

Разработка и сопровождение информационных систем, базирующихся на онтологии и Wiki-технологии *

© В.К. Шестаков

Новосибирский государственный университет, г. Новосибирск
zfc@ngs.ru

Аннотация

В работе рассматриваются проблемы разработки и сопровождения информационных систем, базирующихся на онтологии и Wiki-технологии. В частности, предлагается подход к построению информационных систем на основе Wiki-технологии и онтологий предметных областей, а также способ извлечения онтологий из уже существующих Wiki-систем. Описывается реализация такой системы.

1. Введение

В современном, динамично меняющемся мире деятельность отдельных людей, групп и организаций все в большей степени зависит от того, какая информация им доступна, и насколько эффективно они могут ей воспользоваться. В связи с экспоненциальным ростом объемов такой информации очень остро стоит проблема ее организации, упорядочения и повышения эффективности доступа к ней. К настоящему моменту придумано немало разнообразных методов и систем для решения данной задачи, но всем им присущи те или иные недостатки.

Так, с одной стороны, для строгой формализации предметной области существует такое понятие как онтология. Она позволяет задать явную структуру, но не отвечает за представление данных в удобном обычному пользователю виде. Кроме того, для работы с ней нужны определенные специализированные средства и инструменты, а их освоение для неподготовленного пользователя может представлять серьезную проблему, особенно в условиях острой нехватки времени, характерной для динамично развивающегося мира.

С другой стороны, существует такое удобное средство для сбора и хранения информации как Wiki-технология. Она позволяет работать не только с текстом, но и с разнообразным мультимедийным контентом, имеет удобный и понятный интерфейс, проста в освоении. Ее главным недостатком

является отсутствие явно выраженной и удобно представленной структуры. Это значительно затрудняет и усложняет построение информационных систем на базе этой технологии, особенно если они имеют значительный размер вкуче со сложной и разветвленной структурой информации.

В данной работе рассматривается подход к построению информационных систем, позволяющий объединить достоинства Wiki-технологии и онтологий таким образом, чтобы скомпенсировать их недостатки. Он позволяет объединить строгость структуры с простотой доступа, включая гибкие средства поиска и фильтрации информации, а также многообразие связей между данными. Также следует подчеркнуть возможность оперативного добавления и обновления информации.

На этот же вопрос можно взглянуть и с другой стороны. Недостатком Wiki-технологии является то, что они позволяют отслеживать в создаваемых информационных системах только структурную целостность ссылок, не обеспечивая при этом логической целостности и семантической согласованности используемых в них понятий. Идея предлагаемого подхода состоит в создании инструментария, который бы обеспечивал построение Wiki-систем с согласованной системой понятий (семантически согласованных Wiki-систем). Wiki-системы с такими свойствами можно получить, если строить ее на основе логически непротиворечивой онтологии, описывающей предметную область будущей системы.

Кроме того, рассматривается обратный процесс, важный для сопровождения и развития информационных систем: извлечение онтологий из уже существующих Wiki-систем. Но эта задача может иметь также множество других приложений, которые будут рассмотрены в дальнейшем.

2. Обзор существующих подходов

2.1 Обзор средств создания информационных Wiki-систем с использованием семантических технологий

На сегодняшний день существует достаточно много проектов создания информационных Wiki-

Труды 13^й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2011, Воронеж, Россия, 2011.

систем, так или иначе связанных с семантическими технологиями. Хотя эти проекты нацелены на решение различных задач, общим для них является использование онтологий на тех или иных этапах построения или функционирования информационных систем. Так как на данный момент существует большое количество средств создания Wiki-систем, то в этом обзоре мы сосредоточим наше внимание только на тех из них, которые в той или иной мере используют семантические технологии.

