

АНАЛИЗ ИСПОЛЬЗОВАНИЯ ЯЗЫКОВ ОПИСАНИЯ ОНТОЛОГИЙ ДЛЯ ТЕХНИЧЕСКИХ ПРЕДМЕТНЫХ ОБЛАСТЕЙ

ANALYSIS LANGUAGE FOR DEFINING ONTOLOGIES FOR TECHNICAL SUBJECT AREA

*M. Kaung
L. Nay*

Summary. In this article the author raises the problem of using various tools for describing the semantics of technical areas and for creating warehouse, the technical data within the boundaries of which are related semantically. Special attention focuses by the author to the weak and strengths of these approaches.

Keywords: ontology, database, technical subject area, formalized languages, information

Каунг Мьян Хту

*Аспирант, Курский государственный университет
kaungmhtoo16267@gmail.com*

Нэй Лин

*Аспирант, Курский государственный университет
naylynn16@gmail.com*

Аннотация. В данной статье автор поднимает проблему использования различных инструментов для описания семантики технических областей и для создания хранилищ, технические данные, в границах которых связаны между собой семантически. Особое внимание автор акцентирует на слабых и сильных сторонах данных подходов.

Ключевые слова: онтология, база данных, техническая предметная область, формализованные языки, информация.

Единоное информационное пространство технической предметной области требует формализации и структурирования баз данных, что облегчает ее последующее использование. В противном случае гетерогенность хранилищ данных и источников получаемой информации, отсутствие структуры, наличие противоречий или дублирующих элементов баз данных затрудняет эффективное использование информации в пределах технических предметных областей. Эту проблему решает онтологический подход, разработанный в сфере инженерных знаний, объектом изучения которого становятся методы и методологии построения онтологий, а также функционирование процессов разработки онтологий, особенностей протекания их жизненного цикла, использования комплектов инструментов и поддерживающих их языков [1].

Под языком описания онтологии понимают формальный язык, который предназначен для создания информационных моделей в рамках конкретной предметной области. При этом, с точки зрения разнообразия средств и подходов к данной проблеме выделяют три основные концепции:

1. Формализованные технические языки на базе естественных.
2. Машинно-ориентированные языки.
3. Универсальные языки [2].

Создание формализованных технических языков обусловлено проблемой глобализации промышленных комплексов, характерной для современного производства наукоемкой и высокотехнологичной продукции.

Это напрямую связано с возникновением проблем коммуникативного характера у сотрудников, которые являются носителями различных языков. Используемый в настоящее время практически повсеместно английский язык не может являться универсальным средством деловой коммуникации, что обусловлено наличием ряда недостатков, характерных для естественных языков, к числу которых можно отнести, прежде всего, возможность неоднозначной трактовки текста, что недопустимо для сферы профессиональных коммуникационных связей [3].

Именно поэтому актуальным вопросом является разработка проекта по внедрению формализованного языка на базе естественных. В настоящее время в данной сфере осуществлен ряд таких успешных проектов. Их объединяет такая характеристика, как набор правил формализации, который используется для обеспечения однозначности и полноты семантики текста. Можно выделить следующие характеристики формализованных языков на базе естественных:

- ◆ использование семантической единицы только в одном значении, которое было оговорено заранее;
- ◆ использование исключительно простых предложений;
- ◆ запрет на постановку подлежащего в конце предложения;
- ◆ запрет на изменение инвертированного порядка частей составного сказуемого;
- ◆ запрет постановки прямого дополнения перед сказуемым;

- ◆ запрет на использование причастных оборотов в качестве синтаксических оборотов [4].

Благодаря формализации языка значительно облегчается автоматическая обработка текста. Это доказывают примеры регламентированных языков, созданных под эгидой ассоциации АЕСМА (Европейской ассоциации аэрокосмической промышленности), а также языка Gellish: STE (Simplified Technical English), представляющего собой формализованные спецификации написания текстов для повышения их семантической прозрачности.

Для STE характерен ограниченный набор используемых времён, форм глаголов, а также определенных список разрешенных для использования и запрещенных слов. Предназначение спецификаций, в первую очередь, связано с написанием инструкций.

ASD (AeroSpace and Defence Industries Association of Europe) STE100 ранее был известен как АЕСМА Simplified English. Он представляет собой регламентированный язык, разработанный при участии АЕСМА и АИА (Американской ассоциация аэрокосмической промышленности). Данный язык используется в целях написания документации по эксплуатации объектов аэрокосмической и оборонной промышленности.

Для ASD STE100 характерна строгая грамматика, ограниченный лексический состав и жесткие синтаксические правила. Этот стандарт в настоящее время используется членами АЕСМА и АИА, в частности, представителями компаний British Aerospace, Airbus, The Boeing Company, Lockheed Martin, Rolls Royce, Dassault и Saab Aerosystems и других [5].

