http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (18): 4673-4679, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.4673.4679

Collaborative Research of Seismic Multi-source Data Based on Ontology

[1]Zhuowei Hu, [2]Changqing Liu and [1]Lai Wei
[1]College of Resources Environment and Tourism, Capital Normal University, 100048, China
[2]Beijing GEOWAY Software Co., Ltd., 100043, China

Abstract: Our country is a country prone to earthquakes and earthquake caused significant damage to our economic property and people's lives and safety. development of database technology and geographic information technology provide important technical support for earthquake relief in data management and analysis, but with the ever-expanding volume of data, the important issue currently facing the earthquake relief work is how to dig out useful data from huge amounts of data in earthquake relief, to achieve collaboration between basic geographic data, remote sensing image data, socio-economic data, seismic case data, seismic risk zoning and other data. Ontology has a lot of advantages in terms of solving heterogeneous data. This article will introduce ontology to the field of earthquake rescue data collaboration, through ontology to represent semantic of the knowledge and concepts of earthquake multi-source data applications, so as to solve problems of heterogeneous metadata and heterogeneous properties, while according to the relationship between the data and the earthquake relief needs to establish the rule base, through Jena inference engine to achieve intelligent reasoning data,from the massive earthquake data to mine applicable data for the specific problems of earthquake relief, to achieve collaboration of multiple sources data in earthquake relief.

Key words: Ontology, earthquake, multi-source, collaboration

INTRODUCTION

The earthquake caused major damage to the world economy of property and personal safety. In the process of resisting earthquake catastrophe, people continue to discover, acquire and accumulate knowledge and experience of various disaster reliefs, therefore expanding the amount of data led to the emergence of data access and use issues (Liu, 2011). Traditional based on the keyword data search results can't meet people's needs and the query results sometimes can't be used directly, while the need to manually determine the outcome of the information, screening, find out the data retrieved in line with their intentions and then based on the filtered data for various processing and knowledge acquisition, applications for disaster relief. How to retrieve appropriate relief data and access to knowledge we want from a massive earthquake multi-source data, to achieve synergy between basic geographic data, remote sensing data, socio-economic data, seismic case data, seismic risk zoning and other data in earthquake relief is a problem needed to solve. Ontology has many advantages in solving data semantic differences and heterogeneous data (Bi and Jing, 2004). Thus, ontology technology can solve the problem of data collaboration in seismic rescue. In this study, on the basis of the seismic catastrophe rescue data types, data content, data characteristics, data applications, data processing method, information extraction methods disaster, disaster relief scenario, data processing and other research cases, to design and construct ontology that in the field of earthquake relief multi-source data coordination, to design and construct repository of data Reasoning and application, to achieve intelligent query reasoning in seismic multi-source data through ontology technology (He *et al*., 2003). Overall technical route is as follows.

DESIGN AND CONSTRUCTION OF SEISMIC MULTI-SOURCE DATABASE

Consider the data in the earthquake disaster has the features of many kinds of data, large amount of data, high requirements of data access performance, higher system stability and security of the database, this study selected Oracle 10G database management system and a spatial database engine ArcSDE. During data storage, for non-spatial data directly stored in an Oracle database table, the system uses ADO.NET way for a visit. For spatial data, such as remote sensing data, basic geographic data, achieved access and management through ArcSDE .

Logical schema design of database: Earthquake rescue data types include basic geographic data, remote sensing data, socio-economic data, seismic case data and seismic

Corresponding Author: Zhuowei Hu, College of Resources Environment and Tourism, Capital Normal University, 100048, China

risk zoning data, documentation, data, user data, meta data and so on. Based on the analysis of data content and the situation of data collecting and managing, create a different database for data organization and management, respectively, the entire data is logically divided into three libraries: The main database, the system databases and meta-databases, which are used to store entity earthquake relief data, user data and metadata. Logical database design as shown.

Physical schema design of database: According to the format of multi-source data in Earthquake disaster, all the data is divided into three categories: Vector data, raster data and attribute data, based on different data formats, using different data storage.

To achieve efficient access and management of spatial data, create three table spaces, named: T BS_SPATIAL, TBS_METADATA and TBS_TMP, respectively, spatial data storage, attribute data and temporary data to improve data access efficiency.

