

МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ НА СПЕЦИАЛИЗИРОВАННОМ ИНТЕРНЕТ-ПОРТАЛЕ В ОБЛАСТИ СОПРОТИВЛЕНИЯ МАТЕРИАЛОВ

Л.С. ГЛОБА, Р.Л. НОВОГРУДСКАЯ

Рассмотрен подход к созданию специализированного Интернет-портала знаний в области сопротивления материалов. Для построения модели представления знаний использована онтология, которая позволяет обеспечить единообразное представление знаний и данных, а также их семантическую связность. С помощью понятий онтологии построены внутренние хранилища портала, организовано его информационное наполнение, навигация и поиск.

ВВЕДЕНИЕ

В настоящее время существует огромное количество инженерного оборудования и программного обеспечения для решения специализированных задач разного рода. Однако они являются либо дорогими, либо закрытыми. Важные результаты как теоретические, так и практические, которые получают исследователи, остаются сосредоточенными лишь в конкретной научной среде. Происходит процесс сосредоточения знаний — они становятся доступными узкому кругу ученых и специалистов. Поэтому актуальной задачей становится предоставление возможности использования этих знаний для большего количества исследователей, для которых они предназначены.

Кроме того, недостатком процесса доступа к знаниям является плохая систематизированность и слабая структурированность больших объемов информации в сети Интернет, ее распределенность по разным Интернет-сайтам, электронным библиотекам и архивам.

Из этого следует необходимость создания подхода направленного на построение специализированного Интернет-портала знаний, ориентированного на работу с большим количеством разнообразных информационных и вычислительных ресурсов определенного направления, который не только даст возможность систематизировать и осуществлять поиск информации, но и реализовывать определенные расчетные задачи.

В последнее время при построении порталов научных знаний все чаще используются онтологические модели. В качестве примера можно предложить порталы знаний в области археологии и компьютерной лингвистики [1, 2], при построении которых в качестве модели представления знаний была использована онтологическая модель. Отличительной чертой предложенного подхода является введение в общую онтологию портала дополнительной компоненты — онтологии инженерных расчетов, позволяющей не только описывать инженерные расчеты и их информационное окружение, но и давать возможность конечным пользователям решать свои конкретные задачи в области сопротивления материалов. В результате наличия данной

онтологии предоставляется возможность формально описать и установить взаимосвязь между предметной областью (элементами онтологии инженерных знаний и онтологии предметной области) и непосредственно сервисами, с помощью которых выполняются расчетные задачи.

ПОСТАНОВКА ЗАДАЧИ

Приведем формализованную постановку задачи.

Дано: D — данные, I — информационные ресурсы, $Camp$ — расчетные задачи.

Получить: $\Pi = \{M \cup Inf_{str} \cup Search(Inf_{str}, Camp) \cup Serv\}$, где Π — портал знаний в области сопротивления материалов; M — модель представления знаний на портале $M = O_{портала} = \{C \cup R \cup F \cup D \cup T \cup A\}$; Inf_{str} — структурированная информация, которую мы получаем с помощью функции структурирования; $Search(Inf_{str}, Camp)$ — качественный поиск; $Serv$ — сервисы, с помощью которых выполняются расчетные задачи.

ОНТОЛОГИЯ ПОРТАЛА ЗНАНИЙ

Онтология портала знаний формально может быть задана как: $O = \{C, A, R, T, F, D\}$, где C — множество классов, описывающих понятия некоторой предметной или проблемной области; A — множество атрибутов, описывающих свойства понятий и отношений; R — множество отношений, заданных на классах; T — множество стандартных типов значений атрибутов (string, integer, real, date); F — множество ограничений на значения атрибутов понятий и отношений; D — множество экземпляров классов [3]. Таким образом, онтология портала представляет собой иерархию понятий (или классов), связанных отношениями. Различные свойства каждого понятия описываются на основе атрибутов понятий и ограничений, наложенных на область их значений. Множество отношений состоит из следующих видов отношений: R_{AS} — ассоциативное отношение, R_{IA} — отношение «часть–целое», R_n — отношение наследования, R_{CD} — отношение «класс–данные» — $R = \{R_{AS}, R_{IA}, R_n, R_{CD}\}$. Множество ограничений используется не для всех атрибутов, а лишь для тех, значения которых должны лежать в некоторой области и не могут быть меньше/больше заданной величины, либо определяются каким-либо другим правилом. Так, например, на значение атрибута дата начала класса мероприятия онтологии инженерной деятельности накладывается ограничение, $T(\text{дата}) = \text{date}$, $F(T) > 0$.

