
Управление, вычислительная техника и информатика

УДК 002.53:004.89

РАЗРАБОТКА СИСТЕМ УПРАВЛЕНИЯ ЗНАНИЯМИ НА ОСНОВЕ ЕДИНОЙ ОНТОЛОГИЧЕСКОЙ БАЗЫ ЗНАНИЙ

А.Ф. Тузовский

Институт «Кибернетический центр» ТПУ
Томский научный центр СО РАН
E-mail: TuzovskyAF@kms.cctpu.edu.ru

Предлагается создание системы управления знаниями на основе единой модели знаний организации, описываемой в виде системы дополняющих друг друга онтологий. В систему входит базовая онтология организации и набор онтологий областей знаний. Описан вариант построения такой модели знаний и предложена структура системы управления знаниями на ее основе.

Системы управления знаниями

Знания являются важнейшим ресурсом современных организаций, и эффективность их деятельности, в значительной степени, зависит от управления процессами работы со всеми видами знаний, т. е. от управления знаниями [1, 2]. Управление знаниями выполняется с использованием как организационных, так и технологических средств. Технологической составляющей управления знаниями является система управления знаниями (СУЗ) – информационно-коммуникационная система, объединяющая и интегрирующая обработку как явных, так и скрытых знаний организации. СУЗ поддерживает сети работников со знаниями в их деятельности по созданию, сбору, сохранению, оцениванию, структурированию, визуализации, распространению, поиску и применению знаний.

В данной статье предлагается выполнять создание СУЗ на основе онтолого-семантического подхода с использованием единой онтологической базы знаний и описания в ней всех элементов знаний организации.

Знания организации

Термин знания широко используется, но часто достаточно расплывчато, в среде руководителей организаций и специалистов по управлению информацией. Существует огромное количество определений этого термина, имеющих различное происхождение и в различных контекстах. Одним из часто используемых определений знания, с точ-

ки зрения менеджеров, является следующее: «Знания – это комбинация данных и информации, к которой добавлены мнения, навыки и опыт экспертов, что дает в результате ценный актив, который может быть использован при принятии решений. Знания могут быть явными (формализованными) и/или скрытыми (неформализованными), индивидуальными и/или коллективными» [3]. То есть, знания организации (или контент организационной памяти) размещаются в головах людей, в различных физических объектах, например, таких как печатные материалы, аудио-, видеоматериалы, мультимедиа-инструменты, а также в различных объектах информационной системы, например, в программах, электронных документах, мультимедиафайлах, базах данных. Все эти элементы являются объектами знаний (ОЗ), т. е. конкретными порциями информации, которые взаимосвязаны друг с другом и при правильном применении помогают решать задачи организации.

Все эти объекты сложным образом взаимосвязаны в сеть знаний (СЗ), эффективное использование которой создаёт конкурентные преимущества организации. Следовательно, управление знаниями должно влиять на процессы работы с этими комплексными взаимосвязями и сетями и улучшать их использование.

Все знания также могут быть разделены на декларативные и процедурные. Декларативные (статические) знания хорошо описываются сложной сетью понятий и отношений между ними (семантические сети, фреймы, онтологии). Процедурные

знания описывают порядок выполнения работ и имеют более тесную связь с деятельностью организации, но их формальное описание является затруднительным. Однако процедурные знания, достаточно хорошо могут быть проиндексированы с использованием декларативных знаний, что позволяет создать единую систему описания как декларативных, так и процедурных знаний.

Для того, чтобы определить структуру взаимосвязей между элементами знаний организации, необходимо выделить (абстрагировать) понятия из содержания элементов знаний (документов, опыта сотрудников) и структурировать (организовать) их формальным способом, путем задания взаимосвязей между этими понятиями. Одним из наиболее разработанных в настоящее время способов описания знаний в виде множества понятий и взаимоотношений между ними являются *онтологии* [4, 5].

Онтологию можно определить как знаковую систему $O = \{C, R, P, H, F, G, L, A\}$, где C – набор понятий онтологии, причем для каждого понятия $c \in C$ в онтологии существует, по крайней мере, одно утверждение; R – обозначает бинарный характер отношений между понятиями онтологии, фиксирующие пары области применения (domain)/области значений (range), то есть пары (D, R) с $D, R \in C$; P – множество свойств понятий; H – фиксирует таксономический характер отношений (связей), при котором понятия онтологии связаны нереклексивными, ациклическими, транзитивными отношениями $H \subset C \times C$. Выражение $H(C_1, C_2)$ означает, что понятие C_1 является подпонятием (производным понятием) C_2 ; F и G – функции ссылок такие, что $F: F^{LC} \rightarrow 2^C$ и $G: F^{LR} \rightarrow 2^R$, то есть F и G связывают наборы лексических единиц $\{L_j\} \subset L^C$ наборами понятий и отношений, на которые они соответственно ссылаются в данной онтологии, при этом одна лексическая единица может ссылаться на несколько понятий или отношений и одно понятие или отношение может ссылаться на несколько лексических единиц; A – набор аксиом онтологии.

