Construction and Practice of Knowledge Service System: A Case Study of Marine Knowledge

Lili Song*, Shengwen Cao*, Mogeng Xu*, Xiaoyi Jiang*
*National Marine Data and Information Service, MNR
Tianjin, China
Email: 547177942@qq.com, cao309@163.com, 974160215@qq.com, jiangxiaoyi@nmdis.org.cn

***Abstract*—In the era of big data, integrating, sharing, mining, and analyzing multi-source heterogeneous massive data to obtain useful knowledge has become an important research content and key issue that urgently needs to be solved in various fields. This paper analyzes and puts forward the marine knowledge and organization and management methods, focuses on the construction and practical results of marine knowledge service platform, and discusses key technical issues such as the integration of marine domain knowledge resources and the construction of knowledge maps. The research results can not only provide personalized and accurate knowledge services for all kinds of users involved in the sea, but also provide reference for the construction of knowledge service system in the ocean and related domain knowledge.**

Keywords-knowledge service; knowledge service system; marine knowledge service; knowledge graph.

I. INTRODUCTION

This paper is based on the previous work originally presented at the Sixteenth International Conference on Advances in DBKDA 2024 [1]. The key technologies of marine knowledge service was added in Section IV.

With the rapid development of Internet, 5G, artificial intelligence, Internet of things and other technologies, as well as the acceleration of digital economy in the world, we are already in the era of "information explosion" or even "post information explosion", and the amount of information in global science and technology, economy, society and other fields is growing at a geometric level. According to the statistics and predictions of the international authoritative agency Statista, the global data volume is about to experience a larger explosion, and the global data generation will reach 2142 ZB by 2035 [2]. With the development of various detection equipment and computing simulation, scientific research has entered the fourth normal form of "data intensive discovery" [3]. Ecosystems, Earth observation, oceans, life and health, astronomical observation and other professional disciplines have generated and accumulated massive scientific data through continuous observation/monitoring [3][4]. The data obtained through continuous, multi-source, and three-dimensional observation and monitoring methods using remote sensing, communication, and other technologies in the fields of Earth observation and ocean observation is growing at a daily terabyte rate [5]. As professional data enters the era of big data, how to analyze and mine a large amount of data to obtain useful knowledge has become an important research content and a key issue that urgently needs to be solved.

The ocean is an important strategic space for coastal countries around the world to strive for their interests and development. With the further development of economic globalization and regional integration, the impact of the ocean on national economic development, political diplomacy, and national security is becoming increasingly significant. With the deepening of the strategy of strengthening the marine power and the Belt and Road Initiative, marine information plays an increasingly important role in supporting national strategic planning and the development of marine undertakings. The demand of the state, the government, high-end research experts, the public, especially engineering and scientific personnel, for marine information sharing and services is growing. Therefore, the construction and application of marine knowledge service system is carried out, Realize the integration and precise services of knowledge resources such as literature data, scientific data, expert institutions, and internet data in the marine field, and provide support services for various users to carry out marine research, marine management, and marine strategic planning.

The rest of this paper is organized as follows. Section Ⅱ presents the related work of this paper. Section III presents the classification and management of marine knowledge resources. Section IV presents the construction and functions of marine knowledge service platform, including knowledge content and organization, platform functionality, and effectiveness. Section V presents the key technologies of marine knowledge service. Section VI summarizes lessons learned, conclusions, and future work.

II. RELATED WORKS

In recent years, professional domain knowledge service has gradually become the development strategy of national science, technology and culture. In September 2015, the State Council issued the "Action Outline for Promoting the Development of Big Data" [6], which clearly proposed to carry out "knowledge service big data applications,

