

ПОСТРОЕНИЕ МОДЕЛИ ЗНАНИЙ ОРГАНИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ ОНТОЛОГИЙ

CONSTRUCTION OF THE ORGANISATION KNOWLEDGE MODEL USING A SYSTEM OF ONTOLOGIES

А.Ф. Тузовский (TuzovskyAF@kms.cctpu.edu.ru)

С.В. Козлов (KozlovSV@kms.cctpu.edu.ru)

Институт «Кибернетический Центр» ТПУ, Томск

В докладе предлагается модель знаний организации описывать в виде системы дополняющих друг друга онтологий. В систему входит базовая онтология организации и набор онтологий областей знаний. Описан подход к построению модели знаний и предложена структура системы управления знаниями на ее основе.

Управление знаниями

Начиная с середины 90-х годов, знания начали рассматриваться в промышленных компаниях в качестве важнейшего ресурса, ключевого фактора успеха и нового источника дохода. [1,2]. Хотя компании управляли своими людскими и интеллектуальными активами задолго до этого, научная дисциплина Управление Знаниями (УЗ) была сформирована в это время для решения новых специфических задач бизнеса. К сожалению, не существует общепринятого определения понятия «Управление Знаниями», что в основном связано с различными точками зрения практиков на понимание термина «знание». В данном докладе под Управлением Знаниями понимается комплексный набор мероприятий направленных на поддержание в организации системного порядка работы с информационно-знаниемными ресурсами и специалистами для поиска, накопления и облегчения доступа к знаниям, повторного или многократного их использования [2]. Как показывает опыт ведущих компаний это позволяет организациям обучаться более эффективно и максимизировать получаемый эффект от индивидуальных и коллективных источников информации и знаний.

В момент появления в сфере бизнеса дисциплины Управление Знаниями, множество исследований в области Искусственного Интеллекта были направлены на решение проблем представления знаний (Knowledge Representation) и выполнения логического вывода на знаниях, которые объединяли ранний опыт концептуального моделирования с использованием формальных логик. Центральным элементом этих исследований являлось понятие «онтология», которое представляет собой совместно используемое, формальное описание некоторой предметной области [3]. Термин «совместно используемое» опирается на соглашение сообщества экспертов по конкретному описанию их предметной области, а термин «формальное» указывает на представление этого соглашения в некотором, понимаемом компьютером, формате.

Исследователи онтологий рассматривают Управление Знаниями, как одну из наиболее важных областей применения онтологий.

Использование Онтологий

Для пояснения выгоды от использования онтологий в Управление Знаниями, требуется поиск таких процессов и сценариев работы со знаниями, которые могут напрямую выиграть от применения онтологий. Для этого рассмотрим более тщательно основные свойства онтологий. Онтология, исходя из наиболее часто цитируемого определения в области искусственного интеллекта, является совместно используемой, формальной классификацией предметной области [3]. Онтологии являются моделями данных, обладающих двумя специфическими особенностями, которые ведут к понятию совместного понимания или семантике:

- 1) Онтологии строятся на основе совместного понимания предметной области в рамках сообщества. Это понимание представляется соглашением экспертов по поводу понятий и отношений, которые имеются в предметной области (человеческий фактор в управлении знаниями, основанном на онтологиях).
- 2) Онтологии используют способ представления, который может обрабатываться компьютером (т.е., записываются с использованием формальных языков, таких как RDFS или OWL [4,5]), что дает возможность компьютерам работать с онтологиями. К таким действиям относятся передача онтологий между

компьютерами, хранение онтологий, проверка согласованности онтологий, выполнение логического вывода на онтологиях и с помощью онтологий (компьютерный фактор в управлении знаниями, основанном на онтологиях).

Эти факторы дают возможность поднять уровень интеллектуальности программных систем, путем включения в них знаний о предметной области в форме онтологий. Отметим, что эти два фактора имеют различную важность. В то время как компьютеры могут манипулировать и выполнять логический вывод на знаниях о предметной области с помощью онтологий, понимание того, что обозначают эти символы и правила, остается функцией человека. Поэтому онтология не может существовать без сообщества, которое ее поддерживает (онтология сплетает понимание предметной области человеком и компьютером).

