

ПРИМЕНЕНИЕ ОНТОЛОГИЙ В ОБЛАСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ ОБЕСПЕЧЕНИЯ РЫНКА ТРУДА

Волошина Н.А., Касумов Т.А.

Харьковский национальный университет радиоэлектроники

61166, Харьков, пр. Ленина, каф. Искусственного интеллекта, тел. (057) 702-13-37,

E-mail: natavoloshina@gmail.com

The given work is devoted a question of employment of graduates of high schools in sphere of information technologies. As a way of the decision of this problem the ontology which will contain knowledge from a subject domain of employers and high schools is offered. The system models an abstract image of the employee on the basis of company requirements, and a concrete image of the graduate proceeding from knowledge which are received by the student. And the further comparison of these two images on presence of the general qualities.

В настоящее время человек все более и более стремится к структурированию своих знаний, так как он хочет получить наибольшую результативность, оптимальное и эффективное использование представленных ресурсов, максимальную качественность. Подобной структурности позволяют добиться онтологии. Прогрессивные проекты в самых разнообразных областях все чаще отказываются от использования баз данных в пользу баз знаний[1]. Но это лишь одна из ступеней на пути к повышению эффективности трудовых затрат по сбору знаний, так как довольно часто в не крупных проектах время использования онтологии соизмеримо со временем и усилиями, потраченными на саму разработку данной онтологии. И вот здесь играет важнейшую роль многократное повторное использование уже структурированных знаний.

Одним из способов повторного применения разработанных онтологий является перенос собранных в ней знаний в иную предметную область путем интеграции ряда косвенно связанных онтологий в обобщенную, которая в свою очередь будет охватывать как все области, представленные в онтологиях, участниках интеграции, так и знания в смежных областях, многократно увеличивая количество решаемых вопросов.

Вопрос дальнейшего трудоустройства – важнейший при выборе специальности, так как именно его задают абитуриенты при поступлении в ВУЗ. Многие хотят знать конкретные рабочие места, где они смогут применить свои полученные знания и навыки. Нахождение квалифицированных сотрудников в сфере информационных технологий так же довольно острая проблема, это обусловлено большой динамикой развития данного рынка. Крупные компании имеют даже специальные отделы по поиску новых сотрудников. Данная работа позволит в какой-то мере решить эти вопросы.

Основными задачами представленной работы являются анализ онтологий работодателей и дисциплин, которые преподаются в высших учебных заведениях, с точки зрения необходимости проведения интеграционных работ по созданию онтологической системы, которая основывается на знаниях предметной области этих двух онтологий и будет отражать актуальную ситуацию в рамках системы «ВУЗ-выпускник-работодатель».

Одним из достоинств онтологии является наличие для них инструментального программного обеспечения, обеспечивающего общую доменно-независимую поддержку онтологического анализа. Существует целый ряд инструментов для онтологического анализа, поддерживающих редактирование, визуализацию, документирование, импорт и экспорт онтологий разных форматов, их представление, объединение, сравнение.

В представленной работе используется TopBraid Composer, эта среда относится к сложным инструментальным средствам[2]. TopBraid Composer позволяет создавать и редактировать файлы в поддерживаемых форматах RDF, RDFS, OWL, правила SWRL и SPARQL запросы с помощью удобного графического интерфейса. Помимо того, может быть использована, как среда исполнения для выполнения правил, запросов, механизмов рассуждений. С помощью доступных механизмов TopBraid Composer позволяет выполнять классификацию, а так же проводить проверку логической целостности на

основе OWL. Для выполнения правил и вывода дополнительных отношений между ресурсами в TopBraid Composer используется внутренний движок правил Jena. Предоставляется возможность выполнения SPARQL запросов только над явно заданными данными, или с учетом имеющихся правил, выполняя логический вывод во время исполнения запроса. Имеется возможность выполнения SPARQL запросов как над RDF хранилищами, так и над интегрированными реляционными базами данных. TopBraid Composer позволяет расширять свою функциональность с помощью специально определенных плагинов. TopBraid Composer реализован как плагин к Eclipse и использует некоторые достоинства данной платформы для предоставления таких возможностей, как графическая визуализация и UML импорт, упрощенный вариант интеграции онтологии в программные разработки, основанные на языке Java. Используемый движок Jena представляет собой открытую java среду для создания приложений семантического веб. Среда Jena предоставляет программное окружение для RDF, RDFS, OWL и SPARQL, а также включает основанный на правилах движок логического вывода. Она включает RDF и OWL API, чтение и запись RDF в форматах RDF/XML, N3 и N-Triples, хранение в памяти и в базах данных, а также движок запросов SPARQL[3].

