

# Marine Metadata Lifecycle Modeling and Its Applications

Wang Yan<sup>1,2</sup> Le Jiajin<sup>3</sup> Huang Dongmei<sup>2</sup> Zhang Yun<sup>2</sup>

<sup>1,2</sup> Glorious Sun School of Business and Management Donghua University

<sup>3</sup> College of Information Technology Shanghai Ocean University

<sup>4</sup> School of Computer Science and Technology Donghua University  
Shanghai



**ABSTRACT:** *Marine metadata and metadata standards are important for effective marine massive data management and Heterogeneous data exchange. This paper clarifies the dependencies and scopes of marine metadata in various lifecycle stages and proposes a marine metadata standard framework under lifecycle concept. Then a marine metadata lifecycle management model is given for effective marine metadata processing. In this way, standards and management system' efficiency is improved for lifecycle framework and model's timeliness, maneuverability and relative accuracy. At last, applications of the metadata standard framework and management model in lifecycle are implemented.*

**Keywords:** Marine Metadata, Lifecycle, Standard Framework, Management Model

**Received:** 12 November 2015, Revised 11 December 2015, Accepted 19 December 2015

© 2016 DLINE. All Rights Reserved

## 1. Introduction

Metadata concept was proposed in 1960 by Jack Myers [1], and referred to some ability and mode for effectively describing the data. Metadata plays an important role in information processing and knowledge analyzing, classifying and managing. Metadata is usually defined as the data structure of the data that can support the users' information consumption and the data management system.

Marine data is multidimensional, multisource, and great capacity data. The characteristic differences in different regions and fields result in the heterogeneity of marine data. So data management mode is a combination of overall management and hierarchical management, and storage mode is a combination of centralized storage and distributed storage.

Marine metadata is the main tool to resolve the efficacy problem of marine data management. With the marine metadata, management, use and maintenance of marine resources become more effective. Marine metadata and metadata standards can optimize the efficiency of information exchange in the different fields, the efficiency of data access on heterogeneous multi-platform and form a unified data exchange format.

Currently the mainstream international marine metadata standards are MEDI[2] proposed by International Oceanographic Commission (IOC), ODAS[3] presented by Marine Climatology Group (CMM), EDIOS proposed by EuroGOOS (European Global Ocean Observing System)[4] and ARGO floats metadata standards [5]. All these standards target at the different marine applications, such as ODAS mainly serve climate research applications, EDIOS is used for real-time observation data of oceanographic and ARGO is for the descriptions of buoy objects.

This article designs marine metadata standard framework through analyzing metadata features and its lifecycle feature, and clarifies the dependencies and scopes of the various marine metadata in lifecycle. Then a marine metadata lifecycle process and function management model of marine data is put forward in different stages. Finally, marine applications of metadata standard framework and management model in lifecycle are implemented.

## 2. Marine metadata standard framework in lifecycle

Information's lifecycle refers to the process of the information's value change process along with time. Information data are usually the most frequently used and of the highest value at the beginning when they are generated. And then these data's quantity increases unceasingly, the accesses to the data are less frequent, and their values also reduced gradually over time [9]. Lifecycle metadata model has been researched in [10,11,12,13,14], which the paper [12] focused on the creation of GIS system metadata information lifecycle at different stages. A lifecycle model of metadata has been presented in [13] which reveals the metadata working procedure. According to the concept of lifecycle metadata system evolution procedure can be divided into four roughly groups: requirement assessment and content analysis, system requirement specification define, metadata system design and realize, metadata system service and evaluation.

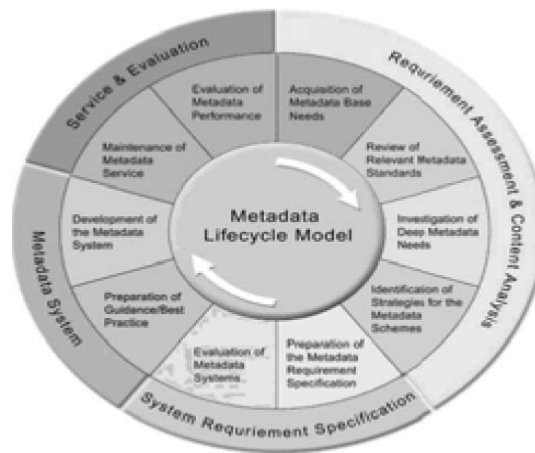


Figure 1. General Metadata lifecycle model [13]

For marine data, there is also a lifecycle concept of information processing. The lifecycle of marine metadata is actually longer and more widely than marine data, because marine metadata are generated before marine data, and extended to the whole marine data process of usage and the management.

