Implementation of Content Patterns in the Methodology of the Development of Ontologies for Scientific Subject Domains

Yury Zagorulko[(✉)], Olesya Borovikova, and Galina Zagorulko

A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences, Acad. Lavrentiev avenue 6, 630090 Novosibirsk, Russia
{zagor,olesya,gal}@iis.nsk.su

Abstract. The paper considers an approach to the development and implementation of ontology design patterns used in constructing the ontologies of scientific subject domains. Using such patterns allows us to provide a uniform and consistent representation of all the entities of the ontology under development and to save resources and avoid errors typical for ontological modeling. The paper pays a special attention to the development and implementation of this kind of ontology design patterns as content patterns playing an important role in the methodology proposed by the authors for the development of scientific subject domain ontologies.

Keywords: Ontology · Ontology design patterns · Content patterns
Methodology for the development of ontology of scientific subject domain

1 Introduction

Ontologies are widely used to formalize the knowledge contained in scientific subject domains (SSD). The ontology not only helps to produce a convenient representation of all the necessary concepts of a modeled domain, but also helps to ensure their uniform and consistent description.

As ontology development is rather complex and time-consuming process, various methodologies and approaches are proposed to facilitate it. Recently, the approach based on the application of ontology design patterns (ODPs) obtained new attention [1, 2]. According to this approach, each ODP is a documented description of proven solution for a typical ontological modeling problem. ODPs are created in order to simplify and facilitate the process of building ontology and to help developers avoid the typical errors uring ontological modeling.

Although the use of ontology design patterns makes it possible to save human resources and improve the quality of the ontology being developed, there exists presently only one methodology for constructing ontology, namely the eXtreme Design methodology (XD methodology) [2] proposed within the framework of the project NeOn [3], which openly declares the application of the ODPs.

Among other tools using ontology design patterns we may point, for example, a plug-in for the ontology development in the NeOn project, and another plug-in for the WebProtégé ontology editor [4].

However, these tools cover only some of the possible tasks dealing with patterns. For instance, there are absolutely no tools to support the construction, search and retrieval of the patterns from an ontology, and a very few tools supporting the collection, discussion and dissemination of the patterns. The catalogs of ontology design patterns [5–7], which have been developed recently, can be attributed to the latter group of tools.

The present paper considers an approach to the implementation of this kind of the ODPs as content patterns that are playing an important role in the methodology proposed by the authors for the development of the ontologies for scientific subject domains [8, 9]. Used in this methodology patterns appear as a result of solution of the ontological modeling problems that the authors of the paper encountered in the process of developing ontology for a wide range of scientific subject domains [10–14].

The rest of the paper is structured as follows. The second chapter contains a brief introduction to the ontology design patterns. The third chapter is devoted to the issues of content patterns development and implementation. It presents a fragment of the OWL implementation of a content pattern for describing scientific activity. The main advantages of the application of ontology design patterns for developing the ontologies of scientific subject domains are discussed in the Conclusion.

2   The Ontology Design Patterns

As mentioned above, ontology design patterns are created to describe solutions to typical problems arising in ontology development.

Depending on the type of the problems for which the patterns are created, they distinguish among structural ODPs, correspondence ODPs, content ODPs, reasoning ODPs, presentation ODPs and lexico-syntactic ODPs [1].

The most popular patterns for the development of ontologies directly by knowledge engineers are structural ODPs, content ODPs and presentation ODPs. Let us consider them in more detail.

Structural patterns either fix ways to solve problems caused by the limitations of the expressive capabilities of ontology description languages or specify the general structure and the kind of ontology. Patterns of the first type are called logical patterns, and patterns of the second type are referred to as architectural patterns. The latter type contains proposals for ontology organization in general, including, for example, structures such as taxonomy and modular architecture.

The content patterns define ways of representing typical ontology fragments that can serve as a basis for building the ontologies for a class of subject domains. Based on these patterns, you can describe more complex (composite) content patterns, supplementing them with missing components and introducing additional relationships. However, it should be remembered that the use of composite patterns may generate assembly and integration problems.

The presentation patterns define recommendations for the naming, annotating, and graphical representation of ontology elements. The application of these patterns should make ontology to be more "readable", convenient and easy to use.