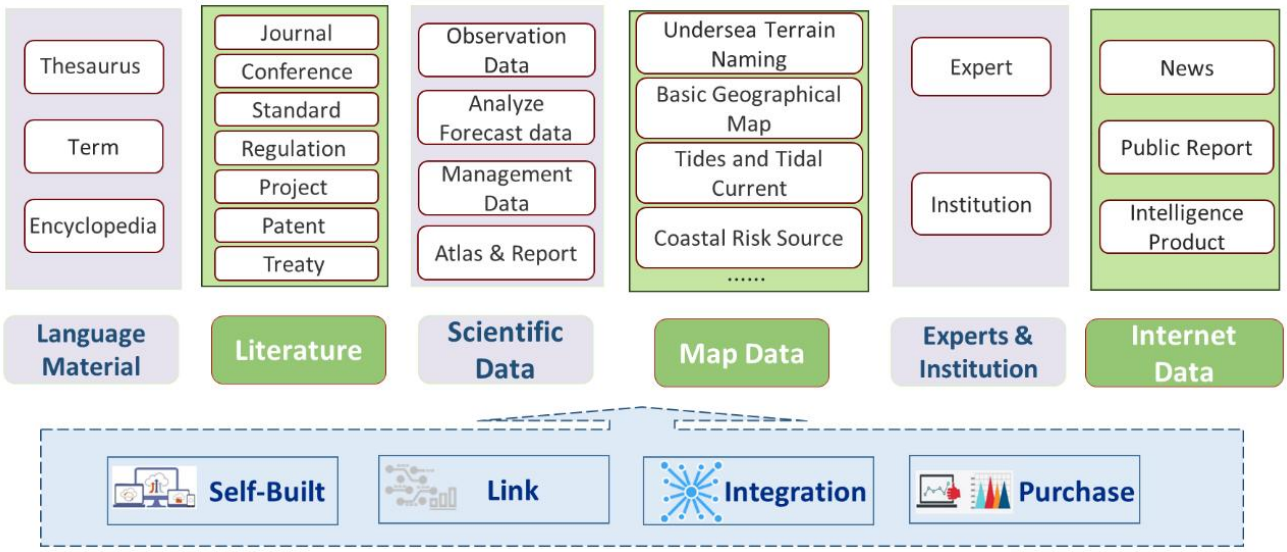

Figure 1. Classification of marine knowledge resources.

establish a national knowledge service platform and knowledge resource service center. ” In 2016, the construction of the Knowledge Resource Service Center by the China Press and Publication Research Institute was approved to promote the national knowledge resource database engineering project. The Chinese Academy of Engineering launched the construction project of “China Engineering Science and Technology Knowledge Center” in 2012. In 2017, the project was included in the big data project supported by the state. With the overall goal of building a national “engineering science and technology think tank”, the project is committed to meeting the needs of the state, the government, high-end researchers, the public, especially engineering and technology personnel, and connecting China’s engineering science and technology fields, including agriculture, forestry, meteorology, chemical industry, marine. With the goal of deep knowledge mining, we aim to build various professional knowledge service systems in the field of engineering and technology based on massive data from metallurgy, geographic resources, and other fields. We aim to build a knowledge integration body with the richest engineering and technology information resources, the widest application range, and the strongest practicality in China. Various professional fields have also conducted research and practice in the construction of knowledge systems, the development of knowledge service standards, the processing of knowledge resources, the construction of knowledge service platforms, and the exploration of knowledge service models [7]-[9].

The definition of knowledge is currently not clearly defined. As the most important concept in the field of philosophy, Plato believed that a statement that can be called knowledge must meet three conditions: it must be verified, correct, and believed by people [10]. The term 'knowledge service' originates from the field of enterprise knowledge management. Although current technology and theory connected to knowledge graphs are progressing rapidly, the integration of knowledge graphs with the marine domain is not comprehensive enough to use relevant technology to obtain further information in the marine domain. In fact, large-scale research on knowledge graphs in the marine domain, embracing various facets of oceanography, has only arisen in recent years [11].

This paper focuses on the construction and practice of marine knowledge service system and investigates the key technical problems of marine knowledge resource integration and knowledge map construction. The research results not only provide personalized and precise knowledge services for various marine-related users, but also provide reference for the construction of knowledge service systems in marine and related fields.

III. The Content and Organization of Marine Knowledge

Knowledge resources are the important basis of marine knowledge service. Determining the classification and management methods of marine knowledge is the foundation for building a marine knowledge service system.

A. Classification of Marine Knowledge Resources

Marine knowledge resources mainly include literature data, scientific data, map data, basic corpus data, Internet data, expert institutions and other comprehensive data. The acquisition methods of knowledge resources mainly include self-establishment, purchase, link, integration and so on. The classification of marine knowledge resources is shown in Figure 1. At present, a sustainable and updated marine knowledge resource system covering 85 sub-categories and 6 major categories has been integrated and formed.

