

Research on Intelligent Acquisition of Shared Data Based on Government Information Resource Directory

Jingwen Li^{1, b}, Shaofu Lin^{1, 2, c}, Xiaofeng Jia^{3, a}, and Song Gao^{3, d}

¹School of Software Engineering, Beijing University of Technology, Beijing 100124, China.

²Beijing Institute of Smart City, Beijing University of Technology, Beijing 100124, China.

³Beijing Information Resource Management Center, Beijing 100101, China.

^aCorresponding author: jiaxf@bjeit.gov.cn

^bjingwen76453@163.com, ^clinshaofu@bjut.edu.cn, ^dgaos@bjeit.gov.cn

Keywords: Government information resource directory, data sharing, intelligent acquisition

Abstract: At present, the state continues to pay attention to the integration, convergence, opening and sharing of government information resources, and the Internet + government has made rapid progress. However, the government data sharing and opening is not thorough, government information exchange mechanism is not smooth, and there are still information islands between departments. The key issue is that government information resource directory and data sharing application are still “two layers of leather”, which restricts the sharing of government information resource. Aiming at these problems, this paper carries out research on intelligent acquisition of shared data based on government information resource directory. Combining five application scenarios, on the basis of solving the problem of disjointing between directory and shared data, this paper further studies the intelligent acquisition of shared data, so as to realize seamless connection between directory and data, and realize efficient, automatic, and intelligent data acquisition.

1. Introduction

Since the 18th National Congress of Chinese Communist Party, in accordance with the overall requirements of the “four comprehensive” strategic deployment, the “five-in-one” strategic layout and the five development concepts, the states has successively issued a series of documents , such as: the “Interim Measures for the Administration of Government Information Resources Sharing”, “Integration and Sharing of Government Information System”, “Guidelines for the Compilation of Catalogues of Government Information Resources”, to make overall deployment for the construction of government information resources sharing. It is clearly stated that by 2020, the departmental sharing of government data and social opening should be basically realized. Foreign countries have carried out many useful explorations in the informatization of government governance, such as: the British government promotes a “paperless” society, the US government has increased the construction of a database integration platform, and the Canadian government has fully promoted the full coverage of the administrative information network, which has provided experience for China to improve government information construction.

At present, China has entered the stage of e-government comprehensive construction, and has also achieved initial success in the information resource directory system and the convergence of government data. However, current difficulties in sharing government information resources, difficult business collaboration, indiscriminate sharing between departments, and information islands still restrict the further development of Internet + government. Solving questions about how to make organic connection between directory and data resource management, and how to open up the barrier from the data management to data acquisition to realize the real-time online access is the key and promise of intelligent acquisition in the future. Therefore, research on seamless docking technology between government information resource directory and data open sharing has important scientific research value and social significance.

2. Demand analysis

The construction of government informatization in our country has gone through three development stages: “office automation” in the single-machine, “government internet access” and now “information resource application”. With the development of time and technology, China's government information resources sharing has achieved some results, but still faces some challenges and changes, as highlighted in the following two aspects:

(1) Data: Society has gradually stepped into the era of big data. Data has gradually changed from the original single, small and fixed data to big data with complex, diverse, changeable and massive characteristics. In the field of government affairs, with the changes in the characteristics of data itself, the form of sharing management has also changed from simple data exchange to data aggregation of various departments, and from different departments of decentralized sharing of systems to the overall aggregation of data sharing platform and cloud.

(2) Cataloguing: Over the years, starting from business sorting, cataloguing in accordance with "business-matter-resource-field" has become an obstacle to the current directory and data management work. Directory and data "two layers", directory application effect is poor. With the development of big data, the data volume of departments is becoming larger and larger. All departments have changed from sorting out their own business functions for cataloging to cataloging according to the data categories collected and aggregated, from manual input to automatic system cataloging. The above problems and changes have seriously restricted the sharing and opening of government data. The data collection standards of various departments are not unified, the problem of heterogeneous data is serious, the information distribution is scattered, cross-departmental and cross-level business coordination is difficult, and fusion analysis is difficult. As a result, the island of information is created under the current situation.

Under the current changes, it is worth studying how to better connect the directory with data and support the sharing and opening of data resources. Therefore, this paper from the technical level researches on the combination of directory and data sets to solve the problem of disjointing between data and directory, research intelligent acquisition of shared data based on government information resource directory, further to solve the shared problems of low efficiency, slow and complicated process, so as to realize relevance, connect, integration and interaction between government information resource directory and data management, open, sharing and application.

