

# Research of LANCANG River Basin Data Information Sharing System Based on Metadata and Dataset Management

Wei Wang, Lei Shi, and Xing Gao

The State Key Laboratory of Resources & Environment Information System Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing, China  
Email: {wang\_wei, shil}@lreis.ac.cn

**Abstract**—Metadata is simply called “data about data” and is the key technology of information sharing. It is also the hotspot information sharing technology. In this paper, based upon the contrast of the national and international metadata standards, it points out that sharing data is the prerequisite and there are drawbacks in traditional data sharing mode. It is necessary to set up sharing mechanism of data management on the basis of metadata and dataset. We adopt three levels of metadata management Metadata Section, Metadata Entity and Metadata Element. We establish the metadata standard template, including Identification Information, Data Quality Information, Spatial Reference Information, Content Information, Distribution Information and Metadata Information sharing system of Digital LANCANG River Basin. The core is metadata management system of the metadata information sharing system and the base is to establish the metadata standard of the metadata management. We set up different metadata warehouse of different metadata management system. Then we define the metadata views to carry out the interface between other information sharing system and users. The metadata information sharing system has characteristics: distributed data organizing and management; distributed data sharing; distributed fast data index and data accessing between multi-operating systems. Through the construction of the integrated management information system of LANCANG River Basin water among different departments of information sharing resources, it effectively uses resource information’s scattered in various sectors and improves the data and useful information.

**Index Terms**—metadata; information sharing LANCANG river basin

## I. INTRODUCTION

In recent years, the data integration and sharing system research has been a hot topic in the field of data management research. In the process of “The LANCANG River expedition data integration and sharing system” development, the need for a unified query of heterogeneous data sources, such as remote sensing data,

geographic data, statistical data of society and economy, water resource data, the data of land use, cultural heritage data and so on. Heterogeneous data source, referring to data sources is structured and semi-structured data [1], can be a traditional relational database systems, can also be a file system, and even the Web data source. Data integration and sharing system, provides a unified interface and sharing for multiple heterogeneous data sources, make integrated query easy for user.

The integrated system used the “middle layer” integrated way, The process is middle layer packaged XML including semantic checking and query after receiving the query request, and then follow the XML format to query, finally collect the results returned by the query, and returned to the user interface layer after integrated conversion[2]. Integrating the multiple data sources into a global management, unified model, the integration and sharing of user-oriented system is completely transparent to the user operation.

## II. METADATA AND RDF

Metadata described the information resources or data object data, the definition of metadata is “Data about Data” usually. The aims of using metadata are identify resources, evaluation of resources and tracking of resources in the course of change, in order to achieve effective to find and locate information resources, the integration organization and management of effective using resources[3],[4]. With the increase of the kind of data, the amount of data increases, there is an urgent need for a reasonable method to avoid the repetitive data assumptions in order to coordinate the sharing of resources between different data department. At this point, the importance of the concept of metadata was only people gradually recognize.

There are many the international metadata standards for scientific database, Typically, such as the United States Federal Data Committee (FGDC) formulated Standards for Digital Geographic Metadata (CSDGM), National Aeronautics and Space Administration (NASA) formulated Directory Interchange Format (DIF), U.S. Agency for International Earth Science Information Network (CIESIN) formulated metadata standards, Dublin Core metadata standard in the UK, ANZLIC metadata

standard of core elements of Australia and New Zealand, And the International Organization for Standardization formulated ISO15046-15 "geographic information - metadata standards" etc. The development and implementation of above standards used within the scope of a national or an international organization, and you can query on the Internet [5].

**A. Metadata Content**

Metadata standard system content is organized in accordance with the section, compound elements and data elements, which are inclusion relationship in turn, the former includes the latter, that the latter is in turn composed of the former. Specifically, it divided into eight basic parts and the four reference section, composed of twelve parts, and standardized contents including identification information, the data quality information [6], the data set inheritance information, spatial data represents information, spatial reference system information. Entity and attribute information, issue information, as well as space of metadata reference information, etc. There are also four part must be referred to the information in section of standardization, they are reference information, a time range of information, contact information and address information.

The metadata standard content system is achieved through metadata NMS, the system mainly consists permissions validation function (server side validation), input and legitimacy checksum function (client side validation) query functionality (server-side queries), return and display functions. Using the spatial metadata NMS as the space station Shareware exchange, can basically achieve the space information network share [7], [8].

The contents form of the metadata shown in Fig. 1.

Metadata is data about data, including the data information as below.