Marine literature data mainly covers Chinese and foreign journal papers, academic dissertations, conference papers, sea-related projects, patents, treaties, standards, policies and regulations, etc. in the field of marine science and

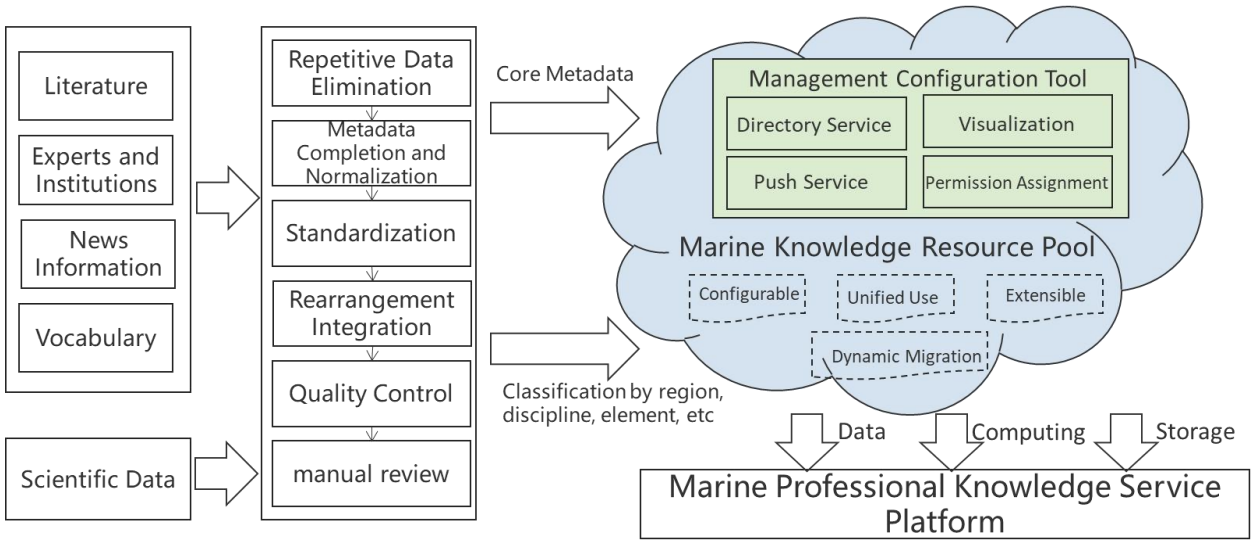

Figure 2. Flow chart of marine knowledge resource management

technology. The data sources are commercial purchase, Ministry of Natural Resources, Ministry of Science and Technology and other ministries and official websites of the government. Basic corpus data includes marine thesaurus, entries and encyclopedias, etc., and the data comes from existing self-built resources in the marine field. Internet data includes news information, public reporting and intelligence products.

Scientific data is the most complex and data intensive part of the marine knowledge resource system. The ocean itself is a huge, complex, and nonlinear system, with extremely complex phenomena and processes, and vastly different spatiotemporal scales [3]. Therefore, marine scientific data is divided into measured data, analysis and prediction products, and comprehensive management data according to their sources and processing methods. Actual measurement data is a standard data obtained through a series of processing, integration, and quality control of ocean data obtained by different observation methods. It is classified into disciplines such as ocean hydrology, ocean meteorology, ocean acoustics, ocean optics, ocean chemistry, ocean biology, seabed sediment, and ocean geophysics. Analysis and forecasting products refer to ocean hydrological and meteorological basic field data and graphic products produced at different time scales, spatial resolutions, and levels. The time scales include calendar year/year, quarter by quarter, month by month, day by day, hour by hour, and even more refined time scales. The spatial resolution includes 1 degree, 1/2 degree, 1/4 degree, 1/8 degree, 1/16 degree, and special customized products in some regions (such as Bohai Bay, Liaodong Bay, etc.). According to different processing methods, products are divided into conventional statistical products, reanalysis products, forecast analysis products, and live analysis products. Comprehensive management data includes marine comprehensive management data such as sea area management, island management, marine economy, forecasting and disaster reduction, and seabed terrain naming, usually in the form of thematic atlases and reports. Map data is to classify, extract and summarize the parts of scientific data with spatial information according to special topics, including basic geographical base map, submarine terrain naming, tidal current stations, coastal risk source and other data. Marine scientific data comes from satellite and ocean stations, buoys, ground wave radar and other field observation stations, as well as daily business management.