Meanwhile the spatial data by the way of data sets and data directories to organize spatial data, create two Raster Catalog (raster catalog) and a Feature Dataset (feature set), named Images, BasicGeoInfoRas and FeatureDatas_All, are used to store remote sensing data, raster formats basic geographic data and vector data, through data collection or data directory to access and manage these spatial data.

Database construction: Before building a database of data, the need for data collection and processing including basic geographic data, remote sensing image data, socio-economic data, seismic case data, seismic risk zoning data and so on.

Creating data are created by Oracle Database Configuration Assistant, a database named: Disaster, data tables, users and a table space are created by the client tools PL/SQL Developer.

Total created three users named Disaster Administrator, Disaster User and Disaster SDB User. Disaster Administrator have database administrator privileges, can manage the entire database; disaster User users can manage attribute data, metadata; DisasterSDBUser possible for spatial data management.

ONTOLOGY CONSTRUCTION OF APPLICATION DOMAIN IN SEISMIC MULTI-SOURCE DATA

Ontology design of earthquake domain: This case involved areas of expertise of Ontology need to be first determined, because There is no existing ontology so there is no need to consider reuse Ontology, then traverse and enumeration the important terms design for earthquake rescue data, clear hierarchy and subordination of data partitioning, define the attributes of various types of data and finally create instances of all classes.

Data applications in earthquake relief, we first need to define the following questions: (1) Earthquake epicenter, magnitude and other basic information, (2) What losses earthquake could cause, (3) What are the seismic data from multiple sources and (4) What are the characteristics of seismic data from multiple sources? On the basis of clearing the above questions, we have to define the scope of the field and ultimately determine the contents of Ontology in this case, including 5 areas: Seismic basic attribute information, the type of earthquake damage, earthquake rescue data types, spatial and attribute features of earthquake rescue data, applications of seismic multi-source data .

After determining the scope of applications of the earthquake relief data, the author access to a large number of relevant documents and information and under the expert guidance, enumeration the important terms covered five aspects involved in application areas of the seismic rescue data then builds a glossary in this area.

Ontology modeling in earthquake domain: Because of the large and complex nature of Ontology works, during building and editing ontology, modeling tool and modeling language of Ontology model is a key part to efficiently create. This article chooses open source ontology model software Protégéas as geographical names ontology modeling tool.

After starting the software, the new OWL/RDF Files, select the OWL-DL language, create a blank file, the system includes a super parent owl in the OWL file : Thing is the super parent of other kind. Choose Create Subclass On the basis of the super parent and then define a subclass name nameplace_indival and reference_frame, successively create other subclasses, until the last sub-class is created. Select the Property tab; you can create object properties and data properties in the property schema box. Select the Object properties tab (Object property), enter the property name in the name box and define the domain and range of properties. Select the Properties tab of data (Data property), enter the property name in the name box and define the type of attribute. Creating individual of geographical names entity. Select Individual tab, select the type of geographical names entity in the class structure box in the left side of Indicial, then created individuals.

CONSTRUCTION OF REASONING REPOSITORY OF SEISMIC MULTI-SOURCE DATA

On the basis of ontology of applications of seismic multi-source data , according to the relationship between the data features and data applications to build a data reasoning repository for implementing collaboration of earthquake rescue data.

Construction of reasoning repository of Earthquake rescue data: Comprehensive research in the aspects of the types of multi-source data, the use of multi-source data, the demand of data in disaster relief in the earthquake disaster is the premise of repository construction, the results of this survey as a basis for repository construction.

Currently the multi-source data used in earthquake is also widely applied, such as extraction of Pour damaged houses, extraction of damaged roads, lake extraction. Different data according to their different spatial features and property features have different practicability; and data in different roles in earthquake relief, its spatial feature and attributes features have different needs.

Reasoning mechanism of earthquake multi-source data: Ontology is a formalized expression between concept and the relationship of concept, so the ontology model contains semantic relationships between concepts and concepts. An important aspect of geographic ontology applications is spatial reasoning. Since ontology is based on the logic, which makes the implicit concepts and the implicit relationships from data source can be found. Departure from a clear concept being queried, in accordance with the structure feature of ontology, successively find descendants concept or ancestors concept as the answer of their potential and support queries. With Ontology reasoning capabilities, the relationship between the elements and attributes of local data sources can be obtained based on the relationship between concepts and concepts in the ontology mapping. The fundamental role of ontology reasoning is reflected in the data integration in the heterogeneous, distributed environments, it can improve the recall rate and precision rate of query results. Ontology Reasoning needs the reasoning tools-reasoning machine.

