

МУЛЬТИДИСЦИПЛИНАРНАЯ СИСТЕМА УПРАВЛЕНИЯ ИНФОРМАЦИОННЫМИ РЕСУРСАМИ РАЗЛИЧНЫХ УРОВНЕЙ ОБЩНОСТИ¹

И. Л. Артемьева, Т. Л. Гаврилова, В. В. Грибова, А. С. Клещев, М. А. Князева,
Н. Ю. Никифорова, В. А. Орлов, М. Ю. Черняховская, Е. А. Шалфеева

Рассмотрен подход к решению проблемы управления созданием, хранением и поддержкой в актуальном состоянии информационных ресурсов, а также обеспечением доступа к ним не только пользователей, но и программ через сеть Интернет.

ВВЕДЕНИЕ

В последние десятилетия было выдвинуто несколько перспективных проектов, развитие которых, после демонстрации потенциальных преимуществ, затормозилось. Самым старым из них является проект массового создания экспертных систем, более новыми — проекты семантического Интернета и агент-ориентированного программирования. Практическое использование научных результатов, полученных при выполнении этих проектов, пока проблематично.

Система, основанная на знаниях, состоит из двух частей — оболочки и базы знаний, а экспертная система — основанная на знаниях система, база знаний которой имеет высокий уровень компетентности. В начале 1980-х гг. казалось, что массовой разработке практически полезных экспертных систем препятствует недостаточная производительность компьютеров. Однако настоящим камнем преткновения для этой технологии оказалась высокая трудоемкость создания и сопровождения компетентных баз знаний. Система, основанная на знаниях, без гарантий компетентности ее базы знаний и сопровождения разработчиков (поддержания на современном уровне) в течение достаточно длительного времени, не имеет шансов быть коммерческим продуктом. Главная проблема — кто и как будет создавать и сопровождать компетентные базы знаний для экспертных систем.

¹ Работа выполнена при финансовом содействии программы № 16 Президиума РАН, проект «Теоретические основы интеллектуальных систем, основанных на онтологиях, для интеллектуальной поддержки научных исследований», и программы № 16 ОЭМПУ РАН, проект «Синтез интеллектуальных систем управления базами знаний и базами данных».

Интернет уже давно стал огромным хранилищем полезной информации, ориентированным на людей. Однако чем больше такой информации в Интернете, тем труднее найти в этом хранилище релевантную информацию. Проект семантического Интернета направлен на решение этой проблемы путем хранения в Интернете онтологий, в терминах которых представлена информация, и благодаря унификации ее представления. Однако и здесь возникла похожая проблема — кто и как будет создавать и сопровождать онтологии и обеспечивать унифицированное представление информации в Интернете.

Проект агент-ориентированного программирования направлен на массовую разработку агентов — программ, функционирующих в сети Интернет, способных взаимодействовать друг с другом и выполняющих полезную работу по заданиям других агентов. Взаимодействие агентов также должно осуществляться в терминах онтологий, хранящихся в Интернет. И здесь возникает та же проблема, что и для семантического Интернета.

Решение этих проблем можно найти, если воспользоваться двумя основополагающими принципами [1, 2]:

- информацию в Интернете следует представлять таким образом, чтобы она была в равной мере доступна как людям, так и компьютерным программам;
- процессы создания и сопровождения информационных ресурсов следует считать самостоятельными, не связанными с процессами их обработки.

В настоящей работе рассматривается подход к проблеме управления информационными ресурсами, основанный на этих принципах.



1. ИНФОРМАЦИОННЫЕ РЕСУРСЫ РАЗЛИЧНЫХ УРОВНЕЙ ОБЩНОСТИ

Под информационными ресурсами различных уровней общности будем понимать языки спецификации онтологий (вообще говоря, любые искусственные языки), онтологии различных уровней общности (онтологии верхнего уровня; метаонтологии; онтологии предметных областей; онтологии знаний и данных и т. п.), знания и данные [3].

Традиционным средством представления информации являются искусственные языки, обладающие конкретным синтаксисом, задаваемым грамматикой, и точной семантикой, задаваемой, как правило, неформально. Простейшим из них является вербальное представление. Будем называть информацию (некоторую идею) вербализуемой, если она может быть представлена в виде отображения некоторого конечного множества терминов в множество значений. Само такое представление (отображение) будем называть вербальным представлением информации. Неявное определение смысла терминов, использованных в вербальном представлении информации, может быть дано с помощью концептуализации — множества всех имеющих смысл (несущих информацию) вербальных представлений с этим множеством терминов. Явное представление концептуализации называется онтологией [4]. Явное представление некоторого собственного бесконечного подмножества концептуализации называется знаниями о совокупности информации, представляемой этим подмножеством.

