8. Шевченко Е.В., Стукач В.Ф. Технологическая платформа – модель инновационного взаимодействия государства, промышленности, науки и образования // Международный научно-исследовательский журнал. – Екатеринбург, 2021. – № 4-4 (106). – С. 183-187.

УДК 004.822:004.89

DOI: 10.26898/agroinfo-2021-73-77

*Загоруйко Ю.А., Ахмадеева И.Р., Шестаков В.К.,
Институт систем информатики им. А.П. Ершова СО РАН, г. Новосибирск*

ПОСТРОЕНИЕ ОНТОЛОГИЙ НАУЧНЫХ ПРЕДМЕТНЫХ ОБЛАСТЕЙ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ ПАТТЕРНОВ ОНТОЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ

В статье представлен подход к разработке онтологий научных предметных областей, базирующийся на системе паттернов онтологического проектирования. Такие паттерны предназначены для формального описания решений типовых проблем, возникающих при разработке онтологий. Их применение позволяет не только обеспечить согласованное представление всех сущностей онтологии и избежать многих ошибок онтологического моделирования, но и сэкономить ресурсы, затрачиваемые на разработку онтологий.

Ключевые слова: научная предметная область, онтология, онтологическое моделирование, паттерн онтологического проектирования.

В настоящее время наиболее эффективным и популярным средством концептуализации и формализации научных предметных областей (НПО) являются онтологии [1]. Формализация семантики предметной области в виде онтологии служит не только целям компактного и непротиворечивого ее описания, она также формирует понятийный базис для представления всей совокупности знаний о ней. Например, в системе информационной поддержки научной деятельности в терминах онтологии может быть описана семантика используемых в ней данных и информационных ресурсов, а в экспертной системе и СППР – экспертные правила, прецеденты и другие компоненты базы знаний.

Разработка онтологий научных предметных областей является довольно сложным и трудоемким процессом. Для его упрощения и облегчения предложены различные методы и подходы. В последнее время интенсивно развивается подход, использующий паттерны онтологического проектирования (Ontology Design Patterns, сокращенно – ODPs) [2-4]. Согласно ему паттерны онтологического проектирования (паттерны ОП) представляют собой документально зафиксированные описания проверенных на практике решений типовых проблем онтологического моделирования.

Паттерны ОП имеют в качестве своих прародителей шаблоны проектирования (design pattern), широко используемые в разработке программного обеспечения. Аналогично шаблонам проектирования, паттерны ОП предназначены для описания решений типичных проблем, возникающих при разработке онтологий [4].

В зависимости от проблем, для решения которых предназначены паттерны ОП, 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). На рисунке 1 представлена упрощенная типология паттернов ОП, предложенная в проекте NeOn [5].

При разработке онтологий инженерами знаний в основном используются структурные паттерны, паттерны содержания и паттерны представления.

Структурные паттерны либо фиксируют различные способы решения проблем, вызванных ограничениями выразительных возможностей языков описания онтологий, либо задают общую (модульную) структуру и вид онтологии. Паттерны первого типа называются логиче-

скими паттернами (Logical ODPs), паттерны второго типа – архитектурными паттернами (Architectural ODPs).

Паттерны содержания задают способы представления типовых фрагментов онтологий, на основе которых могут строиться онтологии конкретных предметных областей.

Рисунок 1. Типология паттернов онтологического проектирования

Паттерны представления определяют рекомендации (правила) по именованию и аннотированию элементов онтологии, применение которых должно повысить читаемость онтологии, а также удобство и простоту ее использования.

Лексико-синтаксические паттерны применяются для автоматизации процесса построения (пополнения) онтологии на основе текстов на естественном языке. Они задают отображения языковых структур в элементы онтологии (классы, свойства, индивидумы).

В настоящее время создано и развивается несколько каталогов паттернов ОП. Наиболее представительный из них размещен на портале Ассоциации ODPA (Association for Ontology Design & Patterns) [6], созданном в рамках проекта NeOn [5].

Для поддержки процесса построения онтологий научных предметных областей нами разработана система паттернов онтологического проектирования [7], включающая четыре типа паттернов: структурные логические паттерны, паттерны содержания, паттерны представления и лексико-синтаксические паттерны. При этом одна часть этих паттернов является универсальной, другая часть ориентирована непосредственно на представление научных знаний. Все упомянутые паттерны ОП реализованы на языке OWL [8].

Необходимость в использовании структурных логических паттернов возникла из-за недостатка в языке OWL выразительных средств для представления сложных сущностей и конструкций, актуальных при построении онтологий НПО, в частности, областей допустимых значений, многоместных и атрибутированных отношений (бинарных отношений с атрибутами).

Рассмотрим структурный паттерн, разработанный для представления атрибутированных отношений (см. рисунок 2).

Следует заметить, что на рисунках, задающих графическое представление паттернов, приведенных ниже, классы обозначаются в виде эллипсов, а их индивиды (экземпляры) и атрибуты – в виде прямоугольников. Связи между объектами показываются сплошной прямой линией, а связи с атрибутами – прерывистой. При этом обязательные классы, атрибуты и индивиды представляются заштрихованными фигурами, обведенными жирной линией. Для представления обязательных связей также используется жирная линия.