КОМПОНЕНТЫ ОНТОЛОГИИ ПОРТАЛА «ПРОЧНОСТЬ МАТЕРИАЛОВ»

Разработанная модель представления знаний включает такие онтологии, как онтология науки (Онтология инженерной деятельности, Онтология инженерного знания и Онтология инженерных расчетов) и Онтология предметной области. Эти онтологии и образуют онтологию портала —

$O_{\text{портала}} = \{O_1(O_2, O_3, O_4)O_5\}$ (рис. 1). Экземпляры классов понятий и отношений, определенных в онтологии портала, образуют его информационное наполнение.

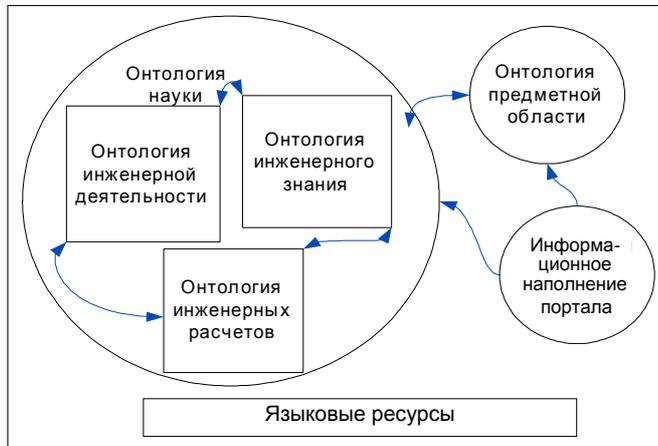


Рис. 1. Компоненты модели представления знаний на портале

Исходными данными для модели представления знаний, характеризующими предметную область, являются языковые ресурсы, представленные в виде коллекции документов.

Онтология науки — O_1 — состоит из трех онтологий — $O_1 = \{O_2, O_3, O_4\}$.

Онтология инженерной деятельности (ОИД) — O_2 — включает общие классы понятий, относящиеся к организации научной деятельности [4], такие как Персона, Организация, Мероприятие, Деятельность, Местоположение, Литература, Документация, Учебные материалы, Публикация, Информационный ресурс. Эти классы связаны отношениями различных типов. Различные свойства каждого понятия описываются на основе атрибутов понятий и ограничений, наложенных на область их значений. Исходя из этого, ОИД может быть описана следующим образом: $O_2 = \{C_{O_2}, A_{O_2}, R_{O_2}, T_{O_2}, F_{O_2}, D_{O_2}\}$.

Онтология инженерного знания (ОИЗ) — O_3 — содержит следующие метапонятия, задающие структуры для описания рассматриваемой предметной области: Назначение исследования (позволяет типизировать проводимые исследования), Метод исследования и Объект исследования (задают типизацию методов и объектов исследования и структуры для их описания), Научный результат (служит для типизации и описания результатов научной деятельности), Исследовательское оборудование (организовывает типизацию оборудования для проведения исследований). Существуют связи между классами заданные отношениями различных типов. Исходя из этого, ОИЗ может быть описана следующим образом: $O_3 = \{C_{O_3}, A_{O_3}, R_{O_3}, T_{O_3}, F_{O_3}, D_{O_3}\}$.

Онтология инженерных расчетов (действий) (ОИР) — O_4 — объединяет классы, которые описывают расчетные возможности, реализованные на портале: Расчет, Результат, Сервис, Параметры. Данные классы связаны отношениями различных типов. Свойства и характеристики каждого понятия описываются с помощью атрибутов понятий и ограничений, наложенных на область значений. Исходя из этого, ОИД может быть описана следующим образом: $O_4 = \{C_{O_4}, A_{O_4}, R_{O_4}, T_{O_4}, F_{O_4}, D_{O_4}\}$. ОИР играет важную роль при

построении общей онтологии портала. С ее помощью становится возможным описать не только доступные пользователю на портале расчетные задачи как сервисы вычислений, но и установить связи между конкретными сервисами (набором сервисов) для реализации конкретных бизнес-процессов, а также связи сервисов с информационными ресурсами, что необходимо для решения производственных задач пользователя.

