

ТОКМАКОВ Дмитрий Иванович – руководитель Центра информационных технологий ФГОУ ВПО Восточно-Сибирской государственной академии культуры и искусств (Улан-Удэ)

ИСПОЛЬЗОВАНИЕ СРЕДСТВ ЯЗЫКА RDF В АННОТИРОВАНИИ ИНТЕРНЕТ-РЕСУРСОВ

Для развития интерактивных средств доступа к данным в интернет далеко не всегда достаточным является использование традиционного подхода к представлению информации в виде гипертекста и гипермедиа с использованием простых языков разметки типа HTML. Решение этой проблемы, как правило, связано с использованием различных процедурных включений, а именно Java-скриптов и Java-апплетов (на стороне клиента) и механизмов CGI-скриптов и сервлетов (на стороне сервера). Таким образом, в рамках интернета становится возможным представлять не только пассивную текстовую и мультимедийную информацию, но и сложные процедуры, взаимодействующие с пользователем с достаточно развитой функциональностью. Зачастую возникает необходимость осуществления доступа не только к одному централизованному ресурсу, но и потребность объединения нескольких информационных ресурсов для совместного функционирования. Для создания таких систем недостаточно использования традиционных Интернет-технологий, а требуются более сложные механизмы удаленного взаимодействия.

Основная на сегодняшний день единица обмена информации в интернете – HTML-страница – полностью ориентирована на понимание человеком:

- HTML-страница хранится на сервере.
- HTML-страница пересылается по запросу пользователя на его персональную машину.
- HTML-страница демонстрируется персональным компьютером пользователю в графическом виде, удобном для понимания человеком.

Этим функции компьютера, как правило, и ограничиваются. Компьютер не может понять и оценить содержимое страниц, поскольку они рассчитаны на интеллект человека. По этой причине зачастую очень трудно осуществлять интеграцию различных Web-ресурсов, так как для этого необходимо научить программные службы данных ресурсов понимать размещенную на них информацию и взаимодействовать друг с другом.

Таким образом, понимание семантики информации является важнейшим условием создания нового поколения Интернет-служб. Только это и позволит отыскивать нужную информацию в лабиринтах глобальной сети, организовывать взаимодействие различных ресурсов между собой. Следовательно, необходимо предложить такой способ описания информации, который бы позволил однозначно и недвусмысленно представлять ее в формализованном виде. Это позволит программам не только читать, но и понимать содержимое Интернет-ресурсов.

1. Обзор проектов, посвященных представлению знаний в Интернет-ресурсах на основе аннотирования Web-страниц

В настоящее время в области представления знаний в Интернет-ресурсах ведутся активные исследования. Их целью является организация интеллектуального поиска в интернете и автоматизация накопления новых знаний. Основная идея заключается в аннотировании Web страниц дополнительной информацией об их содержании, которая в дальнейшем может быть использована для определения степени соответствия данного ресурса поисковому запросу. Кроме решения задач поиска информации, становится возможным решение задач интеллектуального аннотирования, автоматического извлечения знаний из текстов на естественном языке, построение мультиагентных систем и многих других.

Рассмотрим ряд проектов, посвященных данному направлению.

Проект SHOE[1] представляет систему поиска информации в интернете на основе специализированных баз знаний (онтологий) и вывода на базах знаний. Для работы нужно с помощью средств системы создать онтологию (описать классификацию понятий из предметной области, связи между ними и другими объектами) и аннотировать HTML страницы, используя атрибуты, описанные в этой и/или других онтологиях. Пользователи могут сами создавать онтологии и расширять конкретную онтологию новыми понятиями.

Онтология здесь включает “is a” иерархию классов и/или категорий, атомарные связи между категориями, множество правил вывода в форме упрощенных Хорновских формул. В этой системе центральный администратор онтологий не определен. Как следствие, давая запрос, пользователь не может знать все термины, которые требуются для аннотации HTML- документов. Поэтому ответ на запрос может не содержать важную для пользователя информацию.