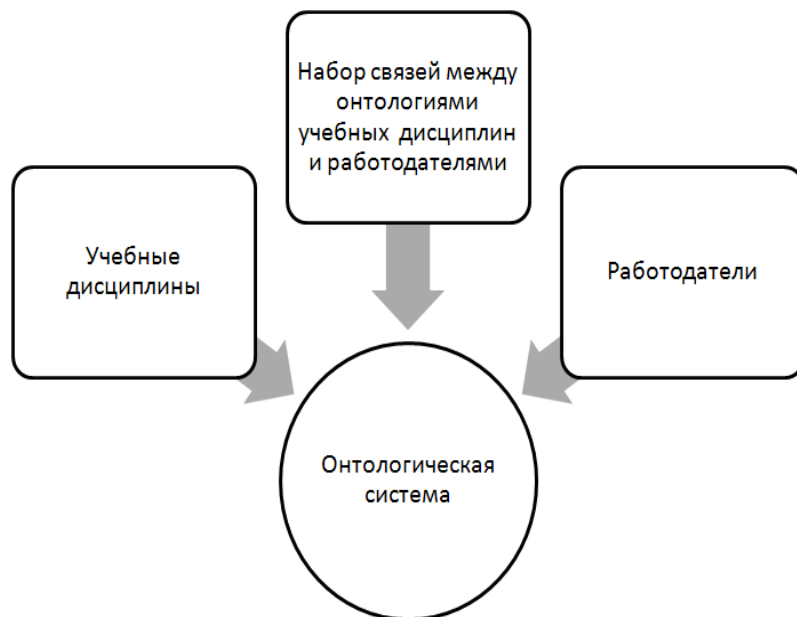


Рисунок 1 – Составляющие части объединенной онтологии

Задача объединения (интеграции) онтологий формулируется следующим образом: даны две адекватно представляющие семантические модели соответствующих предметных областей (ПрО) онтологии, требуется создать объединенную онтологию, которая представляет все понятия исходных онтологий, а также необходимые дополнительные ограничения и взаимосвязи (Рисунок 1).

Онтологическая система содержит в себе ключевые понятия из каждой области знаний. Из ПрО работодателей взяты следующие ключевые классы: «вакансии», «требования», «работодатели». Класс «Вакансии» описывает имеющиеся на текущий момент у компаний свободные рабочие места. Класс «Требования» – знания и технологии, которыми необходимо владеть будущему сотруднику по той или иной специальности, эти знания и технологии будут фигурировать в вакансиях. Экземплярами в классе «работодатели» будут компании, которые имеют рабочие места в сфере информационных технологий.

Учитывая специфику необходимого конечного результата, из онтологии учебных дисциплин выбираются не все знания, а только релевантные поставленной задаче. А именно, используются ключевые классы и свойства с точки зрения работодателей, которые ключевые классы и свойства с точки зрения работодателей, и которые играют решающую роль в последующем трудоустройстве. Из сферы образования были взяты три основных класса: «учебная дисциплина», «знания», «умения». Класс «учебная дисциплина» отражает проф-ориентированные дисциплины, которые читаются на текущий момент в ВУЗах. Знания описывают совокупность данных, которые обеспечиваются изучением курса учебной дисциплины и которыми обладает выпускник этого курса. Умения – практические навыки, которые обеспечиваются знаниями. Пересечение исходных онтологий не очевидно, потому необходимо было выделить основные вопросы относящиеся сразу к двум этим предметным областям. Такими вопросами являются: «Где я смогу работать, после того как закончу ВУЗ?» и «Где найти подходящего сотрудника?». Следовательно, выпускников интересуют те рабочие места (вакансии), требованиям которых он будет полностью удовлетворять. А работодателей интересуют кандидатуры, которые смогут выполнять обязанности в рамках предоставленной ими должности.

В итоге создаются два образа – абстрактный образ сотрудника, который удовлетворяет всем требованиям данной вакансии, и конкретные образы выпускников, имеющих определенный набор знаний и умений. Для интеграции онтологий в области ВУЗов были выделены отдельные кафедры и специальности, на основании чего создана база знаний специалистов, студентов-выпускников. Подобный подход позволил упростить и сделать более точным логический вывод.



Рисунок 2 – Степень соответствия знаний выпускника и требований различных вакансий

Абстрактный образ сотрудника, построенный на основании требований конкретной вакансии компании, сравнивает на соответствие с реальными выпускниками, что позволяет определить оптимальные кандидатуры из представленных в базе знаний (Рисунок 2). Таким образом, система состоит из двух онтологий, описывающих вакансии

в компаниях представленных в сфере ИТ и учебных дисциплин с знаниями, которые ими обеспечиваются, и описания связей между этими двумя онтологиями, что на выходе системы позволяет получить логический вывод, который может быть использован для нахождения сотрудников в сфере информационных технологий. Использование модульной системы в подходе к интеграции, позволяет обеспечить распределенное редактирование и пополнение базы знаний каждой из предметных областей – ВУЗ работает с онтологией учебных дисциплин, в то время как компании пополняют базу знаний для онтологии вакансий. Что сохраняет независимость онтологий и баз знаний друг от друга, а все изменения и привязки осуществляются в специальной, отдельной онтологии, которая лишь ссылается на входные онтологии. Разработанная онтология позволяет выпускникам ВУЗов и сотрудникам одной отрасли или корпорации использовать общую терминологию, и избежать взаимных недоразумений, которые могут усложнить сотрудничество и привести к серьезным убыткам (например, организационная онтология четко отражает взаимную иерархию и связи между требованиями и знаниями). Они обеспечат работу со структурированными источниками данных, для которых может быть построена схема данных, то есть, описаны типы данных и связи между ними, и существует формальный способ получения отдельных элементов данных.

Онтология обеспечивает непротиворечивое накопление любого количества информации в стандартной структуре классификации. Такой подход гарантирует однозначную идентификацию ресурсов независимо от различных трактовок их наименований.

Разработанная система может использоваться как справочный материал для различных пользователей. Абитуриенты смогут спланировать свой профессиональный рост от выбора специальности, до конкретных рабочих должностей. Работодатели смогут сузить область поиска интересующих их специалистов, и иметь представление об их знаниях. Сама система может легко интегрироваться в конкретные программные решения, например, Web-сервисы, которые будут использоваться сайтами вакансий или различными кафедрами в качестве ознакомительного материала по специальностям, представленным на ней, или автономную программу-справочник.

Литература:

1. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем . – СПб: Питер, 2000. – 384 с.
2. TopBraid Composer [HTML] (<http://www.topbraidcomposer.com/>)
3. Официальный портал Semantic Web[HTML] (<http://www.w3.org/2001/sw/>)