# Semantic Adaptation Approach for Adaptive Web-Based Systems

Bujar Raufi, Artan Luma, Xhemal Zenuni, Florije Ismaili

Faculty of Contemporary Sciences and Technologies,
South East European University
Ilindenska nn, 1200, Tetovo, Macedonia
{b.raufi, a.luma, xh.zenuni, f.ismaili}@seeu.edu.mk

**Abstract:**The principles of the WWW used today rely solely on links and on content presented through simple HTML tags. The drawback of the above mentioned modus operandi is that it is purely syntactic with no clear emphasis on meaning about data. The Semantic Web approach tries to bridge that gap by providing additional information, mainly conceptual to the data at hand. In this paper we are focusing on extension of the adaptation methods with semantic adaptation dedicated to Adaptive Web-Based Systems. The approach is consisted of five steps based on querying RDF resources through SPARQL every time the user is visiting a particular resource. It is worth mentioning that the retrieved resources represent similar resources relevant to the currently visited resource which guides the user towards the desired information while browsing.

**Key words:** Semantic adaptation, Web based system.

## Introduction

The expansion of the web in the latest years consecutively with the increase of new applications and social web phenomena has made it a place where information is not simply posted, searched, browsed and read, but also has made it a place where it can be adapted in various ways for a single purpose of action, which is meeting user's needs. The second part of the above mentioned clause is being conducted through the use of *Adaptive Web-Based Systems*. Adaptive Web-Based Systems represent systems that tend to arrange its internal link structure, content or both based on user access patterns.
In (Raufi & Georgieva, 2008), a layered framework is introduced that concentrates on the issue of complexity in adaptive web-based systems. The proposed framework guarantees:

- **Flexibility:** that offers quicker discovery of new knowledge and the reuse of the existing.
- **Expressiveness:** this allows extensive usage of semantic web technologies in the sense of good manipulation and reasoning with the knowledge at hand.
- **Interoperability:** offers sharing and accessing of data from other resources for performing adaptation (closed or open corpus adaptation).
- **Modularity:** that allows a certain degree of independence between layers.

The proposed layered framework is consisted of five layers as illustrated in (Raufi & Georgieva, 2008):

- ***Data Layer*** where all the data and the site's internal link structure resides. Also this is the layer where the atomic information is located with the precedence and next links as well as their particular information weights.
- ***Concept Layer*** represents the knowledge representation layer of the system, consisted of concepts and concept relationship. This layer represents the semantic layer of the proposed framework
- ***User Layer*** where user preferences like access patterns and user behaviour are collected and used for performing the adaptation process.
- ***Adaptation Layer*** which performs the adaptation based on knowledge gained from concept layer and user layer.
- Finally the ***Presentation Layer*** is what the user sees at the end as a final product of adaptation. In this stage, a rearrangement of atomic pieces of information or whole pages is done so the adaptation goal posed in adaptation layer can be met.

In (Raufi & Georgieva, 2010) we have introduced the work done in each layer of the above proposed framework. In this paper we are focusing on semantic adaptation aspects for the framework and approaches proposed in (Raufi, Georgieva, Luma, Ismaili & Zenuni, 2010) and (Raufi, Georgieva, Luma, Ismaili & Zenuni, 2010). The rest of the paper is structured as follows: In section 2, a semantic web technology available for adaptation is introduced. Section 3 outlines an approach towards adaptation which utilizes the Semantic Web, whilst section 4 illustrates an example of semantic adaptation with an included case study and finally section 5 concludes this paper with some future direction towards semantic adaptation approaches.

## Semantic Web Technologies and Adaptation

The advent of *World Wide Web (WWW)* has substantially altered the understanding on how people communicate with each other. This change transformed the world and smoothly pushes the web toward knowledge based society. The above mentioned transform also changes the way how people understand computers. While, at the beginning, the computers were used for performing solely numerical calculations, today their main goal is given to data processing even though the latter is not intended to be human understandable. Data mining is a typical example for this. However, most of the web content today

is used for human consumption, disregarding the aspect of what the data conceptually represents in its essence. This disadvantage makes the content usable only for browsing, online ordering, filling forms or searching.
The reason of using Semantic Web techniques for Adaptive Web-Based System can be outlined as follows:

1) Semantic web can be used to describe every document in adaptive web-based system's repository (documents and any other smaller units) according to a certain vocabulary.
2) After their description, these documents can become machine processable and conceptually describable.
3) Numbers of these described documents are scaling optimally with no particular increase in processing power. Such examples have been detected in OpenLink Virtuoso, BigOWLIM, Garlik 4store and many more (W3C Media Wiki - Large Triple Store., 2010)
4) These described documents can be queried through a particular endpoint and presented to the user.