B. Management of Marine Knowledge Resources

Marine knowledge resource management adopts the principles of “unified construction of resources, on-demand use of services, and automatic update of knowledge”, and uses virtualization technology to build marine knowledge storage resource pools and computing resource pools to realize unified allocation, unified use, extensibility, configuration, and dynamic migration of resources, as shown in Figure 2. Due to the large number of sources of marine knowledge resources, metadata information, data formats and storage methods are quite different. Before entering the knowledge resource pool, standardized description and processing of metadata should be carried out according to certain standards, including repetitive data elimination, metadata completion and standardization, data standardization processing, rearrangement and integration, quality control and manual audit. Among them, the metadata of literature, expert institutions, news information, basic corpus, etc. are standardized in accordance with the metadata specification of China Engineering Science and Technology Knowledge Center [12], and the standardization of marine scientific data is implemented in accordance with the standards and technical specifications of marine information metadata, marine data application record format, and marine observation data management methods. The marine knowledge resource pool provides unified permission management and various resource interfaces to realize catalog service, push service, resource visualization, permission allocation and other functions.

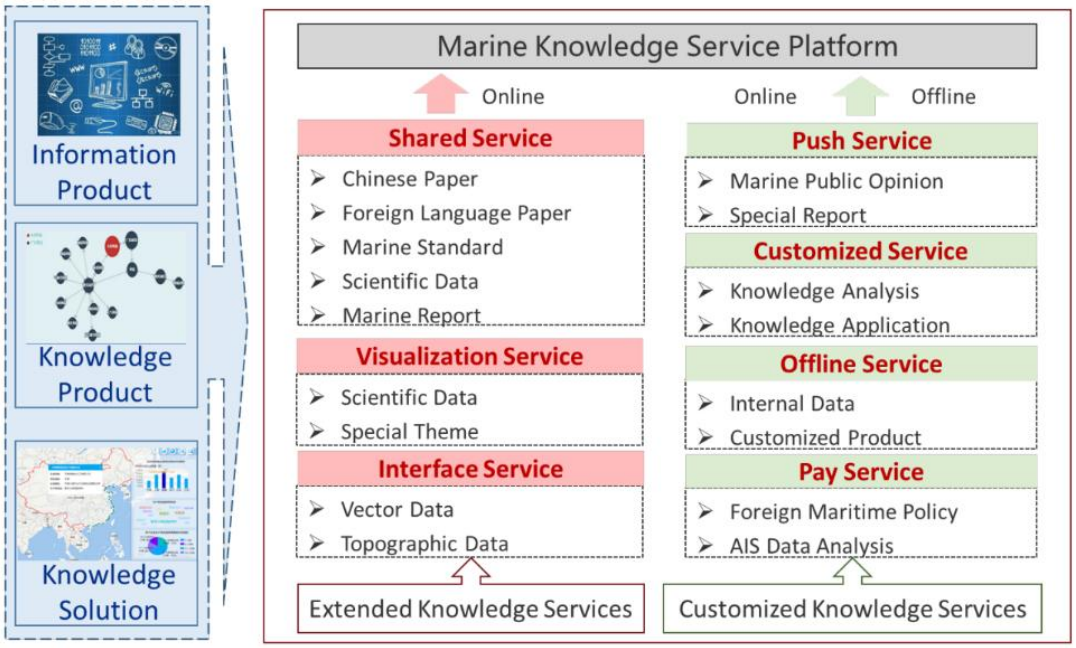

Figure 3. The mode of marine knowledge service, including information, knowledge and solutions

IV. Construction and Functions of Marine Knowledge Service Platform

Knowledge services require specialized and integrated system platform support. The practice of marine domain knowledge service takes the integration and collection of data, processing of information, and analysis and mining of knowledge as the chain to build a resource rich, widely used, advanced and practical marine knowledge system, and realizes the functions of integrated knowledge search, multiple service modes, accurate knowledge application, and comprehensive special topics, providing rich, accurate A three-dimensional knowledge service.

A. One Stop Marine Knowledge Retrieval

Knowledge retrieval is the foundation of knowledge services. Traditional retrieval involves querying based on conditions such as time, space, and subject words. In order for users to quickly find the required data, intelligent retrieval such as full-text and cross language retrieval, as well as building search engines, have become essential functions for many knowledge service platforms or systems. The diversity of marine knowledge resources requires the establishment of a full-text index to achieve a one-stop marine knowledge search. Based on the marine vocabulary (including both Chinese and English), information extraction, filtering, indexing, text classification, word segmentation, and other processing work are carried out on the original corpus. The Solr full-text search engine is used to automatically sort, summarize, and create index files for the processed corpus. Through syntax analysis, internal code conversion Automatic filtering and classification, ultimately achieving classification retrieval, fuzzy retrieval, cross language retrieval, and secondary retrieval, in order to retrieve the knowledge that users need as much as possible. The ultimate goal is to achieve efficient discovery of multi-source heterogeneous marine data resources on the platform, as well as convenient access to original and physical data, providing users with a one-stop intelligent retrieval service of “what you need is where you are, what you see is what you get”.

