METHODS AND IMPLEMENTATION OF THE GEOSPATIAL DATABASES INTEGRATION AND UPDATE TOWARDS E-GOVERNMENT

Jun-san Zhao[a,b], Xue Li[b] , Yaolong Zhao[b], Tao Xu[c], Xiaodong Fu[b]

[a.]Wuhan University, Wuhan 430079,China
[b.]Kunming University of Science and Technology, Kunming, 650093, China;
[c.]Kunming YunJindi GeoInformation Co.Ltd., Kunming , 650106, China

KEY WORDS: Geospatial Data Integration; Data Update; E-Government; Web Service; XML/GML

ABSTRACT:

All the implementations of the geo-spatial information systems depend on the multi-levels, synthetic and integrated databases of spatial and attribute information. These databases have many unique characters that include multiple classifications, multi-scales, multi-sources, different GIS formats, etc. At present most of them is the separate 'Information Island', which cannot meet the needs of the development of the information systems. So the databases must be integrated and updated. This paper will first discuss the principles and methods of the integration and update of the geospatial databases. Then the methods of the information share and exchange of geospatial data are outlined. To elaborate the implementation and the processes of the geospatial data integration and update, the technique routines and solutions are described in the paper.

1. INTRODUCTION

Along with the fast development and application of Geo-Information science and technology, the geospatial information exchange, share and co-processes among the different departments, district, industries and different information systems become very important and popular. Further more, the governments, private agencies and public society have strong desires to get an standard platform to realize the exchange, share, integration and update for the different kinds of geo-spatial databases, so that the geospatial data can be used efficiently and some new applications can be developed on it. Geospatial databases integration and update is also a key problem in the development of electric government (e-government) and enterprises information system construction.

The main objectives of the geospatial data resources integration include: (1) First, to meet the needs of the development of e-government and enterprises information systems by the geospatial data integration and the harmony of the relationship among departments and agencies. (2) Second, to formulate the policies of data share, exchange, and data update and data services. (3). Third, to establish the data standards and the technologic systems of geospatial databases management. (4). Forth, to provide the geospatial data to the government, public and private in multi-scales, multi-times, multi-resources, multi-resolutions and dynamic styles. (5). Fifth, to improve the environment of the study and application of geospatial science and technology. (6). Finally, to promote the development of geospatial data infrastructure construction and the geospatial industry.

2. MAIN CONTENTS OF THE GEOSPATIAL DATA INTEGRATION E-GOVERNMENT

The geospatial data integration towards e-government is a very complicated system engineering related to many technical and non-technical key problems that need to be solved. Generally, it contains the following contents:

- First, investigating, analyzing and standardizing the natural resources data and geospatial data that are stored in different departments of local governments, further more, sorting, standardizing, abstracting and logically concentrating the data.

- Second, relating, structurally optimizing, extending and joining the natural resources and geospatial data for information sharing and integrative applications through the technologies of GIS, Internet/Intranet, WEB, databases and information security. Then building new huge amount of databases or network databases for the e-government.

- Third, constructing a uniform management mechanism for natural resources and geospatial database, which makes it convenient to add, cancel, modify, extend and upgrade the data to databases.

- Fourth, according to the demands and actual requirements of the governments and their underling department, developing an e-government platform so that natural resources and geospatial data can be used in governments' daily work, the examination and approving of the engineering projects, the executing of policy rules, and supporting the decision-making of the governments.

- Fifth, by using the technologies of spatial analysis, data mining, etc. to build forecasting models and decision-making supporting systems based on the databases of natural resources and geospatial information.

- Finally, under the permission of the government policy and law to release the natural resources and geospatial information to the public through networks and Web GIS etc. to provide synthetic information consultation services to the personnel, enterprise and society.

3. GENERAL THOUGHTS OF THE GEOSPATIAL DATA INTEGRATION

The natural resources and geospatial data exist in different formats and types such as vector, grid, multimedia, text and so on. While integrating the multi-resources spatial information, the database engines are used. The huge amount of spatial and attribute are stored and managed by relational database systems. By expending the disks' array and other storages the functions of the relational databases can be exploited to realize the storage and query of the enormous amount of spatial data. The data don't need to be separated and can be stored in the databases as a whole. For the data exchange in the different structure, the data exchange platforms should be developed based on the sufficient understanding to the formats of the data to be accessed. To develop the platform, the definition and standards of the geographic data, such as OGC (OpenGIS Consortium), GML (Geography Markup Language) should be constituted. The XML is used to define the data coding, and the geographic data and attribute data can be freely exchanged among different systems by the data exchange platform developed.