Marine metadata standard framework is a set of rules and methods for developing metadata standard and metadata system, it is the abstraction of metadata that define metadata's function, data structure, format, semantic method, grammar rules and other aspects on a higher level.

Figure2 shows function relationships between different standard framework layers in lifecycle.

The metadata standard framework and the application field apply to data requirement collection stage through the data collection requirement and the usage requirements standard defines. The implementation layer is the layer of data collection and standard implementation, metadata of Digital Ocean, acting on the metadata creation, protection and access phase, and resulting in marine metadata records at last.

And in the application layer, through the user interface, users access data warehouse according to the metadata records. The management layer applies to metadata's whole lifecycle, from generation to maintenance.

Design of marine metadata standard framework is to build the constraint mechanism during all stages of the lifecycle and adapt these constraints changes. So development of marine metadata standard framework should take into account four aspects' constraints: engineering constraints, platform constraints, usage demands and input object characteristic constraints, as shown in Figure 3.

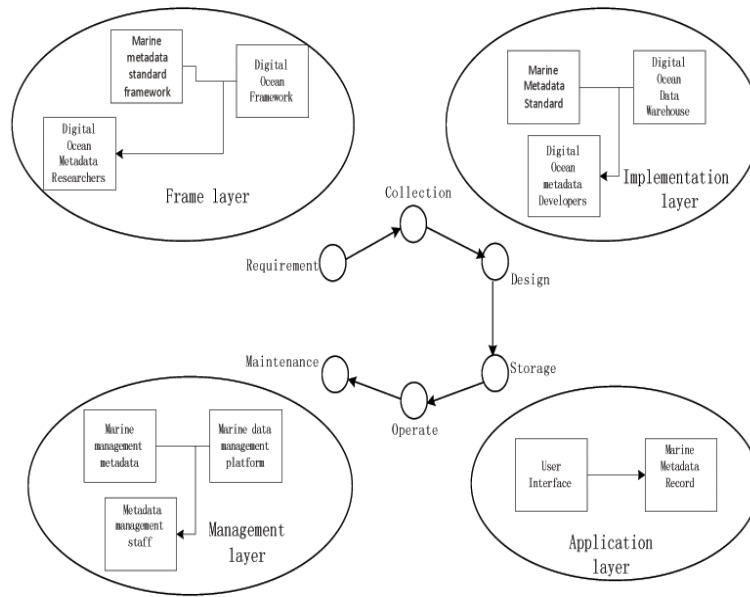


Figure 2. Relationship of marine metadata level in lifecycle

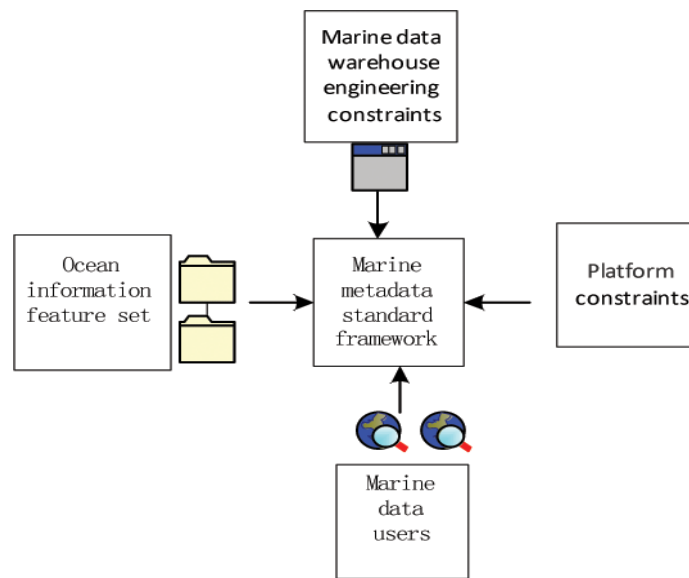


Figure 3. Marine metadata standard framework constraints

Engineering constraints of marine database include the ways of gathering marine information, the process management of database development and constraints of project and users. During the usage, it is necessary to import the user’s features as the use constraint. At the same time, the static environment and the static object constraints are integrated based on software and hardware, such as computing power , storage capacity and many other environment factors .So all these statics and dynamics constraint set are the important framework components.

