Research on the Construction and Information System Framework of Geological Data Standardization

Xiaohong Wu(✉) and Hanrui Sun

Development and Research Center, China Geological Survey, Beijing 100037, China
{wxiaohong,sunhanrui}@mail.cgs.gov.cn

Abstract. In order to promote the transformation and development of geological work, realize the efficient collection of geological survey data and improve the management and service capabilities, it is necessary to carry out the standardization of geological data. The purpose and significance of the standard construction are expounded, the construction framework of the geological data standardization system is proposed on the basis of analyzing the current situation of the standard construction at domestic and abroad and the demand for the construction of the geological survey standard. The relationship between the geological survey data standard and the informatization standard system is expounded, the overall structure of data standardization is analyzed, the overall structure of the data standardization system with the data model as the core is formed, and the main operation process of the system is described. Problems that need to be further solved are put forward from the aspects of docking the development needs of geological survey business and the consistency of old and new standards.

Keywords: data standard · standardization · information system framework · geological big data · data model

1 Introduction

The standardization of geological data is an urgent problem to be solved in the process of the transformation and development of geological survey. It is the key to the high-quality development of geological surveys and the improvement of social service capabilities. It is not only conducive to promoting the aggregation, management and service of geological data, but also to promoting the healthy development of geological industry. The standardization of geological data can give full play to the Co-ordinate, coordinate and supporting role of standardization and informatization work on geological surveys, and is the key to promoting the development of geological information services and industries.

The standardization is inseparable from the support of information systems. The standardization system provides data management services for organizations, standard information, metadata information, geographic elements, data models, etc. which involved in the standardization process; at the same time, it provides a basic information platform for standard management to promote the implementation of standards.

G. Guan et al. (Eds.): ICBDSS 2022, AHCS 8, pp. 81–89, 2023.
https://doi.org/10.2991/978-94-6463-064-0_10

2 Purpose and Significance of Standard Construction

Standards are the technical support for economic activities and social development, standardization plays a fundamental and leading role in promoting the modernization of the national governance system and governance capacity [1]. From the ancient Chinese "cars on the same track, books on the same text", to modern industrial large-scale production, it is a vivid practice of standardization [3]. There are standardized practices and applications in all aspects of the geological survey industry.

The definition and interpretation of the concept of standards in China is based on the definition given in the GB/T 20000.1-2014, namely: "A document for common and repeated use by standardizing activities, developed by consensus in accordance with prescribed procedures, and providing rules, guidelines or characteristics for various activities or their outcomes."

The standardization of geological survey data is to realize the efficient collection and management of geological survey data, improve the quality of data services, and promote the high-quality development of geological work.

3 Status Quo of Standard Construction at Domestic and Abroad

In the international standardization process of the geographic information industry, the most influential are the standardization activities carried out by ISO/TC 211 and OGC. Many geographic information standards developed by these two organizations have been recognized and widely used, and are implemented as supported functions by many well-known GIS vendors.

The scope of ISO/TC211 is standardization in the field of digital geographic information. The goal of the work is to develop a structured set of standards for objects or phenomena related to locations on the earth to standardize geographic information data management (including definition and description), acquisition, methods, tools and services for processing, analyzing, accessing, expressing, and transferring data in digital or electronic form between different users, systems and locations.

Combined with applicable information technology and data standards, it provides a framework for the development of industrial applications using geographic data [6].

OGC is an international consortium dedicated to the discovery, accessibility, interoperability and reusability of geospatial information and services [7], providing a unified specification for the interoperability of data and services between geographic information systems, and for GIS software developers. Provide a detailed public guideline to enable the developed software to achieve the interoperability of geographic data and geographic resource processing [4].

In order to realize the docking with international standards, China has gradually made efforts in organizational management and policy mechanisms. The National Geographic Information Standardization Technical Committee established in 1997 is responsible for the planning, coordination and technical centralized management of national standards in the field of geographic information; The National Standardization Management Committee established by the state in 2001 is used to coordinate, guide and supervise industries, localities, groups and enterprises Standard work.

China has successively issued relevant policies and regulations. In December 2015, the General Office of the State Council issued the "National Standardization System Construction and Development Plan (2016–2020)", which is the first national special plan in the field of standardization in China; in October 2021, the Central Committee of the Communist Party of China and the State Council issued the "Outline of National Standardization Development", which is used to coordinate and promote the development of national standardization.

In the field of geological survey, China has compiled many database construction specifications to guide data production [5]. There are more than 120 standards and project standards, and a standard system covering all stages of original data, result data, data management and data services has basically been formed.