The naming patterns include conventions on the naming rules for namespaces, files, and ontological elements. For example, when declaring an ontology, it is welcome to use capital letters and the singulars in the name of a class, name of the parent class in the names of its subclasses (as a suffix), and small letters in the names of properties.

The annotation patterns specify annotation rules for ontology elements, such as providing classes and their properties with comments and labels in several languages.

At the moment, there is no common standard for describing patterns, but they are most often described in the format proposed by the Association for Ontology Design & Patterns (ODPA) [6]. In accordance with this format, a pattern description includes its graphical representation, text description, a set of scenarios and usage examples, links to other patterns in which it is used, as well as information about the name of the pattern, its author and scope.

The methodology of extreme ontology design [2] also suggests providing each content pattern with a set of Competency Questions. These can be used both at the stage of developing patterns and for finding patterns suitable for the development of a specific ontology.

The development and use of content patterns relies on the five basic operations [1]: import, specialization, generalization, composition and expansion.

Import is including a content pattern in the ontology under development. This is the basic mechanism for reusing such patterns. The elements of imported patterns cannot be modified.

Specialization is creating sub-classes of a content pattern's class and/or sub-properties of some content pattern's properties.

Generalization means that a content pattern *CP1* is a generalization of another content pattern *CP2* if *CP1* imports *CP2*, and at least one ontology element from *CP1* generalizes an element from *CP2*.

Composition is associating classes (properties) of a content pattern with the classes (properties) of other content patterns by means of axioms.

Expansion is adding new classes, properties and axioms to ontology with the aim of meeting the requirements not addressed by the reused content patterns.

3   Development and Application of Content Patterns in Building the Ontologies of SSDs

When creating ontology of SSDs, it is important to ensure the possibility of a uniform and consistent presentation of appropriate scientific concepts and their properties. One way to solve this problem is to use content patterns describing the concepts typical for most scientific subject domains and scientific activities performed in them. A set of such patterns is provided by a methodology for constructing ontology developed within the technology for creating thematic intelligent scientific Internet resources [8].

It should be noted that this methodology, in addition to the content patterns, also includes the presentation patterns and structural logical patterns used for constructing content patterns [9].

The need to use structural logical patterns in this methodology was attributed to the absence in the OWL language [15] some expressive means for representing complex entities and constructions required to build the SSD ontologies. For example, logical structural patterns help to represent attributed relationships (binary relations with attributes) and domains (sets of elementary values).

The patterns of latter type were introduced to specify the ranges of values called domains in the relational data model and characterized by their name and a set of atomic values. Domains are convenient to use in the descriptions of possible values of the properties of a class, when the entire set of such values is known in advance. Using domains not only allows users to control the input of information; it can also make this operation more convenient by providing users with the opportunity to select property values from a given list of values (see Subsect. 3.3).

Because of a limited length of the paper, in this section we will only dwell on the development and implementation of content patterns.

3.1   Development of Content Patterns

Within the framework of the methodology for constructing SSD ontologies, we have developed content patterns for presenting the following concepts: the Division of Science, Object of Research, Subject of Research, Research Method, Scientific Result, Person, Organization, Activity (Scientific activity), Project, Publication, etc.

When developing content patterns for each of them, we defined a set of Competency Questions describing their. These questions helped us to identify the composition of the obligatory and optional ontological elements of the pattern and describe the requirements for them. The requirements were presented in the form of axioms and limitations.

Let us consider some content patterns in more detail. First let us note that in the pictures of the patterns presented below classes are shown in the form of ellipses, while their individuals and attributes are in the form of rectangles. An *ObjectProperty* connection is shown by a solid straight line and a *DataProperty* connection, by a dash line. At the same time, mandatory classes and attributes (individuals) are represented by figures surrounded by a thick line.

As a rule, scientific activity is realized through projects, programs, expeditions and examinations.

The graphical representation of a pattern intended for the description of scientific activity is shown in Fig. 1. This pattern sets the following requirement: when any kind of activity is described, it is necessary to give a reference to the object of research to which this activity is devoted, and to the division of science in which it is being carried out.