B. Diversified Knowledge Service Methods

The mode of knowledge services is divided into three levels: information services, knowledge products, and knowledge solutions. From the perspective of user needs, it is divided into extended knowledge services and customized knowledge services [13]. Extensive knowledge services are not targeted at specific users or problems, but mainly focus on providing general data and knowledge content in the field, such as ocean knowledge sharing services, visualization services, interface services, etc. Customized knowledge services refer to providing corresponding knowledge content or solutions based on user needs, such as push services, customized services, offline services, and paid services. The ocean knowledge service method is shown in Figure 3.

In extended knowledge services, ocean knowledge sharing services provide data or products to users through search and retrieval, online computing, download and use, etc. For self-built data resources, users are provided with free downloads, paid services, offline applications, etc. For commercial purchases of resources, a small amount of information transmission services is provided in the form of products. Visualization service is to intuitively display scientific data and thematic services through histogram, pie chart, bar chart, facing chart, scatter chart, radar chart,

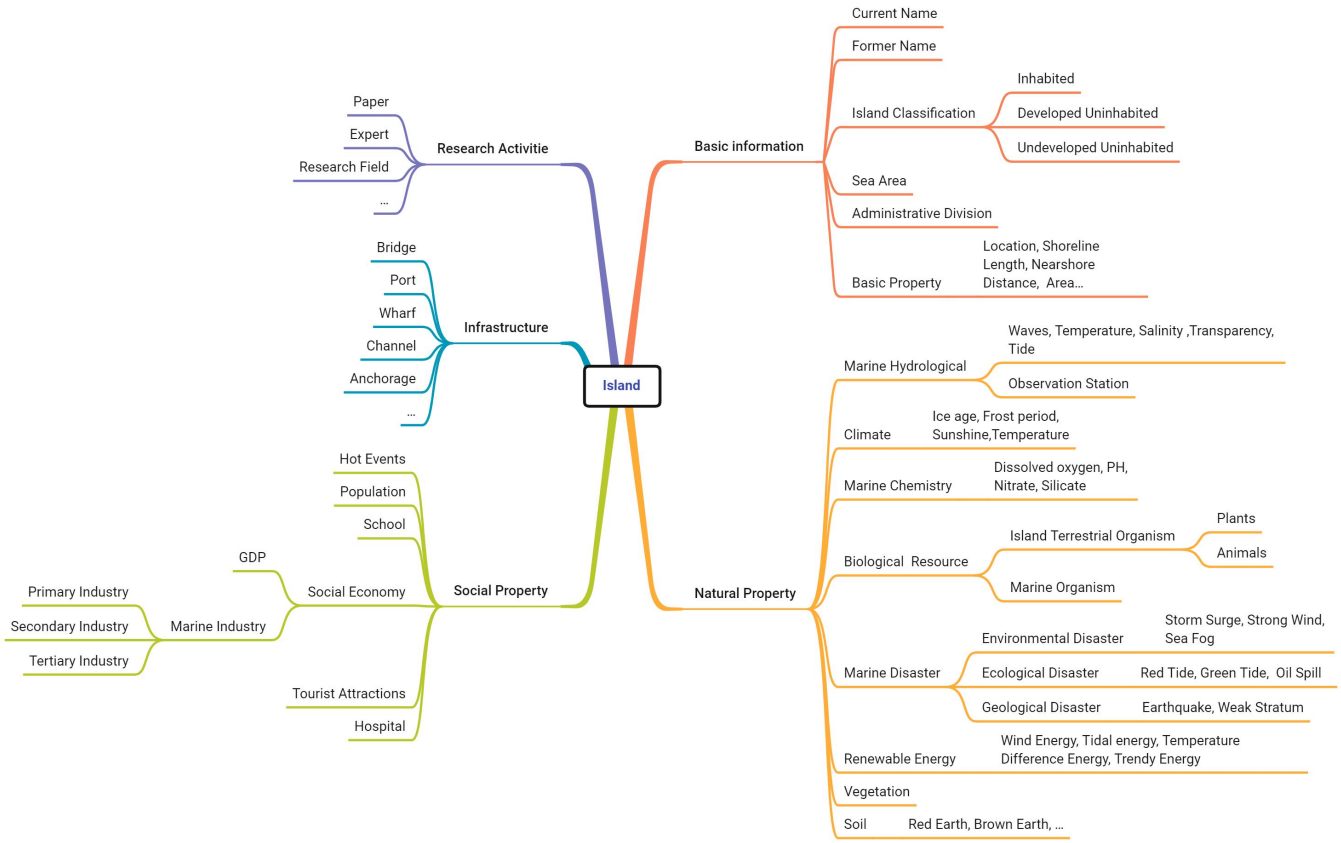

Figure 4. Part of Island knowledge classification system

bubble chart and other chart forms, as well as curve chart, profile map, contour line, contour surface and other two/three-dimensional map display forms, which is conducive to users' intuitive understanding of the characteristics of marine knowledge resources and provides technical support for users ' scientific research, marine engineering, etc. Interface services are services that publish vector maps, terrain maps, etc. in the form of interfaces for users to call.

C. *Precise Knowledge Application*

The marine knowledge service platform focuses on the research fields of national strategy, maritime emergencies, and marine hot spots, studies theme-oriented automatic aggregation technology of knowledge resources, and develops 16 marine thematic knowledge aggregation products such as the 21st century Maritime Silk Road, sea level rise, blue economy, and undersea topography and place names, providing comprehensive, in-depth and accurate thematic knowledge services. Deep processing, indexing and knowledge-based organization of marine data resources are carried out, and knowledge applications such as trend analysis, scholar analysis and institutional analysis are developed. Knowledge maps of typical scenarios such as island development and utilization and site selection for marine ecological restoration are studied and constructed, so as to realize multi-dimensional disclosure and semantic correlation of marine knowledge and effectively support decision-making and application in the field of marine management, as shown in Figure 4.

D. *Marine Think Tank Support Services*

