

УДК 004.89:004.4

А.Ф. Тузовский, А.В. Черный

## Использование семантических технологий при создании социальных сетей

Статья посвящена особенностям создания социальных сетей на основе использования семантических технологий. Рассмотрены предлагаемая архитектура такого рода информационных систем, её возможности, подход к построению онтологий.

**Ключевые слова:** Semantic Web, онтология, социальная сеть.

В 2001 г. Тим Бернерс Ли высказал свое мнение о дальнейшем развитии Интернета, которое теперь называется концепцией семантической Паутины [1]. В предлагаемом подходе вся информация в Сети имеет некоторые смысловые описания, что значительно упростит поиск нужной информации, ее категоризацию. За развитие и поддержку идеи взялся консорциум W3C. Но за прошедшие годы Интернет не претерпел значительной эволюции в этом направлении, несмотря на то, что работать с все возрастающими объемами информации год от года становится труднее. В данной статье описан подход по созданию базиса для развития внутригородской семантической сети на основе городской семантической социальной сети (портала).

### Назначение портала

Основные цели создания городского семантического портала – это, с одной стороны, объединение городских сообществ, описание того, «кто есть кто», с другой стороны – это инструмент для работы с семантической паутиной местного масштаба.

### Архитектура

Основой семантического портала является набор онтологий, каждая из которых направлена на описание узких предметных областей. Онтология базируется на схемах, представленных в моделях Linked of Data (LOD) [2]. Онтология объектов знаний представляет собой набор описаний объектов, которые представляют интерес для портала, например: люди, организации, сообщества, интернет-ресурсы и т.п.

Онтология, основанная на элементах из LOD, используется для уточнения описаний объектов знаний и представления описывающей информации.

Третья онтология содержит лишь набор конкретных экземпляров понятий, описанных в упомянутых выше онтологиях. Такой подход позволяет разделить уровни понятий и экземпляров, что дает возможность сохранить на каждом уровне большую упорядоченность и повысить комфорт работы с этими онтологиями.

С точки зрения архитектуры семантический портал состоит из следующих модулей (рис. 1):

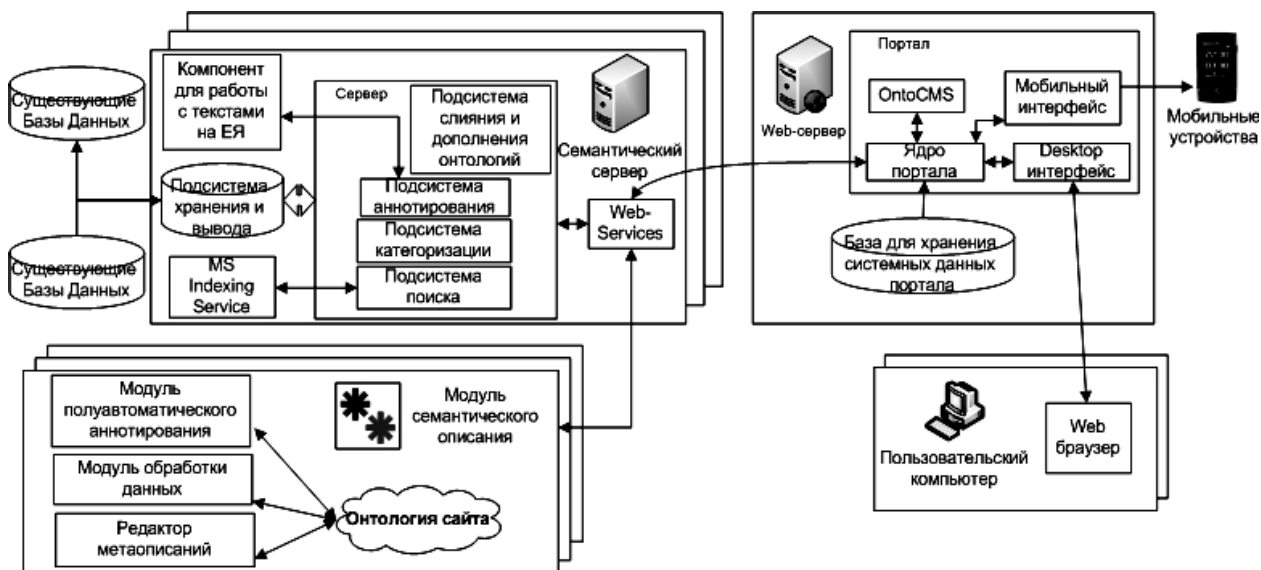


Рис. 1. Архитектура городского портала

Архитектура семантического сервера и Web-сервера аналогична архитектуре этих компонентов в семантической сети организаций [3]. Его основным назначением является вывод и хранение онтологий (A-Box и T-Box), обеспечение функционала полуавтоматического аннотирования, полнотекстового поиска и прочие сервисы. Web-сервер отвечает за формирование мобильного и настольного интерфейсов пользователя, и обеспечивает функциональность системы управления содержимым, основанной на семантике.

Особый интерес в рамках статьи представляет модуль семантического описания. Это свободно распространяемый набор компонентов для описания знаний на уже существующих сайтах, которые такого описания не имеют. Задача модуля – разметка существующих страниц в соответствии со стандартами семантической разметки страниц.

