Marine Ecological Knowledge Management System Based on Ontology Repository

Hongyan Yun[1], Jianliang Xu[2], Moji Wei[3] and Zhenbo Guo[4]

[1,4]*College of Information Engineering, Qingdao University, Qingdao, China, 266071*

[2]*Department of Computer Science, Ocean University of China, Qingdao, China, 266100*

[3]*Information Research Institute, Shandong Academy of Sciences, Jina, China, 250014*

[1]*yunhy2001@163.com,* [2]*xjl9898@gmail.com,* [3]*weimoji@126.com,* [4]*gzb@qdu.edu.cn*

Abstract

This paper researches semantic retrieval and ontology rule reasoning for marine ecological knowledge management system. Upper ontology plays a critical role in ontology development by giving developers a guideline to view the target domain. By fixing a viewpoint of device-function, we present a domain upper ontology for marine ecosystem, and then construct marine ecology ontology and ontology repository. Based on constructed marine ecology ontology repository, authors design and develop marine ecological knowledge management system, which implements main functions include marine ecological knowledge navigation, term query, knowledge retrieval and ecological crisis early warning. Marine ecology ontology application system provides users a semantic support platform for marine scientific research and cooperation, and it also verifies the rationality and feasibility of ontology modeling method and validity of constructed marine ecology ontology.

Keywords: *Ontology Repository, Marine Ecological Knowledge Management System, Semantic Retrieval, Ontology Reasoning, SWRL Rule*

1. Introduction

The 21st century is said to be the century of ocean. Ocean is an open and diverse system, which determines marine science is multidisciplinary and interdisciplinary. Ocean is also a very complex ecosystem which plays an important role in the Earth's environment. After many years' research, many valuable documents were left, such as observational data, experimental results, and papers as well. However, due to diversity of data sources and representation, there are lots of problems in the utilization of material, such as inconsistency in the format of data and inconsistency in the methods of collection and statistics, one thing has different terminologies to express, one term expresses different things, and so on. How to integrate and use these heterogeneous resources are crucial for scientific research and crisis management, government, aquaculture and other marine agencies.

Ontology is a formal, explicit specification of a shared conceptualization. Domain ontology is used to capture, describe and express the knowledge of related fields, and provide a common understanding of knowledge in this area, and determine a common recognition of the vocabulary in the area [1]. Thus, ontology can solve problems of semantic heterogeneity and knowledge sharing and reuse issues, so we can share data sharing among different agencies.

Marine ecosystem is a complex system which is composed of marine biological communities and their habitat environment [2]. Upper ontology plays a critical role in ontology development by providing developers a guideline to view the target domain [3]. Referring to engineering field literature and Kitamura's device-function knowledge representation theory that the "function" concept is used to describe marine ecology functional process [4]; fixing a viewpoint of device-function that a domain upper ontology is presented, and then marine ecology ontology repository is constructed.

The aim of constructing domain ontology repository is to share, reuse and reasoning. Based on constructed ontology repository, authors research semantic retrieval and ontology rule reasoning, and develop a Marine Ecological Knowledge Management System, which implements main functions such as marine ecological knowledge navigation, term query, function query, semantic query answering and marine ecological early crisis warning, *etc*.

2. Ontology Repository and Knowledge Management

Knowledge management deals with domain knowledge acquisition, accessing and maintenance. Ontology provides common concepts, properties, semantic relations between concepts to represent domain. Ontology repository usually uses these ontology concepts to represent domain's facts. For example, constructed marine ecology ontology includes formal definition of these concepts such as "marine organism", "plankton", "red tide plankton", "dinoflagellate" and "sea water" *etc*. Marine ecology ontology can't give crisis warning about "Red tide occurred in Qingdao Sea", while repository can represent some facts knowledge by reasoning. Ontology focuses on describing domain's concepts and terms, while repository usually focuses on representing domain's facts knowledge and specific application.

When representing ontology and repository in a same language, there is no distinct limit, the difference between ontology and repository is which part of constructed repository can be shared and reused, and which part focuses on specific application. Marine ecology knowledge repository includes Marine Ecology Ontology, SWRL (Semantic Web Rule Language) [5] Rule sets, and Marine Ecological Metadata. The architecture of marine ecology knowledge repository is shown in Figure 1.
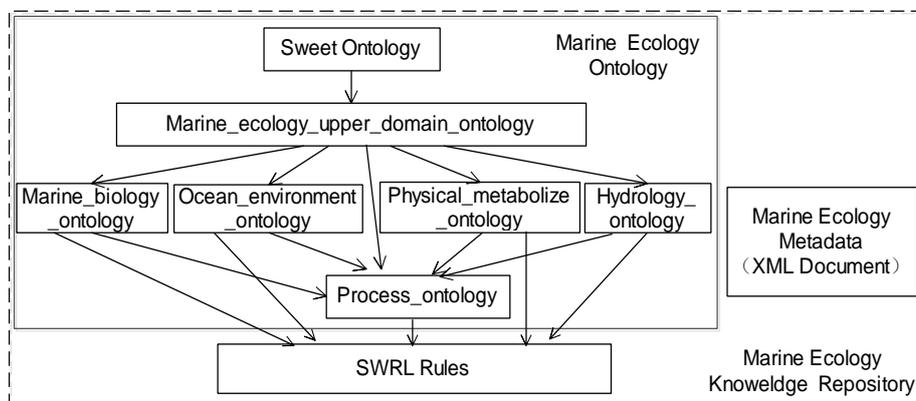