Сначала рассмотрим проект из Санкт-Петербурга [14], в котором в одном из первых начали применяться онтологии при построении таких информационных систем, как корпоративные Wiki-порталы. Исходя из того, что онтологии позволяют создавать более концептуально-ясные и целостные модели предметных областей, разработчиками было решено привязать контент Wiki-системы к онтологии, чтобы избежать проблем с плохой структурой портала, т.к. она в этом случае будет такого же качества, как и определяющая ее онтология. В рамках этого проекта был разработан прототип инструментария *OntologieWiki*, который содержит внутри себя Wiki-движок и для каждого концепта онтологии автоматически создает Wiki-страницу, в которую эксперт вносит информацию об описываемом понятии. При этом конечные пользователи видят онтологию, представленную в виде графа, могут осуществлять навигацию по ее понятиям и видеть содержимое Wiki-страниц, соответствующих выбранному понятию.

Недостатком *OntologieWiki* является то, что онтология, лежащая в основе созданной с помощью нее Wiki-системы, не может быть изменена или дополнена средствами самой Wiki-системы.

Наиболее распространенным инструментальным средством для создания Wiki-систем является популярный Wiki-движок *MediaWiki* [16]. Именно для этого универсального движка больше всего строится расширений, позволяющих использовать средства семантических технологий. Одной из первых была создана надстройка над *MediaWiki*, получившая название *Semantic MediaWiki* [19]. Это расширение допускает возможность импорта онтологий [15] и позволяет использовать в Wiki-системе структуры и связи, определенные в ней. Однако предоставляемый *Semantic MediaWiki* модуль импорта онтологий на данный момент находится в стадии бета-версии и не включен в состав новых версий данного расширения. В связи с этим реализацию импорта онтологий пользователи *Semantic MediaWiki* должны выполнять самостоятельно.

Другим расширением *MediaWiki* является *BOWiki* [1] — основанный на онтологии семантический Wiki-движок. Первоначально он базировался на *Semantic MediaWiki*, а потом был реализован как «чистое» расширение *MediaWiki*. Достоинством *BOWiki* является то, что он позволяет пользователям рассматривать сущности,

описанные Wiki-страницами, в качестве экземпляров онтологических категорий, определять новые отношения в Wiki-контенте, связывать Wiki-страницы, в том числе n-арными семантическими отношениями, запрашивать Wiki-страницы, удовлетворяющие определенным условиям, допускает импорт еще нескольких био-онтологий для ограниченного использования и обеспечивает экспорт Wiki-контента в OWL-представление.

BOWiki разрабатывался как средство для совместного создания и интеграции знаний в области био-информатики, поэтому его функционирование основано на онтологии высокого уровня *General Formal Ontology (GFO)* [6] в OWL-версии. В своей работе он использует также две онтологии, основанные на GFO, — базовую биомедицинскую онтологию (*GFO-Bio*) и онтологию функций (*Ontology of Functions* или *OF*). Как было сказано выше, возможность использования других (внешних) онтологий существенно ограничена, поэтому *BOWiki* не может рассматриваться в качестве инструмента для построения Wiki-систем произвольного вида.

Также для *MediaWiki* существуют расширения, работающие совместно с *Semantic MediaWiki*, и добавляющие дополнительную функциональность по импорту семантической информации. Одним из таких является *RDFIO* [23]. Оно позволяет импортировать произвольные RDF-триплеты [21]. Но это расширение пока имеет статус бета-версии и еще не готово для массового использования. Еще одним подобным расширением является *LinkedWiki* [22]. Оно применяется для получения данных с внешнего сервиса и отображения этих данных в Wiki-системе. Это расширение не может изменить структуру существующей системы, а только дополняет ее, поэтому имеет довольно ограниченную область применения.

2.2 Обзор подходов к извлечению онтологий из Wiki-систем

Существует много подходов к извлечению знаний из Wiki-систем в виде онтологий, в частности, из Википедии. Одними из первых решением этого вопроса занялась группа немецких ученых из Института информатики общества Макса Планка [11]. Их целью было автоматическое построение общей онтологии значительного объема, включающую факты, извлеченные из Википедии с высокой точностью. При этом предполагалось совместно использовать знания из Википедии и *WordNet* [4]. Эта задача была успешно решена, в результате была получена онтологическая база знаний *YAGO*. В данном подходе источниками информации в Википедии служат не сами статьи целиком, а система категорий этой энциклопедии и так называемые «шаблоны-карточки» («*infoboxes*») — стандартизованные таблицы, содержащие основную информацию о предмете, описываемом в статье. При этом из Википедии извлекаются индивиды, большая часть отношений и категории-