Figure 2 shows a pattern designed to describe a scientific result. This pattern sets the following requirement: when we describe a scientific result, we should give a reference to the activity under which it was obtained.
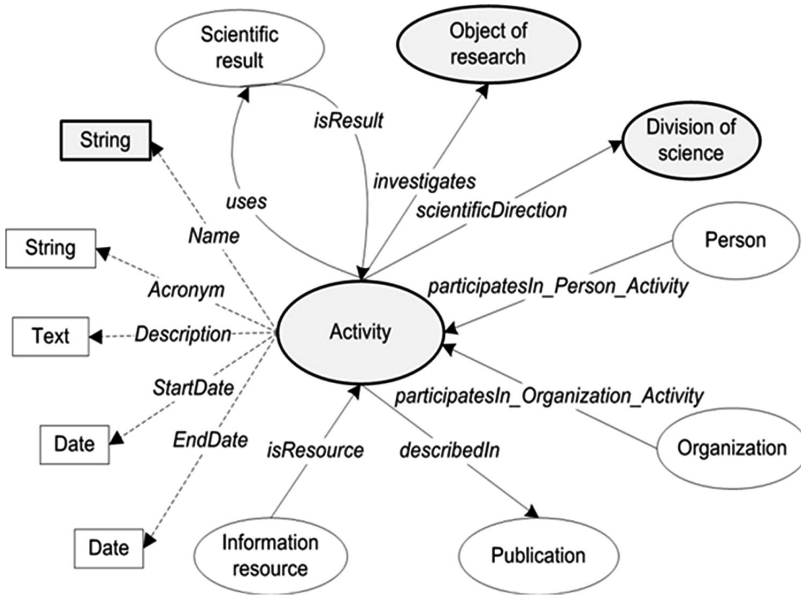
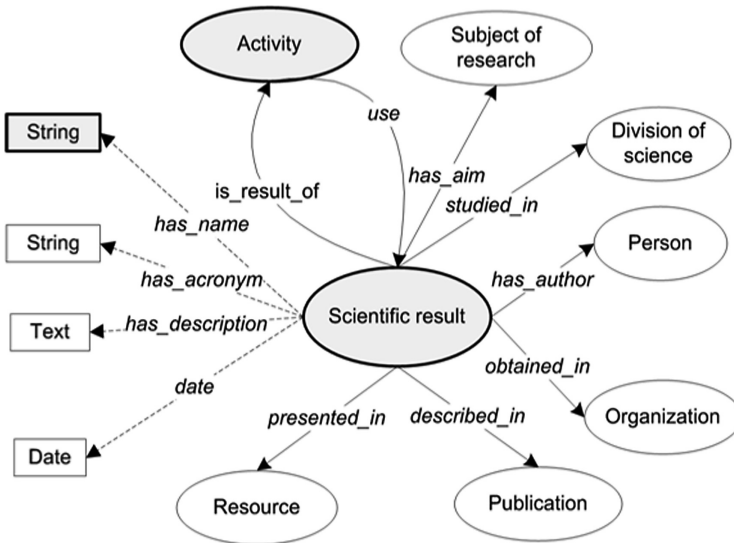**Fig. 1.** Pattern for the description of a scientific activity.

Fig. 2. Pattern for the description of a scientific result.

Figure 3 shows a pattern for describing the object of research, which includes the following classes as obligatory: Subject of Research, Activity, and Division of Science. The instances of these classes should be connected with the object of research by such

relations as *has_aspect, investigated_in* and *studied_in*, respectively. The object of research can be structural (i.e. include other objects of research).
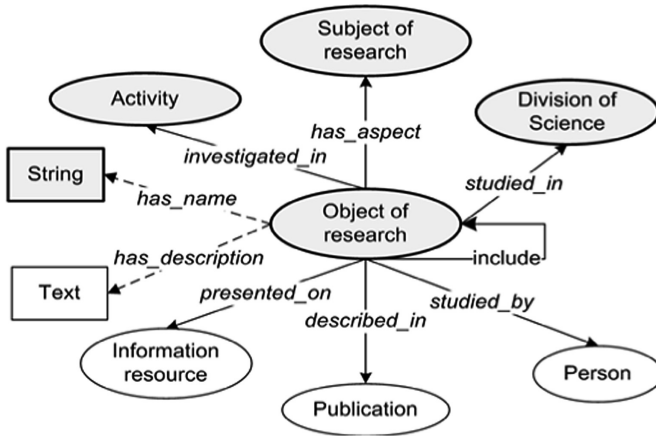

Fig. 3. Pattern for the description of an object of research.

