

УДК 004.822:004.89

Подход к реализации паттернов содержания при разработке онтологий научных предметных областей

Загорулько Ю.А. (Институт систем информатики СО РАН, Новосибирский государственный университет),

Боровикова О.И. (Институт систем информатики СО РАН)

Рассматривается подход к разработке и реализации такого вида паттернов онтологического проектирования, как паттерны содержания. Использование таких паттернов при построении онтологий научных предметных областей позволяет, с одной стороны, обеспечить единообразное и согласованное представление всех сущностей разрабатываемой онтологии, с другой – сэкономить человеческие ресурсы и избежать типичных ошибок онтологического моделирования.

Ключевые слова: *онтология, паттерны онтологического проектирования, паттерны содержания, научная предметная область.*

1. Введение

Онтологии широко используются для формализации знаний в научных предметных областях. С помощью онтологии можно не только удобно представить все необходимые понятия моделируемой области, но и обеспечить их единообразное и согласованное описание. (Заметим, что под научной предметной областью (НПО) здесь понимается предметная область, охватывающая некоторую научную дисциплину или область научных знаний во всех ее аспектах, включая характерные для нее объекты и предметы исследования, применяемые в ней методы исследования, выполняемую в ней научную деятельность и полученные в рамках ее исследований научные результаты.)

Разработка онтологии любой предметной области является довольно сложным и трудоемким процессом, поэтому для его облегчения предлагаются различные методологии и подходы. В последнее время интенсивно развивается подход, базирующийся на применении паттернов онтологического проектирования (Ontology Design Patterns или ODP) [3, 5, 12]. Согласно этому подходу ODP представляют собой документально зафиксированные описания проверенных на практике решений типовых проблем онтологического моделирования. Они создаются для того, чтобы упорядочить и облегчить процесс

построения онтологий и помочь разработчикам избежать типичных ошибок онтологического моделирования.

Несмотря на то, что использование паттернов онтологического проектирования позволяет сэкономить человеческие ресурсы и повысить качество разрабатываемых онтологий, в настоящее время только одна методология построения онтологий, а именно методология экстремального проектирования онтологий XD (eXtreme Design methodology) [9], предложенная в рамках проекта NeOn [14], открыто заявляет об использовании ODP.

Заметим также, что существует не так много инструментов разработки онтологий, поддерживающих использование паттернов онтологического проектирования. К ним, например, относится плагин для инструмента разработки онтологий проекта NeOn, а также плагин для редактора онтологий WebProtégé [13]. Однако эти средства покрывают только часть возможных задач, связанных с паттернами. Так, совсем нет инструментов, поддерживающих построение, поиск и извлечение паттернов из онтологий, и очень мало средств, поддерживающих сбор, обсуждение и распространение паттернов. К последним в какой-то мере можно отнести каталоги паттернов онтологического проектирования [8, 11, 15], также активно развиваемые в последнее время.

В статье рассматривается подход к реализации такого вида паттернов онтологического проектирования, как паттерны содержания [16], которые играют важную роль в предложенной авторами методологии разработки онтологий научных предметных областей [2, 3].

Работа выполнена при финансовой поддержке РФФИ (проект № 16-07-00569) и Президиума СО РАН (Блок 43.2. Комплексной программы ФНИ СО РАН II.1).

2. Введение в паттерны онтологического проектирования

Как было сказано выше, паттерны онтологического проектирования предназначены для описания решений типовых проблем, возникающих при разработке онтологий.

В зависимости от вида проблем, для решения которых создаются паттерны, различают структурные паттерны (Structural ODPs), паттерны соответствия (Correspondence ODPs), паттерны содержания (Content ODPs), паттерны логического вывода (Reasoning ODPs), паттерны представления (Presentation ODPs) и лексико-синтаксические паттерны (Lexico-Syntactic ODPs) [12].

Упрощенная версия типизации паттернов, предложенной в проекте NeOn [14], представлена на Рис.1.