3. Overall architecture design

3.1 General idea

This paper aims at the problems of disjointing between data and government information resource directory. Based on the current situation of data aggregation, directory system architecture and cataloguing in a province, using the function of government information resource directory as the main line, combining the application scenarios of data opening and sharing, this paper research and propose a set of technical solutions to support data sharing and intelligent acquisition by directory.

3.2 Overall architecture model

The system technology implementation of the government information resource directory supports data sharing and opening involves 5 application scenarios and 3 technical levels. The overall architecture diagram is shown in Figure 1.

According to the simple to complex order, the system technology implementation of the directory supports data sharing and opening are divided into the following five ways:

- 1) Directory-based data service management (Business layer)
- 2) Directory-based information navigation query (Application layer)
- 3) Directory-based data exchange configuration and management (Data layer)
- 4) Directory-based application system integration and interaction (Application layer)
- 5) Directory-based service integration assembly and interaction (Application layer)

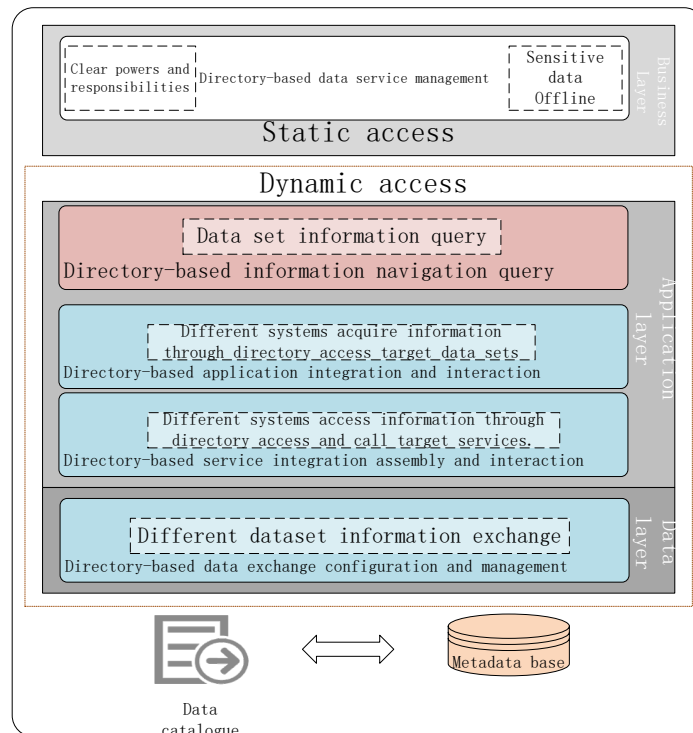


Fig 1. Overall architecture model

This paper will carry out technical research in the above order. This technology system can improve the display form of directory service and content, which not only presents it in human-computer interaction but also realizes the machine-computer interaction. At the same time, based on the directory system, we further designed lots of data exchange modes, such as: “pipelined”, “centralized” and “subscription”, to realize government information resources directory multi-angle and multi-layer supports data sharing and exchange.

4. Technical solutions design

4.1 Directory-based data service management

Directory-based data service management is implemented at the business layer. For government sensitive data, we design it based on the principle that sensitive data is not shared or conditionally shared. Such sensitive data is not allowed to be published or shared online, using metadata to generate a data resource directory of sensitive data. The demand department can understand the existing data and data ownership through directory, and be agreed to obtain offline. The design of directory-based data service management can not only share the sensitive data conditionally, but also clarify the ownership and sharing side of the data resource, and then clarify the rights and responsibilities. The schematic is shown in Figure 2.