Figure 1. Marine Ecology Knowledge Repository

The flow of energy and the cycling of nutrients are two basic processes in marine ecosystem. These two major forces link various biological factors, non-biological factors, and interaction among these factors. Referring to engineering field literature that "function" concept can be used to describe these processes. By referring and expanding

functional knowledge representation frame of Kitamura "device centered" viewpoint [6], we look upon ocean and substance from "device function" viewpoint. In marine ecosystem, marine functions and processes are looked upon as functions, and physical entities (including marine lives and ocean itself) are all treated as devices. Physical entities (namely device) are the material base of marine function and process, so we fix a "device function" viewpoint to construct marine ecology ontology. Marine ecology ontology is composed by 5 levels ontologies as shown in Figure 1. The first level ontology is top ontology; SWEET (Semantic Web for Earth and Environmental Terminology) [7] ontology is adopted as top ontology, which represents fundamental general concepts. Authors use SWEET ontologies' components and expand SWEET ontology to construct marine ecology ontology by defining the basic concepts (*e.g*., time, space, material, energy *etc.,*) and relations among them.

The second level ontology is marine_ecology_upper_domain_ontology which represents domain upper knowledge concepts. That is to confirm "Device function" viewpoint and then to construct domain upper ontology for describing basic principle, important concepts and relations among concepts in marine ecology domain.

The third level ontologies are marine ecology knowledge ontologies, which represent domain knowledge concepts. Because semantic concept granularity of domain upper ontology is too large to perfectly describe marine ecological knowledge, we should construct domain knowledge ontology to describe complete domain knowledge hierarchy. Domain knowledge ontology includes marine_biology, ocean_environment_ontology, physical_metabolize_ material and hydrology_ontology.

The forth level ontology is marine_ecology_process_ontology, which describes domain function and process; that is to construct marine ecology function_process ontology (mainly includes marine ecological system production, consumption, decomposition function, cycle of nutritive materials, and biogeochemical cycle *etc*.,).

The fifth level Rules represent domain general knowledge. Based on constructed domain knowledge ontology, Rules are constructed for ontology reasoning. Authors use SWRL to create rules for describing implicit relations among concepts. This research applies OWL ontology and SWRL rules to implement semantic reasoning for marine ecological crisis early warning.

Knowledge management system provides knowledge acquisition, storage, retrieval, sharing and maintenance functions; its goal is to provide effective knowledge integration, sharing, and coordinate application environment. Namely, the goal of knowledge management system is to share and reuse of knowledge.

Knowledge management system based on knowledge repository has higher reliability and reusability than other knowledge management system; it also has enhanced ability of knowledge management and knowledge acquisition.

3. Architecture and Functional Designing of Marine Ecology Knowledge Management System

Based on constructed marine ecology ontology repository, authors design Marine Ecological Knowledge Management System, which provides an access platform of marine ecological knowledge. This system provides navigation output, knowledge query, semantic query answering and crisis warning functions for users. It adopts layered structure design; system architecture is shown in Figure 2, which includes data level, abstract level and control output level.

Data level is marine ecology knowledge repository, it mainly consists of the following several parts: ① Marine Ecological Metadata. It contains metadata and related information; ② Marine Ecology Ontology. It is a formal definition of marine domain conceptualization model, which includes concepts, properties, relations between concepts,

and instances *etc.*, ③SWRL Rule Sets. It includes semantic reasoning rules and marine ecology crisis conditions. It associates with ontology knowledge repository.

