

# A Framework for Establishing a Semantic Web in a University Website: A Case Study University of Gezira

Gais Alhadi Babikir<sup>1</sup>, Murtada Khalfallah Elbashir<sup>1</sup>, Mohammed A. Saleh<sup>2</sup>

<sup>1</sup>Computer Science Department, FMCS, University of Gezira, Sudan

<sup>2</sup>Department, College of Science and Arts in Ar Rass, Qassim University, Kingdom of Saudi Arabia (KSA)

## Abstract

*The advent of technology has yielded an escalating demand for managing data, information and knowledge effectively. Apparently, the current websites represent the biggest global database that lacks the existence of a proper semantic structure. Therefore, it makes it difficult for the machine to understand the information provided by the users. Unfortunately, this shortcoming is inherited in many universities websites. Indeed, some of these websites place the information incorrectly, untidy, and incomplete, which leads to a delay in ranking the University. The aim of this paper is to design a framework for establishing a Semantic Web in a University Website. This can be accomplished by organizing the information in conceptual spaces according to its meaning by using Semantic Web and Ontology. The main objective of applying the concept of the Semantic Web is to enhance the ranking of the university.*

## 1. Introduction

In this paper, we deal with the Semantic Web technologies to enhance the ranking of the university. Clearly, there is so much data on the web that managing it with conventional tools is becoming almost impossible. So we need new tools and techniques to manage this data effectively. Therefore, to provide interoperability as well as storage between the multiple data sources and distributed systems, and to extract information from the database and warehouses on the web, various tools and techniques are being developed. Consequently, the web is evolving into what is now called the Semantic Web [1]. In recent years, University Web rankings have become very importance around the world. Noteworthy, the hypothesis of the university ranking is that the university's web presence reflects its global performance, the quality of its departments, and its impact on its outputs and international prestige. One of the most important dimensions in ranking a University Website is clarity factor and quality of presentation. The dimension includes qualitative and quantitative criteria. One of the most important dimensions of the web quality is presentation and visibility. So should be placed on the visual form in

the University Website, in terms of a more usable layout and pleasant graphics [2][3]. However, when we evaluate a University Websites the quality, visibility, and quality of presentation must be considered [4]. Also, we have to keep in mind that some of the criteria are depending on the context of the Website.

The contribution of this paper attempts to design a framework for establishing a semantic web in a university website by using some quantitative criteria and Semantic Web. The Semantic Web is not a separate set of new and distinct websites, is an extension of the current World Wide Web (WWW). It builds on the current WWW constructs and topology. Current websites use these constructs to describe information within web pages in ways more readily accessible by outside processes such as search engines.

## 2. Universities ranking

Millions of people visit university web portals looking for information. This could be, for example, students looking for courses description, change in lecture times, lectures tables, laboratory tables, or teacher contact, etc. It is very important that whatever it is the user is searching for is easy to find and the content is easily understood. The importance of university portals opened a new field in website evaluation. Shanghai University published the results of the first global university ranking in 2003. In 2004 a new term was launched called Webometric.

Webometrics ranks universities based on a composite index that considers both the volume of web content (such as the number of web pages and files) and the appearance and impact of these web posts according to the number of external links. In fact, the university ranking system measures the strength of a university's presence on the web through its web domain, sub-pages, rich files, scholarly articles, etc.

### 3. Semantic web

Semantic Web is information that is related to each other in a way that can be easily processed by machines on a global scale. [5] Described the semantic web as an extension of the current web where the information is given a clear meaning. It should be noted that the difference between the semantic web and the present web lies in the representation of information on the web. In the future, the semantic web allows machines to process and infer information at the semantic level. Clearly, the adoption of the semantic web will enable us to store knowledge about web content in an organized manner [6].

The semantic web, also known as Web 3.0 [7], has the effort to enhance the current web so that computers can process the information presented by the clients, interpret and connect it, to help clients to find the required knowledge. The semantic web is intended to form a distributed knowledge-based system. Therefore, the new websites are based on a more sophisticated semantic representation of information, for exceeding the document level, and to understanding and processing by machine [8].