4 Requirements for the Geological Data Standards

The standardization of geological data is not only a requirement for the development of the national situation, but also an urgent need for data management and services in the geological industry.

The "National Standardization System Construction and Development Plan (2016–2020)" clearly proposes to strengthen the formulation and revision of relevant standards for the geological survey industry; The "Outline of National Standardization Development" also proposes to "study and formulate standards for the economical and intensive development and utilization of natural resources such as land and mineral resources, and promote the standardization of green exploration and development of energy resources"; in the "14th Five-Year Plan" geological survey informatization plan, the China Geological Survey required to improve the informatization standard system, and promote the construction of informatization standards in key areas. The following problems still exist in the construction of Geological database.

1. Compared with the international standardization process, the construction of geological survey data standards is in a backward state. In terms of international standards, we generally deal with the follow-up state.
2. The geological data standards are relatively scattered, overlapped and repeated, and have not formed an organic whole. The construction of the standard system needs to be further improved.
3. Many geological database standards are project group standards, with narrow application and low standard repetition rate; there are few general and common underlying standards [4], and the importance and basic standards that require the participation of multiple departments have not been developed.

Since 2017, the China Geological Survey has built a geological cloud, and has built a number of distributed data centers to provide data services for more than 3,000 layers. In order to realize the data interaction and unified services of each data sub-center, it is necessary to improve the construction of the geological data standard system to meet the current and future business management needs of geological surveys.

- Solve the inconsistency in understanding and expression of the same geological phenomenon caused by the intersection of disciplines. By building a unified data model, the semantic unity and expression consistency of the same geological professional data in different fields can be realized, and the aggregation, management and service of geological data can be improved.
- Promote the construction of a standard system, build a relatively complete standard management system and service system, and provide information support for knowledge maps, big data analysis and mining, and AI.

With the continuous advancement of the national standard system construction project, it has become an important task for the standardization of geological survey information to further clarify the standardization needs of the geological survey industry and help build a standard system that meets the actual needs of its own development.

5 Geological Data Standardization System Framework

In order to support the high-quality development of geological surveys and focus on the transformation and development of geological surveys, the China Geological Survey released the "Technical Standard System for Geological Surveys (2021 Edition)", and at the same time, a series of geological data standards have been revised around the " Standard System for Geological Survey Informationization".In order to improve the management and service capabilities of geological data, the development of general data model and thematic data model standards was organized to solve the problem of terminology and semantic consistency of different disciplines and standards, and to form the basic standard of geological data.

5.1 Relationship with Geological Survey Information Standard System

The geological survey information standard system is one of the twelve sub-systems of the geological survey standard system, and the geological survey data standard is an important part of the geological survey information standard system. In order to fully absorb the achievements in the construction of international standards for geographic information, refer to ISO/TC211 and OGC international standards, formulate and improve data model standards for China's geological industry, and serve the transformation of geological survey work. The standard of geological informatization includes four parts: data general standard, data model, data management standard and exchange service standard, and focuses on supporting the construction of standards related to data in the standard system of informatization (Fig. 1).

In the geological survey informatization standard system, the construction of data standards is supported by basic information standards, information security standards and informatization management standards. At the data management level, it is mainly to develop data model standards, including general data models and various professional data models, to solve the problem of unification of industry semantics. The unification of the data model is conducive to realizing the consistency of the collected data and facilitating the data aggregation, management and service of the data sub-centers of the

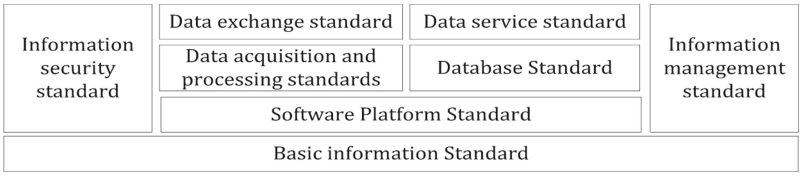

Fig. 1. System Architecture of Geological Survey Informatization Standard

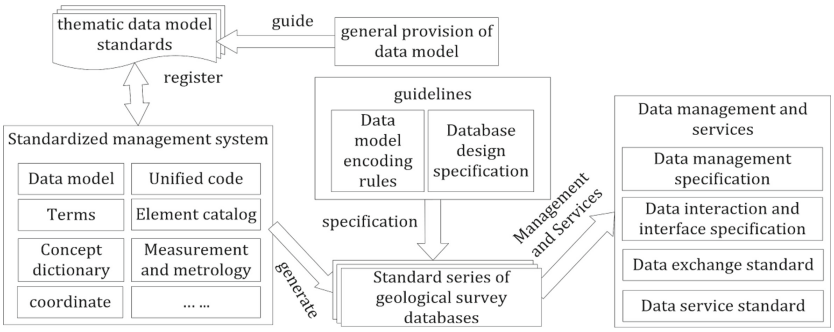

Fig. 2. General structure of geological data standardization

China Geological Survey. At the data service level, the protocols provided by ISO and OGC are mainly used. GML is used for data exchange, WFS and WCS for data access, WPS for data processing, and WMS and WMTS for visualization standards. With other systems, through RESTful interface or Web Service interface, an open system ecosystem is built to realize data sharing and reuse between systems.