Abstract level mainly stores metadata information and ontology information. By analyzing OWL(Web Ontology Language) [8] ontology to acquire ontology information and store corresponding data structure on abstraction level; by dealing with metadata to acquire metadata information and store corresponding data structure on abstract level.
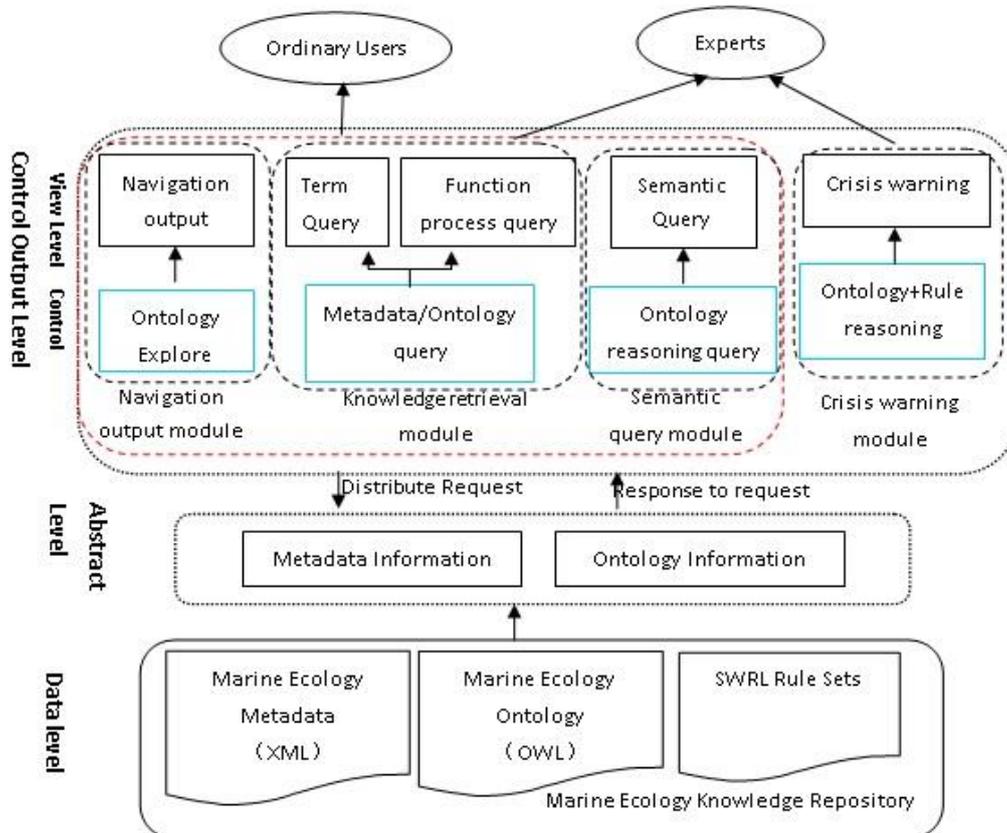
Figure 2. Architecture of Marine Ecological Knowledge Management System

Control output level deals with query and reasoning. It acquires ontology semantic parsing results on controlling output level and implements navigation output, word processing, SPARQ[9] query sentences, Pellet reasoning query, *etc.*, Query part includes term and semantic query answering; it provides term retrieval including hyponym relation, properties, and instances. According to constructed Ontology and SWRL rule sets, by analyzing user submitted data, reasoning part can make an early warning of marine ecology crisis.

4. Function Implementation of Marine Ecological Knowledge Management System

Marine Ecological Knowledge Management System provides a semantic supporting platform for marine ecological service; it implements main functions including Marine ecological knowledge navigation panel, term query, function query, semantic query answering and marine ecological crisis warning, *etc.*

Function implementation adopts a hierarchical thinking; it adopts Jena API to parse semantic OWL ontology and acquire ontology information and store corresponding data structure on abstraction level; it acquires ontology semantic parsing results on controlling

output level, and implements navigation tree output, word processing, SPARQ query sentences, Pellet Reasoner reasoning query, *etc.* This system uses Java as its developing language, and uses Tomcat 6.0 as a backend server, and uses JSP (Java Server Pages) as Web page scripting language, and uses Jena 2.6.3 and embeds reasoning machine Pellet 2.2.2 for OWL ontology parsing and reasoning, and uses DOM4J to parse XML file.

Main implementation methods include: First load ontology model into memory buffer, and then use Jena API to parse ontology. Concrete parsing steps are following: call Model Factory() to create an Ontology Model, and read data from named OWL file into model, and then read model into memory; call ontology interactive method in Jena API, and read class, property and instance of ontology, and then store them into relative data structures for subsequent querying and reasoning function call.

4.1 Navigation Panel of Marine Ecological Knowledge Management System

Navigation Panel provides a visual interface for users to access relevant information about marine ecosystem ontology. It uses Jena API to process ontology information. Main methods of calling Jena are shown in Table 1. Flow-process diagram of navigation module is shown in Figure 3.