Модуль функционирует в соответствии со следующим алгоритмом:

Тексты существующих страниц отправляются на аннотирование на семантический сервер.

После аннотирования модератор сайта может с помощью редактора просмотреть и отредактировать полученное семантическое описание.

Понятия и связи сохраняются в онтологию сайта, что дает возможность размечать подобные тексты в дальнейшем без обращения к семантическому серверу.

При последующем создании страницы модуль может выделить связи и отношения на основе имеющейся онтологии без привлечения семантического сервера, но за выявлением новых связей необходимо вновь его использовать.

Система имеет гибкие настройки, что позволяет использовать возможности встроеного модуля обработки данных для обработки похожих страниц (содержимого каталога товаров, например). Таким образом, заинтересованные сайты смогут получить семантическое описание и будут найдены и обработаны поисковым движком семантического портала, и пользователи портала смогут легко узнать о новостях, товарах и услугах этих сайтов.

#### **Построение онтологии**

Отдельно стоит упомянуть механизм дополнения онтологической модели в предлагаемой системе. Онтология не является закрытой, и у пользователей есть возможность добавлять в нее свои понятия и отношения.

В целях поддержания целостности и порядка в онтологии вновь добавленные понятия и отношения вносятся в отдельные графы онтологии и попадают в основную модель знаний только после верификации менеджером по работе со знаниями. Таким образом, любой документ, пользователь или иной объект знания может быть описан наиболее полно.

#### **Возможности семантического портала**

Семантический портал представляет множество новых возможностей для пользователя. Перечислим возможности, доступные в реализуемом в данный момент портале:

Каждый пользователь, зарегистрированный в системе, может описать себя, причем ввиду гибкости онтологии он может добавить себе нужные свойства, изначально системой не предусмотренные.

Организация может представить себя на портале подробным семантическим описанием товаров и услуг, ею осуществляемых.

Каждый пользователь может зарегистрировать проект, идею, событие и т.п. и получить по ней отзывы и предложения. Также он может описать персонал, необходимый для реализации этого проекта, и расчетную стоимость каждого этапа. Система автоматически уведомит требуемых специалистов о возможности поучаствовать в проекте.

Благодаря модулю семантического описания существующие сайты могут представить содержащуюся в них информацию в более строго описанной форме, после чего у пользователей портала появится возможность их эффективно искать.

Таким образом, пользователь сможет найти всю информацию по интересующей его тематике с учетом расширенных возможностей логического вывода.

#### **Детали реализации**

В качестве основного средства разработки самого продукта используются такие продукты Microsoft: MVC, SilverLight, Asp.NET, MemberShip, Microsoft Visual Studio 2008, Microsoft SQL Server 2008.

В качестве хранилища было решено выбрать систему Virtuoso [4]. В качестве подсистемы работы с текстами на естественном языке используется продукт компании Russian Context Optimizer – RCO [5]. Для модуля семантического описания используются компоненты компании AOL – RML [6].

#### **Заключение**

Создание городской сети, реализующей идеи Web 3.0, является сложной и интересной задачей. С ней связано много трудностей, главная из которых – сложность создания непротиворечивой модели областей знания.

В настоящий момент этот проект реализуется в рамках Томской городской сети.

*Литература*

1. Тим Бернерс Ли, Джеймс Хендлер и Ора Лассила. Семантическая Сеть [Электронный ресурс]. – Режим доступа: [http://ezolin.pisem.net/logic/semantic\\_web\\_rus.html](http://ezolin.pisem.net/logic/semantic_web_rus.html), свободный (дата обращения: 17.11.2010).
2. Сайт проекта Linked Data [Электронный ресурс]. – Режим доступа: <http://linkeddata.org/>, свободный (дата обращения: 17.11.2010).
3. Тузовский А.Ф. Система интеграции информации и знаний с использованием семантических технологий / А.Ф. Тузовский, А.В. Черный // Известия Томского политехнического университета. – 2009. – Т. 315, № 5. – С. 127–132
4. Сайт компании OpenLink Software [Электронный ресурс]. – Режим доступа: <http://virtuoso.openlinksw.com/>, свободный (дата обращения: 17.11.2010).
5. Описание компонента RCO Fact Extractor [Электронный ресурс]. – Режим доступа: [http://rco.ru/product.asp?ob\\_no=1131](http://rco.ru/product.asp?ob_no=1131), свободный (дата обращения: 17.11.2010).
6. Пакет документации к программному продукту RML [Электронный ресурс]. – Режим доступа: <http://www.aot.ru/technology.html>, свободный (дата обращения: 17.11.2010).

**Тузовский Анатолий Федорович**

Д-р техн. наук, проф. каф. оптимизации систем управления НИТПУ  
Тел.: (382-2) 42-14-85  
Эл. почта: TuzovskiyAF@tpu.ru

**Черный Антон Васильевич**

Аспирант каф. оптимизации систем управления  
Национального исследовательского Томского политехнического университета (НИТПУ)  
Тел.: +7-906-954-69-48  
Эл. почта: cherny@tpu.ru

Cherny A.V., Tuzovskiy A.F.

**The use of semantic technologies for social networks creation**

The features of social networks creation based on the using semantic technologies are considered. A suggested architecture of such information systems, its possibilities and approach to ontologies construction are analyzed.

**Keywords:** semantic Web, ontology, social network.