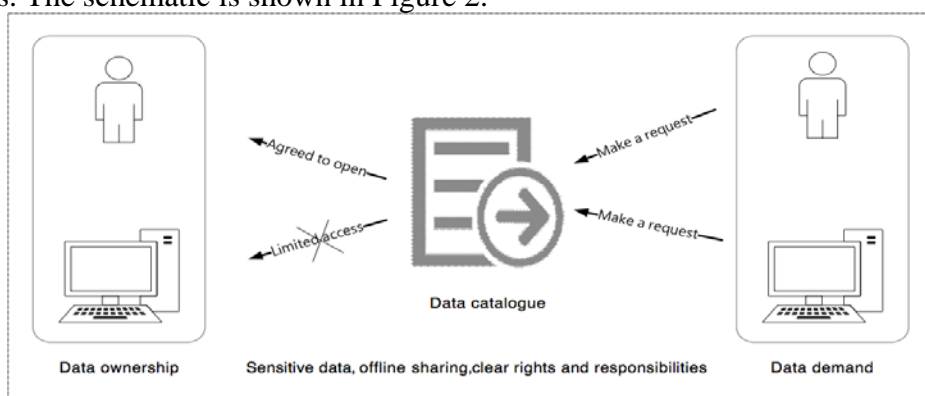


Fig 2. Directory-based data service management schematic

4.2 Directory-based information navigation query

The directory-based information query and navigation service is implemented at the service layer. The government information resource directory management center provides an information inquiry function of human-computer interaction mode, and provides services by means of interfaces. The user logs in to the data directory management service platform according to the authority, and makes a request to the directory management center through the querying service interface, the management center verifies the requester authority, and retrieves the corresponding information resource in the shared repository. If the information is not enough, the requester should make a request, the provider will accept the review and authorize the use after approval. The government information resource directory management center can provide query services such as classified navigation, conditional search and full-text search based on the government information resource directory, and provide directory content browsing function. The user can obtain the required government information resource directory content and electronic information resource content according to the authority, as shown in Fig 3.

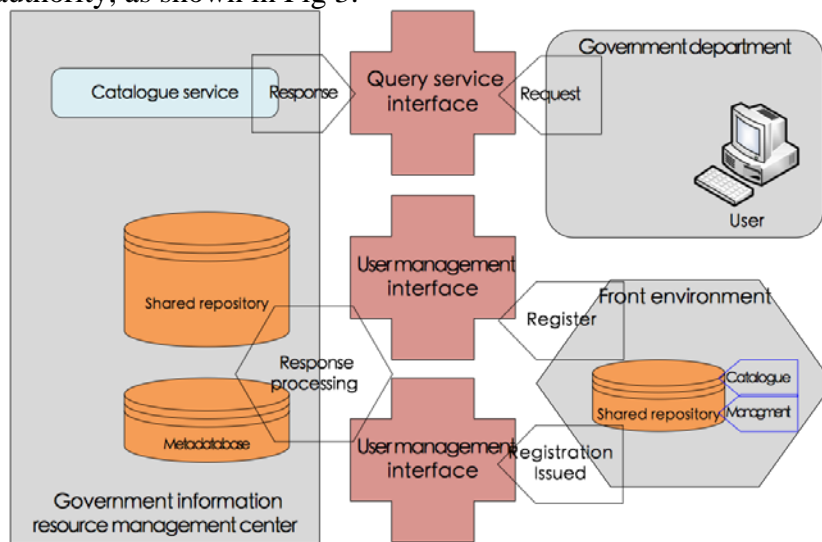


Fig 3. Directory-based information query navigation service schematic

4.3 Directory-based data exchange configuration and management

Directory-based data exchange configuration and management is implemented in the data layer. It is mainly applied to logically aggregated data. The directory supports the configuration and management between data sets (databases or files). After the exchange parties clearly open the shared permissions, they configure between exchanged databases based on the same directory, once the configuration is successful, the directory will no longer be used, and data can be exchanged directly between the data sets (databases). The schematic is shown in Figure 4:

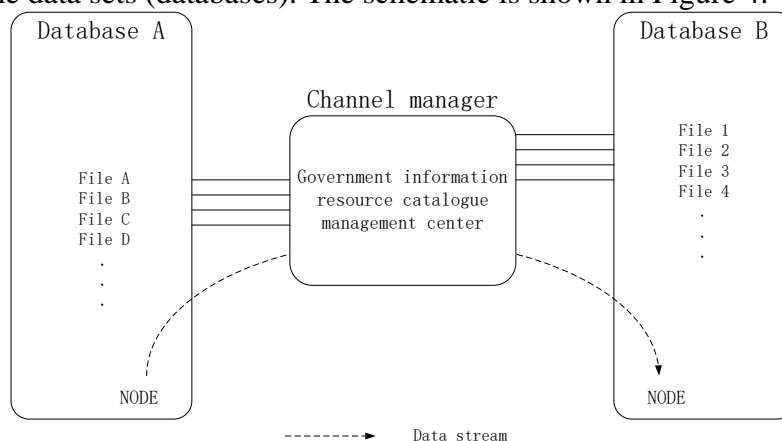


Fig 4. Directory-based data exchange configuration and management schematic

This mode belongs to the central service mode. The channel server is deployed in the management center. The center is responsible for channel management and operation monitoring. Data transmission department is responsible for channel connection, task configuration, monitoring and access management at the sender, and the data receiving department is responsible for these at the receiver. Once the directory-based configuration is successful, the data sending department can extract data from the source database and upload it to the channel. The data receiving department can directly download the data into the database through the channel. This way can effectively decouple and solve the multiple authorization issues of the target data.