Currently ontology reasoning machine has Jess, Racer and Jena and so on. This article is based on Jena application development kit to achieve Geographical Names Search Service.

Jena is an open source project at HP Labs research and development, using Java application development framework aim to building Semantic Web applications. It provides a development environment for programming is RDF, RDFS and OWL - Jena API. For processing RDF API, the framework of the system contains a RDF file format Jena ontology model and contains a parser machine to parser RDF, RDFS, OWL file, persistent storage solutions for RDF model and Ontology processing and operation subsystem, information search SPARQL query language and reasoning subsystem.

Jena can be three aspects for ontology query, followed by the class query, attribute query, individual queries. As the core tools about Jena semantic query processing, it has ability to achieve ontology-based data query services. Except for supporting RDF, RDFS and OWL file format, reasoning also an important subsystem in Jena and its aim to introducing the reasoning and inference mechanism into Jena.

Jena inference engine is accessed through the use of Model Factory class and create a new ontology model (Model) for the related reasoning machine. When users query this new Model, not only will return the original description of the data, but also will return additional descriptive information after performing reasoning machine rules. The core of the reasoning mechanism is InfGraph. Ont/Model API for users to build ontology provide a convenient way to connect with the reasoning machine. Graph-base assertions and Reasoner bound data together and the Graph-ontology definitions refer Reasoner bound together and limit the data structure. Reasoner Registry contains all Reasoner currently used, it is a static class (Li *et al*., 2009).

In fact, Jena itself is not "reasoning machine design expert", reasoning machine contains its own mechanism is CLISP production rules with ontology forward inference system, Just realized the most simple subset of OWL OWL Lite reasoning, operating efficiency is not high. But Jena DIG interfaces can be connected via a professional reasoning machine, such as the Racer reasoning machine, which allows front articulated to the background in the inference engine.

SEISMIC DATA BASED ON MULTI-SOURCE COLLABORATIVE ONTOLOGY PROTOTYPE SYSTEM

Goals of system design: Based on ontology in an earthquake-source data collaboration prototype system is a data management system. System has three layers: the bottom data layer is an earthquake source database; intermediate layer for the data adaptation layer, to achieve data through ontology reasoning, retrieval and collaboration; top level layer of performance, is a data query retrieval system.

This prototype system is mainly to achieve synergy the underlying database contained data through ontology techniques. Costumer can use data query module prototype system input query keywords or a query, such as "Housing down and damage extraction", "road damage extraction", etc., the system automatically based on ontology and knowledge base and knowledge of ontology, the data is automatically reasoning, obtain to meet this condition data and query results back to the user. Coordinate system data between system, format, resolution, etc., for the difference, providing a coordinate system conversion, format conversion, data resampling and other tools to ensure that users get their wishes in the query and consistent with the data structure, thus achieve synergy earthquake-source data.

Principles of system design:

- **Security:** Because the system involves a lot of data, which has a different level of confidentiality, so the system design must take into account data security issues, on the one hand should strengthen user rights management; on the other hand to do data backup and recovery design
- **Friendliness:** Users of this system eventually faced with different professional backgrounds, so this system must be reasonably functional modules organization and interface must simple and easy to operate, so that non-professionals can quickly become familiar with and use the system
- **Scalability:** This prototype system in terms of ontology building, knowledge base construction, function modules, reasoning mechanisms need to constantly improve in many areas, it must be scalable, ontology can continue to improve the construction of the library in future studies, to enrich the knowledge base rules, to enhance the accuracy and efficiency of reasoning and to provide more functional modules and constantly expand and improve the system
- **Practicality:** The ultimate goal of system into practical applications, serving the cause of disaster reduction, so the system described on the basis of research and experimentation, you must have some practical application ability in front, concrete embodiment, when the customer uses the system for data query, system can accurately return data which is user wanted
- **Advancement:** The system uses a technology roadmap previously described, ontology will be applied to the seismic technology collaboration in multi-source data, with some experimental and innovative.

