

# Content Management in Semantic Web

**B. Alkhatib, E. Fakhane**

## Abstract

The Semantic Web is a technique of modern technologies in the world wide web, which it's old do not exceed more than 10 years , and researches still continues until now on the development of this technique, applied it, and the extent to which reflected the pros and cons in the world wide web. This new technology in world of the web would reconsider the content management of the new Web via the means and modern tools consistent with the structure of the Semantic Web, and our goal through this paper is to highlight the content management in the Semantic Web, and find a way to manage this content using entailment technique in semantic web which means extract a new data from existing data.

## Keywords

Content Management, Semantic Web, Web 3.0, Entailment.

## I. Introduction

The researchers is speaking now to move from Web 2.0, static Web which is just a viewer for the texts, images, video, databases and document sharing non-understandable for a computer to Web 3.0, semantic Web, where the related objects, the real world with each other and also can communicate with infinite set of integrated databases developed in a formula understandable to a computer, and also for machines in general, and that can be processed directly or indirectly, in other words make the data understandable and processing by machines.

Once Computers understand what Data means, they can accomplish more intelligent searches on the data, as well as the inference operations and the integration of these data, which saves us a lot of effort. Semantic web provides a solution for this problem by suggesting the data management with help of meaningful relations and tagging. In this paper we present a way to connect to Rdf file as a database, and do some processing operation on this file like insert ,update and delete object, in addition , feed the search operation by doing entailment on data. I consider my work as a continues to other similar work , that is a Semantic Web Solutions by Hafiz Hammad Rubbani which discuss the semantic web inference and gives a good understanding of semantic web technologies, its query language and inference capability. It also discusses how semantics technologies help in data publishing and sharing. He uses some open source tools. Although he uses some programming languages like Java, J2ee to make a content management on data, but he doesn't discuss or implement the entailments on data which gives improvement to Search operations ,and consequently more accurate result. The differential of our work is using and implementing the inference on the semantic web to feed the search operation and add extra meaning to the data.

## II. Method

The content management system (CMS) is consider as a means that facilitate the storage and retrieval of shared information or users' information, and is allow people to give and share data, and means to reduce the repeated input data, and the means of communication between users, the functions of security and respect for human users access individual based on their roles (read, modify, publish, etc.) It is also very important.

Many versions of commercial as well as tools of open-source content management systems available, with regard to the open source tools, and that is very popular is Drupal, which can be used for community and trade sites, forums, blogs, corporate sites, Joomla is a content management system suitable for companies, e-commerce and even personal websites. Moodle (Modular

Object-Oriented Dynamic Learning Environment) is a content management system that is used to create online courses and web sites to learn . MediaWiki originally is used as a site on Wikipedia, but became available for Internet users who want to create collaborative and media websites.

In my work I use :

- 1- Altova Semantic Work as rdf editor to build rdf file which contain ontology description about News including classes, properties and instances.
- 2- Microsoft visual studio as a programing environment tools in asp.net and c#.
- 3- Open Source .Net Library (dotnetRdf) which provide a powerful and easy to use API for working with RDF.

The reasons of choosing dotnetRdf library is:

1. These library is using .Net Environment
2. Full support for Sparql 1.1
3. Have link to object feature.
4. Well designed.

DotnetRdf negative :

1. Partially not fully support RDFa 1.1.
2. Sparql Engine in some query consume memory

dotnetRdf alternative :

1. SemWeb : but it support only Sparql 1.0.
2. Intellidimension Semantics SDK : but it stopped on development since 2010.

## Representing Data

Content management must have the tools that make us representing data in any form we want to show the data.

To represent data, I make a mapping between classes in the rdf file to Table, the class corresponding to DataTable and the property corresponding to Column, as the Figure 1 show:

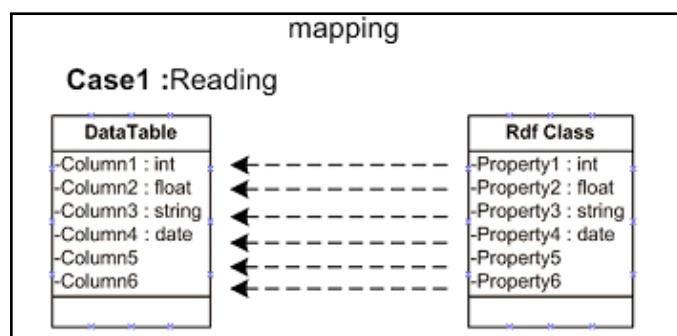


Fig. 1: Mapping

**The modify operations**

Any content management system have the ability of doing modification on the data. We can make the modify operation throw executing Sparql statement like:

Insert Sparql: to add some data to the rdf file.