4.4 Directory-based application system integration and interaction

Directory-based application system integration and interaction are implemented at the application layer, used in data sharing which aggregation by logic and can solve the sharing between heterogeneous data. Further than that, it is beneficial to solve the problem of information resource occlusion and repeated acquisition caused by using different systems. This sharing can be implemented in the following two ways: ftp/file sharing server mode and database sharing data mode.

The ftp/file sharing server mode is: system B accesses the information resource directory and makes a data request, after system A receiving the request and uploads the required file to the shared file server. B receives and downloads, then return the information that situation of monitors or processing information to A. Using this approach needs ensure that there is the common file server and consistent file data format for both systems. The data flow diagram is shown in Figure 5.

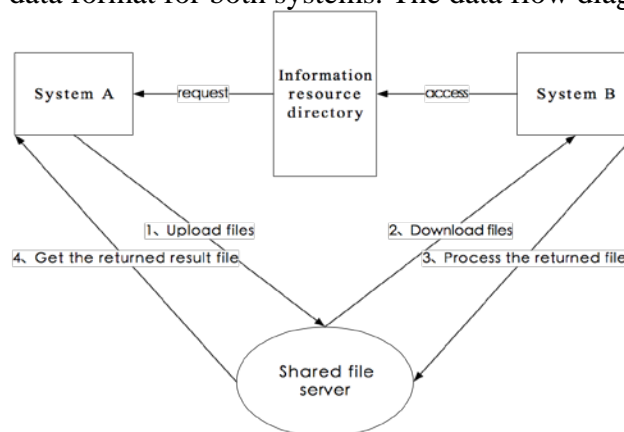


Fig 5. The ftp/file sharing server mode flow chart

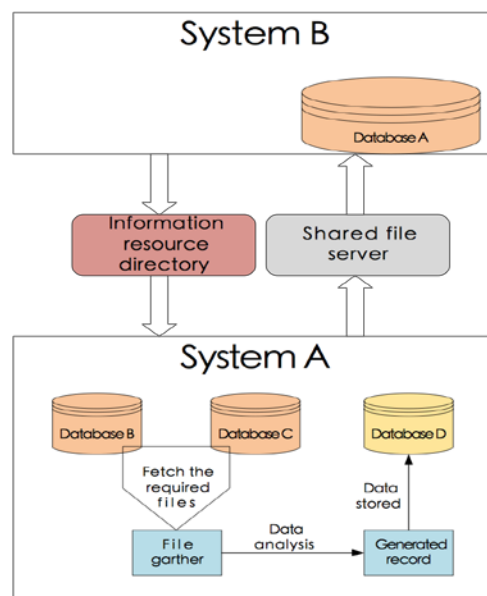


Fig 6. System integration and interaction

Database sharing server mode is: system B can obtain data set description information in system A and access the one or more data set or the result of summary and analysis through the directory.

System B can request the data set information in the system A through directory. If no further analysis and processing is required, A directly uploads the data set information to the shared file server for B to receive. If B needs the record of data application, A responds to the request and performs data analysis through its own platform, and store the results into its own data set(repositories or file), which is transmitted to the shared file server by the data set, and B responds to receive data set information. The interaction between the systems is shown in Figure 6.

4.5 Directory-based service integration assembly and interaction

Directory-based service integration assembly and interaction is implemented in the application layer. We realized the dynamic service interaction of data on the basis of implementing the directory-based application system integration and interaction. System B explicitly puts forward a data service request by accessing directory, after system A responds, A stored the results which analysis in its own system in the form of record or data elements, etc. System A provides the corresponding data service to B according to the analysis result, and B responds and receives this service.

5. Application and practice

5.1 Application preconditions and scenarios

5.1.1 Application preconditions

The technology implementation of data sharing based on the directory should follow basic prerequisites as below:

1) Completing department government information resources directory: compiling and collecting the survey data in a vertical and horizontal way, merging the same metadata items and their metadata information, generating management class metadata. Based on the business type and process, the contents of directory should be adjusted and a department government information resources directory should be formed.

2) Completing theme government information resources directory: based on application needs and oriented by applications, department should determine the theme according to application needs, sort out and integrated business and information resources, extract metadata items and collect metadata information, and form the theme directory.

3) Having the metadata base of information resources: form the metadata base of government information resources through cataloging, submit the metadata to the directory center, and then classify and publish all the metadata.

4) Clear rights and responsibilities: clarify the functions of each department and determine the ownership of the information resources.

5.1.2 Application scenarios

1) Directory-based data service management is presented in static access mode. It does not need to connect to the shared resource repository to fetch the information, the metadata information of the directory is used as the content of the demand. The data requester can understand the relevant information and information ownership department by reading the metadata information of the directory. It is conducive to revealing the relationship between government information resources and government business, government departments and government responsibilities, and clarifying issues such as the ownership, category, how to generate and who to manage government information resources.