System function design: Prototype system contains a database and a data management system. The main database for storing and managing multi-source seismic data; management system mainly for seismic data to manage data from multiple sources, including system management module, data security management module, data storage module, download data queries, data processing module, data display, etc. module. Specific functional modules as shown below.

System architecture: Prototype system uses classic three-layer architecture, the system the bottom data access layer, business logic layer into the middle, the top of the presentation layer.

Seismic data access layer provides multi-database access and management, including additions and deletions connect to the database, investigate spatial data¡¢deletions attribute data and metadata, the system business logic layer data access layer or presentation layer to provide data services.

Business logic layer is mainly for operating synergies specific problems of multi-source seismic data and can also be understood operate paired data layer, as business logic processing of the data, if the data layer is a building block that is the logical layer structures of these building blocks.

Presentation Layer: WEB mainly showing the way, can also be expressed as WINFORM way, WEB mode can also be expressed as: Aspx, if the logical layer is quite powerful and perfect, regardless of how to define and change the presentation layer, logic layer can improve the delivery of services.

Development of software systems, including construction of a database, system development and integration and to achieve specific business functions. The Development of integrated design system and business functions is performed simultaneously and the actual business needs rational design software system structure and function, reflecting the system professionalism. For storage and management of spatial data via ArcSDE and Oracle, to achieve the storage and management of attribute data through Oracle. Based on the Visual Studio development environment using C# language develop data management system, to obtain user management, data import, data query, data export and other functions.

Earthquake catastrophe scenario builds a database management system database. The picture of development and design as shown in Fig. 1-3.