Для описания онтологий в SHOE предусмотрены теги <ONTOLOGY>, <ONTOLOGY-EXTENDS>, <ONTODEF>. Для аннотирования HTML документов существуют следующие теги: <USE-ONTOLOGY>, <META>, <INSTANSE>, <CATEGORY>, <RELATION>, <ATTRIBUTE>. Полная спецификация языка приведена на сайте <http://www.cs.mnd.edu/projects/plus/SHOE/spec.html>

На базе SHOE в University of Maryland at College Park разработан web-crawling робот, обрабатывающий SHOE-документы и помещающий их в свою базу знаний, используя систему представления знаний PARKA.

В SHOE сделана попытка сделать возможным прямое извлечение и классификацию семантически значимой информации из HTML – документов посредством семантического аннотирования документов специальными тегами.

Другим примером использования онтологической системы является инициатива (КА)² [2] (Knowledge Annotation Initiative of the Knowledge Acquisition Community). Это международный проект, целью которого является организация интеллектуального поиска в интернете и автоматическое накопление новых знаний. В этой инициативе выделяют следующие направления:

- Аннотация web-страниц интеллектуальной информацией
- Онтологический реинжиниринг
- Организация интерфейса запросов и вывода по распределенной онтологии.

(КА)² - это открытая инициатива, в рамках которой участники включаются в процесс создания распределенной онтологии и модели извлечения знаний. Ими разрабатывается онтология, которая описывает различные исследовательские группы, работающие в области извлечения знаний. Одной из целей (КА)² является трансформация из представления WWW (как базы знаний) в систему, основанную на знаниях, определение, какие необходимы средства для распределенного проектирования онтологий, и изучение социального аспекта согласования описания онтологии распределенными участниками проекта.

Использование систем на базе онтологий в рассмотренных выше подходах влечет за собой необходимость решения двух больших по временным затратам задач: описание онтологии и аннотирование web-документов.

Аннотация Web-страниц осуществляется за счет расширения HTML специальным тегом <ONTO>, в рамках которого можно задавать онтологии для спецификации WWW страниц. Процесс дополнения Web страниц такой онтологической информацией рассматривается в рамках направления онтологического инжиниринга. В онтологический инжиниринг входит разработка собственной онтологической системы на основе инструментария Ontolingua.

Для поиска в рамках (КА)² предполагается использовать подсистему Ontocrawler, а для организации запросов и вывода разработан программный продукт Ontobroker. Для каждого из этих модулей разработан специальный язык: язык запросов для формирования запросов, язык представления знаний для описания онтологий, язык аннотирования web-документов для расширения стандартного HTML семантической информацией.

Знания, содержащиеся в HTML документах, аннотируются для того, чтобы сделать их доступными как факты для Ontobroker. Документы остаются понятными для стандартных браузеров. В HTML документах объектами аннотирования могут быть:

- объект, определенный URL, может быть определен как пример обычного класса,
- установленное значение атрибута объекта,
- установленная связь между двумя и более объектами.

Среди отечественных разработок можно отметить онтологическую поисковую систему **Jewel** [3], которая для составления онтологических описаний использует фреймо-продукционный способ представления знаний. В основе языка описания онтологий лежит фреймово-продукционный язык JFMDL из состава инструментария JULIA (Java Universal Library for Intelligent Applications)[4]. Язык позволяет производить онтологические описания HTML страниц, используя понятия: категория, условное правило, безусловное правило и концепт.

Под онтологией HTML страницы понимается описание некоторого ресурса, проводимое в терминах общего языка описания онтологий. Для описания онтологии используется надмножество стандарта HTML, в котором расширяется стандартный тег <SCRIPT>, а также вводятся новые теги <USE>, <CONCEPT>, <SET>, <ASSIGN>. Для составления поисковых запросов в системе Jewel применяется специализированный язык. Для реализации пользовательского интерфейса реализованы утилиты администрирования поисковой системы и Java-сервлет, предоставляющий пользователю возможность формулировать поисковые запросы.

Основной причиной, препятствующей повсеместному распространению описанных выше технологий, является отсутствие широкого признания данных технологий со стороны информационного сообщества. Поэтому рассмотренные механизмы вряд ли расширят сферу своего современного применения и получат достойную поддержку в развитии от сторонних поставщиков программных решений.

