Designing and Verifying Application Schema by Applying Standard Element for Managing Ocean Observation Data

Sun-Tae Kim[1], Lee-Kyum Kim[2], and Tae-Young Lee[3]

[1] Korea Institute of Science and Technology Information Daejeon, Republic of Korea
[2] Gwangju University, Gwangju, Republic of Korea
[3] Chonbuk National University, Jeonju, Republic of Korea
stkim@kisti.re.kr, leekyum@kwangju.ac.kr, taehyun@jbnu.ac.kr

Abstract. There is a need to study the OWL-based application schema to ensure interoperable data exchange between ocean-related institutions, and supporting researcher's intelligent data search. In this study, the RDF vocabularies are defined on the basis of the elements derived through element decision study for managing scientific data in the field of ocean observation. The application schema was verified by using the temperature profile data of CTD data observed in the 'Chukchi' sea selected from the data provided by the National Oceanographic Data Center of the US.

Keywords: Scientific Data, Observation Data, Ocean Observation Data, Application Schema, OWL Schema, Metadata.

1 Introduction

It is necessary to establish ocean observation data as Linked Data so as to make the data a model case (Bizer 2009) for connecting structured data on the web and to be published. In this study, an OWL-based application schema will be designed and verified in order to ensure interoperable data exchange between ocean-related institutions and to support researchers' intelligent data search. The determined standard elements are defined on the basis of the elements derived through element decision study for managing scientific data in the field of ocean observation. The higher 21 elements and the lower 173 elements which are the basis for defining vocabularies are found in http://bit.ly/GTuSvi.

2 Designing Application Schema

2.1 Class Design

A vocabulary dictionary is established, which consists of 21 class concepts, 50 object attributes and 92 data type attributes for the metadata standard elements in order to

S. Yamamoto (Ed.): HIMI/HCII 2013, Part II, LNCS 8017, pp. 110–115, 2013.

design an application metadata schema, and to present it at http://bit.ly/uMOxkD. Class design was carried out as follows. <owl:Class> and <rdfs:subClassOf> was used to design the following class. ② and ⑤ define the standard elements "Institution, Person". Each element is the lower class of the common higher element <Agent>, and is defined with the RDF schema vocabulary, <rdfs:subClassOf>, as in ③. ① is the URI declaration for the basic Namespace used in designing RDF OWL schema.

```
<rdf:RDF
omitted…
xmlns="http://www.kisti.re.kr/ontology/scientificdata.owl#" ..... ①
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
<owl:Class rdf:ID="Institution"> ..... ②
<rdfs:subClassOf> ..... ③
<owl:Class rdf:ID="Agent"/> ..... ④
</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Person"> ..... ⑤
<rdfs:subClassOf rdf:resource="#Agent"/> ..... ⑥
</owl:Class>
```

Fig. 1. Design Agent related class

2.2 Designing Object Value-Type Attribute

<owl:ObjectProperty> and <rdfs:subPropertyOf> was used to design the following object value type attribute.

```
<owl:ObjectProperty rdf:ID="hasCreator"> ..... ①
  <rdfs:subPropertyOf> ..... ②
    <owl:ObjectProperty rdf:ID="hasAgent"/> ..... ③
  </rdfs:subPropertyOf>
  <rdfs:comment
rdf:datatype="xsd:string">producer</rdfs:comment>…….. ④
 </owl:ObjectProperty>
 <owl:ObjectProperty rdf:ID="hasPublisher"> ..... ⑤
  omitted…
 </owl:ObjectProperty>
```

Fig. 2. Design hasAgent related object value type attribute

① and ⑤ is to define “hasCreator, hasPublisher” as an object value type attribute to state the producer and publisher information of the standard element <Agent>. Each attribute is the lower attribute of the common higher attribute <hasAgent>, and the higher element was specified with the RDF schema vocabulary, <rdfs:subPropertyOf>, as shown in ③. ④ describes that the ‘producer’ information is stated as the value of object value type attribute hasCreator, using <rdfs:comment> so that humans and systems can understand and process the defined attributes.