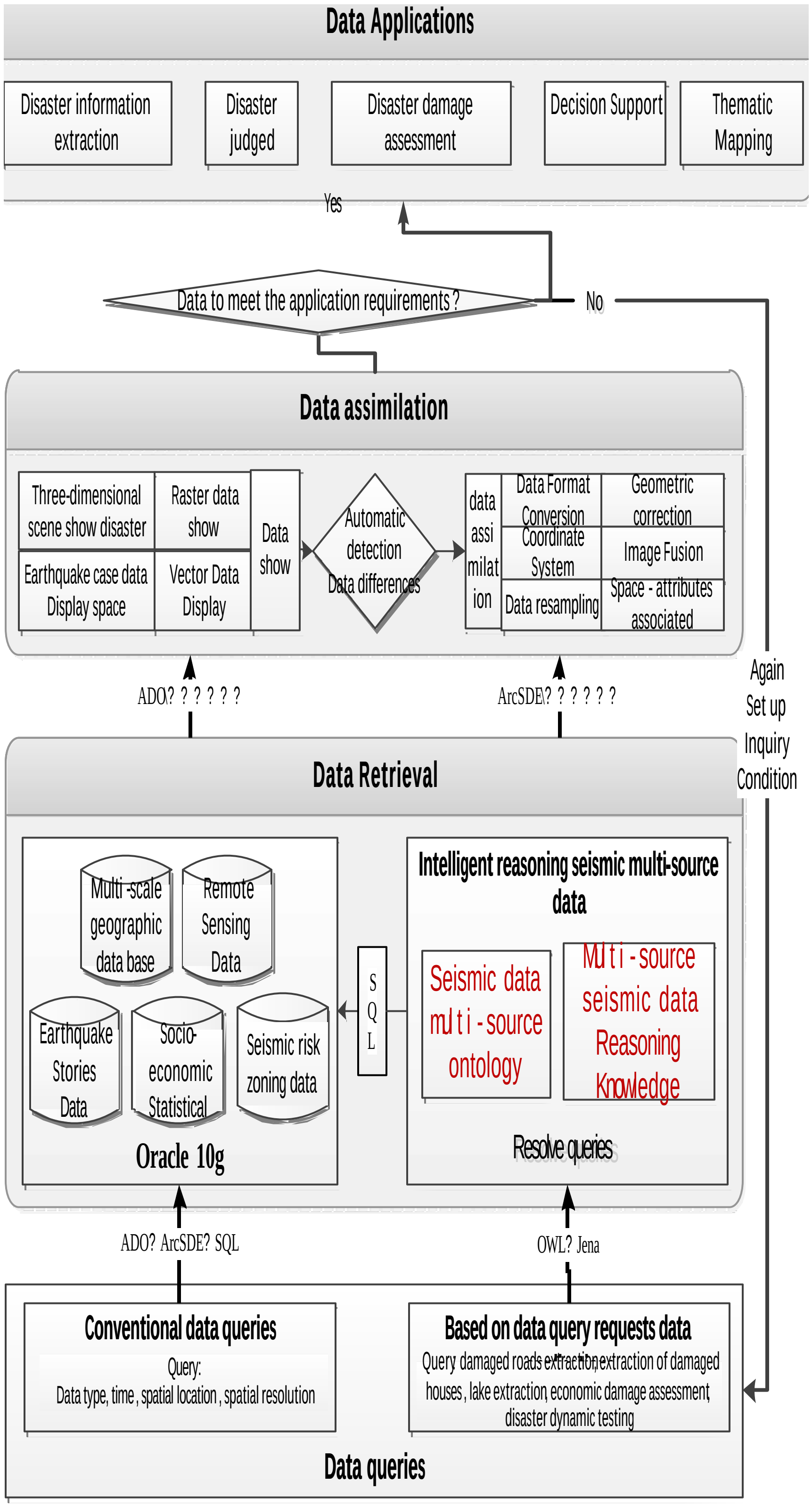

Fig. 1: Software database structure framework

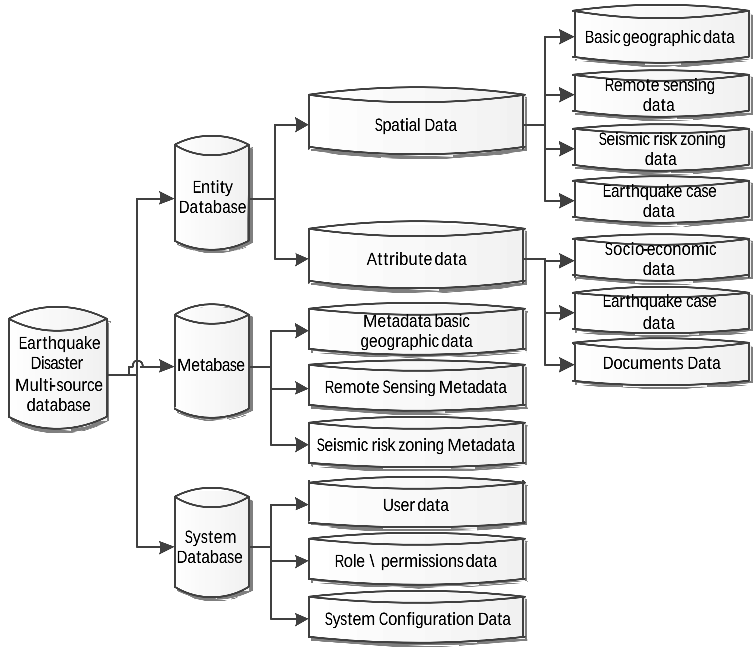

Fig. 2: Earthquake Catastrophe Risk Assessment logical structure of the database

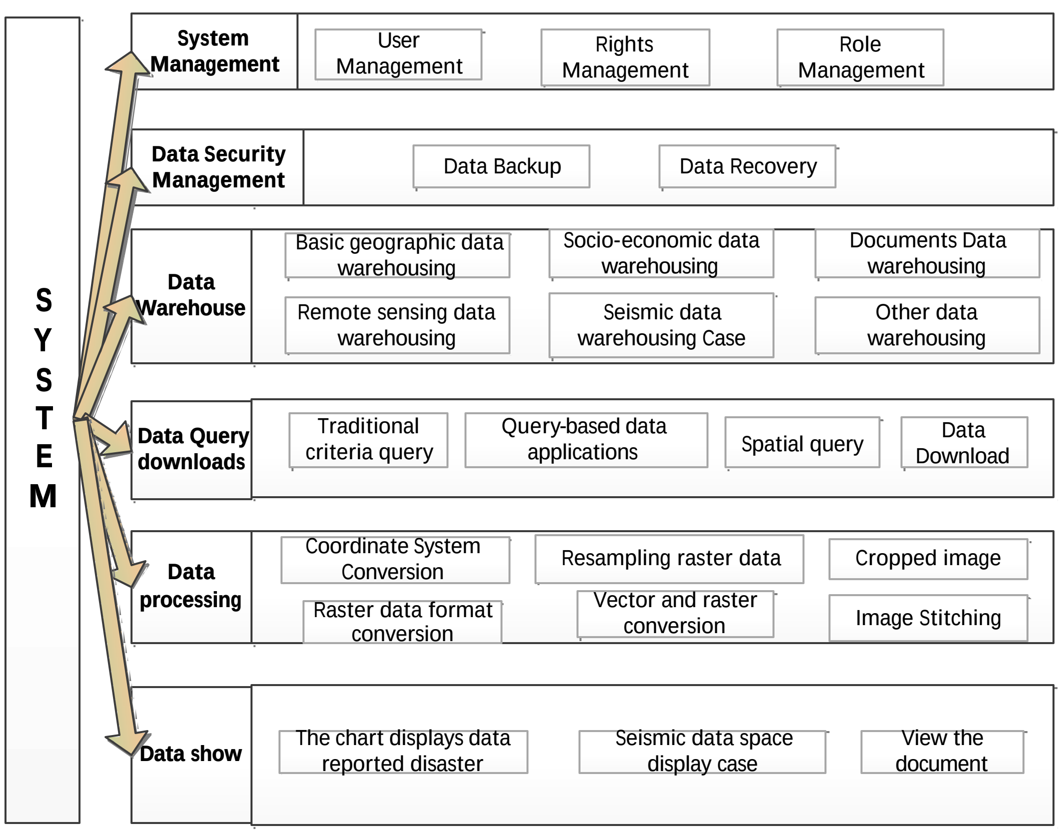

Fig. 3: Database schema

Asia earthquake catastrophe case management system development for building a database design.

CONCLUSION

In this study, collaborative ontology technology to solve the problem of multi-source seismic data, to a certain extent, to avoid the traditional keyword query precision and recall rate issues, while avoiding the traditional data management systems to query the data types in a single data the problem, create multi-source seismic data and data ontology reasoning rules by ontology technology, data query-based applications, according to the rules of inference queries from the database to satisfy the rule data. With the accumulation of business knowledge, knowledge base and inference rules keep getting better, precision and recall rate also has been upgraded. The main results of the work of this study are as follows:

- **Seismic design and construction of multi-source database:** In this study, logical architecture, physical architecture earthquake-source database was designed using Oracle 10g database management system to complete the construction of earthquake-source database. Storage and management of spatial data by which the spatial database engine ArcSDE to achieve
- **Seismic data applications in the field of multi-source ontology construction:** In this study, the application field of seismic data from multiple sources to conduct research important terms of classes and class hierarchies, class features, attributes defined constraints to complete the multi-source seismic data application domain ontology design and finally to build software using protégé ontology