Update Sparql: to modify some data in the rdf file.

Delete Sparql : to delete dome data from the rdf file.

Any insert or update or delete operation must done completely, the modification operation and all the objects related to the modification operation as we will see later.

**Authentication**

Content management system need to make sure that the current user is authenticated or not.

To check the user authenticated or not authenticated, I define a class CUser which has three properties UserName, Password and FullName, if there is an object of class CUser for the user then the user authenticated, if there isn't the user isn't authenticated

**Authorization**

After making sure that the user is authenticated we also need to know if the user have permission to do some processing on the data , and that is authorization.

To check the user authorized or not authorized, I also define a class CRight to represent one right, A property hasRight which is of type CRight in the class CUser show if the user has right or not. If the Class CUser has right then he is authorized ,if he has not then he isn't authorized.

**How to Add an objects of class News(CNews) and linked it**

Suppose we want to add an instance of class CNew which means a New News, How will the added process done, and how it will be linked to, and what are the objects effected by this add operation.

To add a new object of class CNews , it must at first determine the values of object's properties, which is :

- Heading : the head of News.
- Text : the text of News.
- Date : the date that the News done.
- Tags : the News Keywords .
- inCategory : the News Category like (Sport, Political, health ..etc. ).
- aboutEvent: the Event which the news occurred in.
- fromArea : the area which the news came from.

for example we will put a values for these properties :

Properties	Value
Heading	Obama vs Romney
Text	President Obama and Mitt Romney
date	12/16/2006 12:00:00 AM
Tags	Obama Romney
inCategory	Politics
aboutEvent	General
fromArea	NewYork

the object represent as rdf text as follow:

```
<CNews rdf:about="http://www.owl-ontologies.com/news.owl#Obama_vs_Romney">
```

```
<aboutEvent rdf:resource="http://www.owl-ontologies.com/news.owl#General" />
<date rdf:datatype="&xsd;dateTime">12/16/2006 12:00:00 AM</date>
<fromArea rdf:resource="http://www.owl-ontologies.com/news.owl#NewYork" />
<heading rdf:datatype="&xsd:string">Obama vs Romney</heading>
<inCategory rdf:resource="http://www.owl-ontologies.com/news.owl#Politics" />
<tags rdf:datatype="&xsd:string">Obama Romney</tags>
<text rdf:datatype="&xsd:string">President Obama and Mitt Romney were supposed to duke it out over foreign policy.</text>
</CNews>
```

When we add a new object , there may some others objects effected by this adding and consequently must updated too.

The objects which effected by adding object of class CNews is :

- The object associated with news object by inCategory property, and that is Politics.
- The object associated with news object by aboutEvent property, and that is General.
- The object associated with news object by fromArea property, and that is NewYork.

**Politics Object :**

It will modify the object Politics which is of type CPolitics to be contain the news Obama\_vs\_Romny as follow :

```
<CPolitics rdf:about="http://www.owl-ontologies.com/news.owl#Politics">
<categorysNews rdf:resource="http://www.owl-ontologies.com/news.owl#Obama_vs_Romney" />
</CPolitics>
```

**General Object :**

It will modify the object General which is of type CEvent to be contain the news Obama\_vs\_Romny as follow :

```
<CEvent rdf:about="http://www.owl-ontologies.com/news.owl#General">
<eventNews rdf:resource="http://www.owl-ontologies.com/news.owl#Obama_vs_Romney" />
<name rdf:datatype="&xsd:string">General</name>
</CEvent>
```

**NewYork Object :**

It will modify the object NewYork which is of type CCity to be contain the news Obama\_vs\_Romny as follow :

```
<CCity rdf:about="http://www.owl-ontologies.com/news.owl#NewYork">
<areasNews rdf:resource="http://www.owl-ontologies.com/news.owl#Obama_vs_Romney" />
<isCityOf rdf:resource="http://www.owl-ontologies.com/news.owl#USA" />
<name rdf:datatype="&xsd:string">NewYork</name>
</CCity>
```