### 4. University ranking and semantic web

University ranking has become an important component of most search engines to the priorities of search results and to provide the clients with an immediate list of the most relevant results for their query. As more RDF data appear online, the semantic web search engine enables querying large RDF datasets, which requires ranking functionality to present results to the clients in a meaningful way i.e. by prioritizing relevant, important results, which helps to improve the assessment of the website.

### 5. Semantic web technologies

Figure 1 illustrates the semantic web is layers. On the bottom, there is the term identification, which is commonly represented by a URI or an Internationalized Resource Identifier (IRI). Thereon, syntax, and structure have to be defined. Therefore, XML, XMLS, and RDF are commonly used technologies.

Ontologies are the core concept of Knowledge Representation in the Semantic Web. RDFS and OWL are languages that can produce such models. The logic layer, which enables software agents to reason about the represented data, is composed through ontologies, queries, and rules [9].

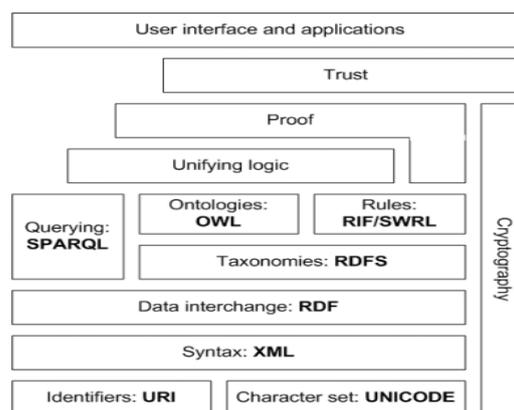


Figure 1. An abstract version of the semantic web layer

The term ontology [10] is clearly used in several disciplines, from philosophy to knowledge engineering, where ontology consists of concepts, properties, relationships between these concepts, and constraints. Ontology is a specification of a representational vocabulary for a website, definitions of classes, relations, functions, constraints, and other objects. Ontologies are used in order to support interoperability and common understanding between the clients and the computer.

The purpose of an ontology is to reach a common and shared knowledge that can be transmitted between users and application systems. Indeed, ontologies play an important role in achieving interoperability across organizations and on the semantic web. Therefore, it is necessary to establish schemas between the various concepts of ontologies [11]. It should be noted that the creation of such correspondence is not an easy task [12]. In fact, the lifecycle of ontologies across the semantic web includes different techniques, starting from manual to an automatic building, merging, mapping, annotation, etc. Each technique involves the specification of core concepts of an ontology [13] [14].

### 6. Methodology

Ontology is the spine of a semantic web system, since it models a domain throughout modelling domain ontology. It is the process that transforms a web application into a semantic web application that makes machine-understandable. Worth mentioning, the proposed model was inspired from [15]. As we can see in the proposed model in Figure 2, the two major activities: Web Pages Generation Track and Web Ontology Generation Track are carried out in parallel. Next, analysis, design, implementation and integration will be performed to produce a final product understandable to the machine as well as a human being.

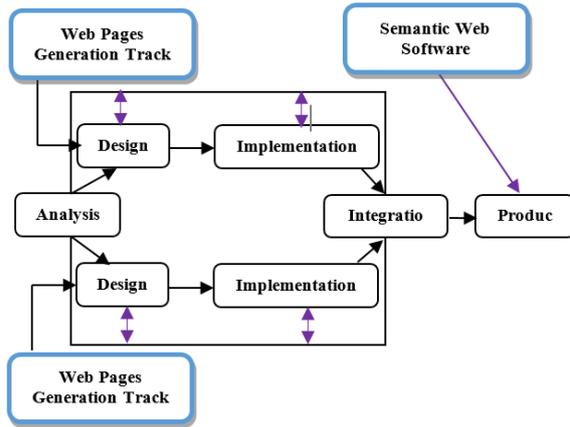


Figure 2. The architecture of proposed model

### 6.1. Analysis phase

The analysis phase defines the new system requirements, without reference to requirements accomplishment. Clearly, the activities of this phase described as follows.

