic Web technology, Linked Data methods, and RDF data standards are found, and the weighted confidence is 100%.

At the same time, we noticed that, in the above weighted association rules, there are 0% cases. But this does not mean that there is no correlation between the two data. Rather, this is a problem arising from the definition of confidence itself. Therefore, in order to explore the situation where there is no correlation between the data, we introduce the concept of data lift. Figure 10 is the weighted relational data lift.

TABLE 2: Corresponding codes of some bibliographies in the library.

Card number	Book number	Data set
120824071023	Tp312/12, Tp312/137.13	D
120824071015	Tp312/137.13, Tp312/342	A
120824071002	Tp312/342, Tp319/145, Tp312/122	B
120824071045	Tp312/486, Tp319/145	C
120824071068	Tp312/122, Tp312/623	A
120824071037	Tp312/623, Tp319/145	D
120824071017	Tp319/145, Tp312/342, Tp312/486	E

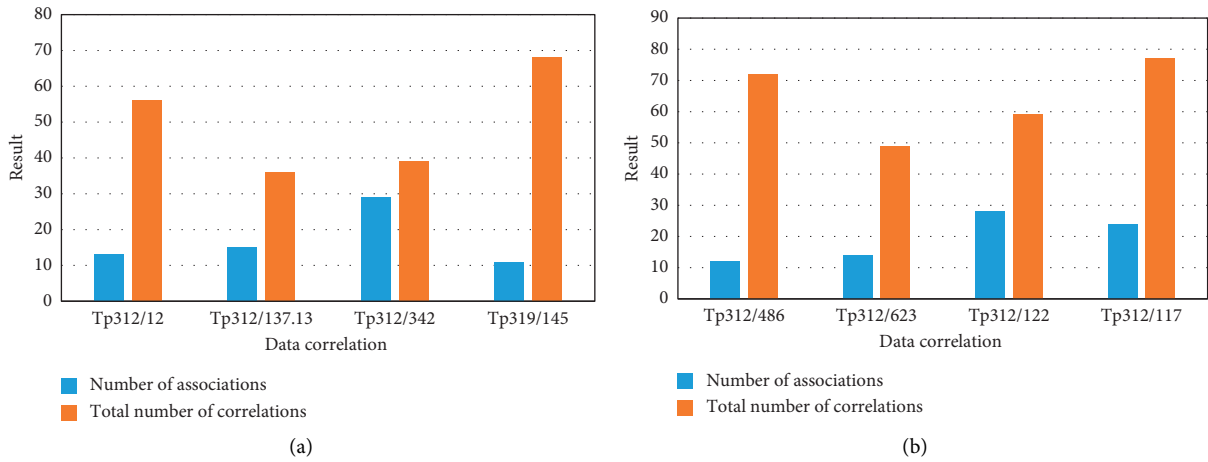


FIGURE 7: Library bibliographic association.

TABLE 3: Transaction set table.

Card number	Classification number	Belonging to the data set
120824071023	A	A, D
120824071002	C	A, B, C
120824071017	D	C, D, B
120824071068	B	B, D

TABLE 4: Association rule confidence.

Association rules	Candidate set	Confidence level
B/C	A, D, B	1
D/B	C, A, D	1
C/D	B, A	0.6
AD/C	A, B	0.7
C/DB	A, B, D	1
A/B	C	0.7
B/C	A	0.5

Figure 10(a) shows that the weighted data lifts are basically greater than 1. This shows that there is a positive correlation between the data. And when the weight is 1, the improvement degree basically reaches 2. Figure 10(b) shows that the unweighted lift fluctuates around a value of 1 when the weight is continuously raised. This shows that the correlation also fluctuates between presence and absence, and the data is relatively unstable. On the other hand, the weighted lift is basically kept at around 2, indicating that it still shows a positive correlation.