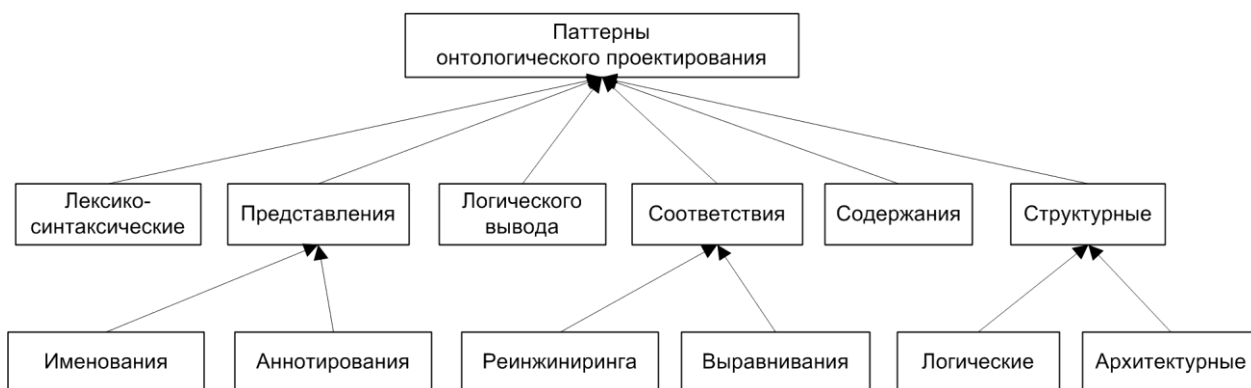


Рис.1. Типы паттернов онтологического проектирования.

Наиболее востребованными паттернами при разработке онтологий непосредственно инженерами знаний являются структурные паттерны, паттерны содержания и паттерны представления. Рассмотрим их подробнее.

Структурные паттерны либо фиксируют способы решения проблем, вызванных ограничениями выразительных возможностей языков описания онтологий, либо задают общую структуру и вид онтологии. Паттерны, предназначенные для преодоления проблемы выразительности языков, называются логическими структурными паттернами (Logical ODP). Паттерны, содержащие предложения по организации онтологии в целом, относятся к архитектурным паттернам (Architectural ODP).

Паттерны содержания задают способы представления типовых фрагментов онтологий, на основе которых могут строиться онтологии целого класса предметных областей. На основе этих паттернов можно описать более сложные (составные) паттерны содержания, дополняя их недостающими компонентами и вводя дополнительные отношения. Однако следует учитывать, что при использовании составных паттернов могут возникнуть проблемы их сборки и интеграции [6].

Паттерны представления определяют рекомендации по именованию, аннотированию и графическому представлению элементов онтологии. Применение этих паттернов должно повысить «читаемость» онтологии, а также удобство и простоту ее использования.

На данный момент не существует единого стандарта для описания паттернов, но чаще всего они описываются в формате, предложенном на портале ассоциации ODPА (Association for Ontology Design & Patterns) [8], созданном в рамках проекта NeOn [14]. В соответствии с этим форматом описание паттерна включает его графическое представление, текстовое описание, набор сценариев и примеров использования, ссылки на другие паттерны, в которых он используется, а также сведения о названии паттерна, его авторе и области применения.

Методология экстремального проектирования онтологий XD [9] предлагает также снабжать каждый паттерн содержания набором квалификационных вопросов (Competency questions). Эти вопросы могут использоваться как на этапе разработки паттернов, так и для поиска нужных паттернов при разработке конкретной онтологии.

Разработка и использование паттернов онтологического проектирования опирается на 5 основных операций над ними [12]: импорт (import), специализация (specialization), обобщение (generalization), композиция (composition) и расширение (expansion).

3. Разработка и применение паттернов содержания при построении онтологий НПО

При создании онтологии НПО важно обеспечить возможность единообразного и согласованного представления используемых в ней научных понятий и их свойств. Одним из способов решения этой проблемы является применение паттернов содержания, описывающих понятия, характерные для большинства научных предметных областей и выполняемой в них научной деятельности. Набор таких паттернов предоставляет методология построения онтологий, разработанная в рамках технологии создания тематических интеллектуальных научных интернет-ресурсов [2].

Следует заметить, что эта методология кроме паттернов содержания включает также паттерны представления и структурные логические паттерны [3, 17], которые используются в паттернах содержания. С помощью структурных логических паттернов, например, представляются атрибутированные отношения (бинарные отношения с атрибутами), многоместные отношения и области допустимых значений, заданные множествами элементарных значений. Последние, следуя традиции, принятой в реляционной модели данных, будем называть доменами.