В качестве технологии, имеющей серьезные шансы на успех, сегодня серьезное развитие получила концепция Semantic Web [5,6], с точки зрения которой, интернет рассматривается как потенциальная база знаний. Для работы со знаниями в сети нужны специальные методы представления и обработки знаний, интерпретации запросов и т.д. В рамках данного проекта комитетом W3C разработан ряд стандартов, повсеместное внедрение которых призвано решить многие проблемы существующего веба. Ключевое отличие семантической сети от нынешнего глобального информационного пространства состоит в том, что компьютеры будут обрабатывать информацию не только как данные, но и как знания.

В феврале 2004 года WWW-консорциумом опубликована спецификация Resource Description Framework (RDF)[7]. В основу RDF положены две идеи: идентификация ресурса с помощью идентификаторов URI [8] и описание ресурсов в терминах простых свойств и их значений. Под ресурсом может пониматься не только некоторый документ или набор документов в Сети, представленный своим URL, но и любой объект, который вы пожелали описать. Базовый строительный блок в RDF — тройка «объект — атрибут — значение», которую часто записывают в виде A(O,V). Часто эти три части также еще называют соответственно: субъект, предикат, объект. Эта запись означает: объект O имеет атрибут A со значением V. Эту связь можно также представить как ребро с меткой A, соединяющее два узла, O и V: [O]-A—>[V]. Данная нотация весьма полезна, поскольку RDF позволяет менять местами объекты и значения. Таким образом, любой объект может играть роль значения, что в графическом представлении соответствует цепочке из двух ребер с метками.

В 2004 году WWW комитетом была утверждена еще одна спецификация RDF Schema [9], которая определяет, каким образом следует использовать RDF для описания RDF-словарей. RDF Schema для RDF задает ограничения и описывает отношения между используемыми терминами.

Таким образом, видно, что с помощью RDF можно достаточно эффективно описывать различные взаимоотношения между различными объектами.

Для использования возможностей языка RDF в аннотировании Интернет-ресурсов, представленных в различных XML-форматах, комитетом W3C ведется разработка стандарта RDF/A[10].

Стандарт RDF/A описывает способ, с помощью которого RDF-данные могут помещаться внутри XHTML-файлов. Для реализации данной возможности для произвольного тега можно определить атрибуты *content*, *property* и *about*. Атрибут *about* подразумевает, что субъект внутри контейнера, определяемого текущим тегом (т.е. тег, внутри которого определен атрибут *about*) - это ресурс, являющийся значением данного атрибута. Атрибут *property* определяет предикат. При этом если наряду с атрибутом *property* определен атрибут *content*, тогда объектом является значение атрибута *content*. В противном случае объектом будет являться содержимое тега. Отдельным случаем является тег *<a>*. В данном теге объектом является значение атрибута *href*, а в качестве предиката выступает значение, определяемое в атрибуте *rel*. Для типизации ресурсов в языке RDF определена конструкция *rdf:type*. Данную конструкцию следует определять как значение атрибута *class*.

Пример использования RDFa в html-документе:

```
<html
  xmlns:vcard="http://www.w3.org/2001/vcard-rdf/3.0#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
.....
<p about="www.vuz.ru/cafedraAIS.html">
<span property="dc:title">Кафедра информационных систем</span> создана в <span
property="dc:date" content="1995-10">октябре 1995 года</span>.
</p>
.....
</html>
```

Приведенный пример html-страницы содержит следующие RDF-тройки:

```
< www.vuz.ru/cafedraAIS.html >
      dc:title "Кафедра информационных систем";
      dc:date "1995-10".
```

Проблемы применения стандарта RDF/A заключается в следующем:

- На текущий момент времени стандарт RDF/A не является окончательным.
- Существующий язык запросов SPARQL к RDF-источникам данных не позволяет напрямую работать с html-ресурсами, аннотированными RDF-разметкой. Поэтому непосредственно использовать существующие инструменты для работы с подобными ресурсами невозможно.

Для решения последней проблемы комитетом W3C разработана технология GRDDL (Gleaning Resource Descriptions form Dialects of Languages) [11]. Цель данной технологии заключается в предоставлении механизмов извлечения RDF-троек из HTML-документов.