листья, а остальная иерархия категорий и отношений достраивается благодаря WordNet [12]. В основе YAGO лежит собственная модель, которая является небольшим расширением RDFS. По заявлению авторов, несмотря на достаточную выразительность модели, она остается разрешимой (в отношении проверки ее непротиворечивости). В проведенном ими оценочном исследовании точность информации, содержащейся в онтологии, составила 95% — намного выше, чем у всех других автоматически построенных общих онтологий. Кроме того, данная онтология имеет гораздо больший по сравнению с упомянутыми выше онтологиями размер. Она используется во многих приложениях и является составной частью многих других семантических проектов. Дальнейшее развитие этой базы знаний — YAGO2 — использует еще один источник данных (GeoNames) и содержит еще более внушительный объем информации [7].

На основе Википедии строят не только общие онтологии, но и специализированные. В частности, группа японских ученых разработала метод для построения крупномасштабной онтологии людей [10]. В нем иерархия категорий, а также сами экземпляры, извлекаются из энциклопедии с помощью машинно-обучаемого классификатора и с использованием японского тезауруса Nihongo Goi-Taikē. Существенной частью метода является использование «похожих категорий» — это категории «родители», «дети» или «сестры», у которых последнее слово совпадает с последним словом целевой категории. Эксперименты, проведенные авторами этой работы, показывают высокую точность и полноту полученной онтологии (они превосходят предшествующие существующие методы), а также существенность использования тезауруса и «похожих категорий». Еще одним преимуществом перед другими методами названа возможность извлечения категорий, не имеющих совпадений в тезаурусе, и категорий с неоднозначными названиями.

Еще одним подходом к построению предметно-ориентированных онтологий занимается совместный проект университетов Инсбрука и Флориды [5]. Его участники утверждают, что основной недостаток существующих онтологий в том, что они разрабатываются только небольшими группами людей, а большинство потенциальных пользователей исключено из процесса разработки онтологий и их мнения не учитываются. Поэтому предлагается использовать стандартные Wiki-технологии, с помощью которых тысячи пользователей уже вовлечены в создание такого огромного по объему источника знаний как электронная энциклопедия Википедия, в качестве среды для разработки онтологий. В работе доказывалось, что URI этой энциклопедии могут служить надежными идентификаторами для онтологических концептов, и демонстрируется применимость этого подхода на практике.

Помимо непосредственного извлечения знаний в виде онтологий, на основе Википедии решают и вспомогательные задачи в этой области. Например, при конструировании онтологий обычно требуется корпус из конкретной предметной области для построения соответствующей иерархии концептов. При этом он должен обеспечивать хорошее покрытие этой области с достаточной степенью качества. Этим вопросом как раз и занимаются ученые из Гонконгского политехнического университета [2]. Они используют в качестве источника Википедию и предлагают новый подход для классификации статей по предметным областям. Основная идея заключается в генерации иерархии предметной области на основе гипертекстовых связей страниц. И только статьи, тесно связанные с этой иерархией, выбираются в качестве кандидатов. Предлагаемый метод пока использует только информацию о категориях Википедии. Далее выполняется ранжирование и фильтрация выбранных страниц. Результаты эксперимента по оценке, проведенного авторами, показывают, что Википедия является хорошим ресурсом для получения корпуса конкретной предметной области сравнительно высокого качества.

