

ТЕХНОЛОГИЯ ПОСТРОЕНИЯ ОНТОЛОГИЙ ДЛЯ ПОРТАЛОВ НАУЧНЫХ ЗНАНИЙ

Введение

В настоящее время очень остро стоит проблема эффективного использования огромного объема знаний и информационных ресурсов, накопленных в различных областях человеческой деятельности. Однако доступ к этим знаниям и ресурсам существенно ограничен из-за того, что они слабо структурированы, плохо систематизированы, рассредоточены по различным интернет-сайтам, библиотекам и архивам. В то же время в процессе научной деятельности у исследователей возникает необходимость в эффективном доступе к публикациям и другим информационным ресурсам, содержащим описания методов и подходов, разработанных в рамках интересующей их научной дисциплины.

Чтобы удовлетворить описанную выше потребность, нами была предложена концепция и архитектура специализированного интернет-портала знаний [Боровикова, Загорулько, 2002; Zagorulko et al., 2005], который должен обеспечивать интеграцию и систематизацию научных знаний и информационных ресурсов определенной тематики, а также содержательный доступ к ним из любой «точки» интернет-пространства. Более того, согласно этой концепции, портал знаний должен обеспечивать доступ не только к собственным информационным ресурсам, но и поддерживать навигацию по релевантным тематике портала ресурсам, размещенным в сети Интернет.

Кроме поддержки гибкого и целостного представления научной дисциплины и различных аспектов научной деятельности (персоналий, организаций, событий, объектов, методов и результатов исследований), обеспечения содержательного доступа к интегрированным знаниям и информационным ресурсам, относящимся к данному научному направлению, важным требованием к portalу знаний является возможность его декларативной настройки на заданную предметную область не только в процессе создания, но и в ходе эксплуатации. Возможность настройки портала позволит отслеживать динамику появления новых знаний и информационных ресурсов по тематике портала и тем самым обеспечит поддержку его актуальности и полезности. Выполнение указанных требований стало возможным благодаря выбору в качестве концептуальной основы и информационной модели портала знаний онтологии.

В этой работе мы используем понятие «онтология» в том смысле, как оно применяется в информатике и искусственном интеллекте [Гаврилова, Хорошевский, 2001]. Мы считаем, что одной из целей онтологии является описание и изучение сущностей, существующих в реальном мире и/или сознании человека. Для систем информатики и искусственного интеллекта, в частности порталов знаний, существует только то, что уже в них представлено или может быть представлено, поэтому мы придерживаемся определения онтологии, данного в работе [Gruber, 1995]. Согласно этому определению, онтология является точной спецификацией концептуализации. Причем под концептуализацией понимается некоторая абстракция, т. е. упрощенное представление мира, построенное для определенной цели. Концептуализация включает объекты, понятия и другие сущности, которые предполагаются существующими в рассматриваемой области, а также отношения между ними.

Следует заметить, что в работе [Ushold, Gruninger, 1996] подчеркивается, что онтология является спецификацией концептуализации, но только в той ее части, которая зависит от определенной области интересов. В работе [Takeda et al., 1998] делается упор на то, что онтологии должны помочь в решении проблем, возникающих из-за того, что в разных областях существуют различные интерпретации одних и тех же терминов. В этой связи онтология рассматривается как соглашение о некоторой области интересов для достижения определенных целей.

Для установления соглашения о знаниях, представленных на некотором, в частности, логическом языке, по мнению N. Guarino [Guarino, Giaretta, 1999], онтология должна характеризовать концептуализацию, ограничивая возможные значения предикатов и функций. В его понимании, онтология – это логическая теория, аксиомы которой ограничивают интерпретации нелогических символов языка.

Основываясь на приведенных выше определениях, можно сказать, что онтология представляет собой точное подробное описание (модель) некоторой части мира (предметной области) применительно к конкретной области интересов. В контексте данной работы онтология будет представлять собой описание некоторой научной дисциплины и связанной с ней научной деятельности, связанной с этой дисциплиной. Задача разработки технологии построения онтологий для порталов научных знаний является весьма актуальной. Описанию нашего опыта создания и использования такой технологии и посвящена данная работа.

Онтология портала знаний

Онтология должна не только обеспечивать формальное представление системы понятий предметной области (ПО) портала, но и поддерживать всю требуемую функциональность, т. е. служить основой для реализации эффективного представления разнообразной информации по тематике портала и содержательного доступа к ней, а также поддерживать интеграцию в информационное пространство портала релевантных ему информационных ресурсов и удобную навигацию по нему. Для эффективного представления знаний о ПО онтология должна обеспечивать описание понятий со сложной структурой и разнообразных семантических связей между ними. Важным требованием к онтологии портала является возможность выстраивания понятий предметной области в иерархию «общее–частное» и поддержка наследования свойств по этой иерархии.

