

СЕМАНТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ДЛЯ ПОСТРОЕНИЯ ОНТОЛОГИЙ ПРЕДМЕТНЫХ ОБЛАСТЕЙ

Волошина Н.А., Нестерчук Ю.Э.

Харьковский национальный университет радиоэлектроники

61166, Харьков, пр. Ленина, каф. Искусственного интеллекта, тел. (057) 702-13-06,

E-mail: natavoloshina@gmail.com ; факс (057) 702-11-13

The given work is dedicated to modern development in the field of information technology. As follows creation ontology by means of semantic modeling. As examples were explored such application domains as profession Intellectual systems decision making of the Kharkiv national university of radioelectronics and requirements of the employers in sphere Information technology. It is described stages of the building ontology and their main advantage.

Семантическое моделирование предметной области (ПО) имеет своей целью эксплицитное, структурированное и формализованное представление специфики исследуемой ПО, основных концептов (сущностей) и взаимосвязей между ними. Система, построенная на основе семантической модели, позволит решать все множество задач данной предметной области. При этом, количество сущностей и их взаимосвязей, составляющих семантические модели, обозримо, а сами модели, как правило, легки для восприятия. На сегодняшний день de facto основным подходом, используемым для построения семантических моделей сложноструктурированных и слабоформализованных ПО, является онтологический подход, активно развивающийся в рамках нового направления инженерии знаний – онтологического инжиниринга (ontology engineering). В компьютерных информационных технологиях под термином “онтология” обычно понимают некоторый механизм или способ описания исследуемой области знаний (или ПО), в частности – базовых понятий этой области и связей между ними.

Так как на сегодняшний день качество подготовки специалистов одна из важных задач высших учебных заведений (ВУЗ) Украины. И одним из показателей высокого уровня образования является соответствие объема знаний и умений, которые дает ВУЗ будущим выпускникам, требованиям рынка труда. В данной работе реализован механизм описания предметной области. В качестве примера рассмотрены и проанализированы две предметные области (были выделены основные понятия и связи между ними): содержательная часть учебного плана специальности «Интеллектуальные системы принятия решений» (ИСПР), знания и умения, которые студенты получают после окончания обучения по выбранной специальности с одной стороны, и, с другой стороны, рынок труда в сфере информационных технологий, а именно, требования работодателей, деятельность которых связана с информационными технологиями. Для получения более точной информации о требованиях работодателей были проработаны требования десяти ведущих компаний в городе Харькове. А именно компании: Sigma, Agile Mind Software, TEAM Offsite Development Center, NIX Solutions, DB Best Technologies, «Stella Systems», Global Logic, CS, Scorto Solutions, TeamDev.

Руководители компаний констатируют, что спрос на IT-специалистов превышает предложение, и найти достойного претендента на вакантную должность крайне сложно. Основная борьба на рынке идет за опытных профессионалов. Сложнее всего найти IT-специалиста, обладающего высокой квалификацией и опытом, уже готового к решению поставленных задач. Говоря об узких специальностях, работодатели чаще всего отмечают нехватку профессионалов в сфере проектного внедрения ERP-систем. Одни из наиболее разыскиваемых на рынке – разработчики на Java.

Предлагаемый анализ будет полезен как выпускникам, которые уже в процессе образования будут иметь представление о тех нишах в сфере информационных технологий, где они смогут применить полученные знания, так и для компаний-

работодателей, которые, получив информацию о знаниях и умениях выпускников, смогут выбрать себе сотрудников.

Для более глубокого анализа, получения более полной семантической информации, структурирования знаний, умений выпускников и требований работодателей данных предметных областей были созданы их онтологии.

Специалисты в области информационных технологий превратили понятие онтологии во вполне практический инструмент для работы со знаниями. Для них онтология представляет собой документ или файл, который в формальном стиле описывает взаимосвязи между объектами и свойства объектов конкретной ПО. При этом онтология получила вполне определенную структуру, состоящую, из двух блоков: таксономии и набора логических свойств предметной области и правил вывода.