Центральное место в данном паттерне занимает служебный класс *Отношение с атрибутами*, с которым связываются классы, моделирующие аргументы бинарного отношения. При этом атрибуты бинарного отношения моделируются свойствами данного класса (в нотации языка OWL либо *DataProperty*, либо *ObjectProperty*). Для представления конкретного типа отношения вводится новый класс, являющийся наследником класса *Отношение с атрибутами*. Подобным образом построен паттерн для многоместного отношения.

Для единообразного и непротиворечивого представления используемых в НПО понятий и их свойств были разработаны паттерны содержания, описывающие основные понятия ба-

зовых онтологий. Благодаря этому, разработка онтологии конкретной НПО в основном состоит в специализации паттернов содержания и построении на их основе фрагментов целевой онтологии.

Рисунок 2. Структурный паттерн бинарного атрибутивного отношения

В систему включены паттерны содержания для представления следующих понятий: *Объект исследования*, *Предмет исследования*, *Метод исследования*, *Раздел науки*, *Научный результат*, *Научная деятельность*, *Проект*, *Персона*, *Организация*, *Публикация* и др.

При разработке паттернов содержания для каждого из них был определен набор квалификационных вопросов, представляющих его содержание [4]. С помощью этих вопросов был выявлен обязательный и факультативный состав онтологических элементов паттерна и описаны требования к ним, которые были представлены в виде аксиом и ограничений.

В качестве примера приведем паттерн содержания, предназначенный для описания выполняемого проекта. Графическое представление этого паттерна приведено на рисунке 3. В этом паттерне отражено требование, состоящее в том, что при описании любого проекта необходимо указать его полное название, дать ссылку на объект исследования, изучению/обработке которого посвящен данный проект, на раздел науки, по теме которого проект выполняется, а также на полученный в ходе выполнения проекта научный результат.

Рисунок 3. Паттерн описания проекта

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

Паттерны представления, входящие в систему, большей частью служат для задания правил именования и аннотирования элементов онтологии, подобных общепринятым правилам разработки онтологии в редакторе Protégé [9].

В предлагаемом подходе, базирующемся на совместном использовании паттернов ОП различного типа, предполагается автоматическое построение лексико-синтаксических паттернов на основе паттернов содержания и структурных логических паттернов.

На данный момент разработаны лексико-синтаксические паттерны (шаблоны) для извлечения информации о научной деятельности из сети Интернет [10]. Каждый такой паттерн представляет собой XML-документ, в котором для класса онтологии, описанного паттерном содержания, указаны текстовые маркеры, сигнализирующие об обнаружении на обрабатываемой HTML-странице описания объекта данного класса, его отношений и атрибутов. В шаблонах для каждого типа извлекаемой информации указываются обработчики, реализующие алгоритмы обхода и анализа соответствующих фрагментов интернет-страниц.

В качестве примера лексико-синтаксического паттерна приведем шаблон для извлечения информации о проекте с сайта (странички) научно-исследовательского проекта. Этот шаблон строится на основе паттерна содержания, описывающего проект, и имеет следующий вид:

```
<Class Name= «Проект» engine = FragmentSearch >
  <Marker Term = «О проекте» PType=Menu FragType=Page />
  <Marker Term = «Проект» PType=Head FragType=Block />
  <Attr Name= «Название» type= string engine = NameEntity >
    <Marker Term = «Проект» PType=link FragType=LinkText/>
    <Marker Term = «Проект» PType=sentence FragType=QuoteText/>
    <Marker Term = «Проект» PType=Head FragType = Head />
  </Attr>
  <Attr Name = «Описание» type=text >
    <Marker Term = «Аннотация» PType = Head FragType=Block />
    <Marker Term = «О проекте» PType=Menu FragType=Page />
  </Attr>
  <Relation Name = «ОписываетсяВ» >
    <Marker Term = «Публикации» PType= Menu FragType=Page />
    <Marker Term = «Литература» PType= Head FragType=Block />
    <Object Name = «Публикация» engine = PublicationList />
  </Relation>
  <Relation Name= «УчастствуетВ_Персона_Проект» >
    <Marker Term= «Об участниках» PType=Menu FragType=Page />
    <Marker Term= «Список участников» PType= Head FragType=Block/>
    <Object Name= «Персона» engine = PersonList />
  </Relation> </Class>
```

Данный шаблон позволяет извлекать такие атрибуты класса *Проект*, как *Название* и *Описание*, а также аргументы отношений *УчастствуетВ_Персона_Проект* и *ОписываетсяВ*, т. е. объекты, описывающие соответственно участников проекта и публикации о проекте.

Параметр маркера Term позволяет задать термин тезауруса, характеризующий смысл извлекаемой информации (например, список публикаций или участники проекта); параметр PType задает тип фрагмента, в тексте которого должен располагаться указанный термин (например, меню или заголовок), а параметр FragType – тип фрагмента, из которого будет извлекаться информация (например, блок или страница).