Так как онтология должна служить целям содержательной декларативной настройки портала на заданную область знаний и поддерживать всю его функциональность, она должна быть организована таким образом, чтобы по ней могли автоматически строиться схема внутренней базы данных портала, формы для заполнения БД портала данными, схема навигации по информационному пространству портала и формы поисковых запросов. Для упрощения настройки портала на выбранную область знаний в онтологии необходимо выделить структуры, независимые от предметной области (ПО) портала. Кроме того, онтология должна обладать такими свойствами, как расширяемость и интегрируемость с уже существующими онтологиями.

Онтология портала знаний, удовлетворяющая описанным выше требованиям, формально может быть задана семеркой вида:

$$O = \langle C, A, R_C, T, D, R_A, F \rangle,$$

где C – множество классов, описывающих понятия некоторой предметной или проблемной области; A – множество атрибутов, описывающих свойства понятий и отношений; $R_C = \{r_c | r_c \subseteq C \times C\}$ – множество отношений, заданных на классах (понятиях); T – множество стандартных типов значений атрибутов (string, integer, real, date); D – множество доменов (множеств значений стандартного типа string); $R_A = R_{AT} \cup R_{AD}$, где $R_{AT} \subseteq A \times T$ – отношение, связывающее атрибуты и типы данных, из которых они могут принимать свои значения, $R_{AD} \subseteq A \times D$ – отношение, определяющее для каждого атрибута его дискретное множество значений (домен); F – множество ограничений на значения атрибутов понятий и отношений.

С содержательной точки зрения определенная таким образом онтология может служить для представления понятий, необходимых как для описания определенной области знаний, так и для выполняемой в ее рамках научной деятельности. Именно такая онтология была

выбрана в качестве информационной модели портала научных знаний. Вводя формальные описания понятий предметной области в виде классов объектов и отношений между ними, онтология портала задает структуры для представления реальных данных и связей между ними. Сами данные на портале представлены как множество связанных информационных объектов.

Каждый информационный объект (ИО) соответствует определенному классу онтологии (является его экземпляром) и представляет собой описание некоторого объекта предметной области. Между информационными объектами могут существовать связи, семантика которых определяется отношениями, заданными между соответствующими классами онтологии. Информационное наполнение портала осуществляется с помощью управляемого онтологией редактора данных, который позволяет создавать, редактировать и удалять информационные объекты и связи между ними. Формы для ввода конкретных ИО и их связей автоматически генерируются по онтологии портала. Таким образом, информационное содержание (контент) портала включает как знания общего характера (представлены в онтологии), так и знания о конкретных объектах и отношениях между ними (представлены информационными объектами и их связями). Для содержательного доступа к этому контенту портал знаний предоставляет развитые средства навигации и поиска.

Навигация по контенту портала осуществляется в соответствии с содержанием онтологии: обеспечивается переход от классов онтологии к их экземплярам, а в дальнейшем переход по онтологическим связям (отношениям) от конкретного экземпляра к спискам связанных с ним экземпляров и т. д.

Поиск информации также базируется на онтологии, благодаря чему пользователю предоставляется возможность задания запроса в терминах предметной области портала. Так, при выборе класса искомых информационных объектов автоматически генерируется поисковая форма, в которой пользователь может задать ограничения на значения атрибутов объектов выбранного класса, а также на значения атрибутов объектов, связанных с данным объектом ассоциативными отношениями. Например, запрос «Найти объекты исследования, датированные бронзовым веком, найденные в экспедициях в период с 1920 по 1990 год» формально будет выглядеть следующим образом:

Класс «Объект исследования»:

Отношение «Датирован»:

Класс «Период»

Атрибут «Название периода» = «бронзовый век»

Отношение «Исследует»:

Класс «Экспедиция»

Атрибут «Дата начала»: (≥ 1920) & (≤ 1990)

Атрибут «Дата окончания»: (≥ 1920) & (≤ 1990)

Технология построения онтологии портала знаний

Технология построения онтологии портала включает язык описания онтологий, методику построения и развития онтологий, а также редактор онтологий. Язык описания онтологий и редактор онтологий выбирались и проектировались таким образом, чтобы они были понятны экспертам, не являющимся специалистами в области информатики и представления знаний, и ими было просто и удобно пользоваться. Кроме того, редактор онтологий строился с учетом его применения для распределенной разработки онтологий.