Knowledge representation through semantic web uses many underlying methods and technologies starting from those used for describing resources on the web (like RDF and RDF Schema), representing knowledge and knowledge flow (like OWL) up to use of rules (like Rule markup languages). All the above mentioned technologies are meant to make the web more machine-processable, which is also a requirement for building adaptive web-based systems.
One promising way of making content for Adaptive Web-Based Systems more machine-processable as well as human readable is the use of method for describing resources on the web with utilization of RDF (Resource Description Framework) (Becket & McBride, 2004) which is more like a data model rather than a language and is based on *object-attribute-value* (also referred to as *resource-property-value*) triples called *statements*.

- ***Resource:*** Represents every object or thing that is describable. Every resource has its own URI (Unified Resource Identifier) that uses the underlying web's standard URL naming to gain access to the resource itself. Even though RDF does not require a strict accessible resource, however the Linked Open Data Initiative (Heath, 2010) that deals with linking web resources highly recommends such an action. Our approach in Semantic Adaptation follows the Linked Open Data recommendation.
- ***Property:*** Represents also a special kind of resources that describe relations among resources (objects). For example, *writtenBy*, *age*, *visited* etc. are typical properties. The same as objects, properties are also represented by URI's
- ***Value:*** Is the actual value of the object itself, such as page name given in the above statement. It is worth mentioning that property value can also be another resource or a simple literal.

In (Raufi, 2011) we propose an architectural model for addressing the complexity of Adaptive Web-Based systems by dividing it into two logical modes: the *offline* and *online* mode.
The *offline mode* orchestrates computer exhaustive tasks such as data mining within web site's repository (clustering and association rule mining), navigation pattern mining and session reconstruction within web server's log files. The reason for performing such tasks offline is the intensive calculation nature of the above mentioned functions considering that dealing with online web application with many requests and clicks during user visits would render the system inefficient. Therefore the offline mode utilizes the above mentioned techniques on the background, issued by the webmaster which performs the overhead tasks, analyzes them and the extracted knowledge is incorporated in the overall adaptation process. The incorporation involves putting additional information on the adaptive presentation panel besides the information retrieved by adaptation algorithm. Based on the above mentioned claim, the offline mode does not obstruct the overall functioning and performance of the system because it is being conducted on the background as a separated task from that of the adaptive web application.
*Online* mode on the other hand considers the adaptation "on-the-fly" presented through the adaptation algorithm elaborated extensively in (Raufi, Georgieva, Luma, Ismaili & Zenuni, 2010). The online mode traces the user visits through the links it clicks and presents adaptive content based on the visited information units (documents or atomic units). The characteristic of the online mode is that it offers two types of adaptive content. The first one is adaptive content delivery through the adaptation algorithm which was based on web site's link structure and document similarity measures and the second one is adaptive content delivery by querying the knowledge repository as presented in figure 1 which in fact is the concept layer of our framework. Querying is conducted on the generated RDF triples from Relational database, represented through RDF Repository. For this purpose SPARQL (Prud'hommeaux & Seaborne, 2008) query language was utilized for querying the knowledge base. The results of SPARQL queries can be presented as result sets, RDF graphs or as RDF triples presented for human consumption through web pages (Adida & Birbeck, 2008), (Prud'hommeaux & Seaborne, 2008).