Конечно же, существуют проекты, в которых данные, извлеченные из Википедии, используются не просто для построения онтологии или для решения отдельно взятой вспомогательной задачи, а применяются как часть крупной информационной системы, тесно взаимодействующей с другими системами (примерами взаимодействия являются автоматический сбор данных из других систем, объединение с другими онтологиями, отображение содержимого в другую систему и т.п.). Для иллюстрации рассмотрим проект под названием «Построение и использование геопространственной онтологии в наблюдающей системе BioCaster», разрабатываемый группой японских и вьетнамских ученых [3]. Для начала они планируют построить геопространственную онтологию, используя информацию из Википедии, а затем использовать ее в своей системе. База данных будет содержать названия стран и самых крупных городов, а также отношения типа «часть-целое» между странами и их частями. Процесс построения проходит в полуавтоматическом режиме: автоматическое извлечение данных, а затем верификация с участием человека. На последнем этапе геопространственная иерархия из Википедии объединяется с онтологией BioCaster. Эта система предназначена для выявления и отслеживания вспышек инфекционных заболеваний на основе новостных сообщений. В дальнейшем эта полученная информация визуализируется с использованием Google Maps.

Еще одним проектом, в котором построение онтологии на основе Википедии используется в качестве составной части, занимаются в Наньянском технологическом университете [8]. В его рамках была предпринята попытка

автоматического построения крупномасштабной мульти-модальной онтологии для классификации веб-изображений. Для текстовой части были использованы преимущества структурных и контентных возможностей Википедии, и объекты реального мира были формализованы в терминах концептов и отношений. Для визуальной части производилось обучение классификатора и по результатам генерировались концепты промежуточного уровня. В дальнейшем использовалось ассоциативное правило поискового алгоритма для улучшения построенной онтологии. Посредством эксперимента была доказана высокая точность работы метода.

3. Построение информационной Wiki-системы на основе онтологии

3.1 Описание работы системы

Предлагаемая в рамках данной работы инструментальная система позволяет на основе ранее разработанной онтологии построить Wiki-систему, структура и содержание которой будет определяться этой онтологией. В дальнейшем построенная Wiki-система может расширяться как обычная Wiki-система с использованием традиционных средств Wiki-технологии.

Общая схема работы системы, реализованной в рамках предлагаемого подхода, объединяющего Wiki-технологии и онтологии, выглядит следующим образом (см. рис. 1). На первом этапе квалифицированный эксперт описывает онтологию предметной области. На втором по ней строится интерфейс на основе Wiki-технологии, состоящий из двух частей: конструкторского и пользовательского. При помощи первого эксперты предметной области могут вносить данные в систему, которые тут же становятся доступны пользователям посредством второго.

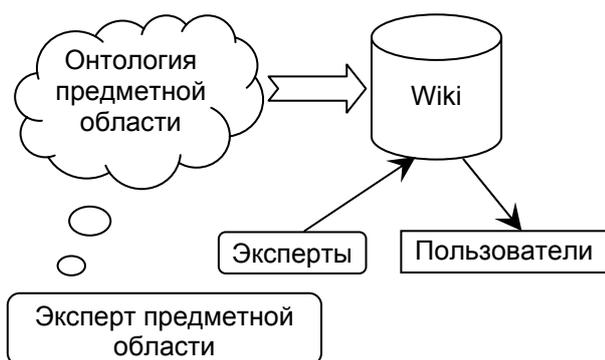


Рисунок 1. Общая схема работы инструментальной системы

Подробная схема работы инструментальной системы представлена на рис. 2. Согласно ей разработка Wiki-системы в рассматриваемой инструментальной системе включает следующие этапы.

На первом шаге онтология, полученная в готовом виде или разработанная в Protégé [20] или каком-либо другом редакторе онтологий, сохраняется или конвертируется в OWL-формат [9]. После этого файл со спецификацией онтологии подается на вход специально разработанному в рамках данного проекта программному модулю Onto2Wiki, который разбирает его с использованием библиотеки RDFLib [18], а затем, при помощи среды Python WikipediaBot Framework [17], создает каркас информационной системы на основе пустого Wiki-сайта, работающего на базе MediaWiki с расширением Semantic MediaWiki (внутренним хранилищем информации в MediaWiki является база данных). При этом в Wiki-систему добавляются нужные страницы, для которых указываются соответствующие атрибуты, расставляются категории и прописываются нужные связи. После выполнения этих действий Wiki-система готова к использованию.