- 1) Data quality information, such as location and attribute accuracy, data integrity and consistency, sources of information and the method used to produce the data.
- 2) Spatial Data Organization Information, that is the mechanism of spatial data in the data set, such as the direct method is used to represent the spatial location (grid or to taste), the indirect method (street address or county codes) as well as the number of objects in the dataset space.
- 3) The spatial reference information, described the meaning of the reference frame in data set, encoding and coordinates. Such as the name of the map projection system parameters, the horizontal and vertical reference, the resolution of coordinates system.
- 4) Entity and attribute information, data and content information, including the entity types, attributes and its range.
- 5) Data distribution information, access to the information of the data sets, including how to contact the distributor, available formats, accessing to data sets online or on physical media, as well as data prices, etc.
- 6) Metadata reference information, metadata reference system and information of relevant units.

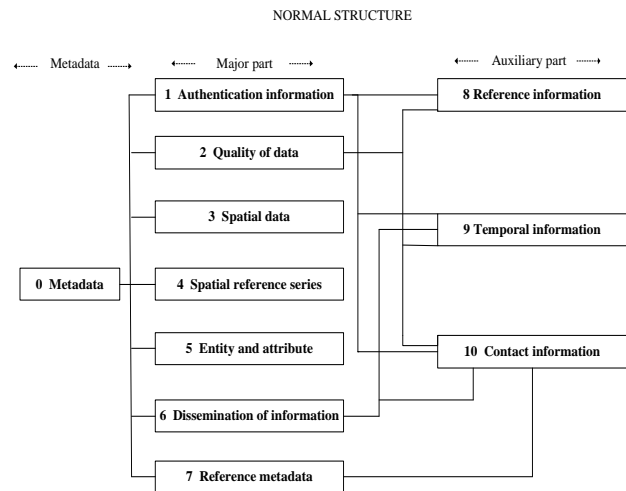


Figure 1. Content of metadata

**B. Metadata Described in RDF Format**

RDF (Resource Description Framework) is a recommended standard developed by W3C. Its purpose is to use the currently existing a variety of metadata standard to describe the various network resources, it will form a man-machine readable system and can be handled automatically by a computer file. The basic model of RDF is composed by nodes, attributes and their values. Node may be any of the network resources [9], such as web, servers or any relevant URI (Universal Resource Identifier) of the resources, or even other metadata. The attributes are formulated characteristics by nodes, and their values can be either automatically (strings, numbers, etc.), can also be other resources or metadata.

**III. THE BASIC STRUCTURE OF SYSTEM METADATA STANDARD**

After comparison of domestic and international metadata standards, there are three levels to manage [10]: subset of metadata, metadata entities, and metadata elements.

Metadata element is basic information unit of metadata, metadata entity is a collection of similar data element, and the subsets of metadata are interrelated metadata entity and the collection of elements. In the same subset, entities are divided into two categories, simple entities and composite entity. The simple entity contains only elements; composite entity contains both simple entity and elements.

In metadata standard of the LANCANG River Basin expedition data integration and sharing system, metadata content is divided into the following sections.

- 1) Identification Information included datasets overview, geographical area, contact information, the limitations of the data set, format, data representation of information, images logos, etc.
- 2) Data Quality Information included data quality and data logs.
- 3) Spatial Reference Information included classification of spatial reference system, reference

description based on the geographical identification code, reference instructions based on the geographical coordinates of the space, geo-referencing system benchmarks, and projection information of geographic coordinate, the vertical coordinate system, elevation system definition, and the depth system definition.

- 4) Content Information included a list of feature classes.
- 5) Distribution Information included method of data storage, distributor, distribution of contact information and network resources.
- 6) Metadata Reference Information included contact information and network resources.

#### IV. INFORMATION SHARING SYSTEM BASED ON METADATA

In the information sharing system, the core part is the metadata management system, and the most critical and basic content of metadata management system is the establishment of metadata standards, as shown in Fig. 2. Creating and defining through metadata standards, which may be needed to establish metabase of the respective metadata management system. It can be defined required metadata view at any time, and it will achieve interface of metadata management system with other systems or user interface.