The pattern presented in Fig. 4 is intended to describe the research methods used in a scientific activity.
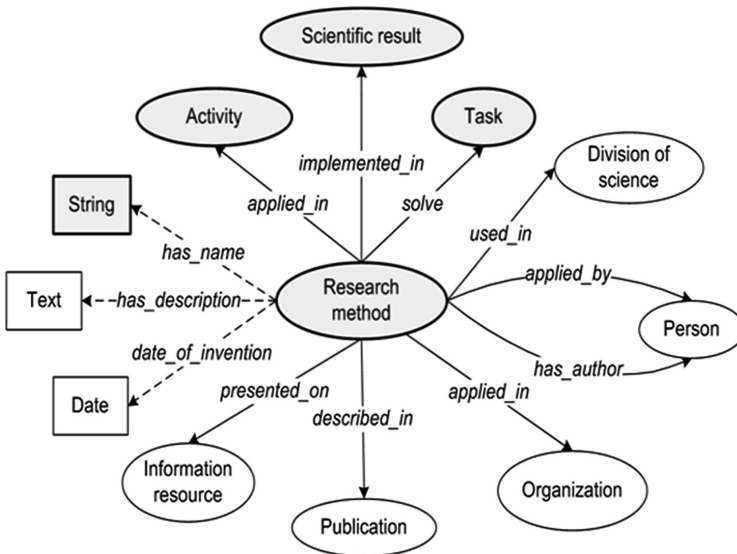

Fig. 4. Pattern for the description of a research method.

Elements of the description of the pattern of a research method are represented by such obligatory ontology classes as Activity, Scientific Result, Task, and others, and such relations as *used_in, implemented_in, solve*, etc.

Figure 5 shows a pattern for describing a publication. This pattern includes many classes but only two of them as obligatory: Person and Division of Science. The instances of these classes should be connected with the publication by the relations *hasAuthorPerson* and *describesDivision*, respectively.
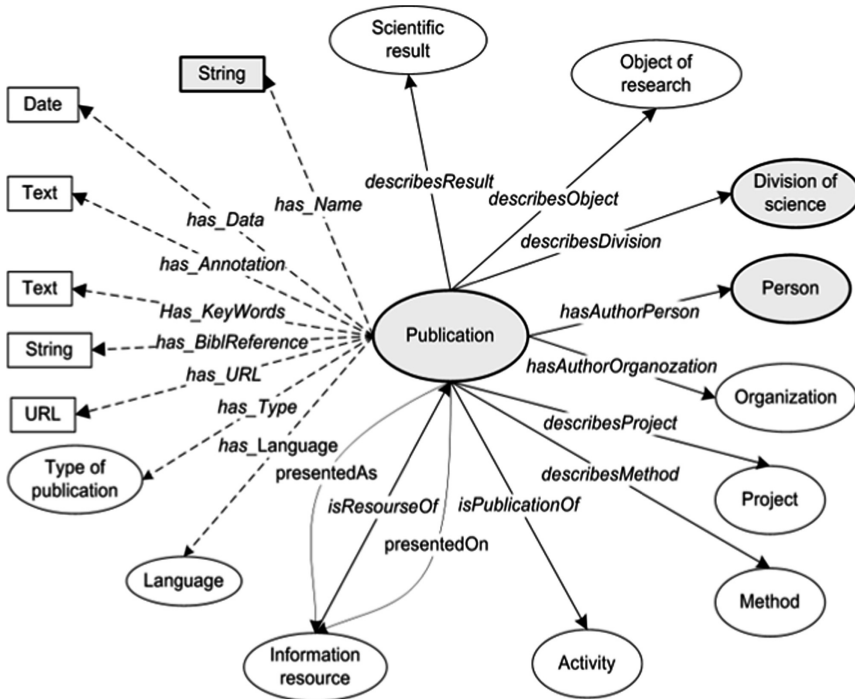

Fig. 5. Pattern for the description of a publication.