#### 6.1.1. Requirement determination

The target humans are identified and grouped into classes, as shown in Figure 3. For each class, which has the same functional requirements, the usability requirements have been applied.

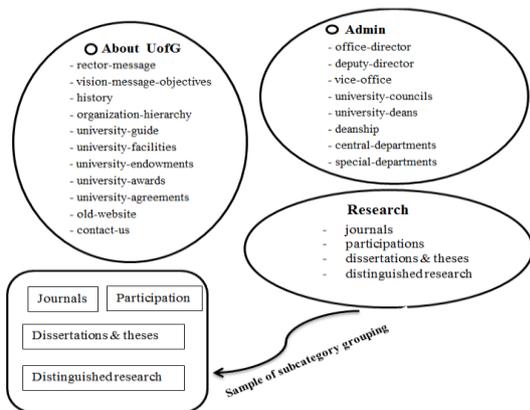


Figure 3. The site map of a university website

#### 6.1.2. Knowledge acquisition

It is the fundamental activity and prerequisite for the ontology construction track. In fact, many relevant sources of knowledge are collected and analyzed, such as university councils, university faculties, distinguished research, and etc., as demonstrated in Figure 4.

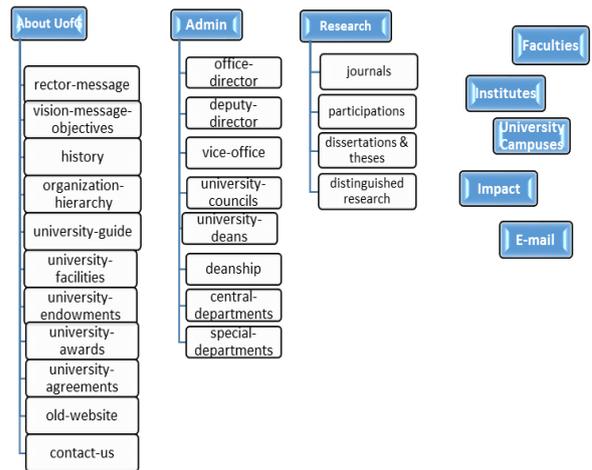


Figure 4. Sketch of a navigation scheme created from a site map of a university website

#### 6.1.3. Requirements modelling

The results of the first activity (i.e., requirements determination) will be rewritten using stripy vocabulary in a formal manner according to the modelling standards, as depicted in Figure 5.

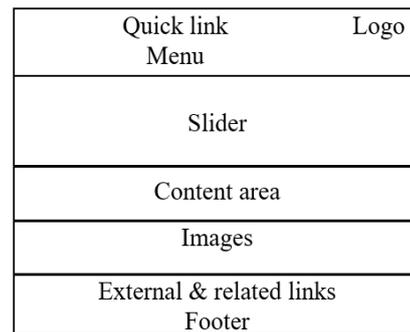


Figure 5. Wireframe of a university website

#### 6.1.4. Formal specification

Here, we organize the results of knowledge acquisition into classes using a controlled vocabulary. As an example, Figure 6 shows ontology for the classes (Faculties, Departments, Staff, Students, Persons, Programs, Subjects and Grade points) in University Website, and Figure 7 presents the taxonomy of persons within a University Website.

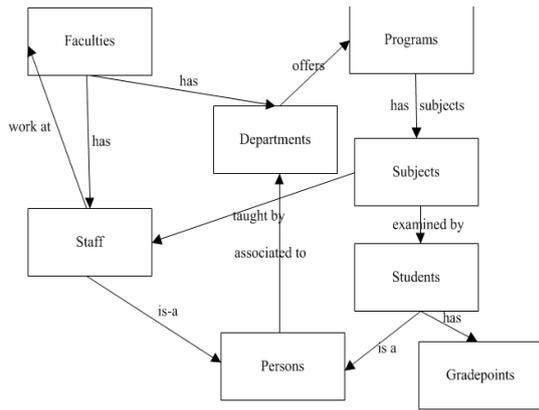


Figure 6. Ontology for classes (Faculties, Departments, Staff, Students, Persons, Programs, Subjects and Grade points)