4. GENERAL TECHNICAL PROJECTS FOR DATA INTEGRATION

Geospatial information integration and sharing can be realized by several technical methods.

4.1 Integrating Data with Database Technology

The spatial data that are scattered in different departments in the governments can be integrated, applied and updated by use of the new spatial database management systems, components or middle ware technologies and communication technologies. For example, we can use Oracle Transparent Gateway, spatial ware and IAS (Internet application Service) to realize the integration of different databases.

4.2 Methods of the Spatial Data Integration Through Web Service Technologies

At present, we can use technologies of Web Service, which is supported by the protocols of XML, GML or SVG, to realize data sharing and exchange among different organizations and departments in governments. Its advantage is that the geospatial data from different formats, GIS platforms and types can be easily accessed under the protocols of XML, GML and SVG. And the geospatial data service will be provided by use of the technologies of Web Service. Its disadvantage is that it's difficult to sustain data exchange timely, and cannot meet the needs of the implementation of complicated e-government systems. According to the results we researched, the middle wares for the query of geospatial data, the operation logical packet wares and data safe control middle wares can be developed under the support of XML and GML. By the combination of these middle wares and databases of administration rules and policies, the e-government running engine should be implementation for the cases examining and approving, monitoring, supervising, analyzing and evaluating and geospatial data exchanging.

4.3 Methods for Data Conversion

The technology of ODBC data interfaces can be used to realize the communications between different databases, and the data can be transferred. The data formats transition and storage from different GIS platform can be completed in the same technologic route. Figure 1 is an example that data of different GIS platforms and different formats are integrated into one database management system (Oracle) through spatial data engine (SDE).

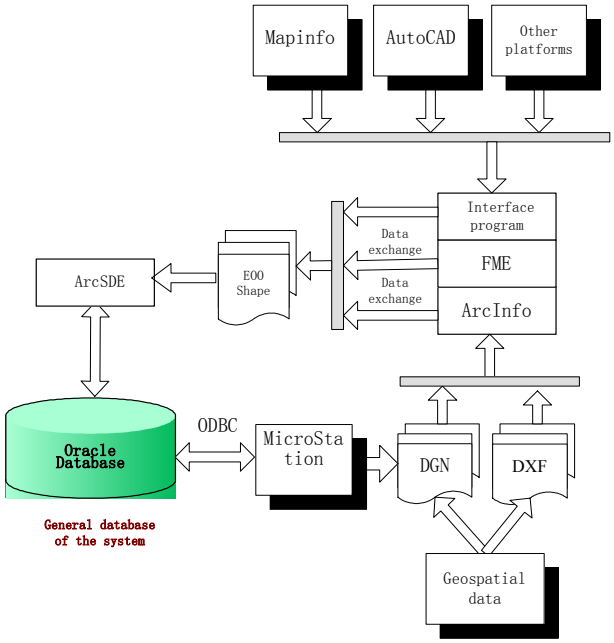

Figure 1. An example of data conversion and integration of geospatial data integration and storage

4.4 Data Organization

In some cases, geospatial data and other correlated thematic data are managed and applied by each separate information system that forms many "information island". To realize the data share and synthetically application, it's necessary that these different data should be uniformly planed, integrated and managed.

While integrating data, we should acquire data from these unattached systems by data copying, abstracting and extracting. The data acquirement and collection should not affect the existing systems' running, not add superabundance load to the existing systems and should insure them running safely. We can achieve data's convergence, integration and management with technologies such as data transmission, database accessing, database management and Internet. In this process, commercial platform software for database management, taking Oracle as an example, are used for the management of geospatial data, attribute data and integrated affair data. The communication of isomerous database systems carries out by use of some commercial software components, such as Oracle Transparent Gateway, or some developed middle wares to store the data into the integrated whole databases and to achieve data share. Oracle Transparent Gateway provides the flexible way to access non-Oracle data, such as SQL Sever, Informix. Figure 2 is an example to use this technology to integrate data such as SQL Sever, Informix etc.

