

УДК 618.3

ОНТОЛОГИИ: СОВРЕМЕННОЕ СОСТОЯНИЕ, КРАТКИЙ ОБЗОР

Н.М. Боргест¹, М.Д. Коровин²

*Самарский государственный аэрокосмический университет имени академика С.П. Королева
(национальный исследовательский университет)*

¹*borgest@yandex.ru* ²*maks.korovin@gmail.com*

Аннотация

В статье делается попытка описать современное состояние и тенденции в области создания и внедрения онтологий. Представлен краткий обзор новейших средств поддержки онтологий и применяемых стандартов для построения онтологий, приведены актуальные положения международной ассоциации по прикладным онтологиям в области оценки функциональных и структурных качеств онтологий разрабатываемых сложных информационных систем в различных предметных областях. Представлены фазы жизненного цикла онтологии как компоненты информационной системы.

Ключевые слова: онтологии, обзор, жизненный цикл, стандарты, информационные системы.

Введение

Международная ассоциация прикладных онтологий подготовила Коммюнике от 15 июня 2013 года по вопросу оценки онтологий в течение всего жизненного цикла [1]. Этот документ, подготовленный в форме рекомендаций, отражает последние тенденции в области создания и поддержки прикладных онтологий как компонентов информационных систем. Использование разрабатываемых семантических моделей предметных областей и происходящих в них процессов позволит создать новый класс интеллектуальных систем с высокой степенью автоматизации. Бурное обсуждение Коммюнике членами Международной ассоциации прикладных онтологий на форуме разработчиков подтолкнуло авторов к написанию краткого обзора современного состояния в области создания и использования онтологий. В оригинале (на английском языке) Коммюнике будет опубликовано в журнале «Прикладные онтологии» («Applied Ontology» Journal). В нашем номере журнала с согласия разработчиков Коммюнике публикуется его перевод на русский язык, выполненный сотрудниками редакции (см. с. 66-74).

1 Стандарты и онтологии

Развитие работ в области информационных компьютеризированных систем и онтологий, накопление успешных практик и обобщающих методик подталкивает специалистов предметных областей (ПрО), баз данных и онтологов к разработке стандартов, фиксирующих полезные начинания. Некоторые примеры таких успешных «рекомендаций» рассмотрены ниже.

1.1 ISO 10303 (STEP)

Стандарты ISO 10303 определяют средства описания (моделирования) промышленных изделий на всех стадиях жизненного цикла. Проект STEP развивается с середины 80-х годов прошлого века. Единообразная форма описаний данных о промышленной продукции обес-

печивается введением в STEP языка Express, инвариантного к приложениям. Первая версия стандарта ISO 10303-11, посвященного языку Express, опубликована в 1990 г. В стандартах STEP использован ряд идей, ранее воплощенных в методиках информационного IDEF1X и функционального IDEF0 проектирования. Но роль стандартов STEP не ограничивается введением только грамматики единого языка обмена данными. В рамках STEP предпринята попытка создания единых информационных моделей (онтологий) целого ряда приложений. Эти модели получили название прикладных протоколов. В качестве альтернативного языка для обмена геометрическими и техническими данными о промышленных изделиях может использоваться язык разметки XML. В 2004 г. компаниями Dassault Systemes и Lattice Technology предложено подмножество 3D XML языка XML, которое получает все большую популярность для межсистемных обменов в CALS-технологиях.

Стандарт ISO 10303 состоит из целого ряда документов (томов), в которых описываются основные принципы STEP, правила языка Express, даны методы его реализации, модели, ресурсы, как общие для приложений, так и некоторые специальные (например, геометрические и топологические модели, описание материалов, процедуры черчения, конечно-элементного анализа и т.п.), прикладные протоколы, отражающие специфику моделей в конкретных предметных областях, а также методы тестирования моделей и объектов. Удовлетворению требований создания открытых систем в STEP уделяется основное внимание — специальный раздел посвящен правилам написания файлов обмена данными между разными системами, созданными в рамках STEP-технологии.

1.2 ISO 22745 (eOTD)

Международная ассоциация управления кодами электронной коммерции (ECCMA) активно продвигает стандарты ISO 22745 (Системы промышленной автоматизации и интеграции), ISO 8000 (Качество информационных технологий) и eOTD (открытый технический словарь).

Стандарт ISO 22745 включает в себя словарь, представляющий собой совокупность терминов, определений и концепций, применяемых для описания отдельных объектов, организаций, адресов, товаров и услуг. В комплексе стандартов ISO 22745 описаны элементы данных, относящиеся к конкретным классам и парам значений свойств.