The concrete framework design need to take account into the concise and accurate principle, the specificity and maximum matching principle, the interoperability and scalability when data access, customer value constraint and implement environmental constraints during data migration and induction process. All these considerations are also fractionalized, analyzed and implemented step by step during metadata lifecycle. As shown in Figure 4, the different principles in the different stages actually reflect the abstraction layer, structure layer and implement layer’s constraints.

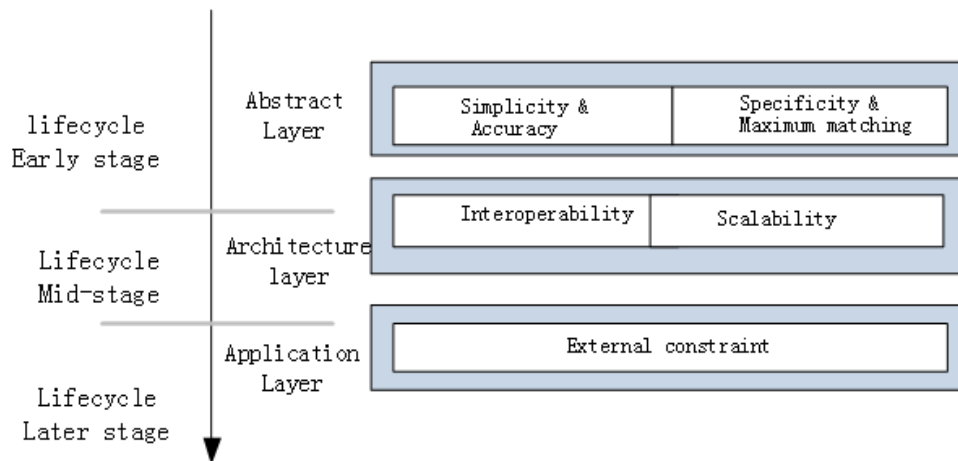


Figure 4. Hierarchy level of metadata standard design principle

### 3. Marine Metadata Model In Lifecycle

#### 3.1 Marine Metadata and Data Lifecycle

Marine metadata play different roles in different lifecycle sectors. Relationships and functions between lifecycle's every process stage and Digital Ocean construction in the whole are shown in figure 5.

In the stage of data collection and processing, descriptive metadata is created which is responsible for the expression and retrieval of information and their standards. In the stage of data collection and processing, descriptive metadata is created which is responsible for the expression and retrieval of information and their standards.

