Research on Geo-information Data Model for Pre-selected Areas of Geological Disposal of High-level Radioactive Waste

M Gao¹, S T Huang², P Wang¹, Y A Zhao¹, H B Wang¹
¹Engineer, CNNC Key Laboratory on Geological Disposal of High-level Radioactive Waste, Beijing Research Institute of Uranium Geology, Beijing 100029, China
²Senior Engineer, National Key Laboratory of Remote Sensing Information and Image Analysis Technology, Beijing Research Institute of Uranium Geology, Beijing 100029, China

E-mail: gaomin820@163.com  hshutao@hotmail.com

Abstract. The geological disposal of high-level radioactive waste (hereinafter referred to "geological disposal") is a long-term, complex, and systematic scientific project, whose data and information resources in the research and development (hereinafter referred to "R&D") process provide the significant support for R&D of geological disposal system, and lay a foundation for the long-term stability and safety assessment of repository site. However, the data related to the research and engineering in the sitting of the geological disposal repositories is more complicated (including multi-source, multi-dimension and changeable), the requirements for the data accuracy and comprehensive application has become much higher than before, which lead to the fact that the data model design of geo-information database for the disposal repository are facing more serious challenges.

In the essay, data resources of the pre-selected areas of the repository has been comprehensive controlled and systematic analyzed. According to deeply understanding of the application requirements, the research work has made a solution for the key technical problems including reasonable classification system of multi-source data entity, complex logic relations and effective physical storage structures. The new solution has broken through data classification and conventional spatial data the organization model applied in the traditional industry, realized the data organization and integration with the unit of data entities and spatial relationship, which were independent, holonomic and with application significant features in HLW geological disposal. The reasonable, feasible and flexible data conceptual models, logical models and physical models have been established so as to ensure the effective integration and facilitate application development of multi-source data in pre-selected areas for geological disposal.

1. Introduction
With the development of spatial information technology, many important achievements have been made in the techniques establishment and application of the spatial data model both at home and abroad. In the field of environmental resources, geological data is already one of the most complex spatial data, however, the data involved in the process of research and engineer of high-level radioactive waste geological disposal repository is much more complicated (multi-source, multi dimension and variability), and the requirements for data precision and comprehensive application is even much higher. Therefore, data model design for pre-selected area of repository is facing more serious challenges. How to fully control and systematic analyse data resources, understand application requirements of geological disposal of high-level radioactive waste deeply, and solve the key problems of multi-source data entity and reasonable classification system, complex logical relationships, effective physical storage structure have become the hot and difficult problems in the research field.

2. Objective of data model design
The objective of geo-information data model design goal is to fully meet the specific requirements of environmental data storage layer, development layer and application layer in repository preselected...
area of repository, to classify and name the data fit in with the relevant codes and standards, to high
efficiency integrate and organize multi-source geo-information data. The geo-information data can
meet the following requirements:
(1) the data model can comprehensively describe the multi-source data characteristics of the
environmental objects in the area of the repository, and fully meet the requirements of data resources
needed by the guidelines of the site selection;
(2) the data model can fully reflect the multidimensional characteristics of data entities in time
dimension and space dimension;
(3) the data model has accurate data entity classification mechanism and classification coding system,
and can be classified according to the different attributes of the data entity;
(4) the logical relationship of the data model is consistent with the spatial and non-spatial attributes,
which is convenient for the system integration of multi-source data;
(5) data model can ensure that the physical storage structure of data is clear and simple, and easy to
compatible and migrate in the database management platform (such as MS, SQL Sever, Oracle,
PostgreSQL, etc.).
(6) the data model can effectively guarantee the data exchange between different application platforms
(such as ArcGIS,QGis, ERSI, etc.);
(7) the data model has sufficient flexibility and scalability to facilitate the expansion of a new data
type and compatible with the external data in the subsequent stages of waste disposal;
(8) the data model can make the development layer easy to control all data objects, and can make the
development of automatic data entities to establish effective data association;
(9) the data model can be conducive to the representation of multi-source, multi-dimensional data to
achieve fast and accurate, and easy to realize the mutual representation of different forms (graphics,
images, text table);
(10) the data model has the relative independence and integrity, and the special data model can be used
independently, which is not restricted by the data model associated with.