На Рис.2 приведен паттерн представления домена, в котором сам домен задается перечислимым классом *Класс1*, включающим конечный набор различных индивидов (объектов или экземпляров класса) *объект1, ..., объектп*, определяющих все возможные значения некоторого свойства объектов класса *Класс2*. При этом класс *Класс1* является наследником служебного класса *Домен*.

На Рис. 3 представлен пример использования данного паттерна для описания домена “Географический тип”.

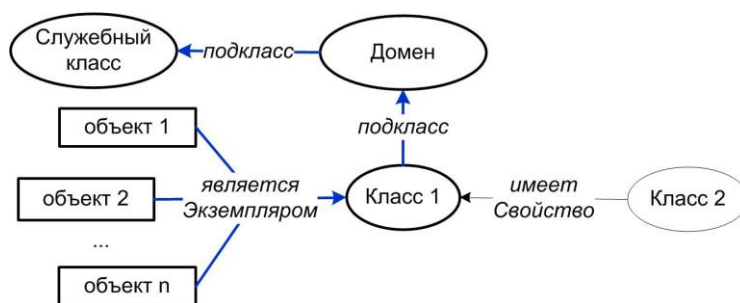


Рис.2. Структурный логический паттерн для представления домена.

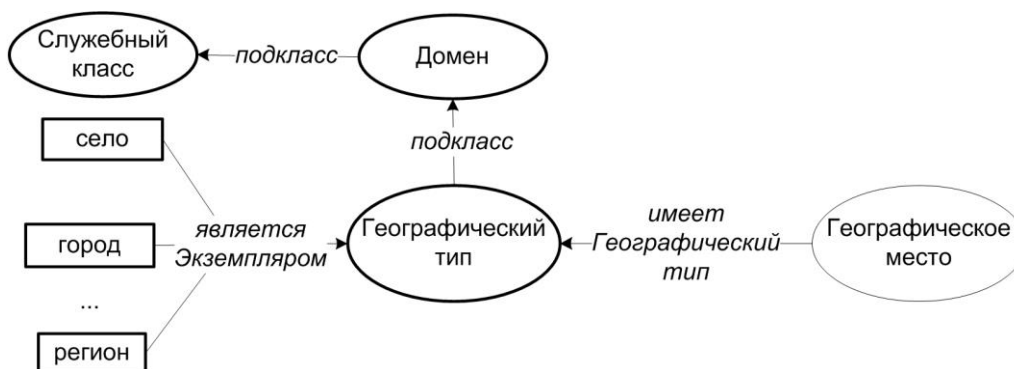


Рис.3. Пример использования паттерна для представления домена "Географический тип".

Заметим, что на рисунках, содержащих графическое представление паттернов, классы обозначаются в виде эллипсов, а индивиды и атрибуты – в виде прямоугольников. Связь типа «класс-класс» показывается сплошной прямой линией, а связь типа «класс-атрибут» – прерывистой. При этом обязательные связи показываются на рисунках линией синего цвета, а связываемые ими классы, атрибуты и индивиды представляются фигурами, обведенными жирной линией.

4.1. Разработка паттернов содержания

В рамках методологии построения онтологий НПО разработаны паттерны содержания для представления следующих понятий: *Объект исследования*, *Предмет исследования*, *Метод исследования*, *Раздел науки*, *Научный результат*, *Персона*, *Организация*, *Деятельность*, *Проект*, *Публикация* и др.

При разработке паттернов содержания для каждого из них был определен набор квалификационных вопросов, представляющих его содержание. С помощью этих вопросов был выявлен обязательный и факультативный состав элементов паттерна и описаны требования к ним, которые были представлены в виде аксиом и ограничений.

Рассмотрим некоторые паттерны содержания подробнее.

Как правило, научная деятельность реализуется через проекты, программы, экспедиции и экспертизы. Графическое представление паттерна, предназначенного для описания научной

деятельности, приведено на Рис.4. В этом паттерне отражено требование, состоящее в том, что любая деятельность должна иметь название и при ее описании необходимо давать ссылку на объект исследования, изучению которого посвящена данная деятельность, на раздел науки, по теме которого она выполняется, и на персону или организацию, входящую в состав ее участников.