It is easy to see that it is impossible to describe any content pattern separately, i.e. not using concepts from other patterns. For example, the pattern for describing the scientific result involves, in addition to the main concept of the Scientific Result, such concepts as Activity, Subject of Research, Division of Science, etc. This unites all content patterns into a single system of patterns, and the concepts represented by them, into a single system of concepts defining a related description of the modeled domain.

In fact, content patterns are fragments of ontology. After the concepts contained in them have been specialized and the patterns have been extended by necessary concepts and properties, they become integral parts of the ontology of the SSD being developed.

3.2   Implementation of Content Patterns

All content patterns of the methodology under consideration are realized by means of the OWL language.

Below is a fragment of the OWL implementation of the content pattern for describing scientific activity in the Turtle format. (Note that the names of the ontological elements of the patterns are taken from the namespace of the ontology inir: http://www.semanticweb.org/IIS/ontologies/INIR_Ontology#.)

Let us describe the content components of this pattern.

First of all, this is the set of classes used in it and the axiom of disjointness defined for these classes (AllDisjointClasses):

```
:InformationResource rdf:type owl:Class ;
:ObjectOfResearch rdf:type owl:Class ;
:Organization rdf:type owl:Class ;
:Person rdf:type owl:Class ;
:Publication rdf:type owl:Class ;
:BranchOfScience rdf:type owl:Class ;
:ResearchResult
[ rdf:type owl:AllDisjointClasses ;
  owl:members
( :Activity
:InformationResource
:ObjectOfResearch
:Organization
:Person
:Publication
:BranchOfScience
:ResearchResult)] .
```

The set of properties of the *Activity* class is as follows:

```
:scientificDirection _Activity_BranchOfScience rdf:type owl:ObjectProperty ;
 rdfs:domain :Activity ;
 rdfs:range :BranchOfScience ;
:investigates_Activity_ObjectOfResearch rdf:type owl:ObjectProperty ;
 rdfs:domain :Activity;
 rdfs:range :ObjectOfResearch;
:describedIn_Activity _Publication rdf:type owl:ObjectProperty ;
 rdfs:domain :Activity ;
 rdfs:range :Publication ;
:participatesIn_Person_Activity rdf:type owl:ObjectProperty ;
 rdfs:domain :Person;
 rdfs:range :Activity;
:isResult_ResearchResult_Activity rdf:type owl:ObjectProperty ;
 rdfs:domain :ResearchResult;
 rdfs:range :Activity;
:isResource _InformationResource_Activity rdf:type owl:ObjectProperty ;
 rdfs:domain :InformationResource;
 rdfs:range :Activity;
:participatesIn _Organization _Activity rdf:type owl:ObjectProperty ;
 rdfs:domain :Organization;
 rdfs:range :Activity;
:Activity_Name rdf:type owl:DatatypeProperty ;
rdfs:domain :Activity;
rdfs:range rdfs:Literal ;
:Activity_Acronym rdf:type owl:DatatypeProperty ;
rdfs:domain :Activity;
rdfs:range rdfs:Literal ;
```

Next, for each instance (individual) of the Activity class, the *FunctionalProperty* property is imposed on requiring only one start date and one end date:

```
:Activity _EndDate rdf:type owl:DatatypeProperty, owl:FunctionalProperty ;
rdfs:domain :Activity;
rdfs:range rdfs:Literal ;
:Activity _StartDate rdf:type owl:DatatypeProperty, owl:FunctionalProperty ;
rdfs:domain :Activity;
rdfs:range rdfs:Literal ;
```

Finally, for each instance of the *Activity* class, by imposing constraints on properties, the following requirements for the number and type associated with this class of concepts are described:

- at least one direction of scientific activity must be specified;
- at least one object of research should be specified;
- activities must involve at least one person;
- an activity must have at least one name.

```
:Activity rdf:type owl:Class ;
owl:equivalentClass [ rdf:type owl:Restriction ;
owl:onProperty :scientificDirection _Activity_BranchOfScience;
owl:someValuesFrom :BranchOfScience] ,
[ rdf:type owl:Restriction ;
owl:onProperty :investigates_Activity_ObjectOfResearch;
owl:someValuesFrom :ObjectOfResearch] ,
[ rdf:type owl:Restriction ;
owl:onProperty [ owl:inverseOf :participatesIn_Person_Activity] ;
owl:minQualifiedCardinality "1"^^xsd:nonNegativeInteger ;
 owl:onClass :Person] ,
[ rdf:type owl:Restriction ;
owl:onProperty :Activity_Name;
owl:minCardinality "1"^^xsd:nonNegativeInteger] ;
```

3.3   Use of Content Patterns

The use of content patterns in the methodology described above is supported by a special editor allowing the users (knowledge engineers) to populate the ontology with the objects (individuals) of classes for which content patterns are implemented.

When populating an ontology with the help of the editor, the user selects from a tree of ontology classes a class (concept) an object of which he desires to create to supplement the ontology, and the editor exploits the name of this class to find the corresponding pattern. After that, the editor, using the information from the pattern, builds a form containing the fields for filling in all the properties of the object created.

It should be noted that when describing such an object, the user should take into account and the editor should control the restrictions on the values and cardinality of its properties. The restrictions are used not only to control the input information, but also as a potential source of the property values of the objects. Thus, the user can select values for each such property, either from the domain description specified by the corresponding structural pattern as the value range of this property or from the list of individuals of the class indicated as the range of this property. Such a list is dynamically formed from the individuals of the class presented in the ontology.

The content patterns considered above were applied to create an ontology of the subject domain "Decision support in weakly formalized areas". The use of these patterns was very effective, since this subject domain includes a rich set of methods and tasks, and also relies on the results obtained in many other scientific disciplines (Theory of Decision Making, Operations Research, Artificial Intelligence, Cognitive Psychology, System Analysis, etc.).

The ontology of this SSD contains detailed classifications and descriptions of tasks and methods for solving decision support problems. It identifies five main groups of methods, namely, modeling methods, data analysis methods, expert methods, reasoning methods, and auxiliary methods.

Figure 6 shows an example of using the pattern for the description of research method to describe the hierarchy analysis method (the analytic hierarchy process) [16] proposed by Saaty in the 1980s.
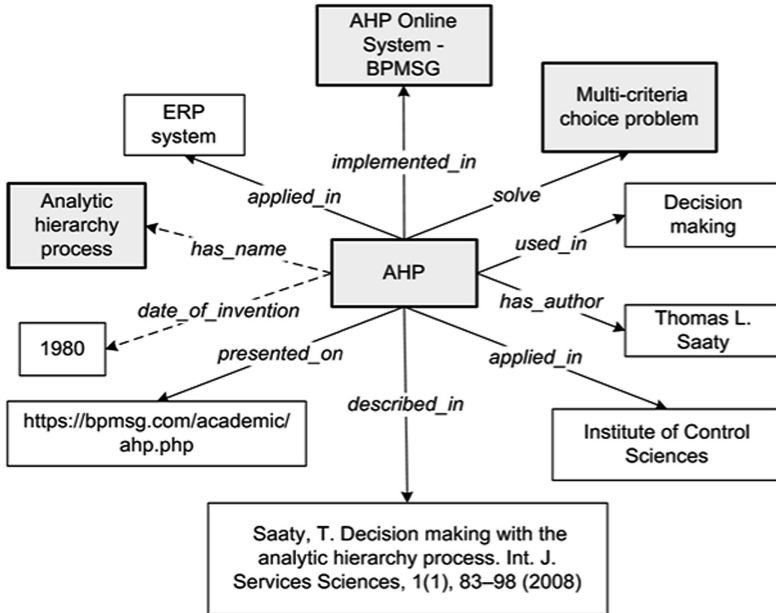

Fig. 6. Example of using the pattern for describing the research method.

4   Conclusion

The paper addresses the issues of applying ontology design patterns to the development of the ontologies of scientific subject domains. It is shown that the use of these patterns provides a uniform and consistent representation of all the entities of an ontology being developed, saving human resources and assuring a lower number of errors committed in ontology development.