Table 1. Main Methods of Calling Jena

| com.hp.hpl.jena.ontology	
OntModel interface	Concrete operate to Ontology model. Include operates such as create and store ontology model, define and query class and property.
OntClass interface	Include relative concrete operates to class, by calling listClasses() in OntModel, return Iterator which composed of all concepts in ontology model, and then call methods in OntClass to operate. Such as acquire, define and delete class; acquire property; acquire, define and delete instance of class, *etc.*
OntProperty interface	Include concrete operate for DatatypeProperty and ObjectProperty, such as acquire domain and range of property, set and judge type of property, *etc.*
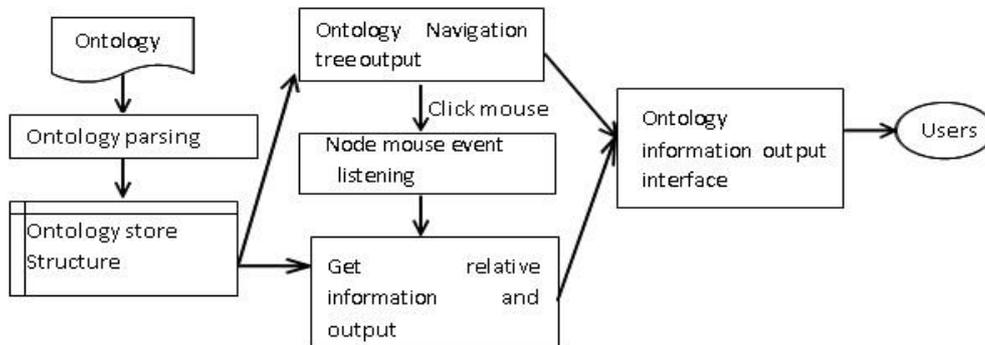
Figure 3. Flow-process Diagram of Navigation Module

We get concept hierarchy relations after Jena API parsing OWL Ontology file. Using concept tree generation algorithm, we can represent ontology concept with navigation tree structure. We adopt development toolkit DTREE (which is provided by Destroydrop) to generate navigation tree. System assigns id for every parsed concept node, and determines its super node's id according to is-a relation, and then forms concept navigation tree. This process includes two methods: ① add (id, pid, name, url, title, target, icon, iconOpen,

open) is for adding tree node; ②openAll( )/closeAll( ) is for spreading/drawing back all nodes. Authors add mouse listening event on concept tree node.

When user click on node of navigation tree, system gets clicking node's id, and transfers the value of id to ontology concept showing interface, and shows relative node information(include super class, subclass, attributes and instances) from ontology information store structure. Navigation panel represents hierarchical relationships in a tree structure among marine ecological terms. It dynamically displays node information according to users' selection. For example, when user access term "pyrrophyta", it shows in Figure 4.
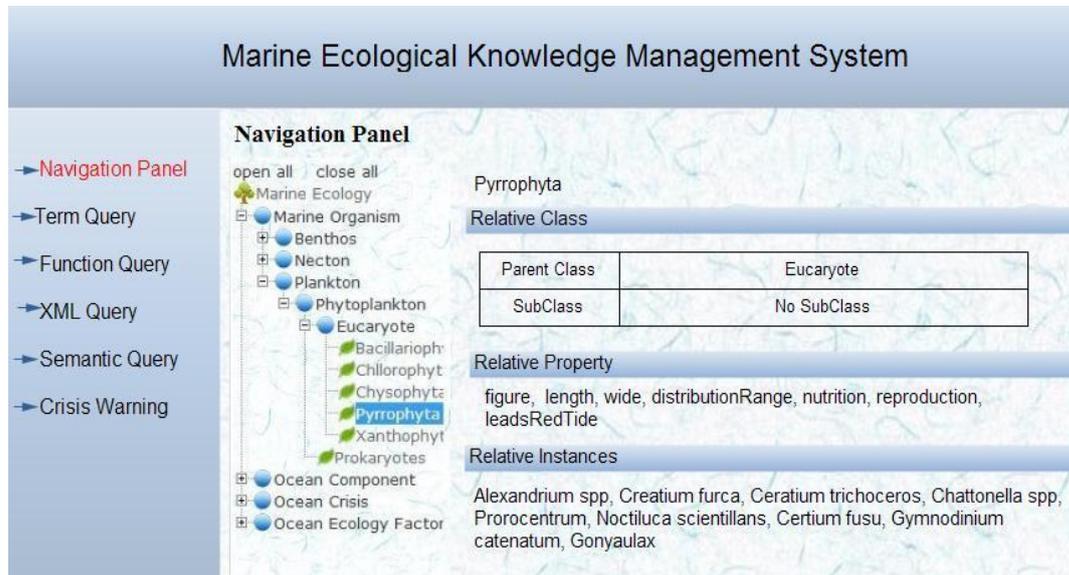
Figure 4. Navigation Panel of Marine Ecological Knowledge Management System