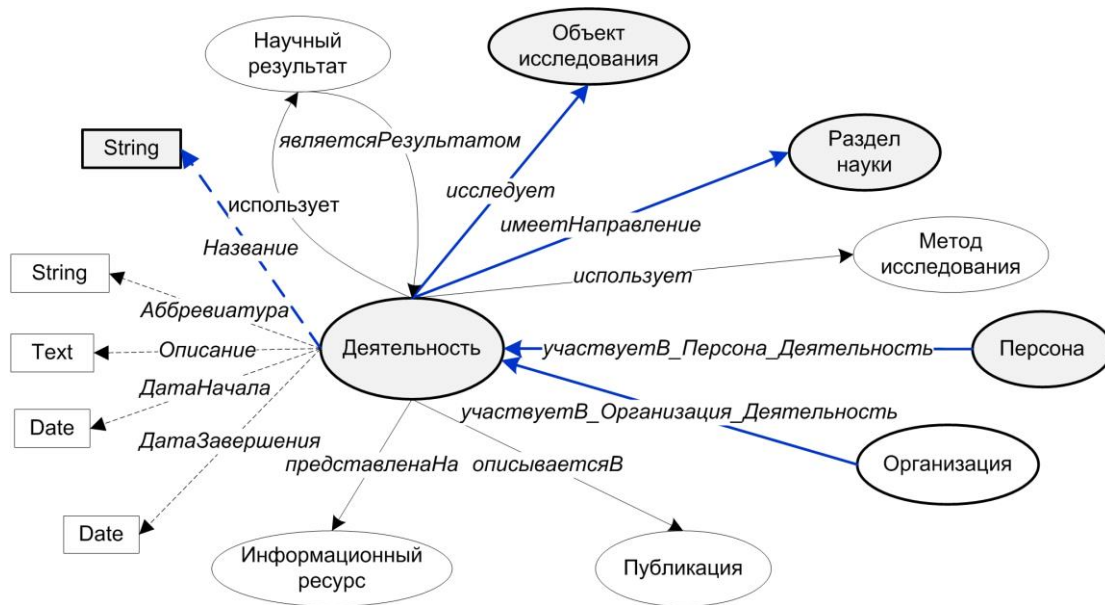


Рис.4. Паттерн для описания научной деятельности.

На Рис.5 приведен паттерн, предназначенный для описания методов исследования, используемых в научной деятельности. В соответствии с ним описание метода должно включать название и описание метода, а также ссылки на решаемые с помощью него задачи, раздел науки, в котором он используется, и автора метода (персону или организацию).

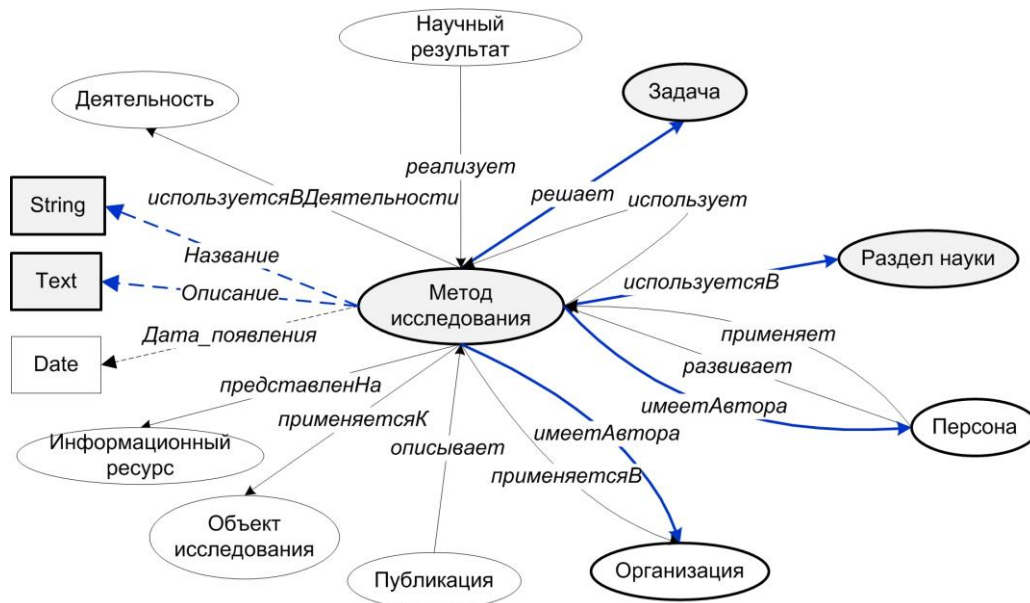


Рис.5. Паттерн для описания метода исследования.