The authors have devised an approach to the implementation of content patterns playing a leading role in the methodology proposed to develop SSD ontologies. This methodology and content patterns described in the paper have shown their practical usefulness in the development of the ontologies of various subject areas.

References

1. Gangemi, A., Presutti, V.: Ontology design patterns. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies. IHIS, pp. 221–243. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-92673-3_10

2. Blomqvist, E., Hammar, K., Presutti, V. Engineering ontologies with patterns: the eXtreme design methodology. In: Hitzler, P., Gangemi, A., Janowicz, K., Krisnadhi, A., Presutti, V. (eds.) Ontology Engineering with Ontology Design Patterns. Studies on the Semantic Web. IOS Press (2016)

3. NeOn Project (2016). http://www.neon-project.org/nw/Welcome_to_the_NeOn_Project. Accessed 30 June 2018

4. Hammar, K.: Ontology design patterns in WebProtégé. In: Proceedings of 14th International Semantic Web Conference (ISWC-2015). CEUR Workshop Proceedings (2015)

5. Dodds, L., Davis, I.: Linked data patterns (2012). http://patterns.dataincubator.org/book. Accessed 30 June 2018

6. Association for Ontology Design & Patterns (2018). http://ontologydesignpatterns.org. Accessed 30 June 2018

7. Ontology Design Patterns (ODPs) Public Catalog (2009). http://odps.sourceforge.net. Accessed 30 June 2018

8. Zagorulko, Y., Zagorulko, G.: Ontology-based technology for development of intelligent scientific internet resources. In: Fujita, H., Guizzi, G. (eds.) SoMeT 2015. CCIS, vol. 532, pp. 227–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22689-7_17

9. Zagorulko, Y., Borovikova, O., Zagorulko, G.: Application of ontology design patterns in the development of the ontologies of scientific subject domains. In: Kalinichenko, L., Manolopoulos, Y., Skvortsov, N., Sukhomlin, V. (eds.) Selected Papers of the XIX International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2017), vol. 2022, pp. 258–265. CEUR Workshop Proceedings, (CEUR-WS.org) (2017). http://ceur-ws.org/Vol-2022/paper42.pdf. Accessed 30 June 2018

10. Zagorulko, Y., Zagorulko, G.: Ontology-based approach to development of the decision support system for oil-and-gas production enterprise. In: Proceedings of the 9th SoMeT_10 New Trends in Software Methodologies, Tools and Techniques, pp. 457–466. IOS Press, Amsterdam (2010)

11. Zagorulko, Y., Borovikova, O., Zagorulko, G.: Knowledge portal on computational linguistics: content-based multilingual access to linguistic information resources. In: Selected topics in Applied Computer Science, Proceedings of the 10th WSEAS International Conference on Applied Computer Science (ACS 2010). Iwate Prefectural University, Japan, 4–6 October 2010, pp. 255–262. WSEAS Press (2010)

12. Zagorulko, Y., Borovikova, O.: Technology of ontology building for knowledge portals on humanities. In: Wolff, K.E., Palchunov, Dmitry E., Zagoruiko, Nikolay G., Andelfinger, U. (eds.) KONT/KPP -2007. LNCS (LNAI), vol. 6581, pp. 203–216. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22140-8_13

13. Zagorulko, Y., Zagorulko, G.: Architecture of extensible tools for development of intelligent decision support systems. In: Proceedings of the Tenth SoMeT_11 New Trends in Software Methodologies, Tools and Techniques, pp. 253–263. IOS Press, Amsterdam (2011)

14. Braginskaya, L., Kovalevsky, V., Grigoryuk, A., Zagorulko, G.: Ontological approach to information support of investigations in active seismology. In: The 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC-2017), pp. 27–29. IEEE Xplore Digital Library (2017). http://ieeexplore.ieee.org/document/8168060. Accessed 30 June 2018

15. Antoniou, G., Harmelen, F.: Web ontology language: OWL. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies, pp. 67–92. Springer, Berlin (2004). https://doi.org/10.1007/978-3-540-24750-0_4
16. Saaty, T.: Decision making with the analytic hierarchy process. Int. J. Serv. Sci. **1**(1), 83–98 (2008)