Использование онтологий для описания модели знаний организации позволяет использовать формальные языки для их описания, например, такие как, RDFS и OWL [6, 7]. А также воспользоваться строгим аппаратом дискретной логики для выполнения операций логического вывода над понятиями и метаданными онтологии [8]. Использование онтологий позволяет интерпретироваться понятия (и соответствующие им термины) как специалистами организации, так и компьютерными программами СУЗ.

Таким образом, предлагаемый подход можно описать следующим образом. Знания организации (ЗО) это множество знаниевых объектов, которые связаны между собой системой взаимосвязанных понятий $ZO = \{OZ_1, OZ_2, \dots, OZ_n\}$. Каждый объект знаний включает множество понятий, и ему в соответствие может быть поставлено некоторое его представлений $OZ_k \rightarrow OZ_i$, которое описывает связь объекта

знаний с множеством понятий, которые в нем содержатся $OZ_i = \Phi_i(C_0, C_1, \dots, C_m)$, C_i – понятия, которые используются для описания OZ_i . Здесь функционал Φ_i задает взаимосвязь объекта знаний с понятиями ЗО (позднее будет уточнен). Понятия, которые входят в состав представлений не являются независимыми, они описываются наборами свойств и связаны между собой различными отношениями $R = \{R_1, R_2, \dots, R_k\}$. Каждое понятие определяет некоторый термин (идею) из области знаний организации и задается в виде набора свойств $C_i = \{p_{i1}, p_{i2}, \dots, p_{ik}\}$. Все множество понятий, связанных с объектами знаний организации описывается в виде одной или нескольких онтологий вида $O = \{C, R, P, H, F, G, L, A\}$, которые представляются на формальном языке описания онтологий.

Использование онтологий позволяет описывать объекты знаний в виде набора метаданных $OZ_i = \Phi_i(C_{i0}, C_{i1}, \dots, C_{im}) = M_i(M_k, M_c)$. Метаданные это данные, описывающие контекст (*context* – от лат. *связь*), и контент (от англ. *content* – содержание) объектов. Контекстные метаданные описывают связь объекта с другими объектами знаний системы, а контентные метаданные описывают содержание объекта (т. е. имеющиеся в объекте знания). Использование метаданных, в особенности контентных (семантических) позволяет эффективно решать такие задачи работы со знаниями, как поиск, категоризация и рекомендация знаний. Контекстные метаданные $M_k = \{p_1, p_2, \dots, p_n\}$ представляют собой набор p_i – значений свойств понятия, соответствующего объекту знаний (документу, отчету, специалисту, подразделению), и эти значения p_i являются либо литералами l_i (константами разного типа), либо ссылками r_i на другие экземпляры метаданных. Второй элемент описания метаданных – $M_c = \{s_1, s_2, \dots, s_m\}$ – это контентные метаданные, которые представляют собой набор утверждений вида $s_i = (c, r, o, v)$, где c – это субъект утверждения (понятие, или экземпляр – контекстные метаданные некоторого понятия), o – объект (экземпляр – контекстные метаданные некоторого понятия), r – отношение между субъектом и объектом, а v – весовой коэффициент, который оценивает значимость данного утверждения для описания объекта знаний.

Онтология вместе с описаниями всех ЗО в виде набора метаданных составляет *онтологическую базу знаний организации*.

Модель онтологической базы знаний организации

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 этой проблемы может быть отказ от глубокой декомпозиции системы и включение в онтологию только наиболее значимых понятий из рассматриваемых предметных областей или 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.

Для решения этой проблемы предлагается создать иерархию областей знаний организации и дать возможность создавать отдельно онтологий разных подобластей, которые могут иметь разную детальность в зависимости от потребностей в их детальном моделировании. В этом случае онтология знаний организации описывается следующим образом $O = \{O_o, \{O_s\}\}$, где O_o – онтология организации, а $\{O_s\}$ – иерархически организованное множество подобластей знаний организации.

Онтология организации описывает основные понятия организации (организационную структуру, сотрудников, внешних агентов, процессы и т. д.), а также объекты, служащие источниками знаний (рис. 1).



Рис. 1. Пример онтологии организации

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.

Онтологии областей знаний связаны с иерархией областей знаний из онтологии организации. Состав и структура онтологий областей не фиксированы и могут меняться по ходу использования модели знаний в СУЗ. Онтологии областей знаний обеспечивают требуемый уровень гибкости модели знаний и возможность её постепенного наращивания.