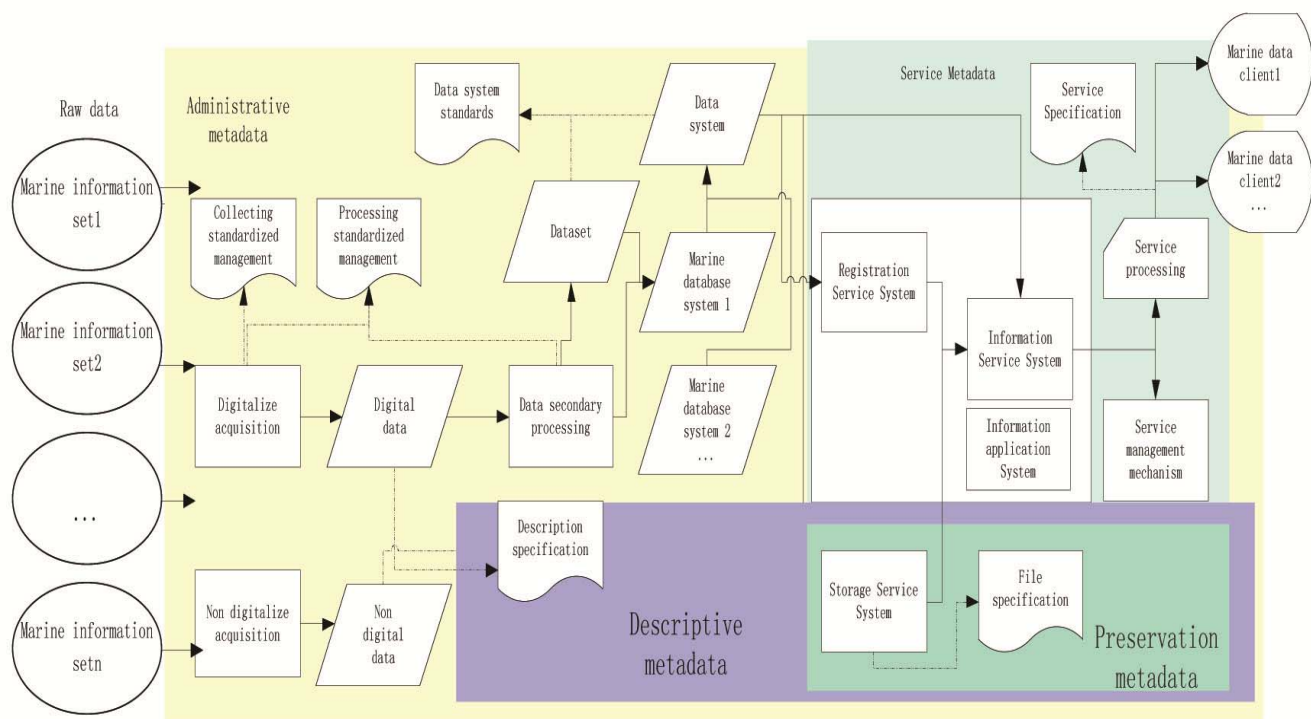


Figure 5. Marine metadata function scope during information lifecycle

Meanwhile descriptive metadata will also extend to the service applications and maintenance phases in the lifecycle. For example, any description and expression's update must comply with the relevant description specifications during the maintenance, update and the quality assessment process. But administrative metadata exists throughout the whole lifecycle of the marine information resources, which is the longest metadata. By specification constraint, administrative metadata functions included data relationship management, data source management, data validation management, data maintenance management and data quality management. Preservation metadata is actually an extensive metadata management that exists for a long time in the lifecycle to save the specific resource object information. Service metadata is for digital resources data services which describe application process, application interface, application permissions and other aspects related information. Another function of Service metadata is to help data exchange between different data system.

### 3.2 Marine Metadata Management Model in Information Lifecycle

In the lifecycle of marine information resources, the different types of metadata are interrelated, interacted and mutual restraint. Although they are different in the location, the shape, the realization of the main functions, the objects, and application details in the lifecycle, their generations are all based on metadata standards, and their ultimate goals are the same that they realize share, exchange and interoperability of marine data resources.

Metadata standards' time model is related to information resources, requirement, acquisition, design, storage, operation, maintenance and other phase. In different stages, the metadata need to comply with the different semantic rules and grammar rules.

Different types metadata associated with different lifecycle stages of marine data can be divided into description metadata, administrative metadata and service metadata. As shown in Figure 6, these metadata are extended in different stages of the lifecycle which reflect different applications and objects' requirements.