На Рис.6 приведен паттерн, предназначенный для описания научного результата. В этом паттерне отражено требование: при описании научного результата необходимо указывать его название и давать ссылку на деятельность, при выполнении которой он был получен, и на публикацию, в которой он был зафиксирован.



Рис.6. Паттерн для описания научного результата.

Нетрудно заметить, что невозможно описать ни один паттерн содержания изолированно, т.е. не используя понятия из других паттернов. Например, в паттерне для описания научного результата (см. Рис. 6), кроме «центрального» понятия *Научный результат*, также используются понятия *Деятельность*, *Предмет исследования*, *Раздел науки* и др. Это объединяет все паттерны содержания в единую систему паттернов, а представляемые ими понятия в единую систему понятий, задающую связанное описание моделируемой предметной области.

Фактически, паттерны содержания являются фрагментами онтологии, которые после специализации содержащихся в них понятий и дополнения необходимыми понятиями и свойствами, становятся составными частями разрабатываемой онтологии НПО.

4.2. Реализация паттернов содержания

Все паттерны содержания рассматриваемой методологии реализованы средствами языка OWL [7]. Ниже приведен фрагмент OWL-реализации паттерна содержания, предназначенного для описания научной деятельности в формате Turtle. (Заметим, что в этом фрагменте используются названия онтологических элементов из пространства имен онтологии *inir*: http://www.semanticweb.org/IIS/ontologies/INIR_Ontology#.)

Опишем содержательные компоненты, которые включает этот паттерн.

Прежде всего, это набор используемых в нем классов и аксиома непересекаемости этих классов (*AllDisjointClasses*):

```
:ИнформационныйРесурс rdf:type owl:Class;  
:МетодИсследования rdf:type owl:Class;  
:ОбъектИсследования rdf:type owl:Class;  
:Организация rdf:type owl:Class;  
:Персона rdf:type owl:Class;  
:Публикация rdf:type owl:Class;  
:РазделНауки rdf:type owl:Class;  
:НаучныйРезультат  
[ rdf:type owl:AllDisjointClasses; owl:members  
( :Деятельность  
:ИнформационныйРесурс  
:МетодИсследования  
:ОбъектИсследования  
:Организация  
:Персона  
:Публикация  
:РазделНауки  
:НаучныйРезультат) ] .
```

Набор свойств класса *Деятельность* (в нотации языка OWL – *ObjectProperty* и *DatatypeProperty*) выглядит следующим образом:

```
:имеетНаправление_Деятельность_Раздел rdf:type owl:ObjectProperty;  
rdfs:domain :Деятельность;  
rdfs:range :РазделНауки;  
:исследует_Деятельность_Объект rdf:type owl:ObjectProperty;  
rdfs:domain :Деятельность;  
rdfs:range :ОбъектИсследования;  
:использует_Деятельность_Метод rdf:type owl:ObjectProperty;  
rdfs:domain :Деятельность;  
rdfs:range :МетодИсследования;
```



```

:описываетсяВ_Деятельность_Публикация rdf:type owl:ObjectProperty;
rdfs:domain :Деятельность;
rdfs:range :Публикация;
:участвуетВ_Персона_Деятельность rdf:type owl:ObjectProperty;
rdfs:domain :Персона;
rdfs:range :Деятельность;
:являетсяРезультатом_Результат_Деятельность rdf:type owl:ObjectProperty;
rdfs:domain :НаучныйРезультат;
rdfs:range :Деятельность;
:являетсяРесурсом_Ресурс_Деятельность rdf:type owl:ObjectProperty;
rdfs:domain :ИнформационныйРесурс;
rdfs:range :Деятельность;
:участвуетВ_Организация_Деятельность rdf:type owl:ObjectProperty;
rdfs:domain :Организация;
rdfs:range :Деятельность;
:Деятельность_Название rdf:type owl:DatatypeProperty;
rdfs:domain :Деятельность;
rdfs:range rdfs:Literal;
:Деятельность_Аббревиатура rdf:type owl:DatatypeProperty;
rdfs:domain :Деятельность;
rdfs:range rdfs:Literal;
:Деятельность_Описание rdf:type owl:DatatypeProperty;
rdfs:domain :Деятельность;
rdfs:range rdfs:Literal;