Методика построения онтологии для портала знаний, с одной стороны, практически полностью определяется ее структурой, а с другой – поддерживается и в то же время ограничивается средствами, предоставляемыми редактором онтологий. В процессе разработки онтологии выделяются и формально описываются классы понятий и их свойства. Свойства каждого понятия представляются с помощью атрибутов и ограничений, наложенных на область их значений, а также бинарных отношений, связывающих данное понятие с другими понятиями. Классы понятий выстраиваются в иерархию с помощью отношения наследования («общее–частное»). Механизм наследования задан таким образом, что понятию–потомку от родительского понятия передаются не только все его атрибуты, но и отношения.

Структура онтологии портала знаний. Для того чтобы онтология удовлетворяла целям портала, она должна быть хорошо структурирована и адекватно отражать его проблемную и предметную область. В связи с этим онтология портала разделяется на предметно-независимые (базовые) онтологии и онтологию предметной области (рис. 1). В качестве базовых выбраны онтологии научной деятельности и научного знания, которые не зависят от предметной области портала.

Онтология научной деятельности фактически является онтологией верхнего уровня и включает базовые классы понятий, относящиеся к организации научной и исследовательской деятельности, такие как *Исследователь*, *Организация*, *Событие*, *Деятельность*, *Публикация*, используемые для описания участников научной деятельности, мероприятий, научных программ и проектов, различного типа публикаций. В эту онтологию также включен класс *Информационный ресурс*, который служит для описания информационных ресурсов, представленных в сети Интернет.

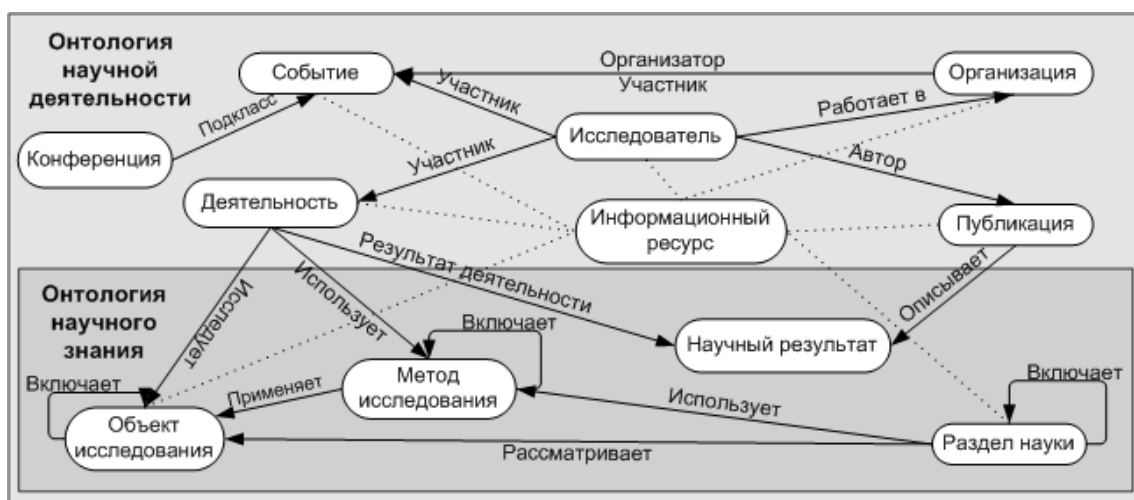


Рис. 1. Базовые онтологии портала знаний

Онтология научного знания, по своей сути, является метаонтологией. Она фиксирует основные содержательные структуры, используемые для построения онтологий более низкого уровня, т. е. онтологий предметных областей, описывающих конкретные отрасли знаний или научные дисциплины. В частности, эта онтология содержит метапонятия, задающие структуры для описания понятий конкретной области знаний, такие как *Раздел науки*, *Метод исследования*, *Объект исследования*, *Научный результат*. Используя эти метапонятия, можно выделить и описать значимые для данной науки разделы и подразделы, задать типизацию методов и объектов исследования, описать результаты научной деятельности.

Понятия онтологии научного знания связаны между собой и понятиями онтологии научной деятельности ассоциативными отношениями. Выбор ассоциативных отношений осуществлялся не только исходя из полноты представления проблемной и предметной областей портала, но и с учетом удобства навигации по его информационному пространству и поиска информации. Перечислим основные ассоциативные отношения:

- «*Научное направление*» – связывает события, публикации, организации, исследователей и информационные ресурсы с разделами науки;
- «*Описывает*» – задает связь публикации с научным результатом, объектом или методом исследования;
- «*Использует*» – связывает метод исследования с деятельностью, исследователем или разделом науки;
- «*Исследует*» – сопоставляет какую-либо деятельность или раздел науки с объектом исследования;
- «*Результат деятельности*» – связывает научный результат с деятельностью;
- «*Является ресурсом*» – служит для связывания информационного ресурса с любым понятием онтологии;

«Работает в» – связывает исследователя с организацией, в которой он работает.