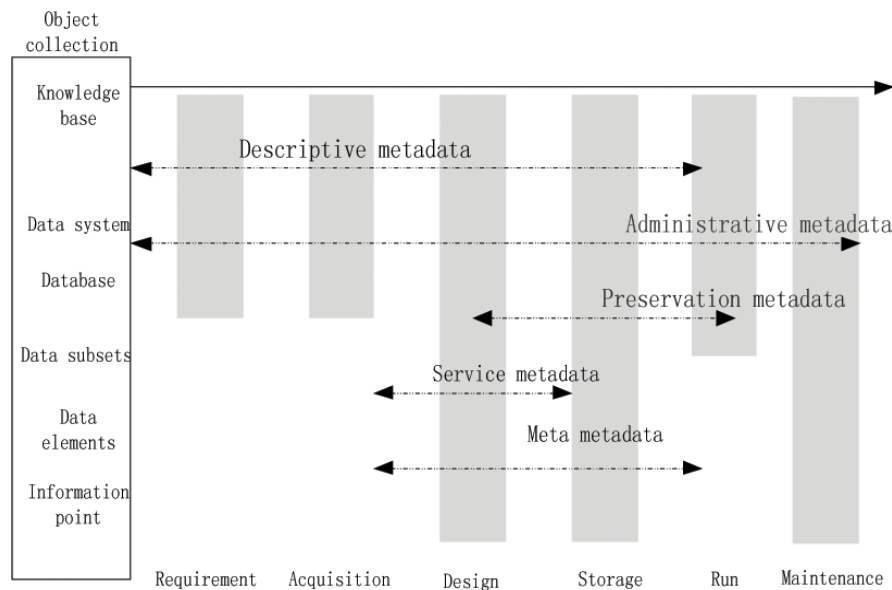


Figure 6. Marine metadata model based on lifecycle

#### 3.2.1 Meta Metadata

As a special kind of metadata, Meta metadata which associate with metadata serves for the management of other metadata. Meta metadata express the metadata through markup language, format language, identifier sign, extension mechanism, conversion mechanism and other information. The UML model and data dictionary of marine metadata description and associated constraints (required or optional) are Meta metadata.

#### 3.2.2 Descriptive Metadata

Descriptive metadata is mainly used to express and identify the contents of the object described, as well as the additional responsibility information during the collection of data. Descriptive metadata of digital marine information is based on the following five levels, as shown in Figure 7.

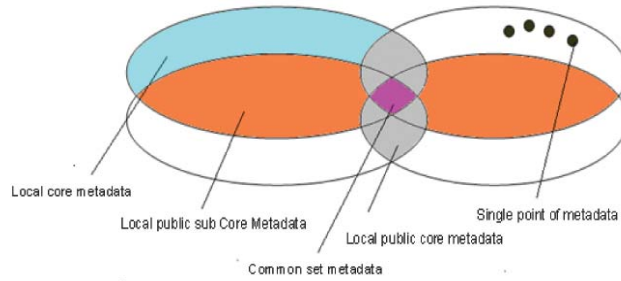


Figure 7. Marine descriptive metadata basic levels

In the figure, the common set metadata is used in all types of resource object which includes most of metadata in ISO 19115, such as identification information, content information, distribute information.

Local public core metadata is based on the digital objects features of marine information, which refer to other metadata standards and is applied to all objects related to the digital ocean information database generally. Local public sub-core metadata can support exchange and communication between different data platforms, so it is defined separately.

### 3.2.3 Administrative Metadata

Administrative metadata is used to manage metadata, whose functions mainly include:

#### 1) Data Relationship Management

Data relationship management is to manage the content object information and the relationships with relevant object components. Management scope includes the basis for creating content and its relationship with other relevant resource objects. For example, in the marine biological data processing, data relationship management involved is shown in figure 8.

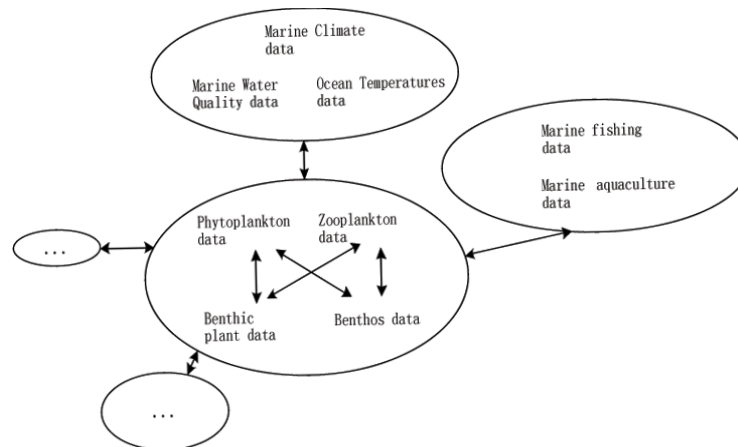


Figure 8. Marine biological data relationships

#### 2) Data Validation Management

Data Validation Management provides authentication mechanisms and criteria for marine data access, identify and mark the verification process when the information is recorded and modified.

#### 3) Data maintenance Management

Data maintenance management serves for data management and process define about maintenance cycle, maintenance standards, maintenance notification, and definitions of information data lifecycle and its various lifecycle maintenance activities criteria.