Онтологии являются более чем простыми словарями точно определенных понятий: их истинная сила заключается в описании связей между сущностями предметной области. Можно рассматривать словарь, как хранилище смысла, но он определяет слова только по их отношению к другим словам. Элемент информации в действительности определен только тем, с чем он связан и как он связан. Вряд ли, есть что-то еще, что может описать смысл. Структура связей является наиболее важным элементом.

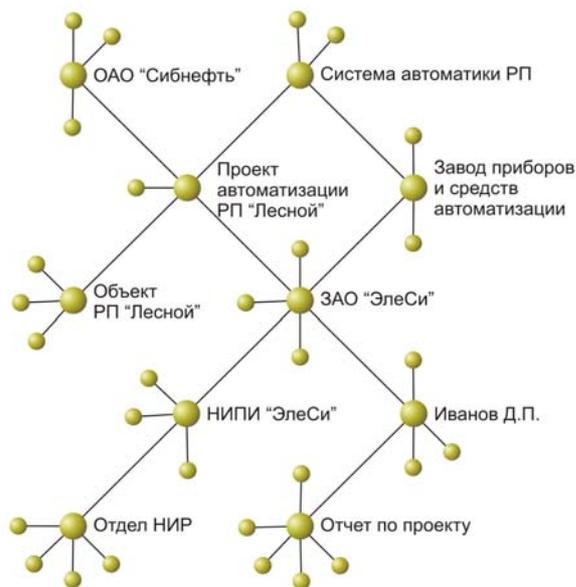


Рис. 1. Отношения между экземплярами онтологии

Полноценные онтологии имеют больше возможностей для описания отношений, чем простые таксономические иерархии, которые имеются в современных решениях по УЗ. Отношения (связи между понятиями) значительно увеличивают количество способов осуществлять навигацию и поиск по предметной области, выполнять анализ, классификацию и визуализацию знаний.

Для того чтобы использовать эту богатую паутину понятий для интеграции информации, разнородные объекты, содержащие знания, должны быть размечены с помощью терминов онтологии. Этот процесс, также называемый семантической нормализацией, позволяет интегрировать концептуально связанные элементы информации, независимо от их форматов и представлений. Хотя эти действия подобны присоединению ключевых слов или классификационных кодов к информации, результирующие описания имеют точную интерпретацию, предоставляемую онтологией.

Модель знаний организации

Для управления знаниями внутри современной организации должна быть создана технологическая инфраструктура – информационно-программная система управления знаниями (СУЗ). СУЗ, основанная на Интернет/интранет и web-технологиях (портал СУЗ), позволяет преодолевать коммуникационные и пространственно-географические барьеры на пути совместного использования знаний. Преимущество web-среды в том, что она обеспечивает удобный доступ к разнородным информационным ресурсам как внутри компании, так и к информационным ресурсам всей сети Интернет.

К сожалению, разнородность затрудняет обнаружение взаимосвязей («знание о том, что мы знаем»), а неструктурированные форматы (спроектированные для чтения людьми) не позволяют выполнять поиск программно. Очевидно, что онтология занимает центральное место в портале СУЗ. Представляя собой не что иное, как единую модель знаний организации, она способствует интеграции разнородных ресурсов в рамках портала СУЗ на концептуальном уровне, обеспечивая единый подход к описанию их семантики.

Модель знаний используется для составления описаний объектов СУЗ и построения поисковых запросов с учетом смысла понятий. Кроме того, она играет важную роль в формировании профессионального языка общения специалистов, помогает ориентироваться в предметной области при обучении.

Создание единой онтологии для детального описания модели знаний организации является весьма трудоемкой задачей. Речь идет о длительном проекте с привлечением коллектива экспертов, знания которых должны охватывать все направления деятельности исследуемой компании.

Решением проблемы может быть отказ от глубокой декомпозиции системы и включение в онтологию только наиболее значимых понятий из рассматриваемых предметных областей. Другим решением является выделение одного из направлений деятельности организации и создание для этого направления детальной, но узко специализированной онтологии. Очевидно, что оба подхода имеют свои недостатки. Первый подход дает слишком грубую и обобщенную модель, а второй не позволяет использовать модель для взаимодействия между всеми подразделениями компании.

Как правило, в пространстве знаний любой организации можно выделить несколько подобластей, которые используются её специалистами. Для каждой из подобластей знаний может быть разработана собственная онтология. Ограничение масштабов модели рамками конкретной подобласти приводит к упрощению модели, позволяет задействовать экспертов узкой специализации. Таким образом, может быть существенно снижена трудоемкость разработки. Согласование этих относительно небольших онтологий может быть обеспечено разработкой онтологии верхнего уровня и программных средств для ее сопровождения.