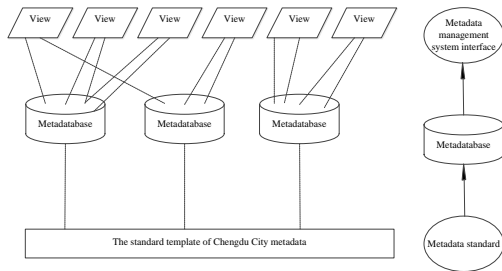


Figure 2. Structure of metadata

##### A. System Features

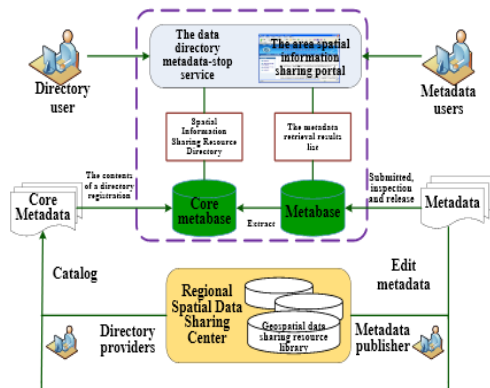


Figure 3. Connection map of information sharing system module

The sharing system characteristics can be seen from the system structure diagram (Fig. 3).

##### B. Distributed Organization and Management

The sharing system implements metadata level access control by establishing the metadata server. The data can be stored on any computer, data server management.

When issued by the user to the metadata server referrals request of the (e.g., query data, etc.), a metadata server from the database of the corresponding metadata record, according to the communication wherein the data recorded in the position information and the corresponding data between servers, and find the data sent to the user. To prevent illegal intrusion, the data stored on the data server will be encrypted so that the data is protected [11]. In order to make full use of network resources, the system needs to consider more flexible logic structure. A metadata server may correspond to multiple data servers, such a system called data sharing system.

##### C. Distributed Data Sharing

The sharing system achieves data sharing by metadata. Users define the metadata information in the log data through a unified interface, and specify the data to be stored in the data server. The metadata server received login instructions, the data is first sent to the specified at the data server; add a new record in the database. After this process is completed, a client of any of the other users on the network can be found by keywords or other search criteria just log data (if the user has access to), to help data providers and data use solve the problem of data conversion, communication and understanding. One of the important data to give full play to the role of metadata so it can be used in many ways, including data document creation, data dissemination, data browsing, data conversion[12],[13]. Metadata have an important role for the promotion of data management, using and sharing. Further, metadata is very important to establish an information exchange network. The network center located in the center metadata database can be connected in real time the sub-node of each of the distribution data element database, helping potential users find the data required by their particular application, data sharing.

##### D. Cross-Platform Data Accessing

In a distributed environment, the problems of cross-platform data access must be solved because of differences of the computer hardware configuration and operating systems. In this regard, JAVA language is the only language of the system development [14]-[16]. JAVA virtual machine (JVM-JAVA Virtual Machine) can be run on different platforms, and JAVA language can also CORBA (Common Object Request Broker Architecture Common Object Request Broker Architecture) to support the distributed network environment, in order to achieve cross-platform data access.

#### ACKNOWLEDGMENT

The authors would like to express appreciations to colleagues in our laboratory for their valuable comments and other helps.

#### REFERENCES

[1] FGDC, Federal Geographic Data Committee Content Standards for Digital Geo- spatial Metadata, 1994.

- [2] D. Fisher, "Terrill tyler distributed metadata management in the high performance storage system," in *Proc. First IEEE Metadata Conference [M] 1 NOAA Auditorium*, Silver Spring, Maryland, April 1996.
- [3] L. Seligman, "Amon rosenthal a metadata resource to promote data integration," in *Proc. First IEEE Metadata Conference [M]. NOAA Auditorium*, Silver Spring Maryland, April 1996.
- [4] P. Drew and J. Ying, "Metadata architecture for multi- system interoperation," in *Proc. First IEEE Metadata Conference [M]. NOAA Auditorium*, Silver Spring Maryland, April, 11996.
- [5] Cinderby: Serve GIS for participation: The future of enviromental GIS?" *International Journal of Environment and Pollution*, no. 3, pp. 304-315, 1999.
- [6] M. Haklay, E. Feitlson, and Y. Doytsher, "The potential of a GIS-based scoping system: An israeli proposal and case study," *Environmental Impace Assessment Review*, vol. 18, no. 5, pp. 439-459, 1998.
- [7] P. A. Hamilton and T. L. Miller, "Differences in social and public risk perception and conflicting impacts. On Point/non-point trading ratios [J]," *Am J Agric Econ*, vol. 83, no. 4, pp. 934-941, 2001.
- [8] M. Iqbal, "Computer-based approach to coastcal zone management in Pakistan with special reference to envirmental impace assement," in *Proc. Pac. Congress Zool Zoological Society of Pakistan*, Kaeachi, Pakistan, 1994, pp. 177-182.
- [9] F. Lasserre, M. Razack, and O. Banton, "GIS-linked model for the assessment of nitrate contamination in groundwater," *Journal of Hydrology*, no. 3, pp. 81-90, 1999.
- [10] Lahmer, Werner, Becker, Alfred *et al.*, "GIS-based approach for regional hydrological modeling IAHS Pulication," *International Association of Hydrological Science*, no. 254, pp. 33-43, 1999.
- [11] A. T. Lombard, P. V. August. *et al.*, "A proposed Geographic Information System for assessing the option dispersion of protected areas in South Africa," *S. AFR. J. SCL/S. -AFR. TYDSKR. WET*, no. 3, pp. 136-140, 1992.
- [12] H. J. E. Rodda, U. Shankar, and S. Demuth, "Application of geographical information system to water quality modeling in New Zealand," *International Association of Hydrological Sciences*, IAHS Publication, 1999, no. 254, pp. 243-251
- [13] A. Thrk, "Visualization in environmental management, Beyond the buzz word, Landscape and urban planning," *Landscape Urban Plan*, no. 4, pp. 253-255, 1992.
- [14] P. Q. Zheng and B. W. Baetz, "GIS-based analysis of development options from a hydrology perspective," *Journal of urban Planning and Development*, vol. 125, no. 4, pp. 164-180, 1999.
- [15] S. Antonie, A. Kazimierz, and Aalewicz, "An interactive reservoir management system for Lake Kariba," *European Journal of Operational Reserarch*, 1998.
- [16] N. Buras, "Scientific all location of water resources: Water resources development and utilization-a rational approach [M]," New York: American Elsevier Publishing Company, Inc, Jan 1972.