Основная идея технологии GRDDL заключается в следующем. Способ извлечения RDF-данных определяется использованием технологии XSLT. XSL-шаблон, описывающий способ извлечения данных, указывается внутри тега **link**, определенного в голове документа (тег **head**). У тега **link** определяются атрибуты **rel** со значением "transformation" и **href** с указанием пути к XSL-файлу. В результате работы XSLT-преобразования формируется документ, содержащий извлеченные RDF-данные, находящиеся в обрабатываемом html-документе. Количество тегов **link**, описывающих способы извлечения данных из документа, может быть произвольным, и, соответственно, может быть сформировано произвольное количество RDF-документов, которые далее можно объединить в единый ресурс.

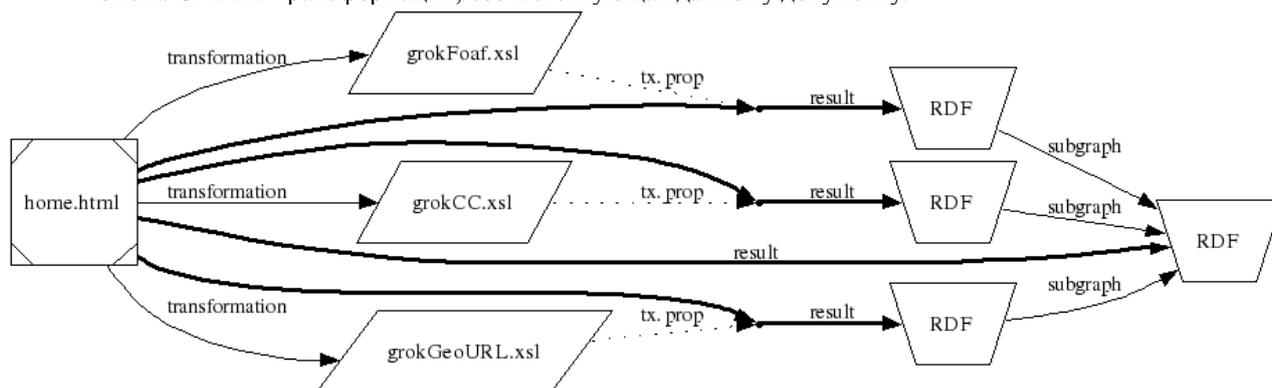
В качестве примера использования технологии GRDDL рассмотрим вариант, приведенный в [11]:

Исходный HTML-документ:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head profile="http://www.w3.org/2003/g/data-view">
  <title>Joe Lambda's Home page [an example of RDF in XHTML]</title>