5. Discussion

Linked Data is a recommended method for publishing and enabling shared data on the web. The emergence of this technology provides us with a lightweight and incremental data interconnection model. Based on Linked Data technology, libraries can make full use of Linked Data standards and specifications to continuously integrate libraries into information flow and become part of Internet information resources. Moreover, after the library takes the lead in

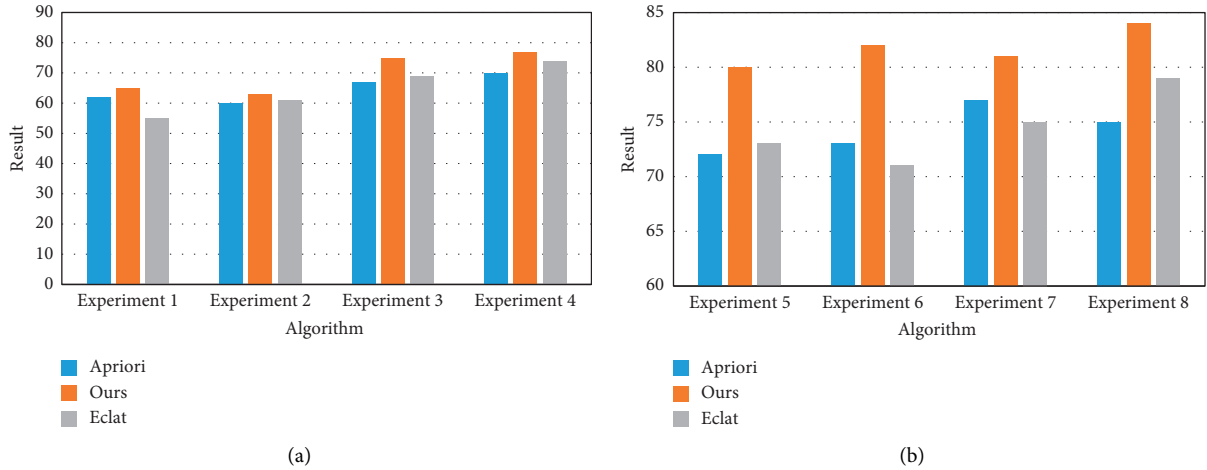


FIGURE 8: Data association degree of different association rules.

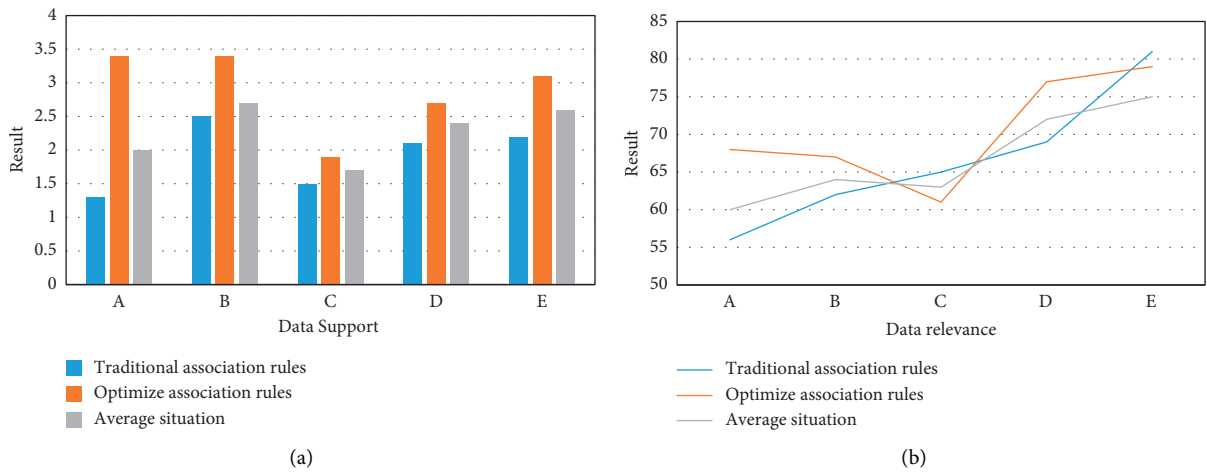


FIGURE 9: Data support comparison.

TABLE 5: Weighted association rules.