3. Contents of data model design
Based on the mastery of data object features and the understanding of application requirements, the
contents of data model design to provide comprehensive concept model and logical model, and then to
carry out thematic data model design on the basis of the comprehensive data model. The specific
contents are as follows:
(1) master data objects and their features: data list was established on the existing data resources, and
non-node database was formed. The basic characteristics of the entity data was described in the data
list, including data source, accuracy, existing form (data format), application mode and etc. The
characteristics of the analysis of new data generated in the follow-up phase was carried out the
investigation of relevant fields in foreign countries.
(2) understand the application demand for data: according to ten factors of site selection criteria, a
detailed understanding of the data acquisition, analysis and application in the whole process of the data
entity characteristics is carried out, and then the specific requirements are finally determined. (3)
classify data entity: the entity data are classified according to the spatial and non-spatial features. For
spatial features, the data entity are classified according to the horizontal and vertical space
classification, and in the coordinates of the spatial points are to three-dimensional coordinate of space
graphics expression. For non-spatial properties, the data entity are classified according to the
classification of the subject topics, and then multi-level classification and coding are established.
(4) propose conceptual model: the conceptual data model is put forward in accordance with data object
and its characteristics, the understanding of data application requirements and the classification system
of the data entities.
(5) confirm the logical relationship: the hierarchical relation in the special subject classification and
code table of entity data is the most important data logical relationship. All data from preselected area
are processed as spatial data, and fuzzy space range positioning is used for no clear spatial location
information documents, pictures and other data entity. Then, the logical relationship of measuring
points and the sample in the data model is determine at last.
(6) establish physical structure: special database storage structure and the corresponding elements are organized according to the subject classification of data entities and existing form and attribute description table, including basic geographic data, basic geology, hydrogeology, geophysics, geochemistry, rock mechanics drilling data, changes in the natural environment, humanistic economic, etc.

(7) analyse the rationality and test feasibility: the actual data are used to test the rationality of the physical storage of data entity and its relationship, including the spatial graph, the data package, the picture, the picture, the document and so on.

4. Data model design ideas
The research describes the data model of pre-selected areas in details. First, the design of multi-source data conceptual model are carried out to comprehensively reflect the characteristics and embody relations of the data required by the sitting criteria. Then, the thematic database are set up with various subjects, as well as the description of the thematic data, in particular with the thematic conceptual model, logical model and physical model, are established according to the specific characteristics and application model of the thematic data in multiple areas. The design of the thematic data model not only consider the comprehensive integrity, but also take the independent use of thematic data into account. Thematic data model is the main unit composed geo-information the data model for pre-selected areas of geological disposal, and drilling, samples, documents and pictures and other data objects content involves many disciplines and belong to various disciplines in the logical structure. But the physical structure and storage and other thematic data are in the same layer from the point view of the simplicity of database management system, therefore, these data are associated with other thematic data through two or three level classification encoding.

The design method is to create directory based on all the selected data, establish unstructured database, make specific advice to the professional staff according to the nuclear safety guidelines HAD 401/06-2013 and make clear to the access to all the data entity form, physical meaning, dependence and application mode etc.. The data model is preliminary designed in MS Access, and then all kinds of thematic data model are put forward to experts in various fields to for review. The experts are centralized organized to demonstrate after corrections and make actual tests for the modified data model. Each thematic database in MS Access are transported to PostgreSQL database platform by PLSQL programming, and each subject database corresponds to each special topic. Eventually, SQL is compiled to a comprehensive test to the data storage structure and data relations.

5. Implementation method
There are 10 factors in the sitting criteria for repository, and each factor have a cross overlap with other factors(as shown in figure 1). In order to effectively and systematically integrate these data, data acquisition method and convenience should be taken into account in the design of database structure to set up conceptual model and logical model. Therefore, data storage with different subjects is conducive to the expansion of database, specific organization and application.