Онтология предметной области (ОПО) — отражает общие знания о предметной области, такие как иерархия классов понятий и семантические отношения на этих классах. Основой ОПО для портала инженерных знаний в области сопротивления материалов послужила системная классификация, состоящая из фрагментов определенной универсальной структуры.

Экземпляры классов понятий и отношений, определенных в онтологии портала, образуют его информационное наполнение.

Исходными данными для модели представления знаний, характеризующими предметную область, являются языковые ресурсы, представленные в виде коллекции документов. Обеспечить автоматическое извлечение знаний из этих данных является главной задачей эксперта при наполнении и настройке системы знаний портала.

ТЕХНОЛОГИЯ ПОСТРОЕНИЯ ПОИСКОВОГО ЗАПРОСА

Поиск информации базируется на онтологии, благодаря чему пользователь имеет возможность задавать запрос не только и не столько по ключевым словам, сколько в знакомых ему терминах предметной области портала. Основными элементами такого поискового запроса являются основные понятия онтологии: ее классы, атрибуты и экземпляры классов, а также отношения, связывающие одни понятия с другими понятиями онтологии [5]. Сформулированный таким образом поисковый запрос представляется простым для задания пользователем, а также полным в плане найденной информации. Например, поисковый запрос: «Найти результаты исследования, проводимые со сталью и направленные на определение ее свойств под воздействием высоких температур, полученные В.А. Гайном в 2008 г. и описанные в книге „Поведение стали“, изданной институтом Материаловедения Германии», формально будет выглядеть следующим образом:

Класс «Результат исследования»

Отношение «проводится с»:

Класс «Объект исследования»

Атрибут «Объект исследования» = «сталь»

Отношение «направлено на»:

Класс «Назначение исследования»

Атрибут «Назначение исследования» = «воздействие высоких температур»

Отношение «полученные»:

Класс «Персона».

Атрибут «ФИО» = «В.А. Гайном»

&

Класс «Результат исследования»

Атрибут «год» = «2008»

Отношение «описанные»:

Класс «Литература»

Отношение «часть–целое»:

Подкласс «Учебные материалы»

Атрибут «книга» = «Поведение стали»

Отношение «выданной»:

Класс «Организация»

Атрибут «Институт» = «институтом Материаловедения Германии».

Покажем как будут выглядеть термины поискового запроса в понятиях онтологии. Приведем их формальное описание. Зададим следующие переменные: $C_1 = \{\text{результат исследования}\}$, $R_{AS_1} = \{\text{проводится с}\}$, $C_2 = \{\text{объект исследования}\}$, $A_{C_2} = \{\text{сталь}\}$, $R_{AS_2} = \{\text{направлено на}\}$, $C_3 = \{\text{назначение исследования}\}$, $A_{C_3} = \{\text{воздействия высокими температурами}\}$, $R_{AS_3} = \{\text{полученные}\}$, $C_{4_1} = \{\text{персона}\}$, $A_{C_{4_1}} = \{\text{В.А. Гайном}\}$, $C_{4_2} = \{\text{результат исследования}\}$, $A_{C_{4_2}} = \{\text{2008}\}$, $R_{AS_4} = \{\text{описанные}\}$, $C_5 = \{\text{литература}\}$, $R_{IA_{56}} = \{\text{подкласс}\}$, $C_6 = \{\text{учебные материалы}\}$, $A_{C_6} = \{\text{поведение стали}\}$, $R_{AS_5} = \{\text{выданной}\}$, $C_7 = \{\text{организация}\}$, $A_{C_7} = \{\text{институтом Материаловедения Германии}\}$.

Для реализации поискового запроса актуально следующее описание классов (согласно описанной выше идеологии: $C_2(A_{C_2}, D_{C_2}, R_{C_2})$, $C_3(A_{C_3}, D_{C_3}, R_{C_3})$, $C_4(A_{C_4}, D_{C_4}, R_{C_4})$, $C_5(A_{C_5}, D_{C_5}, R_{C_5})$, $C_6(A_{C_6}, D_{C_6}, R_{C_6})$, $C_7(A_{C_7}, D_{C_7}, R_{C_7})$ [6]. Для реализации поискового запроса существуют следующие отношения (согласно описанной выше идеологии): отношение вида класс–данные: $R_{IA_{56}} \leftarrow C_5 \subset C_6$, ассоциативные отношения: $R_{AS_1} = \{C_1 \times C_2\}$, $R_{AS_2} = \{C_1 \times C_3\}$, $R_{AS_3} = \{C_1 \times C_4\}$, $R_{AS_4} = \{C_1 \times C_5\}$, $R_{AS_5} = \{C_5 \times C_7\}$.