Для любого искусственного языка может быть определен абстрактный синтаксис, в терминах которого может быть представлен любой текст на этом языке. Такое представление есть вербальное представление этого текста. Таким образом, любая информация, представляемая на некотором искусственном языке, является вербализуемой.

Информационные ресурсы различных уровней общности образуют некоторую иерархию [3]. На ее верхнем уровне находится описание вербального представления в терминах самого этого представления. В терминах этого же представления может быть описан абстрактный синтаксис любого искусственного языка. В терминах абстрактного синтаксиса достаточно мощного языка спецификации онтологий может быть описана любая онтология. В терминах онтологий более высокого уровня общности могут быть описаны онтологии менее высокого уровня общности. В терминах онтологий знаний могут быть описаны базы знаний, а в терминах онтологий данных (схем баз данных) — базы данных. Кроме того, базы знаний могут рассматриваться как описания некоторых совокупностей информации в вербальном представлении.

2. УПРАВЛЕНИЕ ИНФОРМАЦИОННЫМИ РЕСУРСАМИ

Под управлением информационными ресурсами различных уровней общности в настоящей работе понимается деятельность по созданию и хранению информационных ресурсов, поддержанию их на современном уровне, а также организации к ним доступа всех заинтересованных лиц. Главная проблема — найти разумный компромисс между централизацией и децентрализацией этой деятельности. Централизация управления информационными ресурсами делает эту деятельность неудобной для ее участников, а децентрализация ведет к снижению качества информационных ресурсов.

В качестве метафоры для управления информационными ресурсами можно рассматривать научную деятельность. Наука имеет дело с информационными ресурсами различных уровней общности — научными диалектами языка деловой прозы, системами понятий различных уровней общности, системами знаний и данных. Научное сообщество децентрализовано — оно разбросано по разным странам, институтам и лабораториям. Децентрализованно создаваемые информационные ресурсы (например, научные статьи) становятся доступными научному сообществу через научные журналы, редакционные коллегии которых централизованно обеспечивают необходимый научный уровень и новизну этих ресурсов. Монографии и учебники представляют собой согласованное объединение этих информационных ресурсов в единых системах понятий и на принятых научных диалектах. Журналы и научные издательства являются центрами распространения научной информации, а библиотеки — ее централизованными хранилищами. Эта метафора была положена в основу прототипа Интернет-системы управления информационными ресурсами различных уровней общности (Многоцелевого банка знаний), которая в настоящее время проходит испытания и вводится в эксплуатацию в Институте автоматизации и процессов управления ДВО РАН [3].

Все информационные ресурсы, которые управляются средствами системы, физически находятся на выделенном сервере. На верхнем уровне иерархии эти ресурсы объединены в крупные единицы — специализированные банки знаний. Администратор системы принимает решение о создании новых специализированных банков знаний при наличии заинтересованных и компетентных пользователей.

Каждый специализированный банк знаний состоит из основной и экспериментальной областей (студенческий банк — только из экспериментальной области). В основной области находятся проверенные информационные ресурсы, доступные для обработки компьютерными программами. В экспериментальной области ресурсы создаются и

модифицируются. Решение о переводе ресурсов из экспериментальной в основную область (их публикации) принимает администратор специализированного банка знаний по результатам их проверки и оценки качества. Кроме того, каждый специализированный банк знаний состоит из разделов, число и названия которых определяет администратор этого банка.

Связь пользователей с администраторами осуществляется через сайт Многоцелевого банка знаний. Пользователи запрашивают необходимые им полномочия на создание, редактирование или использование информационных ресурсов банка, а соответствующий администратор наделяет их этими полномочиями через административную систему, проверяя необходимый для этого уровень компетенции пользователей.

Создание, редактирование и просмотр информационных ресурсов осуществляются через специализированные редакторы [5]. Использование таких редакторов возможно только в соответствии с полномочиями, которыми наделил пользователя администратор. Каждый такой редактор связан с парой информационных ресурсов, находящихся на соседних уровнях иерархии. В терминах информации верхнего уровня (управляющей информацией) редактируется информация нижнего уровня (целевая информация). Универсальный редактор Многоцелевого банка знаний связывается с некоторой управляющей информацией и, в результате, становится специализированным редактором связанной с ним целевой информации. Чтобы создать новый информационный ресурс, пользователь должен запросить на это полномочие и указать управляющую информацию, в терминах которой он будет формировать этот ресурс. Чтобы редактировать и просматривать информацию, достаточно указать только целевую информацию — управляющая информация определяется единственным образом.