5.1. space and time domain
SRID defined in ESPG is used for reference coordinate system in 6 pre selected area, and there is only one spatial reference system in principle. the million years (MA) and second (s) time units are used to describe the data associated with timing respectively. The time of geological age and geological events occurred is described in the attribute table, as well as time description of the field measurement and real-time monitoring data entities (including data packets, data table).
Figure 1. Data requirements and relationship diagram in Sitting Criteria
5.2. data entity storage and data interface

Spatial vector data is stored as WKB and WKT, and pictures, documents and data packet is stored as BLOB. The BLOB can save all the information of original data package. A pre pyramid image structure is set up to store large amounts of remote sensing data (1GB or more), at the meantime metadata is recorded in the database.

Analysis of data entity is carried out through the description of three levels of metadata. First of all, special database and data elements of the table is getting from the meta database, and then all information of data entities are obtained on the basis of thematic classification code, spatial index, attribute index. All spatial data entities (vector and raster) make good use of Gdal/Ogr data interface, and the special file data package is stored using BLOB data storage (such as logging, spectral and multi-data format).

Base on the above research, the conceptual model (as shown in Figure 2) and logical model of geo-information data model is designed in details. Each thematic data model for preselected area of repository has been taken into account. Based on simplified model of logical relationships of data entities for HLW repository, various logic relation of data entities is set up according to the concrete forms respectively, including two-dimensional horizontal surface spatial graph (as shown in Figure 3), vertical section surface spatial graph (as shown in Figure 4), a one-dimensional curve graph (as shown in Figure 5), point data entity (as shown in Figure 6), documents, pictures and other multimedia data entity (as shown in Figure 7).

![Figure 2. Conceptual data model for Preselected area of HLW repository](image)

![Figure 3. Logic relationship diagram of spatial data entity description for plane thematic map of pre-selected area](image)
Figure 4 Logic relationship diagram of spatial data entity description for section thematic map of pre-selected area

Figure 5 Logic relationship diagram of curve data for preselected area of repository

Figure 6 Logic relationship diagram of survey point, sample and analysis data for preselected area of repository

Figure 7 Logic relationship diagram of multimedia data for preselected area of repository
6. Conclusion

The research has made a solution for a series of key technologies, including the rational classification system, complex logical relation, effective physical storage structure, etc. The research has broken through the traditional industry application of data classification and conventional spatial data organization model, realized in the independent, complete data organization and integration by data entities and spatial relations for the basic as the basic unit with characteristics applied significance in the field of geological disposal, established a reasonable, feasible and flexible multisource data concept model, logical model and physical model, which guarantee the effective integration and convenient application of multi-source data of pre-selected areas for repository.

The integrated data model can easily be extended, and provides an open interface development of new data model for subsequent geological disposal system. It is proved that the established data model can completely meet ten data resources, storage and integrated management needs through the test and practical application, and give a strong support to multi-source data integration and management from storage layer in pre-selected areas for geological disposal repository. Geo-information data model is a relatively comprehensive, system data model in the field of geoscience, even there is no similar precedent both at home and abroad, and also can be extended to other nuclear facilities decommissioning and treatment, monitoring or other research fields.

7. Reference

[1] National Standard of People’s Republic of China: Metadata Standard of geo-information data ISO 19115: 2003;
[2] Geological survey technology standard of China Geological Survey Bureau: Metadata Standard of geology data DD2006-05;
[3] Core information metadata standard: TD/T 1016-2003;
[4] Geological survey technology standard of China Geological Survey Bureau: Spatial database of Digital geological map DD2006-06;
[5] Geological survey work standard of China Geological Survey Bureau: Standard for the evaluation of regional groundwater resources survey and evaluation DD2010-03;
[6] Content Standard for Digital Geospatial Metadata Workbook(For use with FGDC-STD-001-1998)Version 2.0Federal Geographic Data CommitteeMay 1, 2000
[7] Content Standard for Digital Geospatial Metadata. FGDC-STD-001-1998