Открытый технический словарь позволяет точно определить свойства в соответствии с данными ISO 10303, определять информацию и обмениваться данными с партнерами из других стран без искажения смысла данных.

1.3 ISO (P-LIB)

ISO 13584 PartsLibrary - это серия международных стандартов для представления и обмена доступными для компьютерной интерпретации данными о поставляемых компонентах и комплектующих изделиях (узлах, деталях).

Стандарт ISO 13584 (P-LIB) включает в себя 7 разделов:

- общий обзор и основополагающие принципы;
- концептуальная модель библиотеки деталей;
- основные ресурсы;
- логическая модель библиотеки поставщика;
- данные о поставщике;
- программный интерфейс к данным;
- методология структуризации классов (семейств) деталей.

Стандарт ISO 13584 регламентирует:

- средства описания и технологию представления информации о компонентах и комплектующих;
- технологию обработки данных, в том числе хранения, передачи, доступа изменения и архивирования.

В отличие от стандарта ISO 10303 STEP, предназначенного для описания конкретного экземпляра продукта, стандарт P-LIB позволяет описывать классы продуктов (компонентов и комплектующих):

- стандартные детали, определенные международными или национальными стандартами (например, крепеж или подшипники);
- библиотеки (базы) данных о деталях конкретного поставщика.

1.4 ISO 15926

На сегодняшний день одним из наиболее перспективных стандартов организации онтологических баз данных является ISO 15926. Данный стандарт определяет структуру объектов. В нем специфицируется модель данных, определяющая значение сведений о жизненном цикле в едином контексте. Эталонная модель данных (справочная онтология) отражена в библиотеке справочных данных (RDL). Интеграция приложения в информационное пространство требует приведения в соответствие классов и атрибутов прикладной модели этого приложения к классам и атрибутам эталонной модели (операционная онтология) [2].



Рисунок 1 – Структура классов в стандарте ISO 15926

Стандарт ISO 15926 содержит несколько частей.

- ISO 15926-1 содержит обзор ISO 15926.
- ISO 15926-2 специфицирует обобщенную концептуальную модель данных, поддерживающую представление всех аспектов жизненного цикла установки непрерывного производства.
- ISO 15926-4 определяет библиотеку справочных данных, которая может периодически обновляться компетентным органом, назначенным ISO в качестве регистратора, обла-

дающего требуемой инфраструктурой для обеспечения действенного применения данной справочной библиотеки данных.

- ISO 15926-5 специфицирует инструкции для регистратора, касающиеся справочных данных.
- ISO 15926-6 специфицирует сведения, требуемые при определении дополнений к справочным данным, специфицированным ISO 15926-4.

2 Онтологические редакторы и средства поддержки онтологий

Проектирование семантически связанных хранилищ данных – современная быстроразвивающаяся отрасль. К новейшим разработкам в области онтологического инжиниринга относятся:

2.1 15926 Editor

Разработанный компанией Techinvestlab (Россия) программный продукт для работы с данными в формате ISO 15926 позволяет создавать онтологии ПрО в рамках базовой онтологии, изложенной во второй части стандарта ISO 15926. Данный редактор в состоянии обрабатывать любые RDF-совместимые наборы данных. Особенностью данной программы является возможность работы с большими объемами информации, что недоступно большинству открытых онторедкторов. Основные задачи, решаемые ISO 15926, включают в себя получение информации из большинства существующих онтологий в максимально возможном количестве форматов, проверку справочных данных, создание новых справочных данных, в том числе позволяя автоматическую их генерацию из существующих источников [3].

2.2 COLORE

Common Logic Ontology Repository - это проект, разработанный в университете Торонто, целью которого является создание репозитория онтологий верхнего уровня, который бы служил в качестве испытательной площадки при решении задач оценки онтологий и методов их интеграции в рамках логики первого порядка. Языком описания онтологий для данного проекта был выбран недавно принятый стандарт ISO 24707, являющийся языком спецификации онтологий верхнего уровня и баз знаний [4].

2.3 НуQue

Система НуQue, разработанная в университете Карлтона, предназначена для проверки соответствия формализованных гипотез экспериментальным данным (через ряд SPARQL запросов) и определенным формальным правилам. НуQue использует наборы правил, специфические для конкретных ПрО. Результаты работы системы представляются в форматах RDF и OWL [5].