#### 4) Data quality Management

By the evaluation mechanism and application statistics of Administrative metadata, users or experts can assess quality the

content of existing metadata element items.

### 3.2.4 Preservation Metadata

Preservation metadata need to be created in the long-term preservation process of information resources to save specific information of resource object. For the data any records in the management and service process, need to comply with the description metadata, but whether to record, how to record, also belong to the category of preservation metadata. Preservation metadata is required to follow a set of standards to ensure that the user information resources' efficient access and complete reading.

### 3.2.5 Service Metadata

Service metadata reveals and performs digital marine resources services through describing the application process that includes user interface, application interface specifications and application permissions achieved during the marine information resources exchange. Graphic UI and application mode information in the application interface belong to the service metadata.

Marine metadata standard framework and management model in lifecycle provide system frameworks of marine data to help users identify, quantify and evaluate marine data in an orderly manner.

## 4. Applications

In multi-source heterogeneous data management platform construction of Digital Ocean, a marine information metadata standard is designed with its metadata function object set shown in Figure 9 which includes many metadata subsets (UML package).

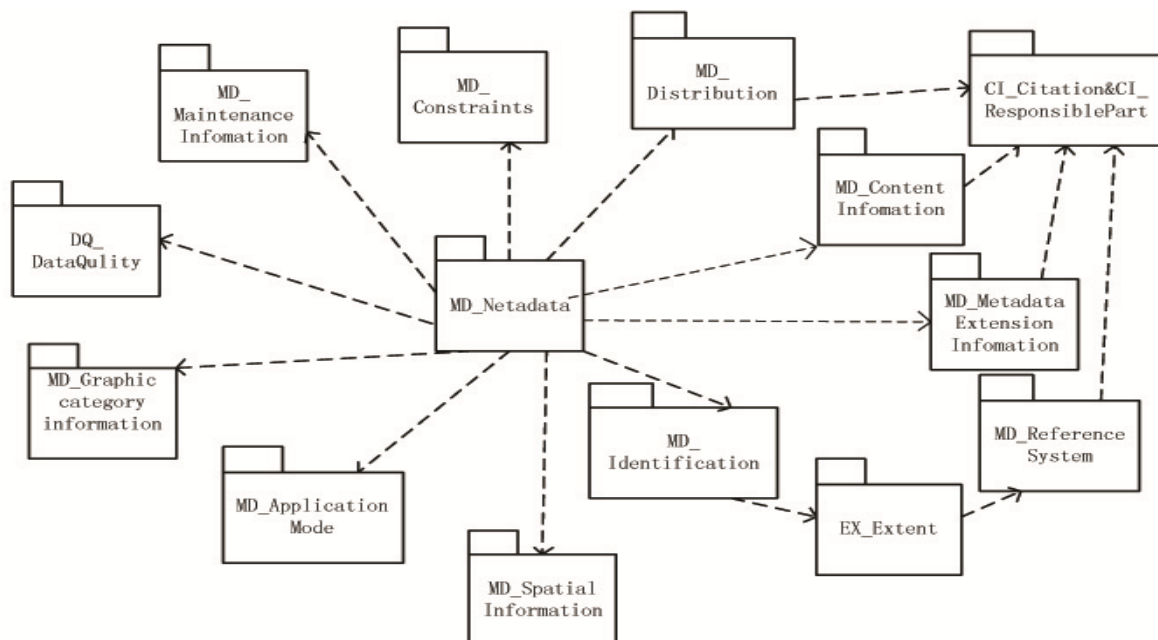


Figure 9. Marine metadata package structure model diagram

MD\_Identification information contains uniquely identifies information of data, such as resource references and summaries information of data sets, MD\_Content information has description information of the data content features, MD\_Distribution contains information resources distributing data and the information of how to obtain the resources, DQ\_DataQuality consists of the data log information and/or data quality reports, MD\_Reference system contains spatial and temporal reference of the instructions, MD\_ extent information includes user-defined extension information, MD\_Maintenance information contains the scope and frequency information of updating data, MD\_Aapplication mode includes the application mode information of the data set, MD\_Graphic category information contains portrayal directory information, MD\_Spatial information includes mechanism information which represents the dataset spatial information,