Ограничение масштабов модели рамками конкретной подобласти приводит к упрощению модели, позволяет задействовать экспертов узкой специализации и выполнять постепенное, модульное развитие онтологии и базы знаний организации. Таким образом, можно существенно снизить трудоемкость разработки СУЗ. Согласование этих относительно небольших онтологий может быть обеспечено разработкой и онтологии верхнего

уровня и использованием методов объединения онтологий отдельных подобластей знаний в одну онтологию [9]. Структура онтологической базы знаний организации показанная на рис. 2.



Рис. 2. Структура модели знаний организации

Структура системы управления знаниями

Структура СУЗ, основанной на единой онтологической базе знаний организации представлена на рис. 3. Она реализуется на основе реляционной базы данных в виде набора таблиц и хранимых процедур для работы с ними, а также в виде библиотек API, которые предоставляют набор методов по работе с онтологией и метаданными.



Рис. 3. Структура системы управления знаниями на основе единой модели знаний организации

Использование онтологической базы знаний позволяет создать набор подсистем, выполняющих семантическую обработку информации и знаний. В частности, подсистема поиска знаний позволяет выполнять контекстные и контентные поисковые запросы к базе знаний, обрабатывать метаданные объектов знаний и отбирать те из них, которые удовлетворяют запросу пользователя. Возможности подсистемы поиска знаний используются другими функциональными подсистемами портала СУЗ, которые предоставляют пользователям различные сервисы, например, такие как навигацию по метаданным объектам знаний и хранилищу документов.

Выполнение контентных запросов к подсистеме поиска знаний связана с оценкой семантиче-

ской близости описаний M_c объектов знаний. Для выполнения логического вывода используются методы дескриптивной логики. Определение близости семантических метаданных объектов знаний сводится к поиску расстояний между наборами взвешенных понятий дерева онтологии. Методы расчета таких оценок описаны в [5, 10]. Использование понятий онтологий и оценок семантической близости позволяет создать единое интеллектуальное пространство организации, в котором размещены все объекты, содержащие знания.

Заключение

Онтолого-семантический подход использован нами при выполнении проекта создания системы управления знаниями компании «ЭлеСи» (г. Томск), которая занимается разработкой средств автоматизации: создано описание общей онтоло-

гии организации и отдельной предметной области «Управляемый электропривод», являющееся частью онтологии «Автоматизация». Общая онтология организации содержит основные понятия, соответствующие объектам знаний организации и ключевые понятия, связанные с ними (Документ, Специалист, Организация, Проект, Процесс, Продукт и т. д.). Разработанная модель включает 24 области знаний, разделенных на группы: «Приборы и средства автоматизации»; «Системы и комплексы АСУ ТП»; «Теория, методы и программное обеспечение для создания средств и систем автоматизации». Детально разработана онтология «Автоматизация», которая описывает 578 понятий и 15 отношений. Количество понятий в разделе «Знания по управляемым электроприводам» – 375, что составляет около 65 % от общего количества понятий онтологии.

СПИСОК ЛИТЕРАТУРЫ

1. Davenport T., Prusak L. Working Knowledge: how organizations manage what they know. – Boston: Harvard Business School Press, 1998. – 200 p.
2. European Guide to good Practice in Knowledge Management – Part 1: Knowledge Management Framework // [Электронный ресурс]. – 2004. – Режим доступа: <ftp://cenftp1.cenorm.be/PU-BLIC/CWAs/e-Europe/KM/CWA14924-01-2004-Mar.pdf>
3. Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). – Томск: Изд-во НТЛ, 2005. – 260 с.
4. Gruber T.A. Translation Approach to Portable Ontology Specifications // Knowledge Acquisition. – 1993. – V. 5. – № 2. – P. 199–220.
5. Тузовский А.Ф., Козлов С.В., Чириков С.В., Ямпольский В.З. Использование онтологий в системах управления знаниями организаций // Известия Томского политехнического университета. – 2006. – Т. 309. – № 3. – С. 180–184.
6. W3C, «RDF/XML Syntax Specification (Revised)» // [Электронный ресурс]. – 2003. – Режим доступа: <http://www.w3.org/TR/rdf-syntax-grammar/>
7. Web Ontology Language. Overview // [Электронный ресурс]. – 2003. – Режим доступа: <http://www.w3.org/TR/owl-features/>
8. Тузовский А.Ф. Работа с онтологической моделью организации на основе дескриптивной логики // Известия Томского политехнического университета. – 2006. – Т. 309. – № 7. – С. 134–137.
9. Тузовский А.Ф. Метод объединения онтологий предметных областей знаний // Известия Томского политехнического университета. – 2006. – Т. 309. – № 7. – С. 138–141.
10. Тузовский А.Ф., Васильев И.А., Усов М.В. Программная реализация основных компонент информационно-программного обеспечения системы управления знаниями // Известия Томского политехнического университета. – 2004. – Т. 307. – № 7. – С. 116–122.

Поступила 17.11.2006 г.