5.2 General Structure of Geological Data Standardization

The most basic of the geological data standards is the data model standard, which is mainly based on the general provision of data model and the special data model standard. The general provision specifies the common guidelines to be followed by various thematic data model standards. On this basis, supplemented by a standardized management system, according to the coding rules and database design specifications, special standards suitable for the professional fields, projects or applications of geological surveys are formulated, so that the special standards established based on the same basic standard have the semantic consistency and same data definition. The professional standards formulated, through the management specifications, development specifications, exchange specifications and service standards in the data standard series, realize the basic management, data exchange and sharing services of geological data (Fig. 2).

Geological data services are developing in the direction of online superposition, calculation, and analysis to realize unified data services. The database standards generated according to the geological data standard framework can ensure that the data extracted from the distributed data centers can complete the rapid data overlay and service integration, generate the special reports or required atlases in a timely and rapid manner,

and realize the data exchange and data sharing between the data centers. Standardization system can ensure the coordination of various standards in the process of standard formulation, including model coordination, UML diagram coordination, terminology coordination, basic content coordination, interdependence coordination, and so on, and achieve the maximum consistency between different standards [8].

5.3 System Framework of Geological Data Standardization

Geographic information interoperability is to establish effective communication between systems, geographic information applications and users. Systems, applications and users are different, and there are four different types of heterogeneity among them: system heterogeneity, syntactic heterogeneity, structural heterogeneity and semantic heterogeneity. Syntactic heterogeneity refers to the physical representation of data, structural heterogeneity is related to conceptual models of geographic elements, and semantic heterogeneity refers to differences in meaning between data that express concepts of reality due to differences in real-world phenomena. The perspective (or context) is caused by abstraction [2]. Resolving semantic heterogeneity through data standards and data models, thereby reducing structural heterogeneity and syntactic heterogeneity, is the key to realizing geographic information interoperability and the foundation for effectively promoting geological information sharing services.

Referring to ISO/TC 211 and OGC standard system framework, a standardized system including metadata, registry, common terms, element data, code table and other unified management functions has been constructed, which can effectively organize unit information, contact information, standard information, terminology Information and metadata information, can solve the problem of the unification of data models in various data standards, and can improve the management and service capabilities of geological survey standards (Fig. 3).

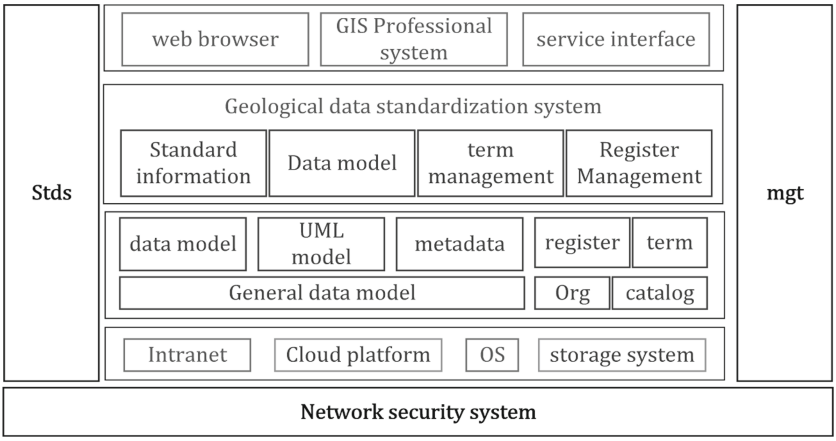

Fig. 3. General structure of geological data standardization

The data standardization system takes the data model as the core. Through the management of standard-related information (such as concept dictionary, registry, metadata, etc.), it effectively integrates the element models of various professional fields to form professional data standards. and realizes data exchange between different systems through the unified coding and service specifications. The data standardization system provides a unified information management platform for data standard management services, systematically manages all data-related standards and specifications in the geological industry, and realizes the management and promotion of data standards.