С помощью заданного формального описания понятий поисковый запрос будет выглядеть следующим образом (таблица).

Т а б л и ц а . Формализация поискового запроса

Поисковый запрос	Описание запроса
$C_1 \vdash$	«результаты исследования»
$c \xrightarrow{R_{AS_2}} A_{C_2},$	«проводимые со сталью»
$\xrightarrow{R_{AS_3}} A_{C_3},$	«определение ее свойств при воздействии высокими температурами»
$\xrightarrow{R_{AS_4}} A_{C_{4_1}}, \& A_{C_{4_2}},$	«полученные В.А. Гайном в 2008 г.»
$\xrightarrow{R_{AS_5}} A_{C_5} \xrightarrow{R_{AS_6}} A_{C_7}.$	«описанные в книге „Поведение стали” выданной институтом Материаловедения Германии»

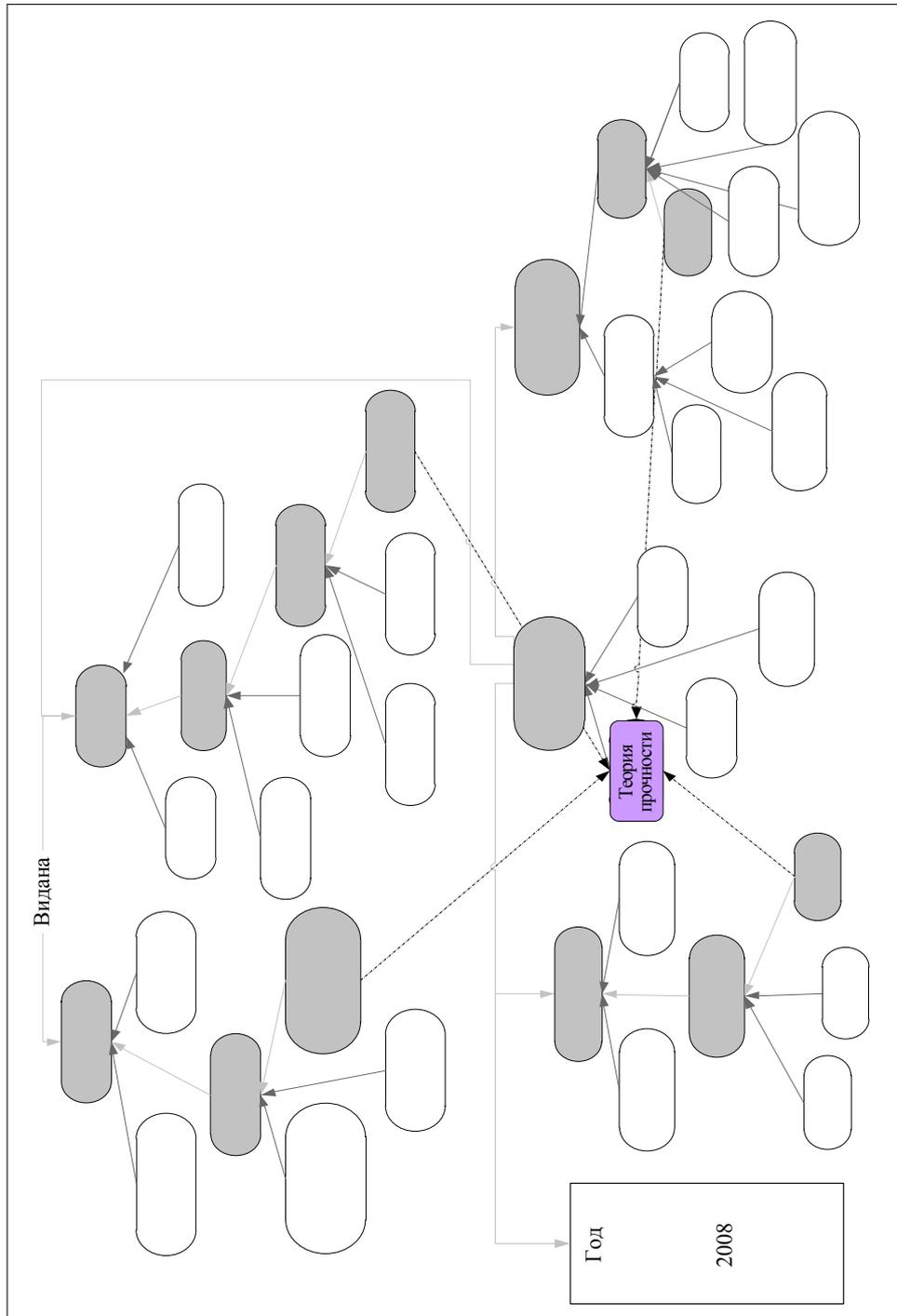


Рис. 2. Процесс поиска по элементам онтологии портала

На рис. 2 изображен процесс поиска по элементам онтологии портала. Экземпляры классов и отношения онтологии, заданные пользователем в терминах поискового запроса, выделены курсивом; элемент, который является результатом поиска, выделен жирным. Класс *Результат исследования* связан соответствующими отношениями с классами *Объект исследования*,

Персона и Литература. В этих классах находим экземпляры, заданные пользователем: Сталь, Гайн, книга «Поведение стали» и Институт материаловедения Германии, находим экземпляр класса *Результат исследования*, который им соответствует. Это — теория прочности.

ЗАКЛЮЧЕНИЕ

За счет структуризации и систематизации информации и вычислительных ресурсов, связей между ними, которые реализуются с помощью построения модели представления знаний на портале, осуществляется эффективный поиск и навигация по информационному пространству портала инженерных знаний.

Формализация знаний делает систему знаний портала легко расширяемой и настраиваемой — в нее могут интегрироваться как новые знания (например, о новых направлениях и методах исследования прочности материалов), так и новые виды информационных и вычислительных ресурсов.