Metadata type based on the lifecycle	Metadata subtype	Marine Metadata Name	Definitions	Constraints / Conditions	Maximum number of occurrences	Data Types
Meta Metadata	-	UML model	Tool of describing marine metadata	M	N	UML diagrams
		Data dictionary	System data of describing marine metadata	M	N	Free text
Descriptive Metadata	Common set metadata	MD_ Identification	Unique identification for basic information of resources	C	Maximum occurrences of reference object used	Aggregation class
		MD_ ContentInfo	Providing data content feature of descriptive information	C	Maximum occurrences of reference object used	Aggregation class
		MD_ Distribution	Providing information of distribution format, distributor and distribution methods for required data sets	C	Maximum occurrences of reference object used	Aggregation class
	Local public core metadata	EX_GraphicExtent	Providing information of geographical coverage data	C	N	Aggregation class
		EX_TemporalExtent	Providing information of time period for data set contents	C	N	Aggregation class
	Local public sub core metadata	EX_DepExtent	Providing information of water depth of data set	C	N	Aggregation class
Administrative Metadata	Data Relationship Management	MD_ ReferenceSystem	Providing information about relevant reference	C	Maximum occurrences of reference object used	Class/string
	Data Validation Management	MD_ Constraints	Restriction information metadata or data sets, including access restrictions and use restrictions	C	Maximum occurrences of reference object used	Class/string
	Data maintenance management	MD_ MaintenanceInfo	Update information of metadata, or data set content	C	Maximum occurrences of reference object used	Class/string
	Data quality management	DQ_ DataQuality	Data quality information in pre-defined data quality range	C	Maximum occurrences of reference object used	Class/string
Preservation Metadata		CI_Citation	Standard reference information	C	Maximum occurrences of reference object used	string
		CI_ ResponsiblePart	Identification and Contact information of responsible and relevant department	C	Maximum occurrences of reference object used	string
Service Metadata		MD_GraphicExpression	File name and file type of graphic data set	O	1	string
		MD_ApplicationMode	Application mode information used to create data set	O	1	Class/string
		MD_MetadataExtensionInfo	User defined Description extended information	C	Maximum occurrences of reference object used	Class/string

Table 1. Marine metadata type table based on the lifecycle



and EX\_Extent describes spatial and temporal extent of relevant entities. CI\_Citation and CI\_ResponsiblePart provide standard methods of references and information index, as well as part information of reference materials.

The UML model, data dictionary and associated constraints of ocean data heterogeneous platform belong to meta metadata. In marine data heterogeneous platform, geographic extent information and temporal extent information belong to the local public child Core metadata, but depth extent information is extensible, so it belongs to the local public child Core metadata. Graphic expression and application mode information is used to describe application interface which belongs to the service metadata. References and responsible unit information describe reference name, date, version, name, and address information, which belongs to preservation metadata, as shown in table 1 in detail.

In Table 1, constraints/conditions indicates metadata element's properties, including mandatory (M), optional (O) and conditions required (C). Maximum number of occurrences indicates that maximum number of instances metadata entity or metadata element.

Take example for the identification information in common set metadata, figure 10 illustrates it plays different roles in the different stages during the information lifecycle in the implementation metadata process. In the requirements phase, it identifies customer browse icon recognition and identification information scope definition description, in the design stage. It identifies the observation information, data and information platform, and gradually clarifies design description. It uses the code table to identify the storage information in the storage stage. In the maintenance phase, it identifies the maintenance information.