Association rules	Weighted confidence level (%)	Association rules	Weighted confidence level (%)
B/A	75	AB/C	100
A/B	100	AD/B	47
A/D	25	BC/D	86
BD/C	100	D/CB	70
A/DB	100	C/AB	100
D/B	72.5	AB/D	0
A/C	50	ADB/C	75

realizing the integration of resources, it can promote the integration and interoperability of other resources on the Internet and continuously promote the development of

Semantic Web technology. At the same time, computers and humans share a single data recognition specification, which will promote great progress in human-computer interaction.

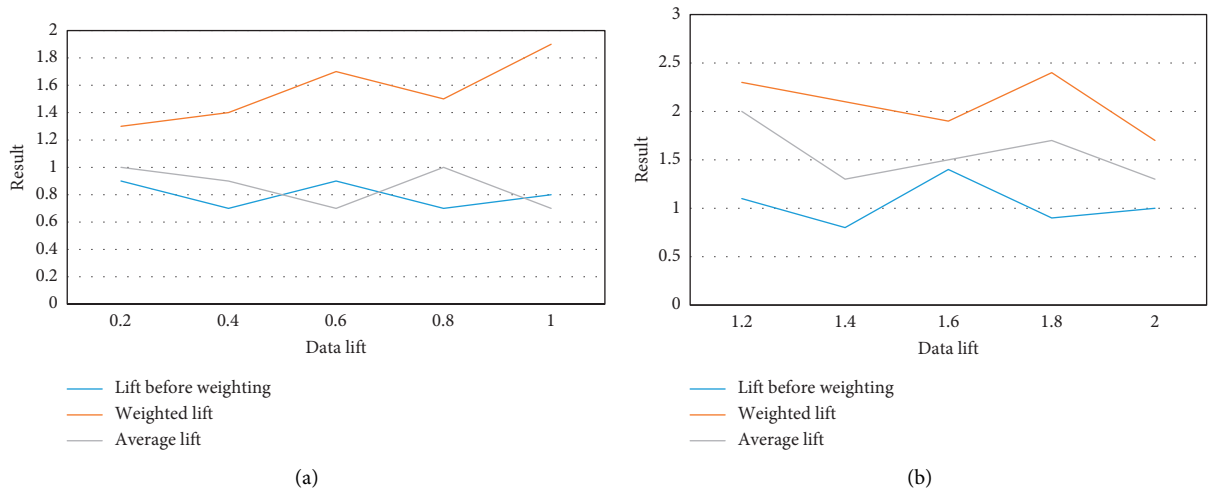


FIGURE 10: Weighted linked data lift.

6. Conclusion

Starting from the technology of Linked Data, the basic characteristics and principles of Linked Data, this paper focuses on the current rules of data association. On this basis, it puts forward the theory of library resource construction based on Linked Data. However, the article only simply optimizes the association rules of the data and does not upgrade the logic of the associated data from the bottom layer, so the optimization still has shortcomings. At the same time, the library resource construction and interconnection model proposed by the article based on this theory also has great room for improvement. In the future, the interconnection of library resources will further drive the development of Linked Data and fill the gaps for this study.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by 2019 Information Science & Library Science Foundation of Sichuan University: Research on Organization Ordering and Semantic Release of Characteristic Collection Resources based on Linked Open Data (No. Sktq201905).