Заметим, что последнее отношение имеет три дополнительных атрибута «должность», «дата принятия» и «дата увольнения», служащих для уточнения, в какой должности работает исследователь и с какого времени.

Онтология предметной области портала знаний описывает научную дисциплину в целом как раздел науки. Она строится для организации эффективного доступа к знаниям и информационным ресурсам по определенной научной тематике и, следовательно, должна соответствовать описанным выше требованиям. Понятия онтологии предметной области, описывая конкретную научную дисциплину, в то же время являются реализациями метапонятий онтологии научного знания и могут быть упорядочены в иерархию «общее–частное».

Как правило, онтология конкретной предметной области включает как минимум четыре базовые иерархии: иерархию разделов, иерархию объектов, иерархию методов исследования и иерархию научных результатов. Построение иерархии понятий является важным моментом при разработке онтологии предметной области. При задании такой иерархии необходимо учитывать не только то, насколько она полно описывает предметную область и связанные с ней информационные ресурсы, но и насколько удобно пользователю с ее помощью осуществлять навигацию по информационному пространству портала и вести содержательный поиск.

Методы развития онтологии. В процессе эксплуатации портала знаний могут появляться новые знания о его предметной области, обнаруживаться пробелы и неточности в знаниях, уже представленных в онтологии. Все это, безусловно, требует изменения онтологии. Однако при редактировании онтологии необходимо следить за тем, чтобы не нарушилась логическая целостность системы знаний портала и не произошла потеря информации.

Изменение онтологии может заключаться в расширении или перестройке ее системы понятий, удалении или переименовании понятий и/или отношений. Вначале рассмотрим случаи, связанные с расширением системы понятий онтологии.

В простейшем случае такое расширение заключается в добавлении нового атрибута какому-либо понятию. Здесь требуется учитывать, что такой атрибут уже мог быть у понятий, являющихся потомками данного понятия. Поэтому нужно просмотреть всех таких потомков и, если необходимо, выполнить переименования соответствующих атрибутов. Добавление нового понятия в самый низ иерархии понятий не требует каких-либо усилий по поддержанию целостности системы понятий, так как при этом новое понятие унаследует все атрибуты и связи вышестоящих понятий. Если добавляется понятие, которое станет корневым в одной из иерархий понятий, необходимо учитывать атрибуты и связи нижестоящих понятий. Возможно, потребуется перемещение части атрибутов и связей в новое понятие, тем более, учитывая перспективу появления новых ветвей иерархии, берущих начало из нового понятия.

Вставка нового понятия в иерархию между двумя «старыми» понятиями также требует некоторых методологических усилий. Чтобы избежать дублирования и возможных коллизий имен, необходимо аккуратно выбрать для него атрибуты и связи из нижестоящих понятий.

При удалении «листового» понятия, т. е. понятия, находящегося в самом низу иерархии, стоит подумать о передаче его собственных атрибутов и связей вышестоящему понятию, чтобы не произошло потери знаний. Нужно иметь в виду, что если на основе удаляемого понятия уже созданы информационные объекты, то, чтобы не потерять данные, необходимо эти информационные объекты привязать к предку удаляемого понятия. Но этого может оказаться недостаточно для сохранения всей информации об этих объектах, если предварительно вышестоящему понятию не будут переданы все собственные атрибуты и связи удаляемого понятия.

Если удаляемое понятие не является «листовым» (рис. 2, а), то перед его удалением необходимо подумать о передаче его атрибутов и связей нижестоящему понятию (рис. 2, б). Информационные объекты, как и в случае с «листовым понятием», должны быть привязаны к вышестоящему понятию и модифицированы в соответствии с его структурой.

Удаление «корневых» понятий онтологии портала, находящегося в эксплуатации или на этапе информационного наполнения, не рекомендуется из-за возможной потери информации.