Создание и редактирование информации возможно в двух режимах — в режиме инженера знаний, когда пользователю доступны все возможности универсального редактора, и в режиме эксперта, когда пользователь обеспечивается интеллектуальной поддержкой, связанной с управляющей информацией. В режиме инженера знаний пользователь создает и редактирует целевую информацию, которая в дальнейшем будет использоваться как управляющая. При этом он связывает с этой информацией спецификацию определенной интеллектуальной поддержки (тексты помощи, ограничения целостности (контекстные условия), процедуры выбора элементов и подмножеств из множеств. В режиме эксперта пользователь получает контекстную помощь, возможность выбора элементов и подмножеств из соответствующих множеств. Целевая информация контролируется на соответствие ограничениям целостности. В обоих режимах возможны ссылки на другие элементы ре-

дактируемой информации, а также элементы других информационных ресурсов. Кроме того, целевая информация, созданная в режиме эксперта, может редактироваться в режиме инженера знаний, чтобы затем использоваться в качестве управляющей.

3. ОБРАБОТКА ИНФОРМАЦИОННЫХ РЕСУРСОВ

Каждый специализированный редактор представляет собой универсальный редактор Многоцелевого банка знаний, интерпретирующий соответствующую управляющую информацию. Все информационные ресурсы, создаваемые такими редакторами, представлены в одном и том же универсальном представлении. Универсальная оболочка информационных ресурсов Многоцелевого банка знаний есть набор операций над информацией в этом представлении в терминах этого представления. Интернет-программа получает доступ к информационному ресурсу Многоцелевого банка знаний через оболочку, если она наделена необходимыми полномочиями.

Если информационный ресурс создан в терминах одной онтологии (управляющей информацией), а должен обрабатываться в терминах другой, то программа обработки должна предварительно решить задачу интерпретации — перевода информации из одной терминологии в другую.

4. ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ РЕСУРСОВ ПРИ РЕШЕНИИ ПРИКЛАДНЫХ ЗАДАЧ

Информационные ресурсы специализированных банков знаний используются при решении прикладных задач. Решатели прикладных задач являются программными компонентами специализированных банков знаний. Они имеют доступ к информации специализированного банка через набор операций универсальной оболочки. При решении прикладной задачи могут использоваться информационные ресурсы других специализированных банков знаний, если решатель прикладной задачи наделен такими полномочиями. Различным группам пользователей специализированного банка знаний может требоваться свой набор информации из этого банка, который определяется областью профессиональной деятельности пользователей данной группы. Например, информационные ресурсы специализированного банка знаний по химии могут включать в себя знания физической, органической, аналитической и других разделов химии. При решении прикладных задач одного раздела химии используются только знания этого раздела, но могут потребоваться и все информационные ресурсы специализированного банка. В этом случае пользователь специализированного банка знаний может получить доступ к требуемому ему набору информационных ресурсов



посредством своего автоматизированного рабочего места, входящего в набор программных средств специализированного банка знаний.

5. СПЕЦИАЛИЗИРОВАННЫЕ БАНКИ ЗНАНИЙ

В настоящее время в рамках Многоцелевого банка знаний начата разработка специализированных банков знаний, предназначенных для решения прикладных задач в предметных областях, — «математика», «химия», «медицина», «преобразование программ», а также для реализации пользовательских интерфейсов программных средств с локальной и распределенной архитектурой, с использованием программных средств Многоцелевого компьютерного банка знаний.

Специализированный банк математических знаний [6—8] предназначен для накопления математических знаний, а также для обучения студентов их формализации и доказательству теорем с помощью компьютеров. Информационное наполнение банка содержит системы аксиом, определения, теоремы и их доказательства, отнесенные к различным разделам математики. Программное наполнение банка состоит из автоматизированного рабочего места студента, включающего в себя редактор математических знаний, подсистемы автоматического и автоматизированного конструирования доказательств и их визуализации.