2.4 Makleod

Данная программа состоит из набора скриптов, призванных оказывать поддержку в решении типовых задач при создании онтологий. В настоящий момент разработчики проекта решают задачу автоматизации операций, не затрагивающих семантику. К таким операциям можно отнести, например, проверку согласованности онтологий. Целевыми онтологиями для данной программы являются онтологии, описанные на языке Common Logic. В ближайшем бу-

дущем планируется объединить данный продукт с проектом COLORE, в рамках которого он будет использоваться для оценки и поддержки онтологий, хранящихся в репозитории [6].

2.5 NCBO BioPortal

Интернет-портал NCBO BioPortal предоставляет доступ к часто используемым онтологиям в биомедицинской сфере и инструментам для работы с ними. Портал поддерживает поиск по нескольким онтологиям одновременно, позволяет создавать и обозревать связи между терминами в разных онтологиях. Портал также позволяет создавать аннотации к текстам в RDF формате [7].

2.6 OntoHub

OntoHub - это программа для управления распределенными гетерогенными онтологиями. Она создана в университете Бремена для поддержки создания и интеграции онтологий, написанных на разных языках. OntoHub поддерживает большинство известных языков описания онтологий наряду с языками формальной логики при создании сложных междисциплинарных связей и отношений в рамках формальной семантики [8].

2.7 OntologyTest

Разработанный в университете Сан-Пабло (Испания) набор инструментов OntologyTest предназначен для проверки соответствия онтологии функциональным требованиям. Программа выполнена в виде набора тестов, которые можно выполнять независимо друг от друга, что позволяет упростить процедуру проверки онтологий как в процессе создания, так и на этапе их эксплуатации [9].

2.8 OntoQA

Программное средство оценки и анализа онтологий, использующее набор метрик для определения и оценки различных параметров онтологии и базы знаний. Метрики подразделяются на две категории – оценка структуры организации онтологии и то, как структурированы сами записи в онтологии. Отличительной чертой данного продукта является его способность взаимодействовать с онтологией, находящейся в стадии разработки, что значительно упрощает процесс анализа [10].

2.9 OOPS

Ontology Pitfall Scanner - это web-приложение, предназначенное для обнаружения потенциальных проблем в логике онтологии, которые могли бы привести к ошибкам в моделировании ПрО. Данная программа позволяет создателям онтологий проверять свои системы на наличие логических ошибок на этапе создания. Примером ошибок, обнаруживаемых данным инструментом, является зацикливание определений двух классов друг через друга, что может привести к проблемам в работе механизма поддержки принятия решений [11].

2.10 OOR

Open Ontology Repository - это онтологический репозиторий, предоставляющий возможности для хранения, обмена, поиска, управления и других дополнительных сервисов для работы с базами знаний [12].

2.11 OpenLinkVirtuoso

Универсальный сервер Virtuoso предназначен для хранения и обработки мультимодельных (RDF и SQL) справочных данных. Он предоставляет платформонезависимую поддержку безопасного хранения, обмена и интеграции онтологических данных. Гибридная архитектура Virtuoso предоставляет функционал классического сервера в области управления данными в реляционных таблицах SQL и управления данными в SPARQL-совместимых RDF [13].

2.12 RepOSE

Repair of Ontological Structure Environment - является системой для поиска и исправления ошибок в онтологиях. Данный продукт способен распознавать и ошибки моделирования ПрО, и семантические дефекты онтологии. Существует несколько редакций RepOSE, одна из них является свободно распространяемой для некоммерческого использования и находится в широком доступе [14].

2.13 SigmaKEE

Sigma Knowledge Engineering Environment - это система для создания, поддержки и оценки теорий, описанных на языке логики первого порядка. Она работает с форматом KIF и оптимизирована под онтологию SUMO [15].

Представленный краткий обзор ряда новейших онтологических систем позволяет заключить, что наряду с разработкой средств создания онтологий появляется все больше инструментов оценки существующих онтологий и средств поддержки их жизненного цикла [1, 16].

Заключение

Представленный краткий обзор и подготовленные в форме Коммюнике рекомендации констатируют актуальность проводимых исследований. Переход к применению семантических технологий при создании информационных систем находит свое отражение в стремительном развитии данной области знаний, как в области стандартов, так и в появлении программных средств создания и поддержки онтологий.