2.3 Designing Data Value Type Attribute

The following details of design show definition of data value type attributes by using <owl:DatatypeProperty> and <rdfs:subPropertyOf>. ① and ⑤ defines data value type attributes corresponding to the standard elements “uniformTitle, alternateTitle”. Each of them is the lower attribute of the common higher attribute <title>, and is defined by using the RDF schema vocabulary, <rdfs:subPropertyOf>, as shown in ③. ④ represents the higher attribute.

② and ⑥ applies “rdfs:domain, rdf:resource” of the RDF schema to represent the attribute so as to define resources that the defined data value type attribute can have as a domain. ⑦ uses the owl:cardinality attribute to define a restriction of the number of the title attribute appearances, which means that classes, such as ScientificData that uses the title attribute, can be stated by using the title attribute only once.

```
<owl:DatatypeProperty rdf:ID="uniformTitle"> ..... ①
  <rdfs:domain rdf:resource="#ScientificData"/> ..... ②
  <rdfs:subPropertyOf> ..... ③
    <owl:DatatypeProperty rdf:ID="title"/> ..... ④
  </rdfs:subPropertyOf>
 </owl:DatatypeProperty>
 <owl:DatatypeProperty
rdf:ID="alternateTitle"> ..... ⑤
  <rdfs:subPropertyOf rdf:resource="#title"/>
  <rdfs:domain rdf:resource="#ScientificData"/> ..... ⑥
 </owl:DatatypeProperty>
 <owl:Restriction> ..... ⑦
  <owl:onProperty rdf:resource="#title"/>
  <owl:cardinality
rdf:datatype="xsd:nonNegativeInteger">1</owl:cardinality>
```

Fig. 3. Design Title related data value type attribute

3 Verifying Selected Element

To verify the OWL application schema, the test for creation of RDF document which follows the schema was conducted. The test dataset, the temperature profile data of CTD data observed in the 'Chukchi' sea selected from the data provided by the National Oceanographic Data Center of the US, were used. All about 10,189 records were collected.

The shape of the collected data is as below. ① is about the platform at which observation was conducted. ② is about the institute which is chare of data creation. ③ is about the project name which supply the fund for observation. ④ is about the depth of the deployment. ⑤ is about start datatime and end datatime of the observation. ⑥ is about latitude and longitude at which observation was taken. ⑦ is about bottom depth of the sea. ⑧ is about number of records which a dataset has. ⑨ is about recording interval. In this example, the observation was taken place per 1 hour. ⑩ is about the records. In this example, the data were collected per 60 minutes which consists of 8774 records.

```
① Platform: moored buoy
② Institute: University of Alaska/Institute of Marine Sciences
③ Project Name:          CHUKCHI94
④ Deployed Depth:         55 Meters
⑤ Start Date/Time:        25-sep-1993 03:00:00
   End Date/Time:         25-sep-1994 16:00:00
⑥ Latitude:             71 03.23 N
   Longitude:           159 31.90 W
⑦ Bottom Depth:           75 Meters
⑧ Number of Records:      8774
⑨ Recording Interval:     60 Minutes

no.  Date        Time     Depth  Temp.  Cond.  Sal.   Sigma T
=========================================
336 25-sep-1993 03:00:00  55.0  5.2198 30.1902 31.0109 24.4906
337 25-sep-1993 04:00:00  55.0  5.3319 30.2183 30.9390 24.4214   ⑩
338 25-sep-1993 05:00:00  55.0  5.2659 30.2401 31.0247 24.4965
339 25-sep-1993 06:00:00  55.0  5.0859 30.1215 31.0571 24.5416
```

Fig. 4. The example of ocean observation data

The figure 4 is showing the procedure of RDF document creation. The crawler ingest the data from the NOAA web site. It collected the data and build the database, simultaneously made a log file as shown below in ①, ② and ③ process. Instance generator read the records from the database as in ④ process and create RDF document which follow the Turtle format as in ⑤. It makes a log file also as Crowler as in ⑥.