**Dr. Wei Wang** was born in 1972 at Zhuozhou County, Hebei Province, China. His education background is as follows.

In 1997 Graduated with B. Eng. of Applied Geophysics in East China Institute of Technology, Fuzhou, Jiangxi province.

In 2008 Graduated with D. Sc. of Geophysics in Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing.

In 2005 Graduated with M. Eng. of Geophysics in China University of Geosciences

Working experiences and publications are as follows.

1997-2002, Geophysical engineer, BGP, CNPC.

2008-2010, Assistant Researcher, ITPCAS, Beijing.

2010, Assistant Researcher, The State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, 11A, Datun Road, Chaoyang District, Beijing, 100101, China

1. Wang W, Yin J J , Liu X W, et al. Equivalent offset method and its application, *Chinese Journal of Geophysics* ,2007,50(6):1580-1587.
2. Wang W, Gao X, Li Y Y, et al. . Crustal velocity structure of S wave beneath Tibetan Plateau with transform function method-Hi CLIMB profile. *Chinese Journal of Geophysics* . 2011, 54(6):766-776 .
3. Wei Wang, Yun Wang, Junjie Yin et al. Error Analysis of the Converted Wave Deduced by Equivalent Velocity Assumption. *Exploration Geophysics* , 2012,43,162-170.

His current research interests are Geological Disaster Investigation and Data Mining.

Dr. Wang was awarded Open Fund (GDL0803) of Key Laboratory of Geo-detection by Ministry of Education, and National Special Fund(41240027)by National Natural Science Foundation of China.



**Dr. Lei Shi** was born in 1981 at Taian County, Shandong Province, China.

2006-2009, China University of Mining & Technology, Beijing. majored in computer application technology

2002-2006, Shandong Institute of Business And Technology, B.A., majored in Computer Science and Technology.

GIS, ecological modeling, remote sensing

1. Lei SHI, Chuanjie XIE, Gaohuan LIU, Qingsheng LIU, Design and Implementation of Taihu Lake Basin Management Information System based on GIS [C]. The 3rd International Conference on Environmental and Computer Science, October 2010, Kunming China
2. Shi Lei, Huang Chong, Liu Gaohuan, Xie Chuanjie, Ran Fankun, Design and Implementation of LANCANG River Basin Management Information System[c]. The 2nd International Conference on Circuits, System and Simulation 2012
3. Chuanjie Xie, Jiaqi Li, Gaohuan Liu, Qing Liu, Lei Shi, Model Component and Interfaces Design In Environmental Modeling System [C]. Proceedings of 2010 3rd IEEE International Conference on Computer Science and Information Technology, July 2010, Chengdu China



**Xing Gao, Ph.D.**, associate professor. Working in Institute of Geographic Sciences and Natural Resources Research, CAS. He is mainly engaged in the work of seismic wave propagation. The main content is, according to the digital seismic network data, using the waveform simulation, inversion etc., geophysical problems related to the study of focus, the earth's internal structure

and medium.