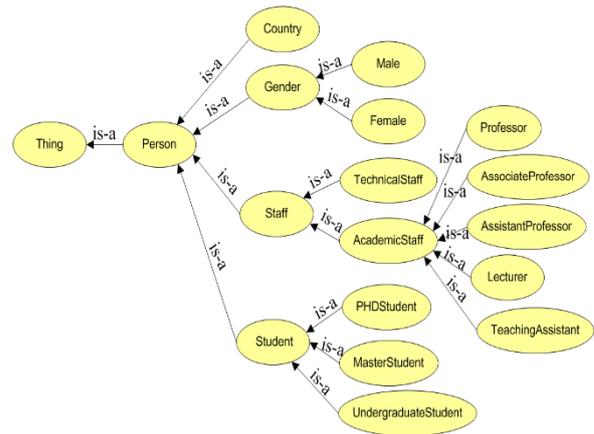


Figure 8. OWL visualization of sub-classes ‘Staff’ in the person ontology within the university website

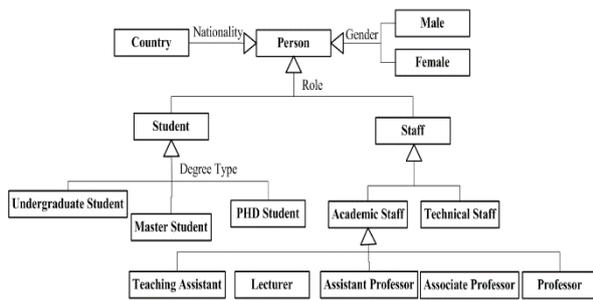


Figure 7. Taxonomy of persons within a university website

### 6.2. Design phase

The design phase defines the requirements of the semantic web application. Clearly, there are two types of requirements, the first one is to make the content understandable to the human, and the second type is to make the content understandable to the machine. Therefore, the classes are classified according to domains, and subsequently an ontology diagram is generated for each one of them.

It should be noted that there are various tools that allows you to graphically create ontology document (quickly and easily) in OWL and RDFS, such as Protégé that was applied in this paper. In addition, there are many tools that allow developers to develop web pages quickly and easily such as Microsoft Visual Studio.

Figure 8 shows the OWL visualization of a section of the person ontology within the university website which covers the sub-class ‘Staff’. An instance of any subclass such as Academic Staff and Technical Staff is created.

### 6.3. Creating website information architecture and content

This step aims to upgrade the design of the university website to cope up with the international standards for the universities websites, by organizing the information in conceptual spaces according to its meaning by using semantic web and ontology.

Obviously, we focus on creating effective site organization and navigation, also known as information architecture (IA) which refers to the structure and organization of the website. In fact, IA describes the ways in which different website pages link to each other and ensures that the information is organized in a consistent on each page. Through the process of developing an information architecture (IA), the site map of the contents of the university will be developed (See Figure 9).

Information architecture in this way helps the visitors to find key information quickly. It also will make the website more coherent and satisfying. Additionally, intuitively organized information architecture to ensure that all of the phases of the website development run smoothly and efficiently.

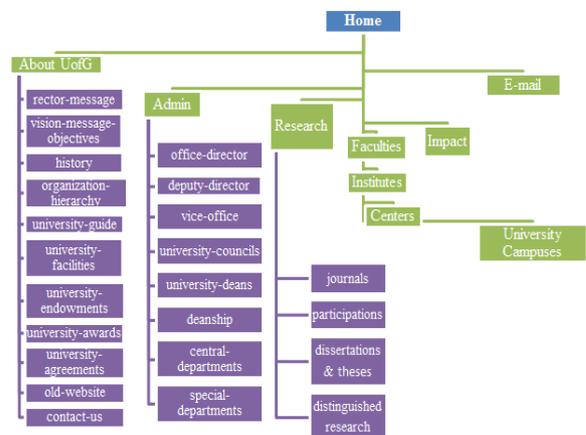


Figure 9. Site map of a university website

## 6.4. Create information architecture

### 6.4.1. Step 1. Defining key stakeholders

We need to clarify the key stakeholders and determining the purpose of the website. The key stakeholders are the people who visit university web portals looking for information regarding students, international student applicants, graduate students, alumni, faculty, academic support staff, all campus users, news media, and external visitors. The university website reflects its global performance, the quality of its departments and services, the impact of its outputs and its international prestige.