Базовая составляющая онтологии – таксономия – определяет классы объектов и взаимодействие между этими классами. Ключевыми понятиями здесь являются понятия подкласса, суперкласса и наследования. Например, класс Нормативные (дисциплины) является подклассом класса Учебные дисциплины (а Учебные дисциплины – суперкласс класса Нормативные). Средствами семантического программирования классы определяются как подмножества наследственно-конечной надстройки, обладающие специальными качествами. Обычно онтология строится как дерево или сеть, состоящая из концептов и связей между ними (рис.1).

Для онтологии предметной области требования работодателей в сфере информационных технологий были выделены такие основные классы, как Работодатели (экземплярами этого класса стали конкретные компании-работодатели города Харькова:), Вакансии (экземпляры – конкретные вакансии, принадлежащие той или иной компании), Требования (основные требования работодателей к знаниям и умениям молодых специалистов). Создание иерархии этих классов происходило с помощью процесса нисходящей разработки. Самыми общими понятиями в данной разработке стали понятия требования и работодатели. С целью более подробного описания предметной области в онтологию требований работодателя был добавлен класс Вакансии.

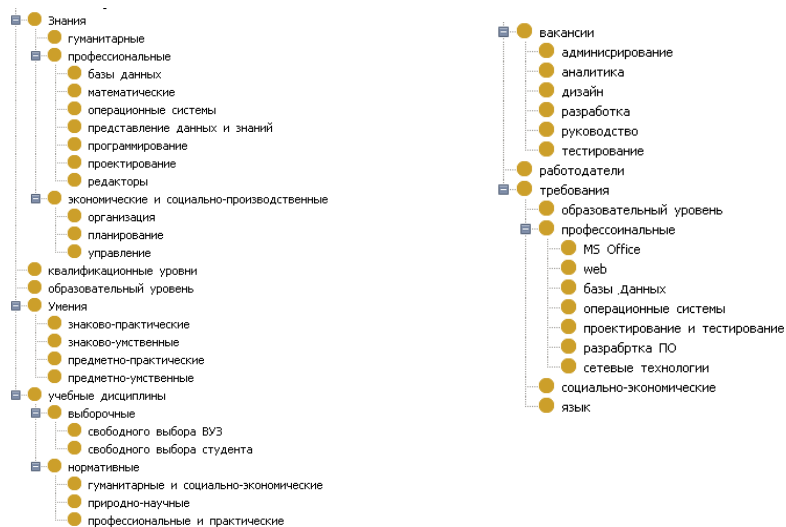


Рисунок 1 – Иерархия классов онтологий специальности ИСПР и требования работодателей в сфере информационных технологий

При разработке онтологии знаний и умений выпускников специальности ИСПР основными классами стали: класс Учебные дисциплины (экземпляры класса – конкретные дисциплины, преподаваемые на специальности), класс Знания (экземпляры – конкретные знания технологий, теории и практики, получаемые при изучении определенной дисциплины), класс Умения (конкретные умения работы с теми или иными технологиями, а также умения применять полученные знания при изучении той или иной дисциплины).

Создание иерархии классов онтологии «Специальность ИСПР» происходило с помощью процесса нисходящей разработки, который начинается с определения самых общих понятий предметной области с последующей конкретизацией понятий. Так, для создания иерархии классов онтологии специальности ИСПР самыми общими понятиями стали понятия знания и умения. После чего проводилась классификация этих понятий, что в итоге позволило создать подклассы классов знания и умения, а также добавить несколько дополнительных классов, таких как класс Учебные дисциплины, которые определяют знания и умения у студентов. Классы Квалификационные уровни и Образовательный уровень добавлены в онтологию с целью более подробного описания предметной области.

После построения иерархии классов отображаются связи между этими классами (рис. 2).