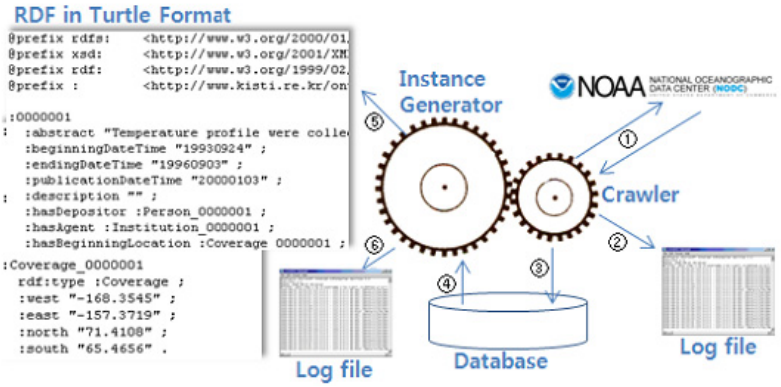

Fig. 5. The procedure of RDF Document Creation

The following shows an RDF document of 'ttl type' created with the instance creator and an RDF document of RDF/XML type created by using the RDF verifier and a converter provided at 'http://www.rdfabout.com/'.

```
# Verify ttl type
@prefix :       <http://www.kisti.re.kr/ontology/scientificdata#>.
:SD0000999
  :title "Chlorophyll data collected by the research vessels … May -
September 2002 " .
  :beginningDateTime "20020414" ;
  :hasBeginningLocation :Coverage_0000999 ;
  :hasSource :Source_0000999_1, :Source_0000999_2;
:Coverage_0000999
  rdf:type :Coverage ;
  :west "-77.76" ;   :east "-65.5" ;   :north "-65.12" ;
<omitted...>

# Verify RDF/XML type
<rdf:Description
rdf:about="http://www.kisti.re.kr/ontology/sd#SD0000999">
   <sd:title>
    Chlorophyll data collected by the research vessels … May -
September 2002
   </sd:title>
   <sd:hasBeginningLocation>
    <sd:Coverage
rdf:about="http://www.kisti.re.kr/ontology/sd#Coverage_0000999">
      <sd:west>-77.76</sd:west> <sd:east>-65.5</sd:east>
<sd:north>-65.12</sd:north>
    </sd:Coverage>
```

Fig. 6. Design Agent related class

The above test demonstrated that the RDF document for applying the OWL-based application metadata profile designed in this study was successfully created by using the data provided by the National Oceanographic Data Center of the US (NODC) as input.

4 Conclusion

In this study, the OWL-based metadata application schema was designed by using the metadata standard element for managing and using ocean observation data. The application schema was verified for the data provided by the NODC. The data of the NODC actually used was verified. We also proved that the selected metadata element and the application schema were actually applicable without modification. When NODC's metadata elements were tested and compared with the metadata elements selected in this study, all CDT data could be described by these elements like 'SourceName', 'Project', 'DeployedDepth', 'DataItem', 'TimeInterval', 'BottomDepth', 'ObservationLocation' etc. And each application schema also described without missing elements.

Therefore, derived metadata standard elements for the ocean observation field are judged to be full of significance and expected to be utilized for metadata management and practical use in the field of ocean observations. The application schema proposed in this study will be useful for managing and using metadata involved in ocean observation.

Acknowledgment. This essay is a modified version and summary of a graduate school thesis, Chonbuk National University (February 2, 2012).

References

1. Bizer, C., Heath, T., Berners-Lee, T.: Linked Data - The Story So Far. Special Issue on Linked Data. International Journal on Semantic Web and Information Systems, IJSWIS (2011)
2. Kim, S.: A Study on Extraction and Design of Standardized Elements on Metadata for Ocean Observational Data. Chonbuk National University (2012)
3. Kanzaki, M.: Semantic Web No Tame No RDF/OWL NYUMON. Hongrung Publishing Company (2005)