### 6.4.2. Step 2. Identifying user's goals and expectations

Students looking for course information, change in lecture times, laboratory tables, e-mail access communication with teacher...etc. Alumni looking for jobs, services, and communication with their colleagues...etc. Staff members are looking for e-mail, academic regulations, theses, and dissertations ...etc. The goal of effective Web design is to anticipate the visitor's needs. To accomplish this, we arranged and labelled information a way that the target audience is expected to see it.

### 6.4.3. Step 3. Defining sites content areas

Defining content areas help us to development navigational structure for the website. In universities, the content that reflects university is its global performance, the quality of its, includes information about the university, vice-chancellor message, vision, message, objectives, history of the university, organization hierarchy, university guide, university facilities, university endowments, university awards, university agreements, old website, contact information, director office, deputy director, vice office, university councils, university deans, deanship, central departments, special departments, university participation, faculties, departments, institutes, centers, students and their activities, alumni & related activities, staff and their services, research, dissertations & theses, journals, university media, and electronic services, university news, the impact of its outputs in community and its international prestige.

### 6.4.4. Step 4. Organizing the content areas

This step organizes the content areas compiled in step 3, into groups of similar or related topics. These groups will be given temporary names that will be later refined as navigation menu items. This activity will help us aggregate and categorize content areas so that navigation is easier for visitors.

### 6.4.5. Step 5. Creating the site map

Now, we create and validate the site map (a visual representation of the content areas). Figure 10 illustrates how to organize the website in a hierarchical way. In this type of structure, pages have a parent/child relationship. Not every page has a child, but all pages have a parent.

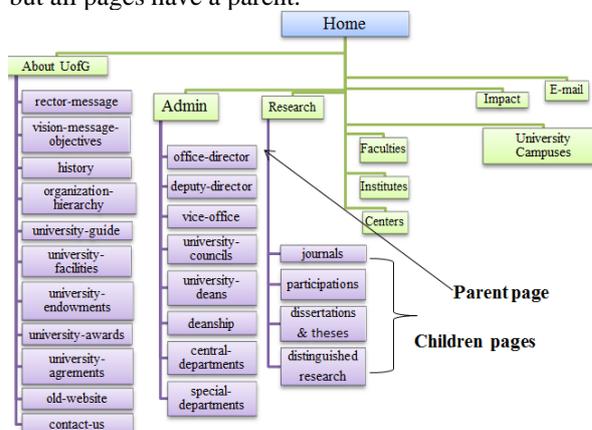


Figure 10. Children and Parent page in site map of a University Website

### 6.4.6. Step 6. Outlining the navigational structure

Here, we take the site map of the university website that we have created in step 5 and draw it to emulate the navigation scheme in figure 4. The subpages are listed under the main content area headings. It should be noted that site navigation should not refer to other sites, pdf files, Microsoft Office documents, or other non-HTML files. Indeed, doing this can be disorienting for the visitor of the website and can be problematic for those with slow connections, and this can reduce the visitors which affects the ranking of the website.

### 6.4.7. Step 7. Labelling the content areas

It is important to give accurate and meaningful labels to content areas. Visitors are clicking on words, so the words need to clearly make sense. In this website, the nomenclature has been tested by showing the navigation scheme outlined in Step 6 for the people how so interested. And they explained it's perfectly clear.

### 6.4.8. Step 8. Creating wireframes

A wireframe is a draw or schema that represents how page areas are organized (as shown in the previous figure 5). In fact, it is important to use wireframes as guides. This will help to ensure that information architecture is not inadvertently revised or obscured later during the design and development processes.

### 6.5. Information architecture standards

The purpose of information architecture is to create an organization for site content that will be as intuitive and easy for visitors to use as possible. To achieve this goal, we improved the structure of the site to make visitors find what they seek and predict what will happen when they click on an item.