Рисунок 2 – Основные классы и связи между ними онтологии «Специальность ИСПР»

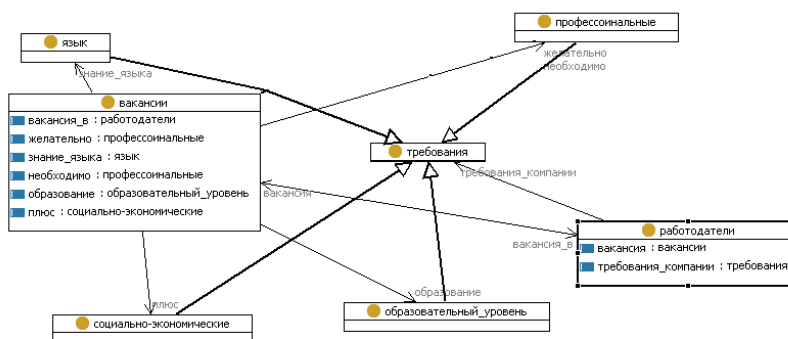


Рисунок 3 – Основные классы и связи между ними онтологии «Требования работодателей»

Основные задачи, которые успешно решаются на базе онтологий, включают предоставление знаний для вывода информации, которые соответствует запросу пользователя; фильтрация и классификация информации; индексирование собранной информации; организация общей терминологии, которой могут пользоваться программные агенты и пользователи.

Онтологии способны точно и эффективно описывать семантику данных ПО и решать проблему несовместимости и противоречивости понятий. Онтологии обладают собственными средствами обработки (логического вывода), которые соответствуют задачам семантической обработки информации.

Онтологии получили широкое распространение в решении проблем представления знаний и инженерии знаний, семантической интеграции информационных ресурсов, информационного поиска и т.д.

Для описания онтологий и работы с ними существуют различные языки и системы, однако, наиболее перспективным представляется визуальный подход, позволяющий специалистам непосредственно «рисовать» онтологии, что помогает наглядно сформулировать и объяснить природу и структуру явлений.

Среди существующих на сегодня решений для работы с семантическим Web можно выделить Protégé [1] – разработанное в отделении Медицинской Информатики Школы Медицины Стэнфордского Университета средство с открытыми кодами, его коммерческий аналог TopBraid Composer [2].

Оба средства, Protégé и TopBraid Composer, обладают схожей функциональностью и пользовательскими интерфейсами. Они позволяют создавать и редактировать файлы в поддерживаемых форматах RDF, RDFS, OWL, правила SWRL и SPARQL запросы с помощью удобного графического интерфейса.

Главным отличием этих двух средств заключается в том, что Protégé является открытой системой с академическими истоками в различных исследовательских проектах, а TopBraid Composer – коммерческая платформа. Поэтому говорится, что TopBraid Composer стабильнее (содержит меньше ошибок), чем Protégé. Кроме того, многие полезные возможности для Protégé предоставляются третьей стороной, а в TopBraid Composer они встроены сразу.

С учетом вышеперечисленных достоинств для реализации онтологий заданных предметных областей была выбрана среда для визуальной разработки OWL - TopBraid Composer.

TopBraid Composer предоставляет полный набор возможностей для покрытия всего жизненного цикла разработки семантических приложений.

В рамках исследования был проведен анализ рынка труда и требований на существующие вакансии в ведущих компаниях в сфере Информационных технологий (ИТ).

Проанализированы учебные дисциплины специальности Интеллектуальные системы принятия решений (ИСПР), знания и умения, которые получают выпускники кафедры Искусственного интеллекта (ИИ).

Результатом стали разработанные онтологии специальности ИСПР и требований работодателей в сфере (ИТ), в которых выделены основные понятия данных предметных областей, построена иерархия классов, определены необходимые свойства классов, а также их экземпляры.

Разработка проводилась с помощью средства визуальной среды описания онтологий – TopBraid Composer.

Литература:

1. The Protégé Ontology Editor and Knowledge Acquisition System [HTML] (<http://protege.stanford.edu/>)
2. TopBraid Composer [HTML] (<http://www.topbraidcomposer.com/>)