4.2 Term Query of Marine Ecological Knowledge Management System
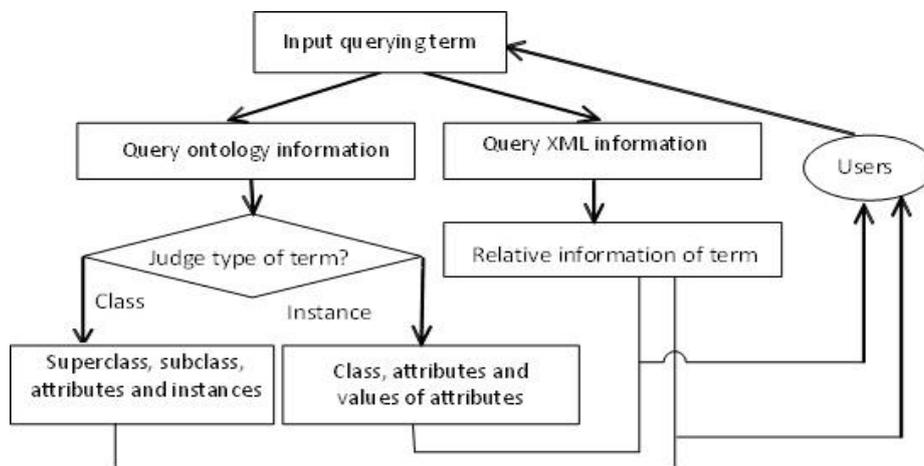
Figure 5. Term Query Flow-process Diagram

There are many concepts in domain ontology. To improve query efficiency in navigation tree, authors provide term query function. When user submits a query term, system can show class, hyponymy, attributes and instance of querying term. Term query processing diagram is shown in Figure 5.

Term query processing is described as following: ① User inputs querying term; ②Acquire relative information matched with querying term from XML document, and store them into data structure; ③Acquire concept information matched with querying term, and then judge the type of concept term; If term is a class, then use methods getSuperClass(), ListSubClasses(), ListdeclaredProperties() and getIndividual() to get term's superclass, subclass, attributes and instances sets, and store them into data structure; If term is an instance, then use method getOntClass() to get belonging class's id, and then use methods isObjectProperty() and isDatatypeProperty() to judge attribute type of term, if term is an ObjectPropert, judge whether object has attribute, if term has attributes, and then return attributes and values of term. ④Merge acquired XML information and Ontology information and then return merged information on interface.

When a user inputs querying term "Alexander algae", querying result is shown in Figure 6.

Figure 6. Term query of Marine Ecological Knowledge Management System

4.3 Function Query of Marine Ecological Knowledge Management System

Aiming at marine ecological function and process knowledge, such as marine ecological production, consumption, and decomposition process, along with cycle of nutritive materials process, this knowledge management system provides function and process knowledge querying. Photosynthesis is a basic function-process in marine ecosystem; its querying result is shown in Figure 7.
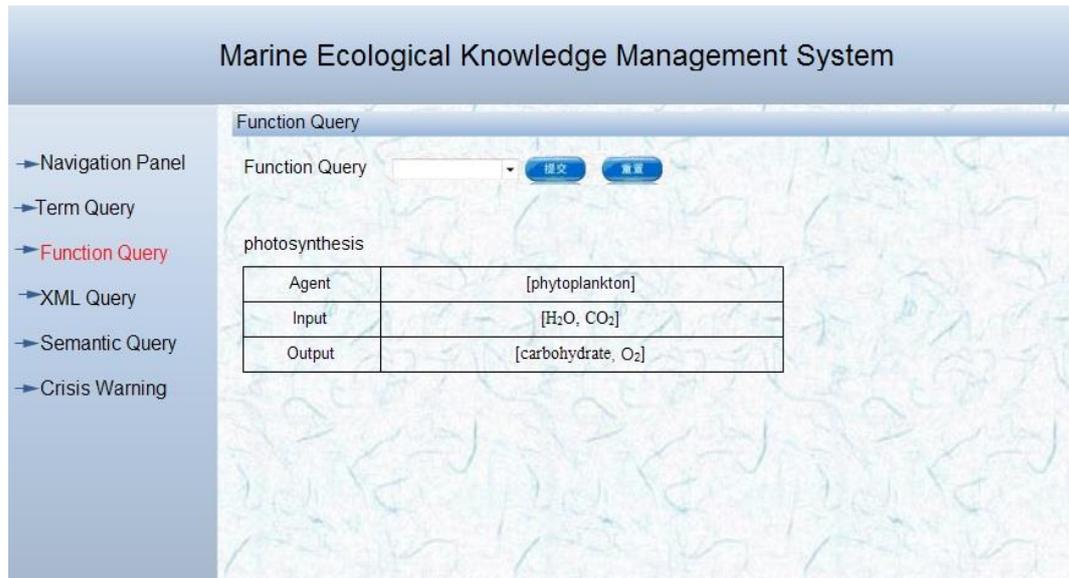
Figure 7. Function Query Instance-photosynthesis