  <link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-
xslts/grokFOAF.xsl" />
  <link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-
xslts/grokCC.xsl" />
  <link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-
xslts/grokGeoURL.xsl" />
  ...
```

Схема GRDDL-трансформации, соответствующая данному документу:



С полученным в результате данной трансформации RDF-ресурсом уже можно полноценно работать, используя все возможности существующих технологий, разработанных для Semantic Web.

2. Метаданные

Данные, доступные из информационных источников, как правило, отличаются синтаксической неоднородностью, разнообразием моделей данных и форм представления данных. Следовательно, необходимо выработать и согласовать унифицирующий подход приведения данных к наиболее распространённым моделям данных и форматам. Учитывая это обстоятельство, для представления сведений о ресурсах стали использовать структурное представление, выделять понятие метаданных, описывающих содержимое ресурса в виде набора именованных значений, в том числе связей с другими ресурсами. Метаданные используются для автоматизированного анализа содержимого ресурса, построения поисковых индексов и позволяют обеспечить достаточно высокую точность и эффективность поиска разнотипной информации.

Проблему интеграции информационных ресурсов на уровне метаданных принято сейчас называть проблемой интероперабельности. При этом возникают довольно сложные логические и даже математические проблемы. Фундаментальное изложение проблем, связанных с созданием и использованием систем метаданных в различных программно-технологических средах, а также подробное описание стандартов, имеющих в области метаданных, содержится в работе [13].

Обзору систем метаданных посвящены работы [12-14], в которых проведено описание и сопоставительный анализ различных систем метаданных. Ряд форматов метаданных, касающихся обслуживания следующих областей знания: культура, образование, наука, описан в работе [12].

Приведем перечень некоторых наиболее известных систем метаданных:

- DC - Дублинское ядро метаданных, универсальная система метаданных, разработанная консорциумом W3C, принятая в качестве национального стандарта в ряде стран [15].

- *LOM – модель учебного объекта* - система метаданных для образовательных ресурсов, разработанная Международной ассоциацией IEEE, принятая в качестве стандарта в России для Единой информационно-образовательной среды [16].
- *GEM* – версия Дублинского ядра метаданных, доработанная для образовательных информационных ресурсов, принятая в качестве стандарта для образовательных ресурсов США.
- *VCard* – система метаданных для представления сведений о физических лицах и организациях, разработанная Международной ассоциацией IEEE.
- *MARC – машиночитаемый каталог* - семейство форматов, широко распространенных для представления библиографической информации, главным образом в мировом библиотечном сообществе. В полном варианте включает до 800 элементов данных. Является стандартом де-факто для российских библиотечных сетей.
- *ONIX* – формат описания информационных товаров в системах электронной торговли (книги, периодика, компактные диски, аудио- и видеокассеты и др.). Основан на системе стандартов ЭДИФАКТ.
- *GILS - Глобальная (правительственная) служба поиска информации*. Система метаданных GILS представляет собой расширение форматов *MARC* для описания информационных ресурсов широкого профиля.
- *EDIFACT* - одна из старейших международных систем метаданных, содержащих правила и структуру описания торгово-транспортных и других коммерческих документов. С 1998 г. поддерживается языком XML. Реализация применительно к торговле книгами и другими информационными товарами – система ONIX .
- Формат Государственного регистра баз и банков данных. Российский стандарт, действующий с конца 1980-х гг. и содержащий систему метаданных для баз данных и других электронных наборов данных. Утвержден Правительством РФ, поддерживается НТЦ «Информрегистр», применяется в ряде отраслей и регионов, а также в странах СНГ для ведения баз метаданных.

Имеется еще несколько десятков менее известных систем метаданных, в той или иной степени претендующих на широкое применение.