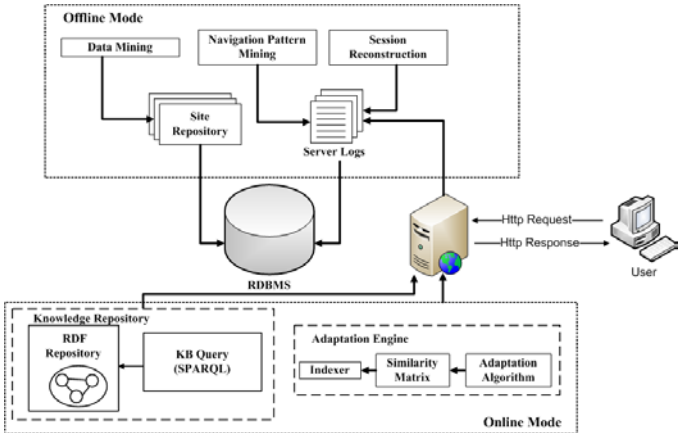


**Figure 1 Architectural Model for Adaptive Web-Based Systems**

Main goal of this paper is to explore the possibilities of utilizing the *Knowledge repository* as given in figure 1, consisted of *RDF Repository* and *Knowledgebase Query (KB)*, for performing RDF-Resource adaptation for users. The mechanism for querying the RDF triples mentioned in the above architecture will be elaborated in section 3.

## Semantic Adaptation Approach for Adaptive Web-Based Systems

Many systems regarding adaptation for semantic web have emerged the recent years like (Mikroyannidis & Theodoulidis, 2007) and (Partarakis, Doulgeraki, Leonidis, Antona & Stephanidis, 2009). However, the approach mentioned on the systems above focus mostly on case of web usage and text mining methodologies for semi-automatic construction and evolution of Web ontologies characteristic mostly for the former and on user interface adaptation which is characteristic mainly for the latter approach.

Developing particular applications for adaptive content presentation on Semantic Web requires a set of particular skills that need to be mastered. The first skill that needs to be considered is the aspect of data representation as recommended from Open Linked Data (Heath, 2010) initiative, which is consisted of four basic rules which are (Berners-Lee, 2006):

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URI's so that they can discover more things.

The second important skill, which is logically derived from rule 3 and 4 given above, is extracting useful information from the visited links and recommending to the users as sources for discovering new resources represented through URI's. This can be achieved in two methods:

- The first method is by simply following the typed links that Semantic Web offers, i.e. each of these links have a clearly defined semantics which allows us to precisely know which links to present to users. For example, `rdfs:seeAlso` and `owl:sameAs` clearly defined semantics and by simply using *Depth-First* or *Breadth-First* search strategies to determine which links should be offered.
- The second method, which will be followed in our approach, is by using a query language SPARQL (Prud'hommeaux & Seaborne, 2008) for querying closed or open corpus repositories in relation to link (resource) that user have already visited.

A typical SPARQL query that can retrieve everything about a particular concept can be written as:

```
SELECT ?propertyValue ?propertyName
WHERE {
      <Resource>
      ?propertyName ?propertyValue.
}
```

Where `<resource>` represent a particular dereferencable URI. Having all these aspects in mind, the proposed semantic adaptation algorithm is consisted of steps given as:

**Step 1:** user visits (clicks) a particular resource while browsing.

**Step 2:** execute a SPARQL query for every resource the user is visiting.

**Step 3:** present similar resources to the one that user have currently visited.

**Step 4:** present the resources to the user in user friendly way

**Step 5:** repeat steps 2-4 every time the user is visiting a resource

The whole semantic adaptation process goes by visiting and providing resources by querying RDF internal or external repositories. It is worth mentioning that resources can be described by generally accepted vocabularies like SKOS (Miles & Bechhofer, 2008), FOAF (Miller & Brickley, 2010) or more general adaptive web-based system ontology (Raufi, Ismaili & Zenuni, 2009). The above mentioned five step algorithm is going to be illustrated with a simple case study outlined in the following section.

## Semantic Adaptation: A Case Study

Let us suppose that a particular RDF resource is being requested over the semantic web application and we are requesting some information about person from DBPedia <http://dbpedia.org/resource/Tim_Berners-Lee>. Executing SPARQL query that will achieve the above mentioned steps between 2 and 4 can be done by utilizing properties from OWL language. One of the properties is the `owl:sameAs` property. The overall query will look like:

```
SELECT ?aliasURI WHERE {
   {
      <http://dbpedia.org/resource/ Tim_Berners-Lee>
      <http://www.w3.org/2002/07/owl#sameAs> ?aliasURI.
   }
union
   {
      ?aliasURI <http://www.w3.org/2002/07/owl#sameAs>
      <http://dbpedia.org/resource/Tim_Berners-Lee>.
   }
}
```

After the execution of the query, the retrieved resources will be outlined as follows:

```
<http://www4.wiwiss.fu-berlin.de/flickrwrappr/photos/Tim_Berners-Lee>
<http://mpii.de/yago/resource/Tim_Berners-Lee>
<http://www4.wiwiss.fu-berlin.de/dblp/resource/person/100007>
<http://rdf.freebase.com/ns/m/07d5b>
```

It is worth mentioning that every time the user is visiting a particular resource, similar resources will be presented to the user as part of the above mentioned algorithm. The sole aspect that needs to be considered is the large number of retrieved resources which often can be enormous; therefore a proper filtering of such resources may be required.