4.4 Semantic Retrieval of Marine Ecological Knowledge Management System

Semantic module deals with users' question query. When user inputs a question query, semantic module first processes word segmentation (Chines word segmentation uses tool named IKAnalyzer) of question, and then matches word from word processing with ontology term. If matched result is failure, then prompt an error and return input question interface; If matched result is success, construct SPARQ query sentence (tripe form) with word processing result for reasoning and return querying result on interface. This module's flow-process diagram is shown in Figure 8.
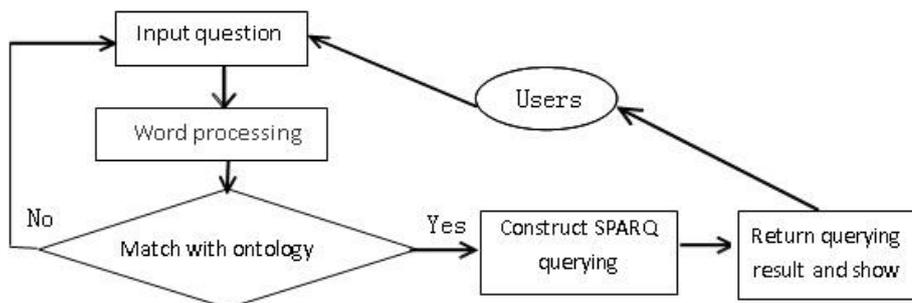

Figure 8. Semantic Retrieval Flow-process Diagram

We use Pellet Reasoner to implement reasoning query. Part of processing code is shown as following:

```
var resultList;
theInstance().create();
createInfModel();
create(queryString);
createPelletExecution();
results=execSelect();
resultList=ResultSetFormatter.toList(results);
```
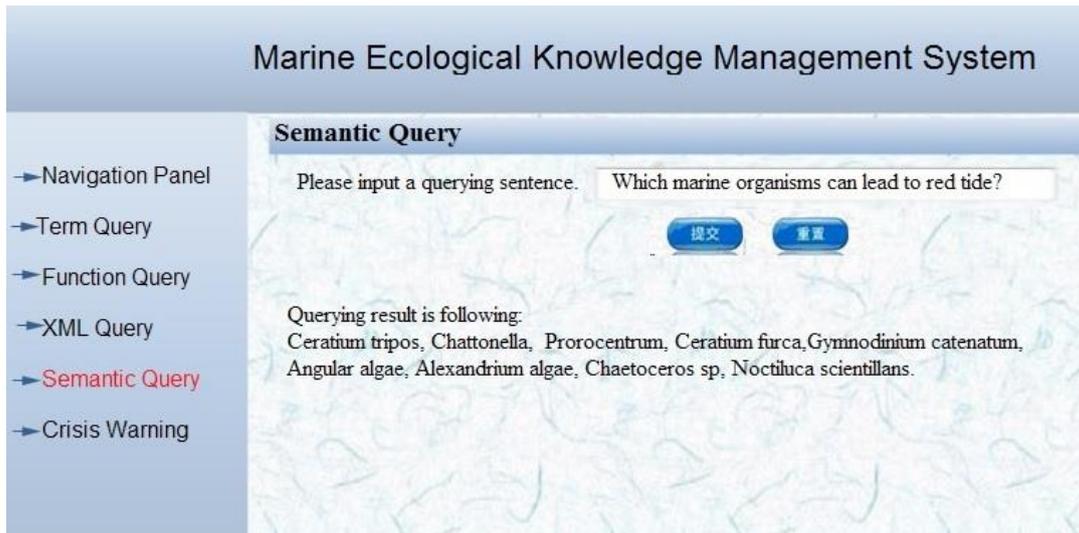
Figure 9. Semantic Retrieval of Marine Ecological Knowledge Management System

When user inputs a question about "Which marine organisms can lead to red tide?", querying result is shown in Figure 9.

4.5 Crisis Warning Case—Red Tide Crisis Early Warning

By building ontology rules and monitoring relative data, this function can provide users early warning about water quality, eutrophication and other marine ecosystem abnormal phenomena.

We construct ontology rules for red tide. Red tides are events which estuarine, marine, or fresh water algae accumulate rapidly in the water column and result in discoloration of the surface water [10]. Most of algae are phytoplankton, among them, the poisonous and harmful red tide alga are in the majority.