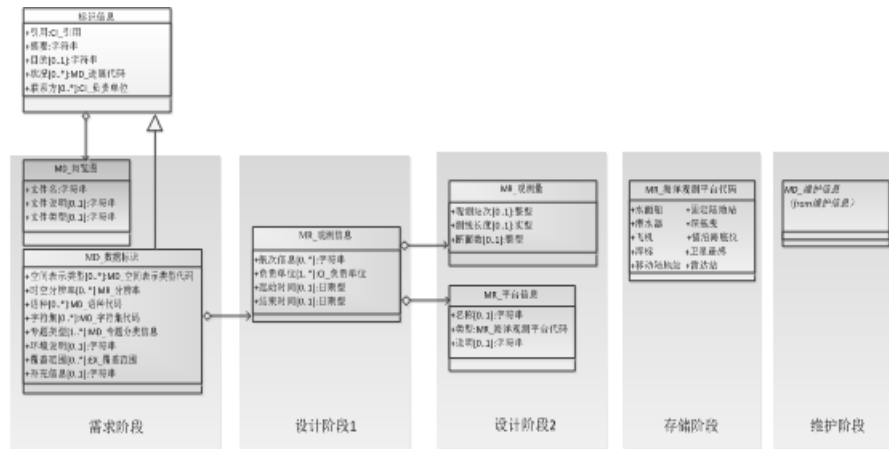


Figure 10. Class diagram of the marine metadata in different lifecycle stages

B/S (browser/server) structure is used in marine metadata management system of integrated management platform. Marine metadata model in lifecycle is used on the server, which tracks the processing chain of information and achieves the management function of information collection, design, operation, and maintenance.

## 5. Conclusions

Metadata can effectively store, manage and make use of data, and it plays a vital role in the management field of marine data and is an important tool for management, utilization and maintenance of the marine information resources. This paper analyzes the current situation of metadata in the field of marine data management research, and then presents a standardized marine framework in lifecycle. And basing on this framework, a marine metadata management model in lifecycle is given to define metadata function scope in different lifecycle stages. By describing metadata's divisions and scopes, metadata management model is used for time domain management and definition of metadata, solves problems of fuzzy management and cross periodic boundary functional attributes defining, and greatly saves development and maintenance cost.

## Acknowledgement

This work is supported by the National Natural Science Foundation of China (Grant No. 41376178), and the Shanghai Science and Technology Committee (Project NO. 11510501300).

## References

- [1] Hatua, S R. (2001). *Metadata: The Foundation Of Content Description*.
- [2] Greg Reed. (2003). Marine Environmental Data Inventory (MEDI), Seventeenth Session of the IOC Committee on International Oceanographic Data and Information Exchange (IODE), Paris, France, 3-7 March.
- [3] Kent, E C., Woodruff, S D., Berry, D I. (2007). Metadata from WMO publication no. 47 and an assessment of voluntary observing ship observation heights in ICOADS[J]. *Journal of Atmospheric and Oceanic Technology*, 24 (2), 214-234.
- [4] Verduin, J., Fischer, J. (2003). EDIOS: European directory of the initial ocean observing system. *Elsevier Oceanography Series*, 69, 265-271.
- [5] Carval, T., Keeley, B., Takatsuki, Y., et al. (2006). Argo data management User's manual. Version, 2: 60.
- [6] Li Xuerong., Li Sha. (2007). Ocean color remote sensing metadata and system design . *Journal of Tropical Oceanography*, 2007, 26 (1): 81-86.
- [7] Shi Lijian., Wang Qimao., Guo Maohua. (2010). Ocean Remote Sensing Metadata and Data System Design. *Remote Sensing Information*, 2010, 3: 117-121.
- [8] Jiang, Y., Lu, L., Guo, Z. (2009). Design of Marine Information Metadata and Directory Service System Based on XML /*Database Technology and Applications*, 2009 First International Workshop on. IEEE, 574-577.
- [9] Somasundaram, G., Shrivastava, A. (2009). Information storage and management [M]. *Wiley Publishing, Inc.* 25-26.
- [10] Barton, J., Robertson, R J. (2005). Developing a metadata lifecycle model [M]/Context: Nature, Impact, and Role. *Springer Berlin Heidelberg*, 249-250.
- [11] Belhajjame, K., Wolstencroft, K., Corcho, O, et al. (2008). Metadata management in the taverna workflow system/ Cluster Computing and the Grid, 2008. CCGRID'08. 8<sup>th</sup> IEEE International Symposium on. IEEE, 651-656.
- [12] Limbach, T., Krawczyk, A., Surowiec, G (2004). Metadata lifecycle management with GIS context *In: Proceedings of the 10<sup>th</sup> EC GI & GIS Workshop, ESDI State of the Art*, Warsaw, Poland.
- [13] Chen, Y., Chen, S. (2005). Metadata lifecycle model and metadata interoperability /5th International Conference on Conception of Library and Information Science.
- [14] Catteau O, Vidal P, Broisin J. (2006). A generic representation allowing for expression of learning object and metadata lifecycle [C]/Advanced Learning Technologies, 2006. Sixth International Conference on. IEEE, 30-32.