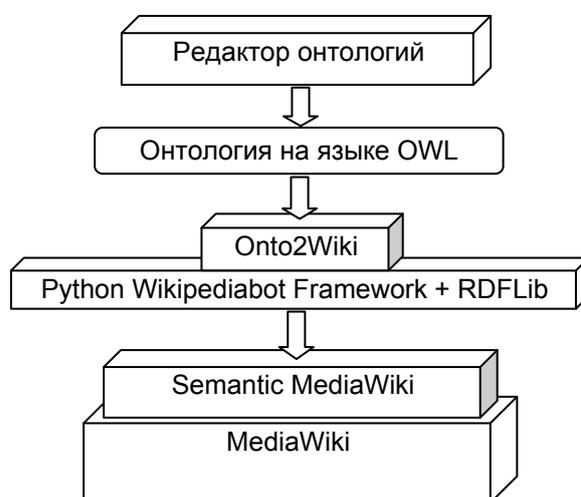


Рисунок 2. Подробная схема работы инструментальной системы

В таблице 1 представлено соответствие между конструкциями языка OWL и семантической Wiki, которое используется при отображении онтологии в структуру и содержание Wiki-системы. В соответствии с этой таблицей для каждого класса онтологии будет создана своя категория, которая будет проставлена на каждой подкатегории, соответствующей подклассу данного класса, и на каждой странице, соответствующей экземпляру данного класса. Также все атрибуты экземпляра онтологии будут указаны в качестве атрибутов на странице, соответствующей этому экземпляру. При этом все отношения онтологии отобразятся в типизированные ссылки [13] между соответствующими страницами и категориями.

Созданную с помощью данного инструментальной системы можно дорабатывать в соответствии с возникающими потребностями, взаимодействуя с ней как с обычной Wiki-системой. Т.е. можно создавать новые страницы, прописывать новые связи, указывать новые категории и т.д.

Таблица 1. Соответствие конструкций языка OWL и Semantic MediaWiki

Конструкция OWL	Semantic MediaWiki
Класс	Категория
Подкласс	Подкатегория
Экземпляр	Страница
Атрибут	Атрибут
Отношение	Типизированная ссылка

Таким образом, особенностью данного подхода является то, что полученная Wiki-система получается полностью независимой от той онтологии, на основе которой была построена. Т.е. изменения, производимые в самой Wiki-системе, не сказываются на исходной онтологии. Значит, для наблюдения за изменением онтологии, теперь уже неважно лежащей в основе Wiki-системы, требуется отдельное средство. Рассмотрению данного вопроса как раз и посвящен следующий раздел.

3.2 Пример работы системы

Рассмотрим работу данной инструментальной системы на примере построения учебной информационной системы «Технология создания экспертных систем». Вот фрагмент построенной онтологии в формате OWL:

```
<owl:NamedIndividual
rdf:about="&www;Технология_создания_экспертных_систем_Class30003">
  <rdf:type
rdf:resource="&www;Ontology1243886372.owl#Литература_по_ЭС"/>
  <rdfs:label rdf:datatype="&xsd:string">Технология_создания_экспертных_систем_Class30003</rdfs:label>
  <название rdf:datatype="&xsd:string">Как_построить_свою_ЭС</название>
  <название_книги rdf:datatype="&xsd:string">Как_построить_свою_экспертную_систему</название_книги>
  <автор_книги rdf:datatype="&xsd:string">Нейлор_К.</автор_книги>
  <упоминается
rdf:resource="&www;Технология_создания_экспертных_систем_Class10004"/>
  <упоминается
rdf:resource="&www;Технология_создания_экспертных_систем_Class10007"/>
  <упоминается
rdf:resource="&www;Технология_создания_экспертных_систем_Class20001"/>
  <упоминается
rdf:resource="&www;Технология_создания_экспертных_систем_Class30005"/>
  <упоминается
rdf:resource="&www;Технология_создания_экспертных_систем_Class37"/>
</owl:NamedIndividual>
```

Подаем эту онтологию на вход системе и после окончания ее работы получаем заполненную Wiki-систему.