Специализированный банк знаний в области медицинской диагностики [9] предназначен для поддержки медицинского образования, медицинских научных исследований, медицинских консультаций и дистанционной диагностики заболеваний. Информационное наполнение банка знаний в области медицинской диагностики включает в себя модель онтологии медицинской диагностики, базы наблюдений и базы заболеваний различных разделов медицины, а также архивы историй болезни. Программное наполнение банка знаний включает в себя редакторы баз наблюдений, знаний о заболеваниях и историй болезни, а также программы оптимизации этих баз и программы медицинской диагностики, предполагающие применение высокопроизводительных вычислительных систем для повышения эффективности логических рассуждений [10].

Специализированный банк знаний по химии [11—13] предназначен для решения прикладных задач с использованием ее онтологии и знаний. Информационные ресурсы банка содержат модель онтологии химии и согласованную с онтологией базу знаний. Программные компоненты — это редакторы информационного наполнения и программы для решения прикладных задач. Классы прикладных задач, которые пользователь может решать с помощью специализированного банка, определяются содержанием модели онтологии и

базы знаний. Примеры: нахождение пути синтеза химических соединений; предсказание физико-химических свойств соединений; определение класса химического элемента, соединения или реакции по описанию свойств объекта классификации; вычисление значений различных характеристик химического процесса в любой момент процесса; проверка непротиворечивости знаний и онтологии относительно результатов химических экспериментов. Первая версия специализированного банка знаний по химии содержит модель онтологии физической и органической химии в объеме университетского курса, а также модели онтологии некоторых разделов аналитической химии.

Общая задача специализированного банка знаний по преобразованиям компьютерных программ [14] заключается в накоплении знаний о преобразованиях программ, а также в обеспечении возможностей проведения компьютерных экспериментов в науке, поддержке активных форм обучения студентов в образовании и макетировании оптимизирующих и распараллеливающих компиляторов в профессиональной деятельности. Информационное наполнение банка состоит из онтологии: знаний о преобразованиях программ; языков программирования; онтологии модели структурных программ; а также баз знаний о преобразованиях программ, архива программ и историй их преобразований. Программное наполнение банка включает в себя средства редактирования информационного наполнения; управляемые базами знаний преобразователи программ, средства визуализации историй преобразований программ и генераторы объектного кода на различные платформы. Также в программное наполнение входит инструментальное средство построения макетов оптимизирующих компиляторов.

Специализированный банк знаний о пользовательском интерфейсе [15] предназначен для проектирования и реализации пользовательских интерфейсов прикладных программ в Многоцелевом банке знаний на основе его информационных ресурсов. Проектирование интерфейса представляет собой разработку его модели средствами редакторов банка. Далее по этой модели автоматически генерируется пользовательский интерфейс, который связывается с соответствующей прикладной программой. Компоненты модели интерфейса — модель системы понятий предметной области, модель прикладной программы, модель выразительных средств интерфейса, модель сценария диалога, а также модели взаимосвязей между ними. Каждый компонент, за исключением модели предметной области, формируется на основе соответствующей универсальной онтологии, хранящейся в банке. Модель системы понятий предметной области является информационным ресурсом того специализированного банка, для которого разрабатывается прикладная программа.

6. ОБЕСПЕЧЕНИЕ КАЧЕСТВА ИНФОРМАЦИОННОГО И ПРОГРАММНОГО НАПОЛНЕНИЯ СПЕЦИАЛИЗИРОВАННЫХ БАНКОВ ЗНАНИЙ

Как уже отмечено, каждый специализированный банк знаний (за исключением студенческого) состоит из основной и экспериментальной областей. Перевод информационных и программных ресурсов из экспериментальной в основную область возможен только в том случае, если их качество удовлетворяет предъявляемым требованиям. Поэтому в рамках работ по созданию и исследованию методов и средств управления информационными ресурсами разрабатываются модели, методы и средства контроля качества информационных и программных ресурсов.

Поскольку специализированные банки предназначены для накопления таких информационных ресурсов (данных, знаний, онтологий и спецификаций языков), которыми можно многократно пользоваться, то необходимо своевременное их оценивание. Важно предоставить создателям информационных ресурсов возможность обнаружить в них дефекты и проконтролировать некоторые свойства моделей данных, знаний, онтологий и языков, тем самым, предоставляя пользователям ресурсов некоторые гарантии проверенности хранимой информации. Особенно важна проверка онтологий и языков, поскольку в их терминах формируется большая часть ресурсов банка. В основе оценивания онтологий положена универсальная классификация их свойств, обеспечивающая систематический подход к оцениванию ресурсов различных уровней общности.