#### 6.5.1. Naming conventions

A logical naming convention helps visitors to know where they are and how to return. So, we use related and easy words for the Web address, menu labels, and page headers. Figure 11 shows good consistency between the Web address, page header, and menu label. This naming convention is especially important for subpages. Not all visitors start from the home page, so subpages need to provide some context, too.

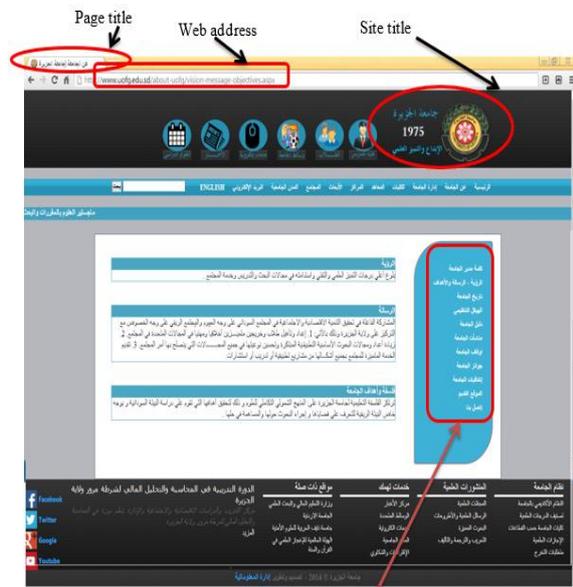


Figure 11. Example of consistent naming conventions

#### 6.5.2. Website title

One of the important things to make visitors know where they are, we set title prominently displayed at the top of every page on a website. In the case of University departments and offices, the site title most often is the name of your office or department. The site title usually is located in the same place on every page, as shown in the previous figure 11.

### 6.6. Improve the presentation of data

We dealt previously with several standards to improve the university website, but still, we have a problem. Clearly, the sheer amount of knowledge on online, alongside its distributed, redundant, and

inaccurate nature, makes the use of knowledge more difficult. So we will deal with this problem by applying a new technology called Semantic Web. The Semantic Web offers visitors the ability to work on shared meaningful knowledge representations on the web. And we will use the ontology of which facilitates the management, student's faculty members, and other stakeholders in the university.

### 6.7. Tools for ontology development

Protégé is the most popular and widely used tool for ontology development [16]. In Figure 12, we use this tool and Microsoft Visio to developing Ontology for the university website.

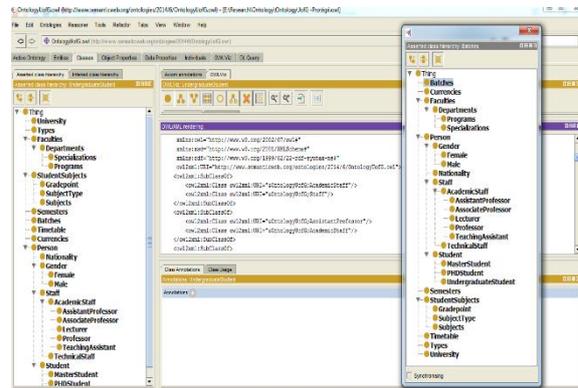


Figure 12. Implementation of University Website Ontology in Protégé

The university website has many faculties and each of these faculties has more than one department, each department offer programs and each program has many subjects. The University keeps track of information about persons either students or staff. All students study at least one subject. These subjects taught by staff and the students are evaluated using grades for these subjects. A grade report must be generated for each student that lists in the grade point table.

## 7. Results

This paper gives a detailed methodology to design the university website according to the Semantic Web or web 3.0 standards to enhance the quality of the information that will be presented on the Website. Microsoft Visual Studio 2017 is used to design the web pages, and the SQL server 2017 is used to host the database. Clearly, in this paper, we designed a framework for establishing a Semantic Web in a University Website by organizing the information in conceptual spaces according to its meaning.