The Sparql Statement of adding the CNews Object will be as follow :

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```

PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX base: <http://www.owl-ontologies.com/news.owl#>
WITH <http://www.owl-ontologies.com/mynews.owl>
INSERT
{ base:Obama_vs_Romney rdf:type base:CNews .
  base:Obama_vs_Romney base:inCategory base:Politics.
  base:Obama_vs_Romney base:text "President Obama and
Mitt Romney were supposed to duke it out over foreign
policy."^^xsd:string.
  base:Obama_vs_Romney base:date " 12/16/2006 12:00:00 AM
^^xsd:dateTime.
  base:Obama_vs_Romney base:heading " Obama v s Romney
^^xsd:string.
  base:Obama_vs_Romney base:tags " Obama Romney
^^xsd:string .
  base:Obama_vs_Romney base:aboutEvent base:General.
  base:Obama_vs_Romney base:fromArea ?ar .
  base:Politics base:categorysNews base:Obama_vs_Romney.
  base:Genera base:eventNews base:Obama_vs_Romney.
  base:NewYork base:areasNews base:Obama_vs_Romney. }
WHERE
  { ?ar base:name "NewYork"^^xsd:string. }

```

### How the search operation on the objects of type CNews will be ?:

In content management system ,we need to do some search operation on the data, this search operation may be on property value, like getting all the news about NewYork city, or search by part of text, like getting all the news that contain a certain text in news text property.

Search can be done for class of type CNews by the following parameters:

- Search by the head of news (Heading).
- Search by the text of news (Text)
- Search by the date of News (Date).
- Search by the News Keywords (Tags).
- Search by the News Category (inCategory)
- Search by the Event which the news occurred in (aboutEvent).
- Search by the area which the news came from (fromArea)

search operation Can be directly, such as news related to a specific event, for example, the following Sparql query will enable us to search for the news that occurred in World War II:

```

SELECT *
WHERE {
  ?CNews rdf:type base:CNews .
  ?CNews base:inCategory ?Category .
  ?CNews base:text ?text .
  ?CNews base:date ?date .
  ?CNews base:heading ?heading .
  ?CNews base:tags ?tags .
  ?CNews base:aboutEvent ?aboutEvent .
  ?CNews base:fromArea ?fromArea.
  ?CNews base:aboutEvent base:World_War_II}

```

In semantic web there is always data hiding behind data, and we can't extract this data through a simple queries, so this data need a specific operation to be available for a simple sparql query this operation called entailment.

Search operation may need to entailment before execute it,

such as searching for news of a continent or a specific country, include all news of its cities, to make this possible, we need to entailment at first to generate the lost triples which feed the search operations:

```

PREFIX base: <http://www.owl-ontologies.com/news.owl#>
CONSTRUCT {?news base:fromArea ?country.
            ?news base:fromArea ?region.}
WHERE {
  ?news base:fromArea ?city.
  ?city base:isCityOf ?country.
  ?country base:isCountryOf ?region
}

```

This statement will generate a new triples, as follow :