```

Далее для каждого экземпляра (представителя) класса *Деятельность* с помощью свойства *FunctionalProperty* накладывается требование к указанию только одной даты начала и одной даты завершения:

```

:Деятельность_ДатаЗавершения rdf:type owl:DatatypeProperty,
owl:FunctionalProperty;
rdfs:domain :Деятельность ;
rdfs:range rdfs:Literal;
:Деятельность_ДатаНачала rdf:type owl:DatatypeProperty, owl:FunctionalProperty;

```

```
rdfs:domain :Деятельность ;
```

```
rdfs:range rdfs:Literal;
```

Наконец для каждого экземпляра класса *Деятельность* путем накладывания ограничений на свойства описываются следующие требования на количество и тип, связанных с этим классом понятий:

- должно быть задано хотя бы одно направление деятельности;
- должен быть задан хотя бы один объект исследования;
- в деятельности должна участвовать, по крайней мере, одна персона;
- у деятельности должно быть, по крайней мере, одно название.

```
:Деятельность rdf:type owl:Class;
```

```
owl:equivalentClass
```

```
[ rdf:type owl:Restriction;
```

```
owl:onProperty :имеетНаправление_Деятельность_Раздел ;
```

```
owl:someValuesFrom : РазделНауки] ,
```

```
[ rdf:type owl:Restriction;
```

```
owl:onProperty :исследует_Деятельность_Объект ;
```

```
owl:someValuesFrom :ОбъектИсследования] ,
```

```
[rdf:type owl:Restriction;
```

```
owl:onProperty [owl:inverseOf :участвуетВДеятельности] ;
```

```
owl:minQualifiedCardinality "1"^^xsd:nonNegativeInteger;
```

```
owl:onClass [ rdf:type owl:Class ;owl:unionOf (:Организация :Персона )]
```

```
[ rdf:type owl:Restriction;
```

```
owl:onProperty :Деятельность_Название ;
```

```
owl:minCardinality "1"^^xsd:nonNegativeInteger] ;.
```

4.3. Поддержка использования паттернов содержания

Методология построения онтологий НПО с использованием паттернов содержания поддерживается нашим редактором данных, который позволяет пополнять онтологию индивидами классов, для которых реализованы паттерны содержания. При этом редактор по имени класса, выбранного пользователем, находит соответствующий ему паттерн и на его основе строит форму, содержащую поля для заполнения свойств объекта этого класса (см. Рис.7).

При создании объекта учитываются ограничения на значения и кардинальность его свойств, заданные в паттерне. Заметим, что такие ограничения используются не только для контроля вводимой информации, но и в качестве потенциального источника значений свойств создаваемого (редактируемого) объекта. Так, в зависимости от типа ограничений пользователь может выбрать значения для каждого свойства либо из описания домена, заданного соответствующим структурным паттерном в качестве области значений этого свойства, либо из динамически сформированного по текущей версии онтологии списка индивидов класса, указанного в качестве области значений этого свойства.

Рис.7. Окно редактора данных

На Рис.7 показано окно редактора данных в момент создания объекта с названием *Метод недоопределенных вычислений* для онтологии НПО «Поддержка принятия решений в слабоформализованных областях» [1, 3, 4] на основе паттерна метода исследования (см. Рис. 5). При создании этого объекта контролируется наличие значений у его атрибутов *название* и *описание*, а также существование ссылок на объекты классов *Персона*, *Задача* и *Раздел науки* (учитываются ограничения на кардинальность связей *имеет Автора-Персону*, *решает Задачу* и *используется В Разделе Науки*). При этом, когда пользователь переходит к заданию автора метода, автоматически формируется список из индивидов класса *Персона* (см. окно с

фамилиями и инициалами в правом нижнем углу Рис.7), из которого пользователь может выбрать нужную персону.

5. Заключение

В статье обсуждены вопросы применения паттернов онтологического проектирования для разработки онтологий научных предметных областей. Показано, что при использовании таких паттернов обеспечивается единообразное и согласованное представление всех сущностей разрабатываемой онтологии, экономятся человеческие ресурсы и сокращается число ошибок при разработке онтологий.

Представлен подход к реализации паттернов содержания, играющих ведущую роль в предлагаемой методологии разработки онтологий НПО. Данная методология и описанные в статье паттерны содержания показали свою практическую полезность при разработке онтологий различных научных предметных областей («Поддержка принятия решений в слабоформализованных областях» [1, 4], «Активная сейсмология» [10] и др.).