References

- [1] C. Hu, B. Murch, A. A. Corcoran et al., "Developing a smart semantic web with linked data and models for near-real-time monitoring of red tides in the Eastern Gulf of Mexico," *IEEE Systems Journal*, vol. 10, no. 3, pp. 1282–1290, 2016.
- [2] M. Vander Sande, R. Verborgh, P. Hochstenbach, and H. Van de Sompel, "Toward sustainable publishing and querying of distributed linked data archives," *Journal of Documentation*, vol. 74, no. 1, pp. 195–222, 2018.
- [3] A. G. Nuzzolese, V. Presutti, A. Gangemi, S. Peroni, and P. Ciancarini, "Aemoo: linked data exploration based on knowledge patterns," *Semantic Web*, vol. 8, no. 1, pp. 87–112, 2016.
- [4] F. Radulovic, N. Mihindukulasooriya, R. García-Castro, and A. Gomez-Perez, "A comprehensive quality model for linked data," *Semantic Web*, vol. 9, no. 1, pp. 3–24, 2017.
- [5] V. Hartmann, T. Lahmer, and K. Smarsly, "Project scheduling under uncertainty and resource constraints," *Pammatone*, vol. 17, no. 1, pp. 839–840, 2017.
- [6] F. Chen, R. Burgos, and D. Boroyevich, "A bidirectional high-efficiency transformerless converter with common-mode decoupling for the interconnection of ac and dc grids," *IEEE Transactions on Power Electronics*, vol. 34, no. 2, pp. 1317–1333, 2019.
- [7] L. Zhao, R. Yu, Z. Wang, W. Yang, L. Qu, and W. Chen, "Development modes analysis of renewable energy power generation in North Africa," *Global Energy Interconnection*, vol. 3, no. 3, pp. 237–246, 2020.
- [8] X. Guo, Q. Qu, X. Guo, W. Yang, and P. Zhang, "Economy supervision mode of electricity market and its incentive mechanism," *Global Energy Interconnection*, vol. 3, no. 5, pp. 504–510, 2020.
- [9] S. Cui, J. Hu, and R. De Doncker, "Fault-tolerant operation of a tlc-mmc hybrid dc-dc converter for interconnection of mvdc and hvdc grids," *IEEE Transactions on Power Electronics*, vol. 35, no. 1, pp. 83–93, 2020.
- [10] T. K. Das, T. S. Kumbar, and T. Ramswaroop, "Use of QR codes for enhancing access to library resources and services: a case study of IIT Gandhinagar library," *Annals of Library and Information Studies*, vol. 67, no. 3, pp. 156–163, 2020.
- [11] V. E. Ikolo, "Doctor's awareness and perception of medical library resources and services: a case study of delta state university teaching hospital (Delsuth), Nigeria," *International Journal of Library and Information Science*, vol. 9, no. 2, pp. 58–71, 2020.
- [12] D. A. Corpuz, "Library resources and functional effectiveness of an academic library: meeting the challenges of the digital age," *Humanities & Social Sciences Reviews*, vol. 8, no. 4, pp. 238–245, 2020.

- [13] B. A-Khateeb, "Determinants of electronic library resources access in Saudi Arabia higher institutions," *International Journal of E-Business Research*, vol. 17, no. 1, pp. 32–41, 2020.
- [14] A. K. Chauhan and R. T. Kumbar, "Awareness about public library resources in rural Varanasi: a study," *Library Herald*, vol. 57, no. 4, p. 448, 2019.
- [15] G. Afebende and G. M. Nna-EtUk, "Information retrieval tools and utilization of library resources by undergraduate students in federal universities in south-south zone, Nigeria," *Journal of Librarianship and Information Science*, vol. 5, no. 2, pp. 36–44, 2019.
- [16] S. A. Sadiku, A. O. Issa, and M. K. Mohammed, "The use of electronic library resources in Nigerian Universities: a review of the literature," *International Journal of Library and Information Science*, vol. 8, no. 2, pp. 40–50, 2019.
- [17] T. Lamanna, "On educating patrons on privacy and maximizing library resources," *Information Technology and Libraries*, vol. 38, no. 3, pp. 4–7, 2019.
- [18] D. Etse and M. S. Boateng, "Management of end-of-life library resources in Ghana: strategies and Sustainability Implications," *The Journal of Academic Librarianship*, vol. 45, no. 4, pp. 368–376, 2019.
- [19] A. L. Association, "15 working librarians to receive awards under a council on library resources fellowship program," *College & Research Libraries News*, vol. 30, no. 7, pp. 225–226, 2019.
- [20] D. W. Yiu and C. Lau, "Corporate entrepreneurship as resource capital configuration in emerging market firms," *Entrepreneurship: Theory and Practice*, vol. 32, no. 1, pp. 37–57, 2008.