## 8. Discussions

In the newly designed website, the information on the university website will be given well-defined meaning, and this will enable computers and people to work in co-operation since the vision of the Semantic Web as a natural extension of the World-Wide Web is to enable the machine to understand the human inputs. In this new Semantic Web, XML is no longer just the universal format for structured documents and data on the Web, but Ontology information that cannot be represented by XML Schema is carried by a new ontology language SWOL. The website is fully functional for as many visitors as possible across a wide range of browsers, devices, and operating systems, without negatively affecting the performance of the website. And the content is viewable, and page design degrades gracefully, depending on how well the visitor's browser adheres to common web standards. Noteworthy, the concepts used in this paper if applied properly can enhance the ranking of a university. The website that is designed according to web 3.0 standards will have a clear presentation of information and the information itself will be current.

## 9. Conclusion

In this paper, we proposed a framework for establishing a Semantic Web in a University Website. In fact, the information in the website has been organized in conceptual spaces according to its meaning by using Semantic Web and Ontology technologies. The University of Gezira website have been upgraded to Web3.0 standards and according to the universities website standards. We expect that the usability of the website will increase to its maximum level, and the site organization will be enhanced to the highest level. Also, the following metrics can be enhanced to their maximum level: contents, compatibility, using file naming rules, using folder naming rules.

## 10. References

- [1] Antoniou, G. and Van Harmelen, F. (2004), A semantic web primer, MIT press.
- [2] E. Pantano, S. Rocco, and V. Milena (2011), 'An exploratory study of the tourist-computer interaction: the role of web site usability on hotel quality perception', International Journal of Digital Content Technology and its Applications, pp. 208-216.
- [3] E. Pantano, and V. Corvello (2010), 'Digital contents management for improving consumers' experience', International Journal of Digital Content Technology and its Application, 4(7), pp.8-10.
- [4] A. Caro, C. Calero, and M.A. Moraga (2010), 'Are web visibility and data quality related concepts?', IEEE internet computing, 15(2), pp.43-49.
- [5] T. Berners-Lee, J. Hendler, and O. Lassila (2001), 'The semantic web', Scientific american, 284(5), pp.34-43.
- [6] R. Agrawal, and R. Srikant (1994), 'Fast algorithms for mining association rules'. In Proc. 20th int. conf. very large data bases, VLDB, vol. 1215, pp. 487-499.
- [7] S.K. Das (2007), Role of Semantic Web in the Changing Context of Digital Environment.
- [8] Colomb, R.M. (2007), Ontology and the semantic web, Vol. 156. IOS press.
- [9] S. Stroka (2005), 'Knowledge Representation Technologies in the Semantic Web', Information Technology and Systems Management, Salzburg University of Applied Sciences.
- [10] A. S. Mihoubi, and M. Simonet (2000), 'An Ontology Driven approach to Ontology Translation', In Proceedings of DEXA, pp.573-582.
- [11] M.M. Taye (2010), 'Understanding semantic web and ontologies: Theory and applications', arXiv preprint arXiv:1006.4567.
- [12] M. Paolucci, T. Kawamura, T.R. Payne, and K. Sycara (2002), 'Semantic matching of web services capabilities', In International semantic web conference, Springer, Berlin, Heidelberg, pp. 333-347.
- [13] A. Gangemi, R. Navigli, and P. Velardi (2003), 'The OntoWordNet Project: extension and axiomatization of conceptual relations in WordNet', In OTM Confederated International Conferences' On the Move to Meaningful Internet Systems', Springer, Berlin, Heidelberg, pp. 820-838.
- [14] A. Gangemi, C. Catenacci, and M. Battaglia (2004), 'Inflammation ontology design pattern: an exercise in building a core biomedical ontology with descriptions and situations', Studies in health technology and informatics, pp.64-80.
- [15] A. Farooq, and M.J. Arshad (2010), A process model for developing semantic web systems', New York science journal, 3(9), pp.43-39.
- [16] N. Malviya, N. Mishra, and S. Sahu (2011), 'Developing university ontology using protégé owl tool: Process and reasoning', International Journal of Scientific & Engineering Research, 2(9), pp.1-8.