According to factors resulted in the occurrence of red tides, we extract relative concepts and attributes from marine ecology ontology, and build corresponding SWRL rules according to red tide occurrence conditions (for example, eutrophication index exceeds the threshold, and so on); and then import relative data from red tide database, including concrete data such as the kind and number of red tide organism, water environment parameters(COD, eutrophication, dissolved oxygen, the PH value *etc.,*) nutrient index, chlorophyll concentration *etc.*

When nutrients contained in seawater is excessive, it will cause phytoplankton bloom in the ocean surface layer and a middle layer, that is, the seawater probably eutrophication. Main red tide organisms are dinoflagellates; and if animal number of feeding dinoflagellates reduced, the number of dinoflagellate will increase; it may leads to occurrence of red tide because of accumulation of dinoflagellate. By analyzing above process, authors defined following conditions as premise [4]:

① Seawaters(?x)∧hasNutrient(?x, ?y)∧ swrlb:greaterThanOrEqual(?y, Upper_Limit)

② Dinoflagellate(?d) ∧ swrlb:greaterThan (?d, Upper_Limit)

③ Dinoflagllate(?d) ∧ eat(?c, ?d)→Consumer(?c,?d)

④ Consumer(?c) ∧ swrlb:lessThan(?c, Lower_Limit)

By analyzing all kinds of factors that may lead to red tides, then this paper constructs corresponding SWRL reasoning rules as following rule-1 and rule-2.

Seawaters(?x)∧hasNutrient(?x, ?y)∧ swrlb:greaterThanOrEqual(?y, 1)→Eutrophication(?x,TRUE)   rule-1

Seawaters(?x) ∧ hasNutrient(?x, ?y) ∧ swrlb:greaterThanOrEqual(?y, Upper_Limit) ∧ Consumer(?c) ∧ Dinoflagllate(?d)∧ eat(?c, ?d)∧ swrlb:lessThanOrEqual(?c, Lower_Limit)→Red_Tide(?x,TRUE)   rule-2

Marine ecology ontology defines concept *Seawaters* and concept *Dinoflagelate*, *Seawater* has property "hasNurtrient" and "Redtide", *Dinoflagelate* has dataproperty "number" and objectproperty "eat". Based on defined concepts, property and relations between concepts, using the built-in predicate swrlb:greaterThanOrEqual and swrlb:lessThanOrEqualmarine, we construct rules named Eutrophication and Red_tide in SWRL Rules Editor as shown in Figure 10.
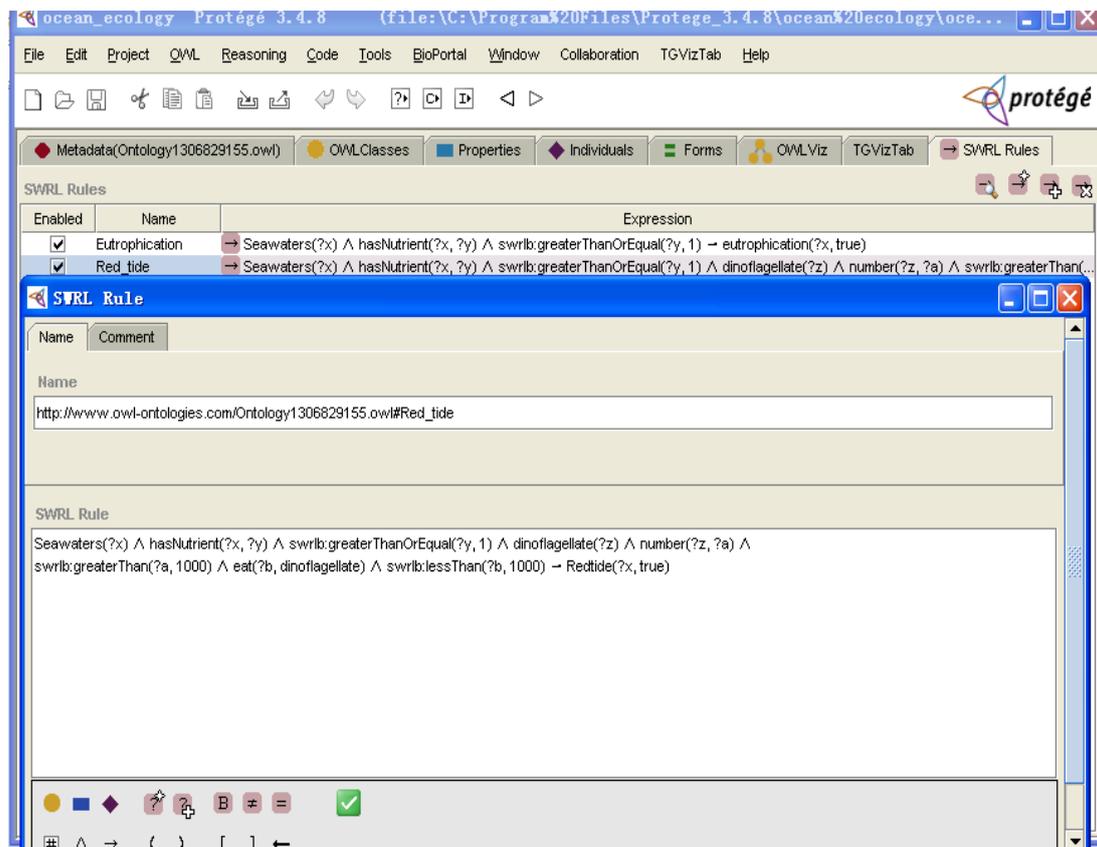