The management process of the geological data standardization system is shown in the figure below. Before formulating a data standards, users should check whether there are similar standards in the standardization system. If there is no corresponding standard and the application is approved, the system can be used to start data standard writing. According to the understanding of the geological business model, the standard compiler selects the corresponding data model and element information from the models that have been stored in the database and passed the expert review, and selects the corresponding reference standards, terms and other information from the system. The relationship is quickly combined into a new professional database standard, and the standard compiler can only modify the partial content according to the business needs. The dashed part in the figure below involves the selection or submission of the data model, which is the core of the standardized system.

Using a standardized system to manage and generate standards can ensure the consistency of core elements in various professional fields, and facilitate the rapid integration and service sharing of similar data in different units (Fig. 4).

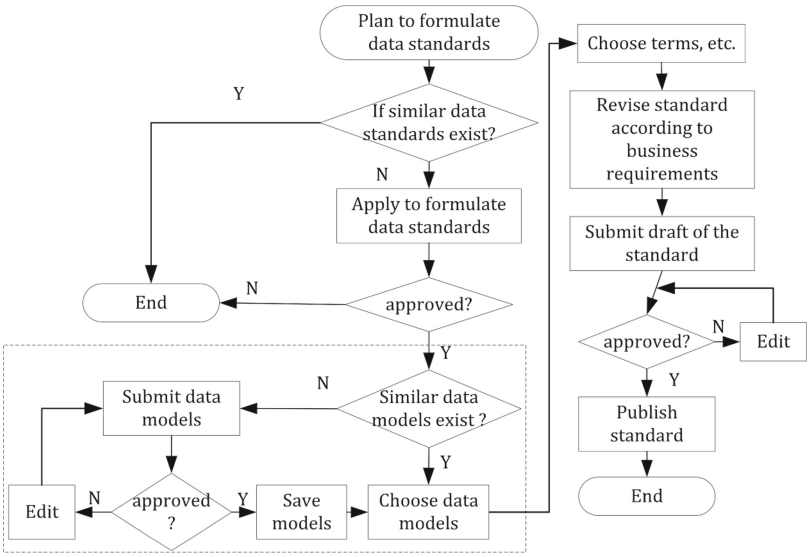

Fig. 4. Management process of the geological data standardization system

6 Conclusions and Existing Problems

Data is the core of enterprise assets, geological data standard is the basis of the data standardization. Data standardization system provides functions such as unified management of data models, execution and analysis of data standards, provides information support for the formulation, management, release, query and execution of data standards, and provides an information platform for the data standardization work of enterprises.

Geological survey has established a relatively mature standard system and service framework in terms of geological data management and services, but with the development of geological survey application requirements and the development of information technology, there are still some problems that need to be solved and improved.

6.1 Actively Meet the Development Needs of Geological Survey Business

With the development of economy and society, the state has put forward higher requirements for geological work. The geological survey business is constantly developing and changing. The establishment of deep-ground, deep-space and deep-sea, monitoring and observation systems, and natural resources surveys are organized and implemented by many different units. These geological work will generate a large amount of new data. The collection, aggregation, storage, management and service of data collection, aggregation, storage, management and service requires professional support software and management tools, and the organization and implementation are carried out in accordance with unified standards to achieve standardized data management and services. The new geological data management model requires continuous improvement of the geological data standard system.

6.2 Data Consistency Problems Caused by Old and New Standards

After decades of development, geological work has accumulated a large amount of original data and result data. These data are generated according to the old data standards and are closely related to the information level and management system at that time. How to establish an effective data interface, realize the effective use of old and new data, maintain data consistency, and minimize the loss of data value is an issue that needs to be seriously considered during the implementation of the standard.

Acknowledgement. This paper is supported by the following projects funding: Geoscience Data Integration and Services (NO. DD20221785).

References

1. Central People's Government of the People's Republic of China, 2021. Development Outline of National Standardization. http://www.gov.cn/zhengce/2021-10/10/content_5641727.htm
2. Geographic information - Reference model - Part1: Fundamentals [S], ISO19101-1, 2014: 10-11

3. xinhuanet, Xi Jinping 's congratulatory letter to the 39th International Organization for Standardization Conference, 2016. http://www.xinhuanet.com/politics/2016-09/12/c_1119554153.htm.
4. Overview of GIS Standardization. Geographic Information World [J], October 2004, Vol. 02 No. 5: 11-15
5. Structure and draft for specification of geological database construction [S], DZ/T 0274, 2015: 6-7
6. International Organization for Standardization, SCOPE. https://www.iso.org/committee/54904.html
7. Open Geospatial Consortium, About OGC. https://www.ogc.org/about
8. Jiang Jingtong. Policy Research on China's Geographic Information Standards. Science of Surveying and Mapping [J], 2008 1, Vol 33 No 1: 21-24