Для систематизации информации о разработанных паттернах ОП и обеспечения к ней содержательного доступа инженерам знаний – разработчикам онтологий НПО – разработан информационно-аналитический интернет-ресурс «Паттерны онтологического проектирова-

ния». Концептуальной основой данного ресурса является разработанная нами онтология паттернов ОП. Эта онтология содержит формальные спецификации понятий данной области знаний, типов паттернов, методов и средств их разработки и применения. По этим спецификациям строятся структуры для представления информации о реальных объектах моделируемой области знаний, интегрируемых информационных ресурсах и методах разработки и использования паттернов. На основе онтологии также организуются удобная навигация по контенту ресурса и содержательный поиск необходимой информации.

В заключение отметим, что использование паттернов онтологического проектирования при разработке онтологии НПО позволяет: обеспечить единообразное и согласованное представление всех сущностей онтологии, сократить число ошибок онтологического моделирования, повысить «понимаемость» онтологии разработчиками, обеспечить возможность коллективной разработки онтологий и сэкономить человеческие ресурсы.

Данный подход успешно применялся при разработке онтологий для таких научных областей, как «Поддержка принятия решений в слабоформализованных областях» [11], «Активная сейсмология» [12], «Поддержка решения вычислительно сложных задач математической физики на суперкомпьютерах» [13] и др.

Работа выполнена при поддержке РФФИ (проект № 19-07-00762).

Использованные источники:

1. Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems. R. Sharman, R. Kishore, and R. Ramesh (Eds.). – New York: Springer, 2007. – 930 p.
2. Gangemi A., Presutti V. Ontology Design Patterns / A. Gangemi, V. Presutti // Handbook on Ontologies. Staab, S. and R. Studer (Eds.). – Berlin: Springer Verlag, 2009. – P. 221-243.
3. Blomqvist E., Hammar K., Presutti V. Engineering Ontologies with Patterns: The eXtreme Design Methodology / E. Blomqvist, K. Hammar, V. Presutti // Ontology Engineering with Ontology Design Patterns. Studies on the Semantic Web. Hitzler, P., and A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti (Eds.). – IOS Press, 2016. – P. 23-50.
4. Karima N., Hamma, K., Hitzler P. How to Document Ontology Design Patterns / N. Karima, K. Hammar, P. Hitzler// Advances in Ontology Design and Patterns, 32. – IOS Press, Kobe, Japan, 2017. – P. 15-27.
5. NeOn Project. – http://neon-project.org/nw/Welcome_to_the_NeOn_Project.html (дата обращения: 01.09.2019).
6. Портал ассоциации ODP (Association for Ontology Design & Patterns). – <http://ontologydesignpatterns.org> (дата обращения: 01.09.2019).
7. Zagorulko Y., Borovikova O., Zagorulko G. Development of Ontologies of Scientific Subject Domains Using Ontology Design Patterns / Y. Zagorulko, O. Borovikova, G. Zagorulko // Communications in Computer and Information Science. – Heidelberg: Springer, Cham, 2018. – Vol. 822. – P. 141-156.
8. Antoniou G., Harmelen F. Web Ontology Language: OWL / G. Antoniou, F. Harmelen // Handbook on Ontologies. Staab, S. and R. Studer (Eds.). – Berlin: Springer Verlag, 2009. – P. 91-110.
9. Noy N., McGuinness, D. Ontology Development 101: A Guide to Creating Your First Ontology / N. Noy, D. McGuinness // – Stanford, 2001. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.
10. Загорюлько Ю.А., Ахмадеева И.П., Серый А.С. Автоматизация сбора информации о научной деятельности для тематических интеллектуальных научных интернет-ресурсов // Аналитика и управление данными в областях с интенсивным использованием данных: XVII Международная конференция DAMDID/RCDL'2015 (Обнинск, 13-16 октября 2015, Россия): труды конференции / под. ред. Л.А. Калиниченко, С.О. Старкова. – Обнинск: ИАТЭ НИЯУ МИФИ, 2015. – С. 105-111.
11. Загорюлько Г.Б. Разработка онтологии для интернет-ресурса поддержки принятия решений в слабоформализованных областях / Г.Б. Загорюлько // Онтология проектирования. – 2016. – Т. 6, – № 4 (22). – С. 485–500.
12. Braginskaya, L., Kovalevsky, V., Grigoryuk, A., Zagorulko, G. Ontological approach to information support of investigations in active seismology / L. Braginskaya, V. Kovalevsky, A. Grigoryuk, G. Zagorulko // The proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russky Island, Russia, 25-29 September 2017. – P. 27-29.
13. Snytnikov A.V., Glinskiy B.M., Zagorulko G.B., Zagorulko Y.A. Ontological approach to formalization of knowledge in computational plasma physics // Third Virtual Workshop on Numerical Modeling in MHD and Plasma Physics: 12-16 October 2020, Novosibirsk, Russia. Journal of Physics: Conference Series, 2020. – Vol. 1640, – P. 012013. doi:10.1088/1742-6596/1640/1/012013