- **Multi-source seismic data repository construction and seismic reasoning multi-source data to achieve synergy:** Meanwhile, based on the survey data, based on the relationship between the data characteristics and data applications to build a data repository reasoning, knowledge base stores the data in a number of inference rules earthquake rescue operations and by inference engine Jena reasoning applied to the data to achieve the realization of seismic data rescue synergy
- **Development of a prototype system:** In this study, the key technologies of seismic data from multiple sources on the basis of ontology-based completed the development of the prototype system, a prototype system to complete the data storage, data retrieval, data processing, data display and other functions. Prototype module provides data retrieval query based data applications, to achieve a collaborative multi-source seismic data

ACKNOWLEDGMENTS

The author thanks the anonymous reviewers for their valuable remarks and comments. This work is supported by 2012 National Science and Technology Pillar Program of China (2012BAH33B03, 2012BAH33B05).

REFERENCES

Bi, S.W. and D.S. Jing, 2004. Information classifying and coding of digital human body. Chinese J. Med. Imaging Technol., 4: 630-633.

He, J.B., X.T. Li, J.T. Bi and Y.R. Cao, 2003. Research on Classification and Coding of Resources and Environment Information and Its Association with Ontology. Geomatics World, 5: 6-11.

Li, H.W., C. Cai and Q.C. Li, 2009. The application of jena in spatial information query and reasoning based on geo-ontology. Eng. Survey. Mapp., 5: 5-9.

Liu, Z., 2011. Analysis on the estimation of earthquake damage and losses of heating system. Master Thesis, Institute of Engieering Mechanics, China Earthquake Administration.