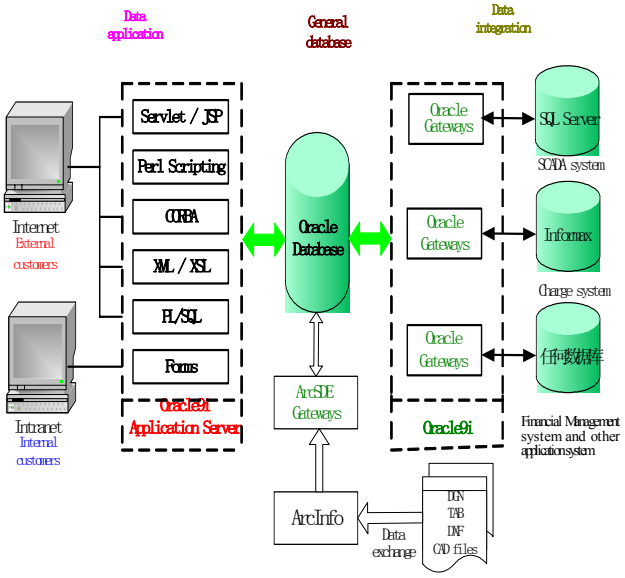

Figure 2. Data organization of data integration system

In the data organization of data integration the metadata and metadata databases must be generated and applied. The major functions of the metadata are to explain the data from different databases, make navigation for databases, transfer data files and manage data. Spatial databases for the e-government are large-scale, distributed, complicated and isomerous. Thus, it's very pivotal to design and apply the standard metadata as well as the meta-databases in e-government.

5. THE SOLUTION OF THE GEOSPATIAL DATA SHARING

5.1 Implementation of the Information Share System

E-government systems based on the databases of natural resources and geospatial data are the synthetic management information systems that use many technologies, such as database, Middle ware, MIS, GIS, computer network and some other information technologies. While designing the systems, it should be sufficiently considered to combine the existing systems to e-government systems to be developed. The integration and conformity of the software systems should be built on the basis of the structures of 3-level or multi-level software framework.

The data layers of the systems include central managing integrated databases, isomerous databases of existing systems and geospatial information databases. Some commercial products of DBMS could be used for the integration of the management of huge amount of geospatial information, e-government databases and other existing system data. The data layer supports the middle logic layer to accomplish the services including data index, query, data transmission and decision-making based on the spatial and attribute data analysis and statistics.

(1) To access the geospatial and attribute data, the data must be formatted, converted and stored in the databases. The spatial data engine is used for the data access by the middle layers. The useful data are selected in accordance with data conversion, data analysis and process. These selected data are integrated, stored and managed directly to Oracle database through ODBC supported by GIS platform, or through software interface, spatial database engine (SDE), and some data conversion tools, such as FME, see figure 3.

(2) The existing data are analyzed and processed according to their different property. Then, they are stored into the integrated databases to achieve uniformed and concentrated data management. The data share and access are carried out by use of the central spatial data middle wares, such as ArcSDE and IAS.

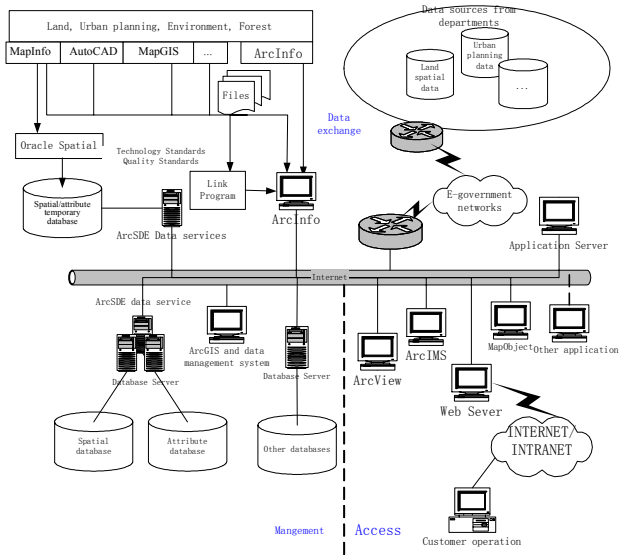

Figure 3. The system structure of the data management and access for the data integration

5.2 Data Management and Access

Figure 3 shows the structures of the data management and access after data integration. The spatial and attribute data are accessed through the upper temporal transition databases or files as showed in Fig. 3. The data conversion, checking-up, input and distribution are realized via transforming programs or the GIS platforms, such as ArcGIS combining with the standard fundamental geospatial information subsystems. By using ArcSDE service, we can acquire geospatial data from the databases that will be provided to management and access program.

Data access can be carried out through the second software development of GIS components or GIS platforms.