```

base:Obama_vs_Romney base:fromArea base:USA
base:Obama_vs_Romney base:fromArea base:America.

```

This triples can feed the graph, and the some search operations can't be done without it, thus we can search for news in the city New York throw searching on news in the country USA or in the continent America.

In the same way, we can search on Football news throw searching on Sport news.

### III. Implementation

We will see how we can implement a simple system uses a semantic web technology.

This section summarizes all the general requirements of the system. this implementation of the study supposed to be used in the project news site, and has been to identify the specific requirements, depending on potential users. Applicable program is supposed to be a web site on the internet complies with a set of requirements.

This news site contains data and regular newsletter structured in an efficient manner, using Semantic Web technologies, and so that the system can meet the functional requirements of the data newsletter, and be flexible enough to meet all the new requirements. Most of these requirements are centered on the publishing data on the Web and retrieve this data, in addition to that the system must be aware extensively to understand the data news in a way so that data returned accurate and true through any queries as possible, also the system must be capable of sharing data with other applications.

#### Functional requirements:

- I must be able to retrieve data news from a certain area, regardless of the city or country or region.
- I must be able to retrieve information newsletter of a specific news category, such as sports, science or business or from a more personalized news categories such as categories that fall under the category Sports such as Football, tennis and cricket.
- I must be able to retrieve data for a particular event news (such as World Cup soccer, the Iraq war).
- The system should be able to retrieve all the news related to a specific region, news related to a particular category.
- The system should be able to publish data on the Web and share data with other applications.
- The system should be flexible enough to accept new news categories.
- Potential Users can write news data within the system if

they have permission, and visitors or anonymous users can browse news for a certain category or a certain area or by some keyword.

**System Layers:**

Three layers have been proposed for the implementation of this application of the Semantic Web, these layers of the system are shown in Figure 2:

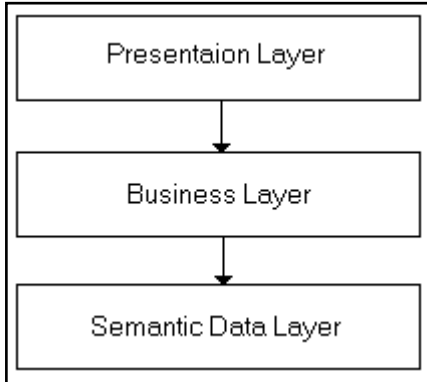


Fig. 2: System Layer

Presentation Layer is an ASP.NET Web pages created within the Microsoft Visual Studio environment and this layer does not deal with semantic data layer directly, but rather through other layer is a layer of operations (Business layer) which in turn deal with the data operations directly such as add, modify, delete and processing. Each terminal will interact only with a business layer, this layer will be responsible for receiving all requests from the terminals and sent it to a semantic data layer, then a business layer send results coming from semantic data layer to the presentation layer (Interface). The business layer is based on the dotnetRDF Library, this library is an open source provide with set of easy and highly effective functions in dealing with RDF under the .Net platform written in C #.

Semantic data layer consists of ontology, which in turn are divided into two section Classes and properties on the one hand and the data on the other hand, all the components of a business layer interact with a semantic data layer.

Data physically stored in the file on the hard disk, rather than memory, the benefit is that the data do not delete when we restart computer.

Some of the classes and methods which defined in business layer are described as follow:

Class	Description
RdfBLL	Convert the Sparql Select result to DataTable with the same column name and type as variables in the Sparql
QueryBLL	This class is responsible for the implementation of the SPARQL query and retrieve results in standard format, so it provides us with methods similar to Sparql query instructions such as Select, Describe, Construct and Ask

CNews	This class will be responsible for all the process operations related to News, such as retrieve news by specific area or category , in addition to insert, update and delete news
NewsInference	This class will be responsible for entailment operation, by instruction Construct

**IV. Testing & Results**

The test of all business component have been done, and all methods return results as expected, here I will show some example about these methods, so I choose the CNews, CUser, QueryBLL and NewsInference classes which contain the most functions, so let's see what methods of every class of these classes have? And how they complete each other.

Let's suppose that some user want to insert, update or delete some news data, these three operation are encapsulating by three methods in the class CNews, which are CNew.Insert, CNews.Update and CNews.Delete.

Before updating data, user needs to browsing data, to see the data that he wants to update.

To represent data from Rdf, and show it as table, I've created two methods in the class RdfBLL, the first is CreateDataTableAs(SparqlResultSet) which create a DataTable according to the Sparql Result passed to it in the parameter.