The platform tracks and compiles important ocean related policies, documents, and reports released by international organizations and institutions in multiple countries and regions around the world, collects, organizes, and analyzes international/regional ocean situations and dynamic information to form international ocean public opinion reports by publishing a series of annual reports such as the China Ocean Economic Statistics Bulletin, the China Ocean Statistics Yearbook, the China Island Statistics Bulletin, and the China Sea Level Bulletin. It also explores areas for tracking and analysis Think tank information service mode combining information reference and data analysis, create a series of marine think tank report products combining " authoritative bulletin, public reports and public opinion analysis " , and provide support for major strategic and theoretical research, marine protection and development, global marine governance, marine scientific research and other fields such as marine power, Maritime Silk Road, etc.

E. *Results of Construction and Operation*

By the end of 2023, a system of sustainable and updated marine knowledge resources covering 85 sub-categories in 6

categories, including scientific data, papers, standards, patent achievements, policies and regulations, and expert institutions in the marine field has been built, totaling about 5 million pieces. It will develop 7 knowledge applications, such as shipping big data analysis and tidal current forecast, and 9 feature topics, such as blue economy and sea level rise, to create an integrated platform for marine information services and knowledge application and provide multi-dimensional and deep marine knowledge mining and analysis services for high-end think tanks, engineers and scientists and the public. The annual user visits are stable at more than 1 million times. Cumulatively, it provides data and report support services for many academicians and dozens of national key research and development programs and local research projects.

V. KEY TECHNOLOGIES OF MARINE KNOWLEDGE SERVICE

In the process of constructing the ocean knowledge service system, it is necessary to study the panoramic integration technology of multi-source marine resources and the construction technology of ocean thematic knowledge graph.

A. Panoramic Integration Technology of Multi-source Marine Resources

Marine data comes from a wide range of sources and has various means of acquisition, featuring multi-modal, heterogeneous and multi-scale characteristics. Therefore, it is necessary to study multidimensional data and knowledge fusion technology for specific marine application scenarios, so as to realize in-depth mining of various resources such as marine literature, scientific data and expert institutions, multi-source heterogeneous data fusion, and cross-disciplinary and cross-domain knowledge horizontal association. In order to realize the panoramic integration of multi-source marine resources, this paper carries out data cleaning and conversion of structured data, semi-structured data and unstructured data by making and revising metadata standards, database construction specifications, interface services and other series of standards and specifications, and forms standardized standard data. Then, through the multi-modal data association and fusion, multidimensional data integration, knowledge fusion and other technologies and methods, finally realize the panoramic integration of marine data, laying the foundation for marine data analysis and mining and knowledge services.