Для оценивания качества программных ресурсов разработана процедура испытаний программного наполнения специализированных банков знаний, определяющая порядок проверки программного обеспечения на соответствие его спецификациям. Перевод программных ресурсов в основные области банков знаний должен осуществляться только в случае успешного их оценивания в соответствии с данной процедурой. В настоящее время проводятся испытания прототипа системы управления информационными ресурсами. Результатами этих испытаний, помимо выявленных дефектов, станут методика испытаний, формы документов по представлению результатов таких испытаний и сами эти документы. Методика испытаний, формы документов и результаты испытаний могут быть полезны при испытаниях следующих версий системы и других подобных систем.

ЗАКЛЮЧЕНИЕ

Рассмотренная система управления информационными ресурсами различных уровней общности состоит из набора средств для создания, со-

проведения и использования информационных ресурсов в Интернете. В системе предусмотрено совмещение централизованного управления информационными ресурсами через администраторов (создание специализированных банков знаний и их разделов, перенос информации из экспериментальной в основную область, наделение пользователей и программ необходимыми полномочиями) и децентрализованного управления этими ресурсами со стороны остальных пользователей системы (создание и редактирование ресурсов, программ их обработки). Разработка системы — первый этап программы работ, направленной на создание и коллективное использование в Интернете информационных ресурсов различных уровней общности, а также программ для их обработки.

ЛИТЕРАТУРА

1. Клещев А. С., Черняховская М. Ю. Современное состояние компьютерной обработки знаний // Сб. науч. тр. — Владивосток, 2001. — С. 273—284.
2. Kleshchev A. S., Artemjeva I. L. Domain ontologies and knowledge processing. Technical Report 7—99, Vladivostok: IACP FEBRAS, — 1999. — P. 25.
3. Орлов В. А., Клещев А. С. Многоцелевой банк знаний. Ч. 1. Концепция и политика. — Владивосток: ИАПУ ДВО РАН, 2003. — 40 с.
4. Uschold M. Knowledge Level Modeling: Concepts and Terminology // The Knowledge Engineering Review. — 1998. — Vol. 13:1. — P. 5—29.
5. Орлов В. А., Клещев А. С. Многоцелевой банк знаний. Ч. 3. Концепция универсального редактора ИРУО. — Владивосток: ИАПУ ДВО РАН, 2003. — 28 с.
6. Гаврилова Т. Л., Клещев А. С. Анализ подходов к решению проблемы правильности математических знаний // Проблемы управления. — 2005. — № 3. — С. 13—19.
7. Гаврилова Т. Л., Клещев А. С. Проблема конструирования правильных интуитивных доказательств в классических аксиоматических системах. Конструирование полных доказательств. — Владивосток: ИАПУ ДВО РАН, 2003. — 24 с.
8. Гаврилова Т. Л., Клещев А. С. Проблема конструирования правильных интуитивных доказательств в классических аксиоматических системах. Модель интуитивного доказательства. — Там же. — 18 с.
9. Черняховская М. Ю. Представление знаний в экспертных системах медицинской диагностики. — Владивосток: ДВНЦ АН СССР, 1983. — 212 с.
10. Москаленко Ф. М. Экспертная система медицинской диагностики, основанная на реальной онтологии медицины, для многопроцессорной ЭВМ // Тр. II Междунар. конф. «Параллельные вычисления и задачи управления» / Ин-т пробл. упр. — М., 2004. — С. 999—1084.
11. Артемьева И. Л., Цветников В. А. Фрагмент онтологии физической химии и его модель // Электронный журн. «Исследовано в России». — 2002. — № 3. — С. 454—474. <<http://zhurnal.ape.relarn.ru/articles/2002/042.pdf>>.
12. Артемьева И. Л., Рештаненко Н. В. Модульная модель онтологии органической химии // Информатика и системы управления. — 2004. — № 2.
13. Артемьева И. Л., Рештаненко Н. В. Специализированный компьютерный банк знаний предметной области «Химия» // Искусственный интеллект. — 2004. — Т. 1. — С. 235—244.
14. Князева М. А. Банк знаний о преобразованиях компьютерных программ // Там же. — Т. 3. — С. 179—186.
15. Gribova V. V., Kleshchev A. S. From an ontology-oriented approach conception to user interface development // International Journal Information theories & applications. — 2003. — Vol. 10, N 1. — P. 87—94.

☎ (4232) 31-04-24

e-mail: kleshchev@iacp.dvo.ru