Figure 10. SWRL Rules Building-Red Tide

We created relative database about factors resulted in the occurrence of red tides. We can import or input red tide data into Marine Ecological Knowledge Management System. When red tide data exceeds threshold according to SWRL rule sets, the system will display crisis warning information such as "The region has been in the state of eutrophication, may produce red tide" *etc*., Crisis warning of red tide is shown in Figure 11.

Figure 11. Crisis Warning of Marine Ecological Knowledge Management System

5. Conclusion and Outlook

Domain ontology repository is to share, reuse and reasoning. Based on constructed ontology repository, authors research semantic retrieval and ontology rule reasoning, and develop Marine Ecological Knowledge Management System, which implements many functions such as marine ecological knowledge navigation, term query, function query, semantic query and marine ecological crisis warning, etc. Query part includes data search and semantic query. Users can retrieval and query the hyponym, properties, relations and instances of the marine ecological ontology. Reasoning part can analyze monitoring data submitted by users and make an early warning of marine ecological crisis, according to OWL Ontology and SWRL Rule Sets.

Marine ecological knowledge management system provides ordinary users a marine ecological knowledge access platform, and provides marine experts a semantic support platform for marine scientific research and cooperation, and it also verifies the rationality and feasibility of ontology modeling method and the validity of constructed marine ecology ontology.

At present, we has constructed relatively rough rules of reasoning, and realized the early warning of red tide. But comprehensive function is still not enough. Along with marine ecological ontology repository is gradually improved, and more reasoning factors are to be added, red tides crisis warning function will be more accurate and practical. Through enriching SWRL rule sets and strengthening ontology rule reasoning, it will provide marine ecological crisis warning more accurately, and also play a positive role in marine ecological environmental protection.

Acknowledgements

References

[1]. J. Xu, J. Xiong and R. Mizoguchi, "Toward an Ontology for Ocean Ecology and Sustainability", Proceedings of the 22nd Annual Conference of the Japanese Society of Artificial Intelligence, **(2008)** June, pp. 11-1, Asahikawa, Japan.

[2]. G. Shen and B. Shi, Editor, "Marine Ecology", (Second Edition), Science Press, Beijing, **(2008).**

[3]. H. Yun, "Ontology Modeling of Marine Ecology from Device-Function Viewpoint", Phd. thesis, Ocean University of China, Qingdao, **(2012)**.

[4]. H. Yun, J. Xu, Z. Guo and X. Wei, "Modeling of marine ecology ontology", Journal of Computer Applications, vol. 4, no. 34, **(2014),** pp. 1105-1108.

[5]. **(2014),** http:.//www.w3.org/Submission/SWRL/.

[6]. Y. Kitamura, T. Sano and K. Namba, "A functional concept ontology and its application to automatic identification of functional structures", Advanced Engineering Informatics, vol. 2, no.16, **(2002),** pp. 145-163.

[7]. R. Raskin, "Guide to SWEET ontologies", NASA/Jet Propulsion Lab, Pasadena, **(2013)**.

[8]. **(2014),** http://www.w3c.org/TR/2009/REC-owl-syntax-20091027/.

[9]. **(2014),** http://www.w3.org/TR/sparql11-query/.

[10]. **(2014),** http:// http://en.wikipedia.org/wiki/Red_tide.
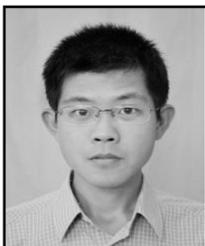
Authors

Hongyan Yun, is presently an associate professor in College of Information Engineering, Qingdao University. Her research interests include knowledge representation and reasoning, semantic web and ontology engineering, information integration and intelligent information system. (yunhy2001@163.com)

Jianliang Xu, is presently a professor in the Department of Computer Science and Technology, Ocean University of China. His research interests include automata theory, computational complexity and semantic web. (xjl9898@gmail.com)

Moji Wei, is presently an assistant researcher in Information Research Institute, Shandong Academy of Sciences. His research interests include knowledge representation and reasoning, ontology theory and application, and data mining. (weimoji@126.com)

Zhenbo Guo, is presently a professor in College of Information Engineering, Qingdao University. His research interests include GIS, Semantic Web and Intelligent Information System. (gzb@qdu.edu.cn)