Среди большого количества разнообразных систем метаданных, используемых в информационных ресурсах в сфере науки, культуры и образования и рассмотренных выше, наиболее распространенными и признанными являются системы метаданных семейства *MARC*, *DC*, а также быстро распространяющаяся в информационно-образовательной среде система метаданных *LOM*. Поэтому при создании сводных каталогов электронных ресурсов разработчикам чаще всего приходится иметь дело с интеграцией описаний ресурсов, представленных в этих системах метаданных. При этом следует отметить, что вследствие сложности форматов семейства *MARC* и их конкуренции между собой их область применения вряд ли выйдет за пределы обмена библиографической информацией между библиотеками.

Кроме собственно систем метаданных, обладающих семантикой, к метаданным относят формальные средства, использование которых предусмотрено языками разметки и протоколами, принятыми в интернете. Это теги `<meta>`, определяемые для языка разметки HTML и для протокола HTTP соответственно спецификациями RFC 1866 и 2616.

По общему мнению, эти теги использовать неудобно, поскольку они не могут применяться для других типов файлов и быстро становятся громоздкими.

3. Технология включения RDF-троек в HTML ресурсы

Серьезной проблемой при создании веб-документов, включающих семантическую разметку, является то обстоятельство, что разработчику приходится заниматься не только вопросами дизайна и верстки документа, но и вопросами корректного аннотирования содержимого документа. Одним из немногих вариантов решения данной задачи является автоматизация данного процесса, тем более что значительное число веб-документов формируется не вручную, а автоматизированным способом с помощью скриптов или иных средств.

В данной статье рассматривается вариант автоматизации формирования веб-документов, содержащих семантическую разметку согласно стандарту RDF/A. Исходные данные представлены в xml-формате. Хотя текстовый формат xml-файла позволяет работать с ним из текстового редактора, целесообразнее предоставить пользователю специальный программный интерфейс. Автором использованы для данной цели средства языка PHP. Задача пользователя заключается в заполнении полей формы необходимой информацией.

В процессе формирования исходных данных следует учитывать необходимость однозначной идентификации объектов, описываемых в документах. Например, сведения об одной и той же персоне могут встречаться в разных документах. Поэтому, для однозначной идентификации данной персоны следует включать в документ не только какие-либо данные о ней (например, ФИО), но и URI-ссылку на персональные сведения данного объекта, хранящиеся в формате VCARD. По этим данным уже можно будет однозначно утверждать, что персона, упоминаемая в разных частях документа или в разных документах, является одним и тем же лицом.

Формирование веб-документов производится автоматизированным способом посредством XSLT-преобразования информации, представленной в XML-формате. Исходные данные подвергаются XSLT-

преобразованию с помощью заранее подготовленных XSL-шаблонов в HTML-файлы с внедренными RDF-данными (рис.1).

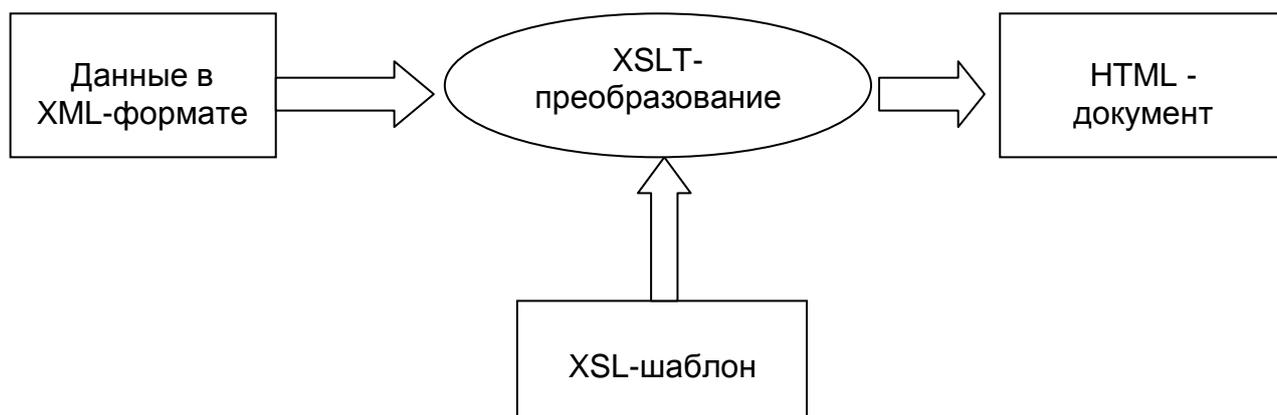


Рис.1 Формирование документов.

В данном случае XSL-шаблон определяет не только внешний вид веб-документа, но и способ внедрения в него RDF/A разметки. В результате получаем HTML-файл, содержащий информацию, доступную не только человеческому восприятию, но и способную быть извлеченной и обработанной различными программными сервисами.

4. Формирование RDF-данных

Для получения данных из RDF-хранилищ используется язык запросов SPARQL, разработанный комитетом W3C. Но, несмотря на то, что полученные HTML-файлы аннотированы RDF-данными, воспользоваться данным языком для выполнения запросов непосредственно к ним нельзя. Следовательно, применить существующий программный инструментарий для извлечения из HTML-документов RDF-троек невозможно.

Одним из возможных вариантов решения этой проблемы является применение GRDDL-технологии для извлечения RDF-троек. Для этого над каждым файлом проводится GRDDL-преобразование и извлеченный набор RDF-троек помещается в RDF-хранилище. В качестве хранилища могут использоваться различные продукты: Joseki RDF-server, Sesame и др. Для каждой RDF-тройки указывается ссылка на ресурс, из которого она была извлечена. Таким образом, информация, занесенная в RDF-хранилище, представлена в следующем виде:

```
<URI-ссылка на ресурс> : <субъект> <предикат> <объект>  
<URI-ссылка на ресурс> : <субъект> <предикат> <объект>  
.....  
<URI-ссылка на ресурс> : <субъект> <предикат> <объект>
```

В итоге в RDF-хранилище формируется своеобразный индекс HTML-файлов, аннотированных согласно стандарту RDF/A. Следовательно, предоставляется возможность организации поисковых запросов с помощью средств языка SPARQL к HTML-файлам, аннотированных RDF/A разметкой. В результате появляется возможность организации эффективного поиска по содержимому различных Интернет-ресурсов, т.к. их содержимое доступно для понимания программным службам. Например, программа может определить, что информация о какой-либо персоне, содержащаяся на различных сайтах, относится к одному и тому же человеку.

5. Проблема организации запросов к RDF-данным

Существенной проблемой организации запросов к RDF-данным является сложность формирования SPARQL-запроса. Непосредственное написание запроса пользователем требует знания пользователем языка SPARQL и структуры данных, к которой подобный запрос должен применяться. Поэтому в качестве возможного решения данной проблемы может являться вариант, когда пользователю предлагается воспользоваться набором готовых шаблонов запросов, в которых достаточно заполнить реквизиты, согласно которым будет происходить отбор данных из хранилища.

Заключение

Предлагаемая технология дает возможность разработать один из подходов к внедрению технологий Semantic Web в описании Интернет-ресурсов. Широкое внедрение данных технологий позволит организовать

более эффективную работу с Интернет-ресурсами в таких областях, как информационный поиск или интеграция данных.

Литература:

1. *Simple HTML Ontology Extentions.* <http://www.cs.umd.edu/projects/plus/SHOE>
2. *Knowledge Annotation Initiative of the Knowledge Acquisition Community.* <http://www.ksl.svc.stanford.edu>
3. Е.В. Сизиков, Д.В. Сошников. Онтологическая поисковая система Jewel для реализации интеллектуального поиска в Интернет- и интранет-сетях. http://www.mai.ru/projects/mai_works/articles/num7/article10/author.htm
4. Д.В. Сошников. Инструментарий JULIA для создания распределенных интеллектуальных систем на основе продукционно-фреймового представления знаний. http://www.mai.ru/projects/mai_works/articles/num7/article9/author.htm
5. Тим Бернес-Ли, Джеймс Хендлер, Ора Лассила. *Semantic Web.* *Scientific American*, 2001. - http://ezolin.pisem.net/logic/semantic_web_rus.html
6. Манцивода А.В. Семантический Интернет - <http://teacode.com/concept/eor/sp2.html>
7. *Resource Description Framework (RDF): Concepts and Abstract Syntax.* -<http://www.w3.org/TR/rdf-concepts/>
8. *Uniform Resource Identifier (URI): Generic Syntax*// <http://www.gbiv.com/protocols/uri/rfc/rfc3986.html>
9. *RDF Vocabulary Description Language 1.0: RDF Schema*// <http://www.w3.org/TR/rdf-schema/>
10. *RDFa in HTML Overview (Editors' Draft)* -<http://www.w3.org/2006/07/SWD/RDFa/>
11. *Gleaning Resource Descriptions from Dialects of Languages (GRDDL).* W3C Working Draft 2 March 2007 - <http://www.w3.org/TR/2007/WD-grddl-20070302/>
12. Соколовский В.В. Сопоставительный анализ стандартов метаданных для сферы науки, культуры и образования / В.В. Соколовский. М., 2006.- 12 с. (Серия «Электронные библиотеки: теория и методика»; Препринт №06.004, март 2006 г. / Российская ассоциация электронных библиотек – НП ЭЛБИ)
13. Когаловский М.Р. Энциклопедия технологий баз данных. – М.: Финансы и статистика, 2002.
14. Антопольский А.Б. Информационные ресурсы России. – М.: ИТЦ «Информрегистр», ИПКИР, 2004.
15. *Dublin Core Metadata Initiative* - <http://dublincore.org/>
16. *IEEE 1484.12.1-2002. Learning Object Metadata standart.* – New York: IEEE, 2002.
17. *SPARQL Query Language for RDF* - <http://www.w3.org/TR/rdf-concepts/>
18. *vCard: The Electronic Business Card* - <http://www.imc.org/pdi/>
19. *The Friend of a Friend (FOAF) Project* -<http://www.foaf-project.org/>