B. Construction technology of ocean thematic knowledge graph

There have been attempts in the application of knowledge graph research in the marine field, such as the construction of islands and reefs, and the construction of polar ontologies [14][15]. However, most research focuses on the construction and analysis of knowledge graphs from a bibliometric perspective [16]-[20], and there is a basic gap in the unified knowledge expression mode across multiple marine disciplines and the construction of thematic knowledge graphs to support management and application decision-making. Therefore, in order to build a map of marine domain knowledge, it is necessary to first build a professional knowledge representation model, and then integrate multimodal data such as ocean maps, texts, and numbers, breaking through the automatic construction and reasoning technology of knowledge maps, so as to build a large-scale thematic knowledge map for marine big data analysis and decision-making. In terms of specific technical routes, the construction of professional knowledge representation models usually includes top-down and bottom-up approaches, as well as a combination of the two. The professionalism of the marine field determines the need for expert knowledge participation in model construction, driven by business scenarios, and based on expert knowledge, a combination of top-down and bottom-up approaches is adopted to construct a knowledge model. Then, knowledge extraction technologies such as graph mapping/D2R, wrapper, supervised learning, and knowledge fusion technologies such as entity disambiguation, entity alignment, and knowledge combination [21]-[24] are used to build a knowledge map of multi-source fusion, multi-scale mapping, and multimodal representation. On this basis, we will carry out research on ocean big data decision support technologies such as multi-source information mixed knowledge reasoning [25], complex network analysis [26][27], and graph computing, achieving the updating and improvement of graphs, and ultimately providing intelligent decision support for ocean management applications such as marine ecology and island development and utilization.

VI. CONCLUSION

This paper analyzes and puts forward the content and organizational management mode of marine knowledge, focuses on the construction and practice of marine knowledge service system from the aspects of knowledge content and organization, platform construction and function, and operation effect, and discusses the key technical issues such as resource integration of marine literature data, scientific data, expert institutions, Internet data, and knowledge map construction. It can provide reference for the development of knowledge services in marine and related fields.

With the advent of the era of big data, it has become a trend in the field of marine data management, mining, analysis and service to quickly discover useful information from numerous data to intelligent and accurate decision-making services. Marine knowledge services are facing huge opportunities and challenges. In the next step, on the basis of continuously promoting the integration and sharing of marine data resources, the construction of marine knowledge service platform will strengthen the development of products for users' personalized needs, take knowledge service as the starting point, promote the transformation of data-information-knowledge-application, promote the realization of industry-university-research-use cooperation demonstration of marine knowledge service, and build the brand of marine knowledge service.

ACKNOWLEDGMENT

This research received founding from China Engineering Science and Technology Knowledge Center Project, grant number CKCEST-2022-1-4. All the data in this research is sourced from National Marine Data Center of China. The authors would like to thank the anonymous reviewers for their insightful comments and substantial help in improving this paper.

REFERENCES

[1] L. L. Song, "Construction and practice of marine knowledge service system," in DBKDA2024, The Sixteenth International Conference on Advance in Databases, Knowledge, and Data Applications, March 2024.

[2] Big Data White Paper (2020), China Academy of Information and Communications Technology, 2020.

[3] S. Liu, G. Chen, Y. J. Liu, F. L. Tian, "Research and Analysis on Marine Big Data applied technology," Periodical of Ocean University of China , vol. 50, pp. 154-164, 2019.

[4] G. R. Yu, H. L. He, Y. K. Zhou, "Ecosystem observation and research under background of big data," Bulletin of Chinese Academy of Sciences, vol. 33, pp. 832-837, 2018.

[5] R. D. Wang, M. X. Gao, L. Shi, C. Wang, B. Xu, "Research and thouthts on the opening and sharing of scientific data under background of big data," China Science and Technology Resources Review, vol. 52, pp. 1-5, 2020.

[6] State Council of the People's Republic of China. Action Outline to Promote the development of Big Data [online]. Available from: https://www.gov.cn/gongbao/content/2015/content_2929345.htm, 2015.08.31.

[7] Q. X. Xie, Z. J. Yong, F. He, Y. N. Lang, "Construction of the general framework of professional domain knowledge service system," Publication and Distribution Research, vol. 8, pp. 35-37, 2018.