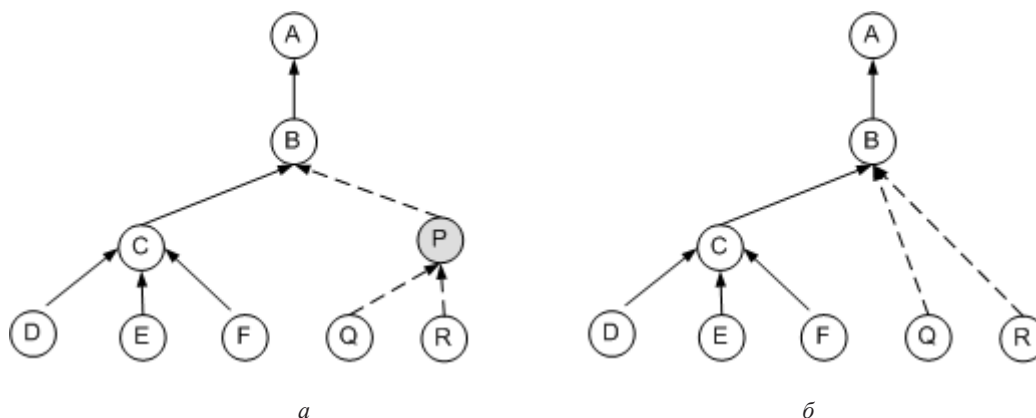


Рис. 2. Удаление «нелистового» понятия

При удалении атрибутов из понятий также нужно учитывать возможную потерю информации. Частным случаем удаления атрибута является его перемещение в понятие более высокого или более низкого уровня, когда выясняется, что данный атрибут является более общим или, наоборот, более специфическим. В первом случае потери информации не происходит, так как перемещаемый атрибут все равно будет унаследован модифицируемым понятием. Во втором случае такая потеря возможна, и нужно принять меры для ее восстановления.

Иногда приходится перемещать понятие внутри иерархии. При этом нужно учитывать, что меняется не только набор наследуемых понятием атрибутов, но и связей. Возможно, что некоторые потерянные в результате этого перемещения атрибуты и связи придется восстанавливать «вручную».

Достаточно интересным случаем представляется перемещение поддеревьев из одной ветки иерархии в другую. Этот случай практически рекурсивно сводится к рассмотренному выше. В большинстве случаев достаточно «привести в порядок» корневое понятие перемещаемого поддерева, а остальные понятия модифицируются автоматически.

Язык описания онтологий. За основу языка описания онтологий был взят уже хорошо зарекомендовавший себя язык представления знаний системы Semp-ТАО [Загоруйко, Попов, 1997; Zagorulko et al., 2001].

Классы понятий на этом языке описываются в следующем виде:

```
class ИмяКласса (КлассРодитель);
    ОписаниеАтрибутов;
constraints
    ОписаниеОграничений;
end;
```

Приведем пример упрощенного описания класса «Человек» и его наследника – класса «Исследователь»:

```
class Человек;
    Фамилия: string;
    Имя: string;
    Отчество: string;
    Пол: Пол;
    Дата рождения: date;
    Дата смерти: date;
constraints
    Дата_рождения < Дата_смерти;
end;
class Исследователь (Человек);
    Степень: Ученая_степень;
    Звание: Ученое_Звание;
    E-mail: string;
    Рабочий_телефон: string;
```

end;

Описание отношения имеет вид:

relation ИмяОтношения (ИмяАргумента1: Класс1;
ИмяАргумента2: Класс2);

МатематическиеСвойстваОтношения;

ОписаниеАтрибутов;

constraints

ОписаниеОграничений;

end;

Отношениям могут быть приписаны следующие математические свойства: *транзитивность, симметричность, рефлексивность*.

Приведем пример описания отношения «Работает в»:

relation Работает_в (кто: Человек; где: Организация);

Должность: Должность;

Дата_принятия: date;

Дата_увольнения: date;

constraints

Дата_принятия > Дата_рождения +18;

Дата_увольнения > Дата_принятия;

end;

Домены описываются следующим образом:

Domain ИмяДомена = МножествоСтроковыхЗначений;

Приведем примеры описания некоторых доменов:

Domain Пол = {мужской, женский};

Domain Должность = {Директор, Ученый секретарь, Руководитель лаборатории, Руководитель отдела, Главный научный сотрудник, Ведущий научный сотрудник, Старший научный сотрудник, Научный сотрудник, Младший научный сотрудник, Техник}.

Редактор онтологий был построен на основе приведенного выше языка описания онтологий. Он строился с учетом возможности применения для распределенной разработки онтологий и реализован как веб-приложение, доступное зарегистрированным пользователям через Интернет. С целью обеспечения распределенной разработки онтологий в редакторе поддерживается механизм делегирования прав экспертам разных уровней.

С помощью редактора онтологий (рис. 3) можно создавать, модифицировать и удалять любые элементы онтологии: классы понятий, отношения, домены.

При создании класса указываются его уникальное имя и набор атрибутов, служащих для задания различных свойств понятий, а фактически описывающих структуру объектов данного класса. Для класса может быть выбран родитель из ранее созданных классов, при этом от родительского класса наследуются не только все атрибуты, но и отношения, а сам родитель связывается с новым классом отношением «класс-подкласс». Для каждого атрибута класса задается имя, область допустимых значений (тип или домен), количество возможных значений (одно или множество), а также указывается обязательность заполнения.