На рис. 3 представлен пример одной из страниц полученной системы. На нем видны атрибуты и типизированные ссылки, находящиеся на странице. А вот так выглядит Wiki-разметка данной страницы:

```
Экземпляр Как построить свою ЭС
название книги: [[название книги::Как построить свою экспертную систему]]
автор книги: [[автор книги::Нейлор К.]]
упоминается: [[упоминается::MYCIN]]
упоминается: [[упоминается::DENDRAL]]
упоминается: [[упоминается::PROSPECTOR]]
упоминается: [[упоминается::PUFF]]
упоминается: [[упоминается::Экспертные системы]]
[[Категория:Литература по ЭС]]
```

Как построить свою ЭС

```
Экземпляр Как построить свою ЭС
название книги: Как построить свою экспертную систему
автор книги: Нейлор К.
упоминается: MYCIN
упоминается: DENDRAL
упоминается: PROSPECTOR
упоминается: PUFF
упоминается: Экспертные системы

Категория: Литература по ЭС
```

Рисунок 3. Пример страницы

4. Извлечение онтологий из существующих Wiki-систем

Как было сказано выше, подход к построению информационных Wiki-систем на основе онтологий предметных областей кроме механизма «отображения» онтологий должен включать и обратную процедуру, т.е. метод извлечения онтологий из Wiki-систем. Причем этот метод может иметь гораздо больше применений, чем это может показаться на первый взгляд.

4.1 Метод извлечения онтологий

Задача извлечения онтологии из Wiki-системы решается с помощью тех же инструментов, которые применялись при ее «отображении». Разница заключается в использовании вместо модуля Onto2Wiki другого программного модуля — Wiki2Onto, также разработанного в рамках данного проекта. При этом все операции выполняются в

обратном порядке. Т.е. сначала модуль Wiki2Onto при помощи Python WikipediaBot Framework извлекает онтологию из Wiki-системы, а затем с использованием библиотеки RDFLib сохраняет в файл на языке OWL. При этом используется та же самая таблица соответствия конструкций языка OWL и семантической Wiki (см. таблицу 1).

Рассмотрим этот процесс более подробно. Онтология из Wiki-системы извлекается в следующем порядке. Сначала извлекаются все классы, при этом каждому классу соответствует одна категория Wiki-системы, а структура вложенности категорий Wiki-системы определяет иерархию классов. Затем извлекаются все страницы как экземпляры соответствующих классов. Для пустых страниц, на которые в Wiki-системе имеются ссылки, заводится специальный служебный класс «Несуществующие страницы». После этого просматриваются все ссылки на каждой странице. Для начала определяется, является ли ссылка обычной или семантической. Если ссылка обычная, то для соответствующего экземпляра класса в OWL-онтологии заводится объектное свойство «Ссылается на» со значением в виде экземпляра, имя которого совпадает с именем страницы, на которую указывает ссылка. Если ссылка семантическая, то она имеет следующую структуру <название свойства, значение свойства>, и для нее сначала определяется тип ее свойства. Если свойство имеет тип «Страница» или его тип не указан, то в OWL-онтологии заводится объектное свойство с соответствующими именем (название свойства) и значением (значение свойства). (Заметим, что по умолчанию свойство ссылки имеет тип «Страница».) Если же свойство имеет какой-то другой стандартный тип, то тип свойства данных в OWL-онтологии определяется согласно таблице 2. Для пользовательских типов свойств создается собственный тип свойства данных.

Таблица 2. Соответствие при извлечении стандартных типов

Тип свойства	Тип OWL
Строка	String
Число	double
Булево	boolean
Дата	dateTime
Текст	string
Код	string
Телефонный номер	Annotation property
URL	Annotation property
Почта	Annotation property
URI аннотации	Annotation property

Так как реализация модуля Wiki2Onto еще не доведена до финальной стадии, то пока поддерживаются не все стандартные типы свойств Semantic MediaWiki, а только те, что представлены в таблице 2.