(1)Spatial data engine: The spatial data engine (SDE) is built for the efficient spatial data access and data sharing that is irrespective to the GIS platforms. The functions includes:

- **Spatial data storage**

To adopt industrialized standard SDE for management of the magnanimity spatial data, and the data are organized and managed by use of Oracle Spatial, Informix Spatial DataBlade, and SDE etc.

- **Middle wares for spatial data share**

To achieve efficient spatial data access for profession use, comprehensive application models and management

mechanisms are provided. The realization of spatial data share is irrespective with GIS platforms.

● **Web service layers of spatial data share**

The Web service of spatial data can be accomplished by Web service technology and spatial data share middle wares. The spatial data can be accessed by different thematic systems through Web.

(2) E-government Running Engines

Depending on the development of e-government, the geospatial data access middle wares and embedded profession logical wares implemented. Combining with the databases of administrative law, rule and policy the e-government running engine is created, which supports the operation examining and approving, supervising, analysis and evaluation in e-government, as well as geospatial data exchange, reporting, query and analysis.

(3) Decision-making analysis and support engines

Engines for decision-making analysis are used to build the platforms for the examining and approving of spatial data related projects, as well as important decision-making analysis served for the governments. These kinds of engines are consisted of knowledge management engine, policy and law matching engine, statistics and analysis engine, target management engine, integration GIS and decision-making system.

The contents of the data cooperation management applications include: (1) To update geospatial data regularly by related departments with "remote automatic updating tools" and data update mechanism; (2) To acquire correlated data from "natural resources and geospatial information databases" according to the requirements, and providing geography information service to relational departments; (3) To update and exchange data daily with system applying departments, and realize real-time updating of fundamental geospatial data simultaneously.

6. THE METHODS OF THE GEOSPATIAL DATA UPDATING FOR E-GOVERNMENT

The e-government-oriented geospatial information databases can be divided to basal geography databases, natural resources databases, e-government operation databases, thematic databases and integrated documents databases etc. According to the relationships of the above-mentioned data and the characters of e-government, the logical arrangements of databases are separated to 5 levels, i.e. total databases, department databases, subdivision-databases — logical layers and physical layers. According to the different use status, the geospatial data can be defined as three classed, they are work or temporal databases, current situation databases and historic databases. The functions of each database can be described as below:

(1) Work databases. To avoid the current situation databases to be destroyed while the users collect and check the data, and operate the databases for the daily work. Work databases are temporal and could be used to permit the databases to be unitary and exclusive.

(2) Current situation Databases. Databases that storage latest data after exactly examination, and should be strictly controlled in the data adding, updating and deleting of databases.

(3) Historic databases: Databases storing GIS spatial data, relational historic attribute data and data update records. These databases assure the data could be traced by historic time. And the historic databases can only be added new data but not modified. They are also the bases of the data query and statistics for the users.

E-government databases support systems application and management that store office automation (OA) data, synthetically file data and system maintenance data.

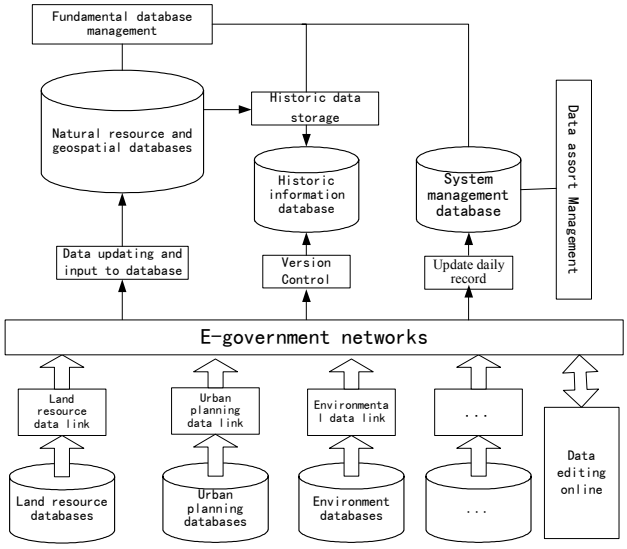

Figure 4. The flow chart of the geospatial data updating for e-government

In terms of the results we analyzed that each departments of government should accomplish the thematic data cleaning, inputting and updating under the requirements of the data center. So the databases departments owned not only meet the their needs and e-government development, but also are convenient to transfer and update the data for the data center. The processes of the implementations of the geospatial data updating for e-government are expressed in Fig.4.

6.1 Periodic Data Updating

To update the large amount of data we can use the data updating software interfaces. For the updating of the small change and less amount thematic data, we can update them online through the networks. While we update the data, the daily work records, maintenance records and the data version should be noted.