## Conclusion and Future Work

In this paper we have introduced a semantic approach towards adapting the content in the conceptual aspect by utilizing semantic web. We have introduced an approach towards adaptation through querying the RDF resources from internal or external repositories. The proposed semantic adaptation approach delivers similar and relevant resources every time the user clicks an RDF resource. With this algorithm, constant and up to date information is being presented to the user.
The future work on presenting adaptive on semantic level would involve:

- Limiting and filtering the number of resource according to their relevance and
- Presenting the content of resources in a more user friendly manner. Fortunately, there is a promising W3C standard technology called RDFa which is designed to present such resources in a manner which is more “human-consumable” (Adida & Birbeck, 2008).

The above presented future guidelines would ensure minimum robustness to the system considering that resources presented will be limited and filtered based on some criteria as well as it will provide readability of resources by presenting them in user friendly trait.

## References

Adida, B, & Birbeck, M (2008). RDFa primer: Bridging the human and data webs. Retrieved October 2008, from http://www.w3.org/TR/xhtml-rdfa-primer/

Becket, D, & McBride, D (2004). RDF/XML syntax specification (revised). Retrieved 2008, from http://www.w3.org/TR/REC-rdf-syntax/

Berners-Lee, T (2006). Linked Data. Retrieved 2011, from http://www.w3.org/DesignIssues/LinkedData.html

Heath, T (2010). Linked open data project. Retrieved 2010, from http://linkeddata.org/

Mikroyannidis, A, & Theodoulidis, B (2007). Heraclitus: A Framework for Semantic Web Adaptation. *Internet Computing, IEEE* , 11(3), 45-52

Miles, A, & Bechhofer, S (2008). SKOS Simple Knowledge Organization System RDF Schema. Retrieved 2011, from http://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html

Miller, L, & Brickley, D (2010). FOAF Vocabulary Specification 0.98. Retrieved 2011, from http://xmlns.com/foaf/spec/

Partarakis, A, Doulgeraki, C, Leonidis, A, Antona, M, & Stephanidis, C (2009). User Interface Adaptation of Web-Based Services on the Semantic Web. *Universal Access in HCI*, 711-719.

Prud'hommeaux, E, & Seaborne A. (2008). SPARQL query language for RDF. Retrieved January 2008, from http://www.w3.org/TR/rdf-sparql-query/

Raufi, B (2011). System Implementation and Adaptation Evaluation in Adaptive Web-Based Systems. *Procedding of 12 International Conference on Computer Systems and Technologies - CompSysTech'11*,

Raufi, B, & Georgieva, J (2008). Adaptive-Web Based Systems: Current Trends and Open Issues. *Proceedings of the 4th International Bulgarian-Greek Conference on Computer Science*, 3, 1004-1009.

Raufi, B, & Georgieva, J (2010). Adaptive Web-Based Systems: From Framework to Implementation. *International Conference on Computer Systems and Technologies CompSysTech'10*.

Raufi, B, Georgieva, J, Luma, A, Ismaili, F, & Zenuni, Xh (2010). An Adaptation Algorithm for Adaptive Web Based Systems based on Link Structure and Document Similarity. *9th WSEAS International Conference on Telecommunications and Informatıcs (TELE-INFO ´10)*, 29-34.

Raufi, B, Georgieva, J, Luma, A, Ismaili, F, & Zenuni, Xh (2010). Generating Content Recommendation for Adaptive Web-Based Systems. *COMPUTERS and SIMULATION in MODERN SCIENCE*, 3,

Raufi, B, Ismaili, F, & Zenuni, Xh (2009). Modeling a Complete Ontology for Adaptive Web Based Systems Using a Top-Down Five Layer Framework. *Proceedings of 31st International conference on Information Technology Interfaces*, 503-510.

W3C Media Wiki - Large Triple Store. (2010). Retrieved May, 2010, from http://esw.w3.org/