2) Directory-based information query navigation service functions are displayed in human-computer interaction mode. Applicable to physical aggregation data, users can query the current single data set (database or data file) information, without connecting to other databases and systems. It facilitates the navigation, retrieval, location and discovery of government information

resources, and facilitates users to query and obtain government information.

3) Directory-based data exchange configuration and management is mainly applied to logically aggregated data. It is applicable to data requirements that cannot be solved by a single database. It can be applied to different databases in the same department or across department. It is conducive to the standardized collection and information sharing of government information resources, and establishes the business logic and sharing mechanism of information sharing and exchange between departments, which is beneficial to business collaboration and information sharing of government departments.

4) Directory-based application system integration and interaction is mainly applied to logically aggregated data, which essentially solves the sharing between heterogeneous data, and is beneficial to solve the problems of information resource occlusion and repeated acquisition caused by using different systems. It is implemented in the application layer and obtains data through the directory. It is mainly used for information resource exchange, analysis and acquisition between different units or different departments.

5) Directory-based service integration assembly and interaction are applied to achieve mutual service between departments, avoiding repeated collection and analysis of information, and effectively solving problems such as “information islands” between departments and units.

5.2 Application practice

This paper uses the function as the main line, puts forward the directory-based data service management, information navigation query, data exchange configuration and management, application system integration and interaction and service integration assembly and interaction, and supports a province to carry out actual data sharing business improve the status quo of directory and data "two layers", to solve the "information island problem" between departments. Specific examples are as follows:

Certain province through the construction of directory management system of big data sharing platform, units authorized users complied, published and update the government information resource of itself, and each units can query all government information resource directory of the whole city to see the data sharing attribute, sensitive attribute, ownership and responsibility. Units cannot obtain the sensitive data online, so as to realize the privacy protection of sensitive information. For shareable data, take the legal person basic information base of the big data platform as an example, through directory provide query online, comparison, statistical analysis and spatial location service of legal person basic information base for authorized users. At the same time, according to the directory sharing attribute, the shared data exchange can be obtained. For the government data of unconditional sharing class, the demand department can directly obtain the data exchange through the Shared exchange configuration. For government data with conditional sharing, if departments have sharing rights, they can obtain data exchange directly through the sharing exchange configuration. If they are not authorized, they can apply to the providing departments through the big data platform for sharing. After obtaining the consent, they can conduct data exchange services.

6. Conclusion

In the field of Internet + government affairs, this paper studies hot issues such as difficulty in sharing government information resources and coordinating business departments and Information Island. This paper also summarizes and proposes a set of technical solutions aiming at five application scenarios of information resource sharing with the function of directory as the main line, and verifies them in practice. In the future, we will continue to improve and modify this technology solutions basis of application practice, increase the research on block chain technology, and further optimize the data resource opening, sharing, and confirmation supported by government information resources directory.

References

- [1] Zhanhui Yan, Limin Li Standardization research on information resource sharing mode between government departments [J]. Chinese standardization, 2018(15):81-85.
- [2] Chunqi Huang. Analysis of policy implementation in the process of information resource sharing by local governments [D]. Jilin university,2018
- [3] Biying Yan. Research on the application of big data in government information resource sharing [D]. Shenyang normal university,2018
- [4] Si An, Xiaomi An. The research progress and thinking of foreign government information resource management [J]. Information science, 2018, 36(03):13-19+80.
- [5] Yongjin Zhang, Meilin Zhang. Integration and sharing of government information systems: process, experience and direction [J]. Chinese administration, 2018(03):22-26.
- [6] Zhaoyan Cui. Big data construction of China's government information system [J]. Electronic technology and software engineering, 2018(08):158.
- [7] Xiaoling Liu. Design and application of data aggregation and sharing platform for municipal government affairs [J]. Fujian computer, 2018, 34(05):138-139.
- [8] Fanglin Zhang, Hao Cui. Construction concept of data file center for government information sharing platform [J]. Beijing archives, 2018(05):23-26.
- [9] Qingguo Meng. Government data exchange and sharing and implementation mechanism based on three-power separation [J]. Software and integrated circuit, 2018(08):30-31.
- [10] Feng Yan. Comparison of data exchange and sharing methods of government departments [J]. Electronic technology and software engineering, 2018(15):165.
- [11] Huiping Zhang, Shuxin Hu. Research on the promotion strategy of "Internet + government service" cross-department data sharing [J/OL]. Intelligence journal: 1-7[2018-12-06].