Домен характеризуется названием и множеством элементарных значений. Для каждого значения домена можно также указать язык (на данный момент – русский или английский), на котором оно было введено.

Отношения могут связывать только уже созданные классы онтологии. Они являются направленными и бинарными и могут иметь собственные атрибуты, уточняющие связь между аргументами отношения. Для более удобного представления информации пользователю портала предоставляется возможность настройки визуализации знаний и данных. При этом для каждого класса в редакторе онтологий задается шаблон визуализации объектов этого класса и шаблон визуализации ссылок на них. По умолчанию, порядок, в котором отображаются атрибуты классов и связанные с ними отношения, в том числе и атрибуты отношений, определяется порядком их задания в онтологии. По желанию пользователя этот порядок может быть изменен.



РЕДАКТОР ОНТОЛОГИИ								
КЛАССЫ ОТНОШЕНИЯ ДОМЕНЫ СОЗДАТЬ КЛАСС СОЗДАТЬ ОТНОШЕНИЕ СОЗДАТЬ ДОМЕН								
КЛАСС: ОБЪЕКТ ИССЛЕДОВАНИЯ  								
<div style="border: 1px solid black; padding: 5px;"> Объект Исследования <input type="checkbox"/> Материальный Объект <input type="checkbox"/> Нематериальный Объект </div>								
Имя атрибута	Тип	Имя домена	Множеств.	Обязат.	Уник.	Ссылка	Кр.ссылка	Порядок
Название объекта	строка		нет	да	да	1	1	1
Описание объекта	строка		нет	нет	нет			2
Дата открытия	дата		нет	нет	нет			3
URL	строка		нет	нет	нет			4
ОТНОШЕНИЯ КЛАССА ОБЪЕКТ ИССЛЕДОВАНИЯ								
Объект-Включает								1
Объект Исследования								нет
Подобен								2
Объект Исследования								нет
Степень схожести	число		нет	нет	нет			
Объект-датируется								3
Период								нет
Точность датирования	строка		нет	нет	нет			
Объект-найден-в								4
Место								нет
ОТНОШЕНИЯ С КЛАССОМ ОБЪЕКТ ИССЛЕДОВАНИЯ								
Применяется-к-Классу-Объектов								1
Метод Исследования								нет
Ресурс-Объекта								2
Информационный Ресурс								нет
Описывает-Объект								3
Публикация								нет
Исследует								4
Деятельность								нет
Объект-Включает								5
Объект Исследования								нет
Рассматривает-Объект								6
Раздел Науки								нет
Подобен								7
Объект Исследования								нет
Степень схожести	число		нет	нет	нет			

Рис. 3. Редактирование онтологии

Пример построения онтологии

Описанная в предыдущем разделе технология была применена для разработки онтологии для археологического портала знаний. Основой для построения онтологии археологии послужила предложенная в работе [Холюшкин, Гражданников, 2000] и развиваемая в настоящее время системная классификация археологической науки, фиксирующая явные и неявные связи между используемыми в ней понятиями. Понятия системной классификации использовались как для построения классов онтологии и доменов их атрибутов, так и для создания экземпляров классов онтологии.

Построенная онтология археологии включает четыре базовые иерархии: разделов археологии, объектов исследования археологии, методов исследования археологии и научных результатов археологии (рис. 4). Эти иерархии опираются на следующие классы понятий, построенные на основе базовых понятий онтологии научного знания.

Раздел археологии. Является корневым классом в иерархии научных направлений в археологии. К разделам археологии относятся, например, такие направления, как *Общая археология*, *Полевая археология*, *Реконструктивная археология*.

Метод исследования археологии. Служит для описания методов исследования, применяемых к определенному типу археологических объектов. В данном классе были выделены подклассы *Подход*, *Принцип*, *Методика*, *Технология*, *Археологический метод*, а также группа используемых в археологии методов, пришедших из других наук, образующая соответственно подклассы *Биологические методы*, *Физические методы* и *Химические методы*.

Объект исследования археологии. Образует иерархию объектов исследования в археологии. В качестве его свойств выступают описание объекта, дата открытия, точность датирования и др. Он содержит в себе подкласс *Нематериальный объект*, разделяющийся в свою очередь на подклассы *Историческая группа людей*, *Историческая личность*, *Историческое место*, и подкласс *Материальный объект*, наследниками которого являются подклассы *Артефакт*, *Комплекс*, *Памятник*.