The second method is FillDataTableFrom(SparqlResultSet) which call the previous method and fill the Data Table with data retrieved by the Sparql Result in the parameter.

User must be authenticated to do any modifying operation on data except browsing data, can be done by anonymous user (visitors).

To ensure the user is authenticated, I have created a method ExecuteAsk(Query) in the class QueryBLL, this method ensure that the user is exist , by checking the triples which contain username and password exist or not, if exist then the user is authenticated and consequently can login.

After ensuring the user is authenticated, he need to be authorized too for doing some processing on data, like insert , update and delete , so the user needs to have a permission to do this operation.

The method IsUserHaveRight(String User, String Right) in the class CUser check if the user have the rights like the right insert, update or delete on news.

We must always keep the integrity of data, for example, before delete any triple in the rdf file, it must ensure that the deleted triple not related to other triple, we can do this by the method HasDetails(String Main) in the class QueryBLL, which check if the current deleted triple has any detail data or related triple to prevent deletion and send a warning message to the user.

When users want to do some search, for example , show me all the news related to America, the result in the related work such as Semantic Web Solutions by Hafiz Hammad Rubbani show only the data associated with America continent, while our system before doing the search operations it does some entailment operations on data, which in its turn add a new triples.

This triples doesn't change the meaning of content, but on the contrary it support the meaning by making the data more clearly

This new triples is very useful for the search operation, because it make the query statement very simple, for example if we want

to show all the news related to America, the system must retrieve all the news about America or the news about any country of America, or the news about any city of America, also if the users want show all the news related to sports category the system must show all the news which fall under the sports category such as Football, tennis...etc.

This search can't be done without entailment operation, which add a new triples to the graph to feed the search operation, the method responsible to execute the entailment is `InferByConstruct()` in the class `NewsInference`, which depend on the properties `IsCityOf`, `IsCountryOf` and `Subclassof` to do the entailment.

## V. Conclusion and future work

We showed in this paper how to do a simple content management in semantic web, using Microsoft Visual Studio, C#, ASP.Net and RdfdotNet Library, and we have applied the general features of the content management system, in addition to entailment operation, which feed the search operations, finally we represent example with classes and its used methods.

This application is expected to be extended in future and a new features will be added, such as put the entailment rules in the external file to be more flexible instead being static in the code.

## References

- [1]. *Integrating Semantic Web Features into Open-Source CMS Solution.*
- [2]. *Introduction to the Semantic Web International Conference on Dublin Core and Metadata Applications, Singapore, 2007-08-31* Ivan Herman, W3C.
- [3]. *Managing Open Source Software Development Projects with Semantic Web Technology – Wikan Danar Sunindyo* Thomas Moser<sup>1</sup> Stefan Biffli.
- [4]. *Metadata and Semantics* Miguel-Angel Sicilia University of Alcalá Madrid, Spain Miltiadis D. Lytras University of Patras Patras, Greece.
- [5]. *Modeling Rules on the semantic web* UNIVERSITY OF BELGRADE FACULTY OF ORGANIZATIONAL SCIENCES MILAN V. MILANOVIĆ.
- [6]. *Programming the Semantic Web* Toby Segaran, Colin Evans, and Jamie Taylor.
- [7]. *SCMS - Semantifying Content Management* Axel-Cyrille Ngonga Ngomo<sup>1</sup>, Norman Heino<sup>1</sup>, Klaus Lyko<sup>1</sup>, René Speck<sup>1</sup>, Martin Kaltenböck<sup>2</sup> <sup>1</sup> University of Leipzig AKSW Group Johannsgasse 26, 04103 Leipzig <sup>2</sup> Semantic Web Company Lerchenfeldergürtel 43 A-1160 Vienna.
- [8]. *ISSN 0103-9741 Monografias em Ciência da Computação n° 08/07 A Semantic Web Application Framework* Leonardo Magela Cunha Departamento de Informática
- [9]. *Semantic Web Solutions* By Hafiz Hammad Rubbani (hammad@itu.dk) Supervisors Henning Niss Thomas Hildebrandt
- [10]. *Semantic Web-Based Content Management System* Jiri Kopecký Supervisor: Ing. Kamil Matoušek, Ph.D.
- [11]. *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management* by Michael C. Daconta, Leo J. Obrst and Kevin T. Smith ISBN:0471432571 John Wiley & Sons © 2003 (281 pages)
- [12]. *Learning SPARQL*- Bob DuCharme.

## Authors' Profile



*Bassel AlKhatib is the web science master director at the Syrian Virtual University and the head of Artificial Intelligence department at Information Technology Faculty at Damascus University. He holds PhD degree in computer science from the University of Bordeaux-France, 1993. Dr. Alkhatib supervises many PhD students in web mining, and knowledge management. He also leads and teaches modules at both BSc and MSc levels in computer science and web engineering in both Syrian Virtual University and Damascus University*



*Mohammed Eyad Fakhane: Web Science master student. Bachelor of information technology degree, SVU 2011.*