Говоря о формализованных языках, нельзя не сказать об STR (Simplified Technical Russian) — упрощённом техническом русском языке. Экспорт высокотехнологичной продукции аэрокосмической промышленности отечественного производства обусловил необходимость создания документации на английском языке, что, в свою очередь, породило проблему перевода документа.

После успешной реализации STE в российской компании BETA AIR в партнёрстве с компанией Tedopres началась разработка языка STR, широко используемого в отечественной аэрокосмической отрасли [6].

Характеризуя Gellish, можно отметить, что данный язык предназначен для полной и однозначной спецификации процессов, продуктов и услуг, как физического характера, так и сферы бизнеса. При помощи данного языка предоставляется информация о продукции на различных стадиях ее производства: приобретения, изготовления, монтажа, эксплуатации, хранения, а также

осуществляется обмен информацией между компьютерными системами. Использование Gellish осуществляется на всех этапах жизненного цикла.

Создание машинно-ориентированных языков было направлено на решение проблем создания машиночитаемых баз знаний, в частности для создания специализированных языков описания их семантики. Начало работы по созданию машинно-ориентированных языков датируется началом 70-х гг. XX века [7].

Специализированные машинные языки онтологического описания выступали в качестве инструментов первых попыток создания онтологий. На основании типа логики применяемой машинными языками логики разработана следующая классификация машинных языков:

1. Фреймово-продукционные машинные языки.
2. Языки дискриптивной логики.
3. Языки логики первого порядка [8].

В качестве примеров можно привести широко распространенные машинные языки, к числу которых можно отнести KIF (Knowledge Interchange Format) [9] и CysL (Cycorp Language) [10].

KIF представляет собой универсальный машинно-ориентированный язык для обмена данными в рамках выбранной предметной области. Для KIF характерна декларативная семантика и логическая всесторонность. Данный машинный язык предусматривает представление произвольных предложений в исчислении предикатов первого порядка, обеспечивает представление знаний, используется для описания объектов, функций и отношений. KIF также называют языком продукционного типа, где запись каждой продукции осуществляется в виде импликации.

CysL представляет собой формальный язык, синтаксис которого базируется на логике первого порядка. Объекты, составляющие словарь CysL, объединяются в выражения, которые используются для создания утверждений в CysL базы данных.

В качестве общего недостатка машинно-ориентированных языков можно назвать сложность их синтаксиса и потребность в специальных инструментах для интерпретации.

Названный недостаток несвойственен для универсальных языков, которые с одной стороны сохраняют описательную полноту для компьютерных систем, а с другой — семантически прозрачны для человека.

В качестве примеров можно привести универсальные языки XML (Extensible Markup Language) и HTML

(HyperText Markup Language). Первый рекомендован W3C (World Wide Web Consortium — консорциумом Всемирной паутины) в качестве языка разметки данных. Спецификация XML связана с описанием XML-документов и частичным описанием поведения XML-процессоров (программ, читающих XML-документы и обеспечивающих доступ к их содержимому).

HTML представляет собой стандартный язык разметки документов во Всемирной паутине. Большинство

веб-страниц содержат описание разметки на языке HTML. Язык HTML интерпретируется браузерами и отображается в виде документа в удобной для человека форме.

В результате анализа можно сделать вывод, что для описания технических предметных областей наиболее целесообразно использовать следующие языки:

1. язык программирования C # и C++.
2. язык программирования Delphi.

ЛИТЕРАТУРА

1. Ontology Summit 2013 Communique. Towards Ontology Evaluation across the Life Cycle. Available at: http://ontology.cim3.net/cgi-bin/wiki.pl?OntologySummit2013_Communi que.
2. Conservative groupoids recognize only regular languages. Available at: http://www.academia.edu/9200925/Conservative_groupoids_recognize_only_regular_languages_Extended_Version_.
3. Gellish A. Product Modeling Language. Available at <http://sourceforge.net/apps/trac/gellish/#a5.1TheGellishlanguage>.
4. Orlando Chiarello, Dr. Knezevic Jezdimir The Role of Simplified Technical // English in Aviation Maintenance. 2013. No. 2. P. 36–38.
5. ASD Simplified Technical English Specification ASD-STE100. Available at: <http://www.asd-ste100.org>.
6. Simplified English. Available at: <http://www.aecma.ru/>.
7. Semantic Information Modeling in Formal English. Available at: <http://www.gellish.net/>.
8. Oscar Corcho A Roadmap to Ontology Specification Languages. Available at: http://www.cs.man.ac.uk/~ocorcho/document s/ekaw00_CorchoGomezPerez.pdf.
9. Казекин М. М. История языков представления онтологий. http://www.ict.edu.ru/ft/006054/2008_4_03-11.pdf.
10. Knowledge Interchange Format. Available at: <http://www.ksl.stanford.edu/knowledgesharing/kif/>.

© Каунг Мьят Хту (kaungmhtoo16267@gmail.com), Нэй Лин (naylynn16@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»