Список литературы

1. Загорулько Г.Б. Разработка онтологии для интернет-ресурса поддержки принятия решений в слабоформализованных областях // Онтология проектирования. 2016. Т. 6, №4(22). С. 485-500.
2. Загорулько Ю.А., Загорулько Г.Б., Боровикова О.И. Технология создания тематических интеллектуальных научных интернет-ресурсов, базирующаяся на онтологии // Программная инженерия. 2016. Т. 7. № 2. С. 51-60.
3. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. Применение паттернов онтологического проектирования при разработке онтологий научных предметных областей // Аналитика и управление данными в областях с интенсивным использованием данных: сборник научных трудов XIX Международной конференции DAMDID/RCDL'2017. Москва: ФИЦ ИУ РАН, 2017. С. 332-340.
4. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б., Шестаков В.К. Использование паттернов для разработки онтологии информационно-аналитического интернет-ресурса «Поддержка принятия решений» // Информационные и математические технологии в науке и управлении. 2017. № 3. С. 144-153.
5. Ломов П.А. Применение паттернов онтологического проектирования для создания и использования онтологий в рамках интегрированного пространства знаний // Онтология проектирования, 2015. Т. 5, № 2(16). С.233-245.

6. Ломов П.А. Автоматизация синтеза составных онтологических паттернов содержания // Онтология проектирования. 2016. Т.6, №2(20). С.162-172. DOI: 10.18287/2223-9537-2016-6-2-162-172.
7. Antoniou G., Harmelen F. Web Ontology Language: OWL // Handbook on Ontologies. Berlin: Springer Verlag, 2004. P. 67-92.
8. Association for Ontology Design & Patterns, 2018. URL: <http://ontologydesignpatterns.org> (дата обращения: 22.10.2018).
9. Blomqvist E., Hammar K., Presutti V. Engineering Ontologies with Patterns: The eXtreme Design Methodology. P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti (eds.). Ontology Engineering with Ontology Design Patterns. Studies on the Semantic Web, IOS Press, 2016. P. 23-50.
10. Braginskaya L., Kovalevsky V., Grigoryuk A., Zagorulko G. Ontological approach to information support of investigations in active seismology // Proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russky Island, Russia, 25-29 September, 2017. P. 27-29. IEEE Xplore digital library. URL: <http://ieeexplore.ieee.org/document/8168060> (дата обращения: 22.10.2018) ISBN: 978-1-5386-1206-4. DOI: 10.1109/RPC.2017.8168060.
11. Dodds L., Davis I. Linked Data Patterns. 2012. URL: <http://patterns.dataincubator.org/book> (дата обращения 22.10.2018).
12. Gangemi A., Presutti V. Ontology Design Patterns // Handbook on Ontologies. Springer, 2009. P. 221-243.
13. Hammar K. Ontology Design Patterns in WebProtégé. Proceedings of 14th International Semantic Web Conference (ISWC-2015). Posters & Demonstrations Track. CEUR Workshop Proceedings, 2015. Vol. 1486. URL: http://ceur-ws.org/Vol-1486/paper_50.pdf (дата обращения: 22.10.2018).
14. NeOn Project, 2018. URL: http://neon-project.org/nw/Welcome_to_the_NeOn_Project.html (дата обращения: 22.10.2018).
15. Ontology Design Patterns (ODPs) Public Catalog, 2009. URL: <http://odps.sourceforge.net> (дата обращения: 22.10.2018).
16. Presutti V., Daga E., Gangemi A., Blomqvist E. eXtreme Design with Content Ontology Design Patterns. Proceedings of the Workshop on Ontology Patterns (WOP 2009), Washington D.C., USA, 2009. Vol. 516. P. 83-97.
17. Zagorulko Y., Borovikova O., Zagorulko G. Pattern-Based Methodology for Building the Ontologies of Scientific Subject Domains. In: New Trends in Intelligent Software Methodologies, Tools and Techniques. Proceedings of the 17th International Conference SoMeT_18. H. Fujita and E. Herrera-Viedma (Eds.). Series: Frontiers in Artificial Intelligence and Applications, Vol. 303. Amsterdam: IOS Press, 2018. P. 529–542.