Благодарности

Авторы выражают благодарность Министерству науки и образования РФ, поддержавшему разработки по созданию интеллектуальных систем управления в реальном времени по государственному контракту Минобрнауки РФ № 07.524.12.4022 и государственному контракту Минобрнауки РФ №14.514.11.4005

Список источников

- [1] Ontology Summit 2013 Communique. Towards Ontology Evaluation across the Life Cycle. Current Version is: v1.0.4 - 2013.05.31 http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2013_Communique. Коммюнике Онтологического саммита 2013. Оценка онтологий в течение всего жизненного цикла. Перевод с англ. опубликован в журнале Онтология проектирования. - 2013. - №2. - с. 66-74
- [2] *Андриченко, А.Н.* Тенденции и состояние в области управления справочными данными в машиностроении/ А.Н. Андриченко // Онтология проектирования. – 2012. - №2. - с. 25-35.
- [3] <http://techinvestlab.ru/dot15926Editor>, актуально на 05.05.2013

- [4] **Schorlemmer, M.** and **Kalfoglou, Y.** Institutionalising Ontology-Based Semantic Integration, *Applied Ontology* (2008) 3:131-150.
- [5] Ontology Description using OWL to Support Semantic Web Applications, Rajiv Pandey, *India International Journal of Computer Applications* Volume 14– No.4, January 2011.
- [6] *OntologySummit 2013 Extrinsic Aspects of Ontology Evaluation Community Input*, 2013.
- [7] *BioPortal: Ontologies and Integrated Data Resources at the Click of a Mouse*. Noy NF, Shah NH, Whetzel PL, Dai B, Dorf M, Griffith N, Jonquet C, Rubin DL, Storey MA, Chute CG, Musen MA. *Nucleic Acids Res.* 2009(37)W170-3.
- [8] **Kutz, O.** **Mossakowski, T.** and **Lücke D.** “Carnap, Goguen, and the Hyperontologies: Logical Pluralism and Heterogeneous Structuring in OntologyDesign”. In: *Logica Universalis 4.2* (2010): Special Issue on ‘Is Logic Universal?’, pp. 255–333.
- [9] **Suárez-Figueroa, M.C.**(coord.): *NeOn Methodology for Building Contextualized Ontology Networks*. NeOn European Project Deliverable 5.4.1 (2008).
- [10] **Abdulaev A.** *Reality, Universal Ontology and Knowledge Systems: Toward the Intelligent World*, 2008. 306 p.
- [11] <http://oeg-lia3.dia.fi.upm.es/oops/index-content.jsp>, актуально 05.05.2013.
- [12] **Baclawski K.** *Ontology repository research issues workshop*, August 2009. Founder, convener and chair.
- [13] Alan Ruttenberg: *Harnessing the Semantic Web to Answer Scientific Questions*. 16th International World Wide Web Conference. <http://www.w3.org/2007/Talks/www2007-AnsweringScientificQuestions-Ruttenberg.pdf>.
- [14] **Ivanova, V., Laurila J Bergman, Hammerling U, and Lambrix P.** *Debugging taxonomies and their alignments: the ToxOntology - MeSH use case*. 1st International Workshop on Debugging Ontologies and Ontology Mappings, 2012.
- [15] **Benzmüller, C., Gabbay, D., Genovese, V., and Rispoli, D.,** (2011). *Embedding and Automating Conditional Logics in Classical Higher Order Logic*. Submitted to IJCAI 2011.
- [16] **Боргест Н.М., Коровин М.Д.** *Онтологии: современное состояние, стандарты, средства поддержки*. Уч. пособие. СГАУ.– Самара, 2013– 84 с.

Сведения об авторах



Боргест Николай Михайлович, 1954 г. рождения. Окончил Куйбышевский авиационный институт им. С.П. Королёва в 1978 г., к.т.н. (1985). Профессор кафедры конструкции и проектирования летательных аппаратов Самарского государственного аэрокосмического университета (национального исследовательского университета). Член Международной ассоциации по онтологиям и их приложениям (IAOA). В списке научных трудов около 100 работ в области автоматизации проектирования.

Nikolay Mikhailovich Borgest (b.1954) graduated from Kuibyshev Aviation Institute named after S.P. Korolyov in 1978, PhD (1985). He is Professor at Samara State Aerospace University named after academician S.P. Korolyov (Aircraft Design Department of SSAU). He is an International Association for Ontology and its Applications (IAOA) member. He is co-author of about 100 scientific articles and abstracts in the field of CAD and AI.

Коровин Максим Дмитриевич, студент Самарского государственного аэрокосмического университета имени академика С. П. Королёва (национальный исследовательский университет). Области научных интересов: САПР, CALS технологии.

Korovin Maxim Dmitrievich, the student of Samara State Aerospace University named after academician S. P. Korolyov (National Research University). Area of scientific interests: CAD, CALS.