Для достижения гибкости и масштабируемости модели знаний при ее формировании целесообразно использовать модульный принцип. Авторами предлагается структура модели знаний организации показанная на рис.2.

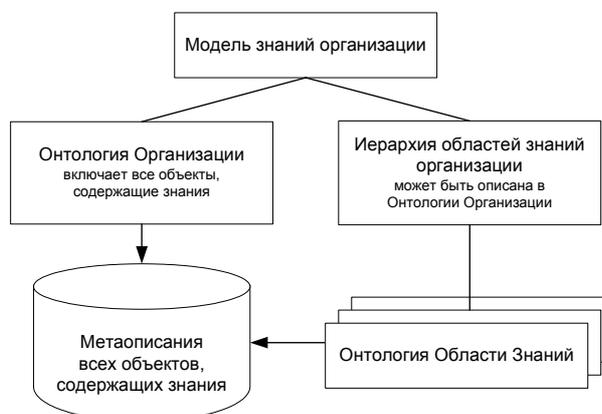


Рис.2. Структура модели знаний организации

Онтология Организации описывает основные понятия компании (организационную структуру, сотрудников, внешних агентов, процессы и т.д.), а также объекты, служащие источниками знаний для компании (рис. 3). Онтология Организации также содержит понятия и отношения, необходимые для формирования иерархии областей знаний и последующего использования этой иерархии приложениями СУЗ. Эта иерархия отражает предметные области, связанные с коммерческой, научной или иной деятельностью компании, то есть те области, знания в которых представляют интерес для компании.



Рис.3. Пример Онтологии Организации

Онтологии Области Знаний (ОЗ) связаны с иерархией областей знаний из Онтологии Организации. Состав и структура Онтологий ОЗ не фиксированы и могут меняться по ходу использования модели знаний в СУЗ. Онтологии ОЗ обеспечивают определенный уровень гибкости модели знаний и возможность её постепенного наращивания.

Для объектов, обладающих знаниями, которые описаны в Онтологии Организации: документам, специалистам, проектам, продуктам и т.д., создаются метаописания, которые используются при работе СУЗ.

Метаописание объекта i имеет следующий вид: $M_i = \{M_{ik}, M_{ic}\}$, где M_{ik} - набор метаданных, соответствующих свойствам понятия онтологии, экземпляром которого является описываемый объект. Структура этих метаданных (свойства понятий) описана в Онтологии Организации. Значениями элементов этих метаданных являются либо ссылки на другие экземпляры понятий, либо литералы. Таким образом, эти метаданные связывают экземпляр объекта – источника знаний с другими экземплярами объектов, описанными в онтологии, в связи с чем, эта составляющая может быть названа «контекстными метаданными». Контекстные метаданные описывают внешние атрибуты объекта и среду, в которой он существует в пределах организации и ее непосредственного окружения.

Вторая составляющая метаописания M_{ic} непосредственно описывает знания, которые в явной или неявной форме заключены в объекте. Эта часть метаописания связывает объект с экземплярами понятий из Онтологий Областей Знаний и может быть названа «контентными метаданными». Контентные метаданные для объекта i могут быть описаны следующим образом: $M_i = \{(tr_{i1}, k_{i1}), \dots, (tr_{in}, k_{in})\}$, где $tr_{ij} = \langle s_{ij}, p_{ij}, o_{ij} \rangle$ – триплет; s_{ij} – некоторое понятие из онтологии предметной области, знания о которой содержатся в описываемом объекте; p_{ij} – отношение определенное в онтологии предметной области; o_{ij} – ссылка на экземпляр понятия в онтологии предметной области; $k_{ij} \in (0,1]$ – коэффициент, обозначающий релевантность триплета tr_{ij} объекту i . При описании триплета метаданных элементы p_{ij} и o_{ij} могут отсутствовать. Процесс составления таких метаописаний называется аннотированием объектов и выполняется с использованием специальных редакторов.

Система управления знаниями

Структура СУЗ, основанной на единой модели знаний организации представлена на рис. 4. Модель знаний представляет собой концептуальную основу СУЗ: определяет множество понятий и отношений, а также правила их совместного использования. База знаний СУЗ (экземпляры понятий) формируется из метаописаний всех объектов, которые могут содержать знания.