[8] R. X. Zhao, J. Li, J. Zhang, D. D. Zhang, "Construction of multi-scenario agricultural knowledge service system," Journal of Agricultural Library and Information Technology, vol. 32, pp. 4-11, 2020.

[9] J. Chen, W. Z. Liu, H. Wu, Z. L. Li, Y. Zhao, L. Zhang, "Basic issues and research agenda of geospatial knowledge service," Geomatics and Information Science of Wuhan University, vol. 44, pp. 38-47, 2019.

[10] Epistemology. Stanford Encyclopedia of Philosophy[online]. Available from: https://plato.stanford.edu/entries/epistemology/, 2005.12.04.

[11] Q. He, C. Yu, W. Song, X. Y. Jiang, L. L. Song, J. Wang, " ISLKG:The construction of island knowledge graph and knowledge reasoning," Sustainability, vol. 15, 13189, 2023.

[12] Metadata Specification of China Engineering Science and Technology Knowledge Center (v2.0 version), Engineering Science and Technology Knowledge Center of Chinese Academy of Engineering, 2020.

[13] X. X. Zhang, "Where knowledge service is going-analysis of five knowledge service modes in press and publication industry," Publishing & Printing, pp. 1-5, 2019.

[14] Y. Zhang, F. Z. Su, Q. Wang, W. Z. Wu, "Island ontology construction and its application in spatial data organization," Geography and Geo-Information Science, vol. 33, pp. 52-58, 2017.

[15] D. M. Huang, Q. Zhang, J. Wang, Q. M. Wei, J. C. Shi, J. G. Zhu, "Research on the expression and the construction of polar science data ontology based on spatial relationship," Chinese Journal of Polar Research, vol. 30, pp. 77-87, 2018.

[16] S. D. Li, H. F. Zhang, M. Xu, "Research progress of marine spatial planning based on CiteSpace," Marine Development and Management, vol. 39, pp. 23-30, 2022.

[17] K. Yi, "Characteristics analysis of marine innovation research at home and abroad based on CiteSpace knowledge map," Journal of Liaoning Normal University (Natural Science Edition), vol. 44, pp. 81-90, 2021.

[18] F. Peng, N. N. Fu, W. Hu, J. L. Hu, "Analysis and enlightenment of knowledge map of marine resources research at home and abroad," Resources Science, vol. 42, pp. 2047-2061, 2020.

[19] Y. M. Wang, J. Y. Yuan, S. Q. Lin, "Analysis of marine ranching based on CiteSpace knowledge map visualization," Journal of Ocean University of China (Social Sciences Edition), pp. 42-55, 2020.

[20] T. Y. Bai, R. F. Yang, D. T. Hui, W. W. He, Y. G. Zhang Y, J. He, "Knowledge mapping analysis and outlooks of the digital ocean research," Bulletin of Surveying and Mapping, pp. 131-136, 2020.

[21] Z. M. Xiong, H. Y. Ma, S. Li, N. Zhang, "Application and Prospect Analysis of Knowledge Graph in Marine Field," Computer Engineering and Applications, vol. 58, pp. 15-33, 2022.

[22] L. J. Fu, Y. Cao, Y. Bai, J. W. Leng, "Development status and prospect of vertical domain knowledge graph in china," Application Research of Computers, vol. 38, pp. 3201-3214, 2021.

[23] Y. C. Liu, H. Y. Li, "Survey on domain knowledge graph research," Application of Computer Systems, vol. 29, pp. 1-12, 2020.

[24] Z. G. Ma, R. Y. Ni, K . H. Yu, "Recent advances, key techniques and challenges of knowledge graph," Chinese Journal of Engineering, vol. 42, pp. 1254-1266, 2020.

[25] L. Tian, J. C. Zhang, J. H. Zhang, W. T. Zhou, X. Zhou, "A Survey of Knowledge Graph-representation, construction, reasoning and knowledge hypergraph theory," Journal of Computer Applications, vol. 41, pp. 2161-2186, 2021.

[26] C. J. Liu, S. Y. Li, H. L. Hu, S. Fang, "Figure Database Research and Application Progress in the Analysis of Complex Networks," Data analysis and knowledge discovery, pp. 1-15, 2022.

[27] W. Yu, M. L. Zhui, H. Y. Lin, N. Yu, X. M. Li, X. P. Yang, "Empirical Research on Event Detection Based on Dynamic Complex Network Analysis," Journal of Information Technology, vol.40, pp. 108-114, 2021.