Научный результат археологии. Служит для описания результатов научной деятельности в археологии, таких как открытия, новые законы, теории, исторические факты. Свойства понятий этого класса представлены атрибутами: описание результата, дата получения и тип

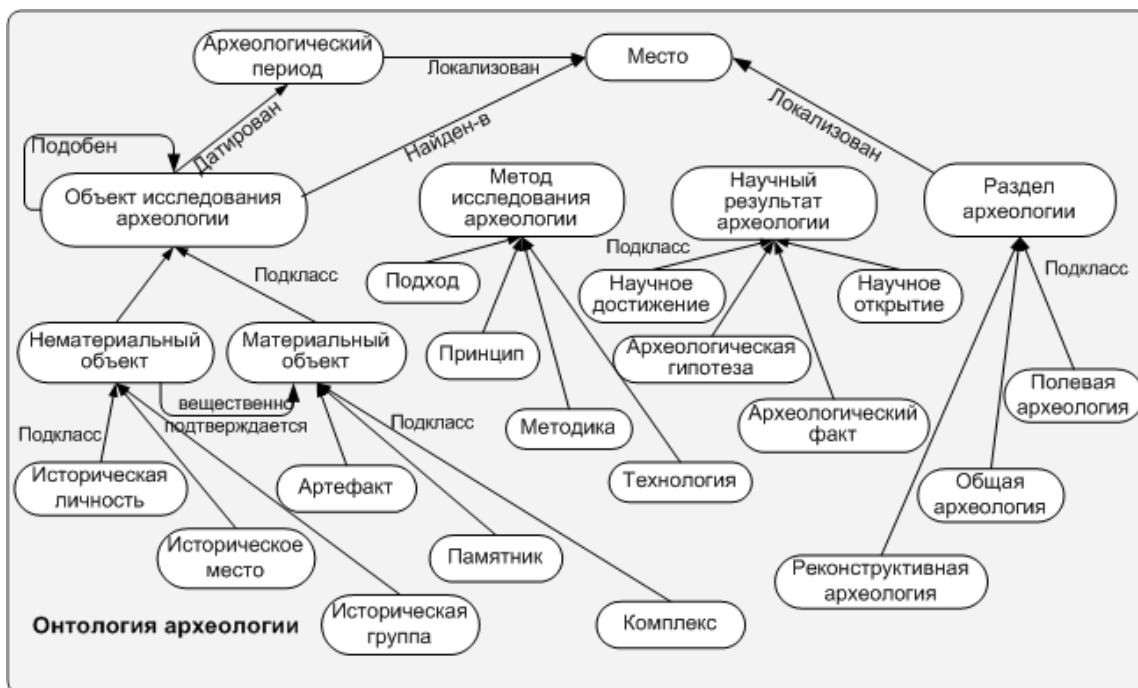


Рис. 4. Фрагмент онтологии археологии.

результата. Данный класс включает следующие подклассы: *Археологическая гипотеза*, *Археологический факт*, *Научное достижение*, *Научное открытие*.

Кроме перечисленных выше классов, построенных на основе метапонятий онтологии научного знания, были введены классы *Археологический период* и *Место*. Класс *Археологический период* является характерным для исторических наук, он служит для датирования объектов исследования. *Археологические периоды* образуют иерархию вложенности и исторического следования и описываются временным интервалом. Класс *Место* служит для указания географического местоположения объекта исследования или организации, а также для географической привязки разделов науки.

Описанные выше иерархии классов связаны между собой и с классами онтологии научной деятельности посредством ассоциативных отношений, одни из которых были введены при описании базовых онтологий, другие – при построении онтологии археологии. Так, иерархия методов исследования связана с иерархией объектов исследования посредством отношения «*Применяет*», а иерархия научных результатов, служащая для типизации и описания результатов научной деятельности, связана с деятельностью с помощью отношения «*Результат деятельности*».

Иерархия результатов исследования связана с публикациями, в которых описывается результат, посредством отношения «*Описывает*». Объекты исследования связаны между собой с помощью отношения «*Подобен*», задающего степень схожести (подобия) объектов между собой, и отношения «*Вещественно подтверждается*», задающего связь нематериального археологического объекта исследования с материальным объектом исследования, который подтверждает его существование. Связь иерархии разделов науки с используемыми в них методами исследования организована с помощью отношений «*Использует*», а с объектами исследования разделы науки связаны отношением «*Рассматривает*». Использование отношений «*Изучает период*» и «*Локализован*» позволяет установить хронологическое и географическое положение раздела науки. Следует заметить, что полная онтология археологического портала включает онтологию археологии и онтологию научной деятельности, рассмотренную выше.