Следует отметить, что Wiki-система, из которой извлекается онтология, не обязательно должна функционировать на расширении Semantic MediaWiki. Однако в случае использования Wiki-системы без этого расширения извлекаемая онтология будет гораздо беднее, так как в ней не будет присутствовать специальная семантическая информация. В частности, нельзя будет извлечь атрибуты, разнообразие отношений также будет невелико. Правда, путем индивидуальной настройки на данную Wiki-систему объем извлекаемой из нее информации можно увеличить.

4.2 Варианты применения метода извлечения онтологий

Самое очевидное из применений это, конечно же, посмотреть на то, как изменилась онтология после того, эксперты поработали над содержимым Wiki-сайта, созданного на базе первоначальной онтологии. Это может понадобиться не только ради обычного любопытства, но и для вполне серьезных целей: например, таких как отслеживание развития проекта, верификация получаемой онтологии, проверка качества и сбалансированности получаемой структуры данных, координация в развитии отдельных частей предметной области. К полученной онтологии также можно применить одну из существующих машин вывода для получения неявных знаний.

Кроме того, извлекать онтологию можно не только из той Wiki-системы, в которую она была предварительно «отображена», но и из любой другой Wiki-системы. Например, мы хотим построить онтологию некоторой предметной области и обнаружили Wiki-систему, содержащую нужные нам сведения из этой области. Вместо того, чтобы вручную строить нужную нам онтологию «с нуля», мы можем извлечь из этой системы ее предварительный, черновой вариант, а затем уже дорабатывать его, что гораздо проще.

Еще одно применение метода извлечения онтологии заключается в объединении нескольких Wiki-систем по близким предметным областям. Непосредственное объединение систем может быть довольно сложным и потребовать очень много ручной работы, потому что крайне затруднительно отследить все связи и пересечения между двумя системами. Используя предлагаемый подход, можно поступить гораздо проще: извлечь онтологию из каждой системы и провести их слияние, а после этого по объединенной онтологии построить требуемую Wiki-систему.

5. Заключение

В данной работе рассмотрен подход к построению информационных систем на основе Wiki-технологии и онтологий предметных областей. В рамках этого подхода предложены метод построения Wiki-систем на основе онтологий и метод извлечения онтологий из существующих

Wiki-систем, разработан прототип инструментальной системы, реализующий данные методы в объеме, указанном в разделах 3 и 4.1.

Разработка метода построения Wiki-систем на основе онтологий позволяет строить Wiki-системы с хорошей структурой и согласованной системой понятий.

В дальнейшем планируется завершить реализацию модуля Wiki2Onto, а также расширить его возможности, в частности, применить методы компьютерной лингвистики для извлечения информации со страниц Wiki-систем.