6.2 Process of Data Updating

Because of the complexity of the spatial data, all the different datasets can't be linked seamlessly by use only one common connection interface system. Thus, we must implement different programs to realize the data integration and updating. The integrated data are stored in the data center after data conversion and other processes through e-government networks.

(1) Data conversion and formatted

The data that need to be updated to central basal database include geospatial data and relational attribute data, metadata

for describing spatial data, non-spatial data such as operation transaction or statistic results. The different updating technologies and methods should be adopted in the processes of data updating.

Metadata are registered and transformed to XML format according to metadata standards. Geospatial data and attribute data in different format, coordinate system and projection system are transferred to uniform format, such as EOO, SHP, VCT etc. according to data exchange standard via conversion programs.

(2) Data report

The Data are reported to data center from local departments through the protocols of FTP can be uploaded to central server memory region via FTP etc. FTP uploading modules can be sent out with data conversion and interface programs. And relational information of data title, updating date and updating person should be noted while reporting data.

(3) Data updating

The updating information is acquired by checking the update data. If data are not reported in time, news should be sent to data managing department to remind them for data reports. If the data have been submitted, the styles of the data update could be choose such as in an automatic input model or a handwork input model according to their extent and the complexity.

6.3 Online Compiling and Updating

For the situation of less spatial data compilation, online compiling and updating are used. Authorized with certain data volume or map layers, systems provide compiling function of the spatial data and attribute data, and submit data after compiling in one time by affair controlling, meanwhile, register the modified metadata information and import relational affair data.

Work, such as noting relational information about data variation, saving metadata, changing data editions, must be done while submitting data, via whether data interface or online compiling.

7. THE TECHNICAL IDEAS OF THE GEOSPATIAL DATA INTEGRATION AND UPDATE

According to the characteristic and requirement of geospatial databases integration and updating for the e- government, the key technical ideas are: the data share and exchange are realized gradually from static to dynamic styles. In the initial stage of the implementation of data center or the data exchange application, the data integration and updating could be achieved by static ways. The method is suitable for large amount of data transmission. After the data resource center system is set up, and the amount of the data to be integrated and exchanged becomes relatively smaller, the data can be transferred and exchanged dynamically through the networks in real time. The controls of the data standards, data formats, ranges and coordinate systems, etc. can be completed in accordance with the characteristics of the data source by the data checks function of the exchange system. Data safety system is established by use of the technologies of the CA, data access pass, authority management and data monitor, etc. The data copy, exchange and backup between different departments are accomplished through the commercial middle wares of data integration or the geospatial data engines developed particularly for the e-government. The datasets exchanges are realized by use of the XML protocols. The data update and real time issue are realized through the news mechanisms.

8. CONCLUSIONS

The paper mainly studies the methods, theories and implementations of the geospatial data integration and update. The main goals are to provide some ideas and solutions for the applications of the geospatial and resource data to the e-government through the technologies of data integration and update. As the complexity of the topics, we just give some initial and general solutions. The integrations and update of the geospatial data sources are the key technological problems in the development of e-government, but they can only be solved gradually accompanying with the progress of the Geo-informatics science and technologies, as well as the widely applications of the geospatial data in the e-government.

REFERENCES

G. Rugh, 1997. Land information management in Queensland. Proceedings of IKUSASA CONSAS'97, Durban, South Africa, 24-28 August, 12 pages.

J.D. McLaughlin, 1991. The Evolution of the Multipurpose Cadastre Concept: A North American Perspective. In Cadastre in Perspective. J.Besemer (Ed.). Apeldoorn: DKOR.

Zhao Junsan and Zhao Yaolong, The Development of Land and Resources Management Information System for the Local Governments in China，GEOINFORMATICS'2002, C32 PP.1-7.

Zhao Junsan ,Zhao Yaolong,,Research of the Land Management Information System Based on Web GIS and Spatial Databases for Provincial and Local Governments in China. Proceedings of DMGIS '200.

Zhao Junsan and W. Barry. 1998. Development of GIS/LIS for Cadastral Survey and Land Registration in China: A Status Report. 10th International Geomatics Conference. June, 1998, Ottawa, Canada.

About the First Author

Zhao Junsan, Mr. Zhao Junsan is a Professor in the Department of Land Information and Surveting, Kunming University of Science and Technology. His research interests are on GIS/LIS, Integrating of GPS, GIS and RS, land and resource information, and cadastral management. Professor Zhao has published more than 40 papers.