Заключение

В статье представлена технология построения онтологий для порталов научных знаний, включающая методологию построения онтологий и эргономичный редактор. Данная технология поддерживает разработку онтологий, по которым автоматически строятся:

- схема внутренней базы данных портала (логическая структура БД и ее ограничения целостности);
- формы для заполнения БД портала данными (информационными объектами, являющимися экземплярами понятий онтологии);
- схема навигации по информационному пространству портала (по отношениям онтологии);
- формы поисковых запросов (по понятиям и отношениям онтологии).

Эта технология обеспечивает возможность декларативной подстройки онтологии в ходе эксплуатации портала знаний, что позволяет отслеживать динамику появления новых знаний и информационных ресурсов по тематике портала и тем самым обеспечивает поддержку его актуальности и полезности.

На основе предложенной технологии совместно с Институтом археологии и этнографии СО РАН был разработан археологический портал знаний [Андреева и др., 2006], обеспечивающий содержательный доступ широкому кругу пользователей к систематизированным знаниям и информационным ресурсам по археологии и этнографии. В настоящее время этот портал содержит более 4 тыс. информационных объектов, связанных более чем 15 тыс. онтологических отношений, и продолжает активно пополняться новыми знаниями и данными.

Хотя первоначально данный подход разрабатывался для организации содержательного доступа к гуманитарным информационным ресурсам, по нашему убеждению, он может быть успешно применен и для других областей знаний. В настоящее время исследуется возможность применения предложенного подхода для разработки интернет-порталов знаний не только для гуманитарных, но и точных наук. В частности, этот подход сейчас применяется для разработки портала знаний по компьютерной лингвистике [Загоруйко и др., 2006], которая представляет собой синтез гуманитарной (лингвистика) и точных (информатика и математика) наук.

Список литературы

Андреева О. А., Боровикова О. И., Булгаков С. В. и др. Археологический портал знаний: содержательный доступ к знаниям и информационным ресурсам по археологии // Тр. X национ. конф. по искусственному интеллекту с международным участием КИИ'2006. М.: Физматлит, 2006. Т. 3. С. 832–840.

Боровикова О. И., Загоруйко Ю. А. Организация порталов знаний на основе онтологий // Тр. междунар. семинара «Диалог 2002» Компьютерная лингвистика и интеллектуальные технологии (Протвино, 6–11 июня 2002 г.). М.: Наука, 2002. Т. 2. С. 76–82.

Гаврилова Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем: Учеб. СПб.: Питер, 2001. 384 с.

Загоруйко Ю. А., Боровикова О. И., Кононенко И. С. и др. Подход к построению предметной онтологии для портала знаний по компьютерной лингвистике // Компьютерная лингвистика и интеллектуальные технологии: Тр. междунар. конф. «Диалог 2006» (Бекасово, 31 мая–4 июня 2006 г.). М.: Изд-во РГГУ, 2006. С. 148–151.

Загоруйко Ю. А., Попов И. Г. Описание сложных предметных областей на основе интеграции средств представления знаний // Тр. междунар. семинара «Диалог'97» по компьютерной лингвистике и ее приложениям. М.: 1997. С. 110–115.

Холюшкин Ю. П., Гражданников Е. Д. Системная классификация археологической науки (элементарное введение в археологическое науковедение). Новосибирск: Изд-во ИДМИ Мин-образования, 2000. 58 с.

Gruber T. Towards Principles for the Design of Ontologies Used for Knowledge Sharing // Int. J. of Human and Computer Studies. 1995. Vol. 43, No. 5/6. P. 907–928.

Guarino N., Giarretta P. Ontologies and Knowledge Bases: Towards a Terminological Clarification // Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing / Ed. by N. Mars. Amsterdam: IOS Press. P. 25–32.

Takeda H., Takaai M., Nishida T. Collaborative development and Use of Ontologies for Design // Proc. of the Tenth International IFIP WG 5.2/5.3 Conference PROLAMAT 98, September 9–12, Trento, Italy, 1998.

Ushold M., Gruninger M. Ontologies: Principles, Methods and Applications // Knowledge Engineering Review. 1996. Vol. 11 (2). P. 93–155.

Zagorulko Y. A., Popov I. G., Kostov Y. V. Subdefinite Data Types and Constraints in Knowledge Representation Language // Bull. of NCC. Ser.: Computer Science. 2001. Vol. 16. Novosibirsk: NCC Publisher, 2001. P. 153–170.

Zagorulko Yu., Borovikova O., Bulgakov S. et al. Ontology-based approach to development of adjustable knowledge internet portal for support of research activity // Bull. of NCC. Ser.: Computer Science. 2005. Is. 23. P. 45–56.

Материал поступил в редколлегию 29.08.2007