Подсистема поиска знаний обрабатывает метаданные объектов и отбирает те из них, которые удовлетворяют запросу пользователя. Возможности подсистемы поиска знаний используются функциональными подсистемами Портала СУЗ, которые предоставляют пользователям различные сервисы (навигация по элементам базы знаний и репозитарию документов).

Работа различных подсистем портала (поиск, категоризация знаний) связана с оценкой семантической близости пар объектов, а точнее, их метаописаний. Для выполнения логического вывода используются дескриптивные логики. Дескриптивные логики описывают знания в терминах понятий и ограничений ролей, которые используются для автоматического вывода классификационных таксономий. Эти логики имеют сильное влияние на современные языки онтологий. Все языки, начиная с OIL, DAML+OIL и OWL получили их формальную семантику от дескриптивных логик.



Рис.4. Структура системы управления знаниями на основе единой модели знаний организации

Определение меры подобия между документами сводится к поиску подобия между наборами взвешенных терминов онтологий. Методы расчета оценки подобия/расстояний между семантическими метаданными можно найти в [6]. Использование понятий онтологий и оценок семантической близости позволяет создать единое интеллектуальное пространство, в котором размещены все объекты организации, содержащие знания. Схема такого интеллектуального пространства показана на рис. 5.

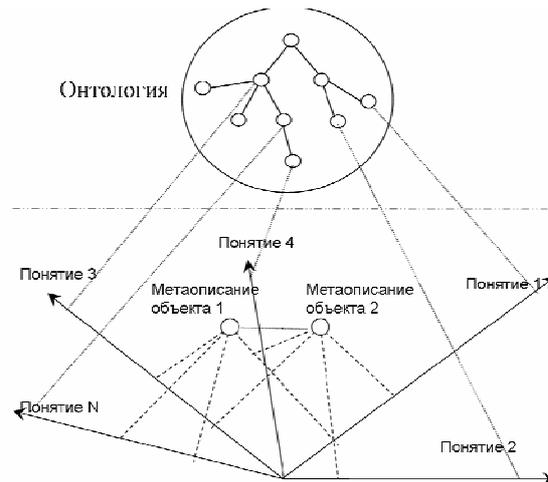


Рис. 5. Схема интеллектуального пространства

Заключение

Данный подход использовался при разработке проекта и базовых элементов системы управления знаниями компании «ЭлеСи» (г. Томск), основной деятельностью которой является разработка средств автоматизации. Совместно со специалистами компании была создана онтология организации и выполнено онтологическое описание одной предметной области. Онтология организации содержит основные объекты – контейнеры знаний организации и ключевые понятия, связанные с ними (Документ, Специалист, Организация, Проект, Процесс, Продукт и т.д.). Разработанная модель знаний компании включает 24 области знаний, разделенных на три группы:

- Приборы и средства автоматизации;
- Системы и комплексы АСУ ТП;
- Теория, методы и программное обеспечение для создания средств и систем автоматизации.

Детально была разработана онтология «Автоматизация», которая описывается с помощью 578 понятий и 15 отношений. В качестве основной предметной областью была выбрана «Знания по управляемым электроприводам», в которой было выделено 375 понятий, что составляет около 65% от общего количества понятий онтологии.

Разработанная онтология используется в составе семантического портала управления знаниями для решения следующих задач:

- в качестве терминологической базы для описания источников знаний;
- для семантического поиска источников знаний (сотрудников, документов, организаций);
- для категоризации всех объектов компании, содержащих знания.

Список литературы:

1. Davenport T., Prusak L. Working Knowledge // Boston: Harvard Business School Press, 1998.
2. Тузовский А.Ф., С.В. Чириков, В.З. Ямпольский Системы управления знаниями (методы и технологии) // Томск: Изд-во НТЛ, 2005.
3. Gruber T.A Translation Approach to Portable Ontology Specifications // Knowledge Acquisition, 5(2):199–220, 1993.
4. W3C, "RDF/XML Syntax Specification (Revised)," // электронный ресурс <http://www.w3.org/TR/rdf-syntax-grammar/>
5. Web Ontology Language. Overview. // электронный ресурс <http://www.w3.org/TR/owl-features/>
6. Тузовский А.Ф., Васильев И.А., Усов М.В. Программная реализация основных компонент информационно-программного обеспечения системы управления знаниями // журнал «Известия ТПУ», №7, 2004, с. 116-122.