Литература

- [1] Backhaus M., Kelso J., Bacher J., Herre H., Hoehndorf R., Loebe F., Visagie J. BOWiki – a collaborative annotation and ontology curation framework // In Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge, CKC Banff, Canada, May 8, of CEUR Workshop Proceedings. Volume 273. Edited by Noy N., Alani H., Stumme G., Mika P., Sure Y., Vrandečić D. Aachen, Germany: CEUR-WS.org, 2007.
- [2] Cui G.Y., Lu Q., Li W.J. and Chen Y.R. Corpus Exploitation from Wikipedia for Ontology Construction // In Proceedings of the Sixth International Language Resources and Evaluation (LREC 2008), Marrakech, 2008, pp. 2125–2132.
- [3] Doan, Son, Ngo, Quoc-Hung, Kawazoe, Ai, and Collier, Nigel, “Building and Using Geospatial Ontology in the BioCaster Surveillance System”, Available from Nature Precedings, <<http://dx.doi.org/10.1038/npre.2008.2110.1>> (2008).
- [4] Fellbaum C., editor. WordNet: An Electronic Lexical Database. MIT Press, 1998.
- [5] Hepp M., Bachlechner D., Siorpaes K. Harvesting Wiki Consensus – Using Wikipedia Entries as Ontology Elements // In Proceedings of the First Workshop on Semantic Wikis – From Wiki to Semantics, co-located with the 3rd Annual European Semantic Web Conference (ESWC 2006), 2006, pp. 124–138.
- [6] Herre H., Heller B., Burek P., Hoehndorf R., Loebe F., Michalek H. General Formal Ontology (GFO) – A foundational ontology integrating objects and processes [Version 1.0] // Onto-Med Report 8, Research Group Ontologies in Medicine, Institute of Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig (2006)
- [7] Hoffart J., Suchanek F.M., Berberich K., Weikum G. YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia // Research Report MPI-I-2010-5-007, Max-Planck-Institut für Informatik, November 2010.
- [8] Huan Wang, Xing Jiang, Liang-Tien Chia, and Ah-Hwee Tan Wikipedia2Onto --- Adding Wikipedia Semantics to Web Image Retrieval // In Proceedings of the WebSci'09: Society On-Line, 18–20 March 2009, Athens, Greece.
- [9] Motik B., Patel-Schneider P.F., Parsia B., eds. OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax // W3C Recommendation, 27 October 2009, <http://www.w3.org/TR/2009/REC-owl2-syntax-20091027/>.
- [10] Shibaki Y., Nagata M., Yamamoto K. Constructing Large-Scale Person Ontology from Wikipedia // In Proceedings of the 2nd Workshop on “Collaboratively Constructed Semantic Resources”, Coling, 2010, pp. 1–9.
- [11] Suchanek F.M., Kasneci G., Weikum G. YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia // In Proceedings of the 16th International Conference on World Wide Web (Banff, Alberta, Canada, May 8–12, 2007). WWW '07. ACM Press, New York, NY, pp. 697–706, 2007.
- [12] Suchanek F.M., Kasneci G., Weikum G. YAGO: A Large Ontology from Wikipedia and WordNet // Web Semantics: Science, Services and Agents on the World Wide Web, Volume 6, Issue 3, September 2008, pp. 203–217.
- [13] Vulkel M., Krutzsch M., Vrandečić D., Haller H. and Studer R. Semantic Wikipedia // In Proceedings of the 15th International Conference on World Wide Web (Edinburgh, Scotland, May 23–26, 2006). WWW '06. ACM Press, New York, NY, pp. 585–594, 2006, <<http://doi.acm.org/10.1145/1135777.1135863>>.
- [14] Муромцев Д.И., Горовой В.А., Малинин А.А., Гаврилова Т.А., Злобин А.Н., Катков Ю.В. Интеграция wiki-технологии и онтологического моделирования в задаче управления знаниями предприятия // Труды 11-ой национальной конференции по искусственному интеллекту с международным участием КИИ-2008 (г. Дубна, Россия). — М.: ЛЕНАНД, 2008. — Т.3. — С. 360-368.
- [15] Описание импорта в Semantic MediaWiki http://semantic-mediawiki.org/wiki/Help:Ontology_import.
- [16] Сайт проекта MediaWiki <http://mediawiki.org>.
- [17] Сайт проекта Python WikipediaBot Framework <http://pywikipediabot.sourceforge.net>.
- [18] Сайт проекта RDFLib <http://rdflib.net>.
- [19] Сайт проекта Semantic MediaWiki <http://semantic-mediawiki.org>.
- [20] Сайт редактора онтологий Protégé <http://protege.stanford.edu>.
- [21] Страница описания RDF <http://www.w3.org/RDF/>.
- [22] Страница расширения LinkedWiki <http://www.mediawiki.org/wiki/Extension:LinkedWiki>.
- [23] Страница расширения RDFIO <http://www.mediawiki.org/wiki/Extension:RDFIO>.

Development and Maintenance of Information Systems Based on Ontology and Wiki-Technology

© V.K. Shestakov

The paper is devoted to the development and maintenance of information systems based on ontology and Wiki-technology. In particular, an approach to creation of information systems based on Wiki-technology and ontology of subject domain as well as a method of extraction of ontologies from already existing Wiki-systems are considered. An implementation of the approach and method are described in detail.

* Автор выражает благодарность своему научному руководителю Ю.А. Загорюлько за интересную постановку задачи и постоянное внимание к работе.