Ближайшими целями авторов является апробация предложенного подхода. Описание онтологии портала на языке OWL. Кроме того, планируется создать метаописатели сервисов, которые будут использоваться для решения вычислительных задач. С помощью метаописателей осуществляется подключение необходимой информации, а также взаимодействие сервисов, необходимых конкретному бизнес-процессу для решения конкретной инженерной задачи.

ЛИТЕРАТУРА

1. Андреева О.А., Боровикова О.И., Загорюлько Ю.А. и др. Археологический портал знаний: содержательный доступ к знаниям и информационным ресурсам по археологии // Тр. 10-й нац. конф. по искусственному интеллекту с междунар. участием КИИ-2006. — М.: Физматлит, 2006. — Т. 3. — С. 832–840.
2. Загорюлько Ю.А., Боровикова О.И. Технология построения онтологий для порталов знаний по гуманитарным наукам // Тр. Всероссийской конф. с междунар. участием «Знания-Онтологии-Теории» (ЗОНТ-07). — Новосибирск. — 2007. — Т. 1. — С. 191–200.
3. Боровикова О.И., Загорюлько Ю.А., Сидорова Е.А. Подход к автоматизации сбора онтологической информации для интернет-портала знаний // Компьютерная лингвистика и интеллектуальные технологии: тр. междунар. конф. Диалог-2005 (Звенигород, 1–5 июня 2005 г.). — М.: Наука, 2005. — С. 65–70.
4. Глоба Л.С., Новогрудская Р.Л. Подход к построению интернет-портала инженерных знаний // Тр. X междунар. науч. конф. им. Т.А. Таран «Интеллектуальный анализ информации». — 2010. — С. 53–62.
5. Боровикова О.И., Загорюлько Ю.А. Организация порталов знаний на основе онтологий // Тр. междунар. семинара Диалог-2002 «Компьютерная лингвистика и интеллектуальные технологии». — Протвино. — 2002. — Т. 2. — С. 76–82.
6. Глоба Л.С., Новогрудская Р.Л. Систематизация информационных ресурсов Интернет-портала «Прочность материалов» // Вісн. Харк. нац. ун-ту ім. Каразіна. Сер. Матиматичне моделювання. Інформаційні технології. Автоматизовані системи управління. — 2011. — Вип. 16, № 927. — С. 27–36.

Поступила 02.06.2010