Employing Semantic Web Technologies to Leverage Learning and Research

Friday Joseph Agbo
MSc Student, Department of Computer Science
University of Ilorin, Ilorin
Email: freddy4uagain [AT] hotmail.com

Kissinger Sunday
Computer Science Unit
Department of Mathematics
UsmanuDanfodio University, Sokoto

Abstract--- **The Web has always been a set of resources (such as web pages, files, etc) connected to each other by hypertext links which are untyped. By untyped we mean that the relationships between the linked resources are not easily discernible and are not necessarily machine-interpretable. Adding meaning (i.e. semantics) to these resources and their linkages will make them more machine-interpretable thereby making it easier to find relevant information from related social structures. In this paper, we present how semantic web technologies offer potentials for improving the interactivity of today's learning and research while putting students/teachers in control of their learning process spanning across disparate tools and services. This, we consider as a Better By Far (BBF) approach to research and general learning. Ontologies and ontology languages are the key concepts used in this study to provide additional meaning to these resources and their links. Also in this paper, popular ontology technological tools like Resource Description Framework (RDF) and Extensible Markup Language (XML) were used to discuss the implementation of Semantic Web (SemWeb).**

Keywords*--- *Semantic Web, Ontology, RDF, OWL, BBF, XML, Agent, Metadata

I. BACKGROUND TO THE STUDY

With numerous documents either in electronic form or printed materials littered around the world, one would think that learning and research about a particular existing area should be fast and easy. But this is not the case as harnessing these documents for effective utilization is needed to achieve that aim. It is with this idea of harnessing abundant available documents that Semantic Web (SemWeb) Model was formed in the early 1960s by the cognitive scientist Allan M. Collins, linguist M. Ross Quillian and psychologist Elizabeth F. Loftus as a form of representing semantically structured knowledge.

In this information technology age, many of the learning materials are in electronic form and reside in many servers around the globe. Accessing these materials sometimes may require some prior knowledge of where in particular to find them since they are readily scattered in different locations. The introduction of SemWeb model in the context of the modern internet extends the network of hyperlinked human-readable web pages by inserting machine-readable metadata about pages and how they are related to each other. This enables automated agents to access the Web more intelligently and perform more tasks on behalf of users. All information displayed on a web pages are structured by a Markup Language and they are useful to human beings alone. SemWeb introduces the concept that will enable learning of the web page contents and to make it machine-readable.

The application of SemWeb Technologies is considered to be a Better By Far (BBF) approach for enhancing learning and research especially among scholars.

A. Agents and Metadata

Agents are intelligent semantic technological system design to study, learn and extract knowledge from web pages, which have the ability of reasoning on this knowledge to deduce inferred knowledge. The agent would do this not by looking at pictures and reading descriptions like a person does, but by searching through **metadata** that clearly identify and define what the agent needs to know. Metadata are simply machine-readable data that describe other data. In the SemWeb, metadata are invisible as people read the page, but they're clearly visible to computers. Metadata can also allow more complex, focused Web searches with more accurate results

B. Semantic Technologies: Ontologies, Knowledge Representation and Reasoning

SemWeb Technologies (SWT) is an intelligent **agent** that searches the contents of several web sites and return results based on the predefined criteria supplied by the user. Put in another way SemWeb has been defined as the conceptual structuring of the web in an explicit machine-readable way (Gómez-Pérez & Corcho, 2002).

Strassner, O'Sullivan, & Lewis (2007) define ontologies as a formal, explicit specification of a shared, machine-readable vocabulary and meanings, in the form of various entities and relationships between them, to describe knowledge about the contents of one or more related subject domains throughout the life cycle of its existence. Semantic technologies, including ontologies and semantic description languages, are quite similar to human thinking and memorization: they allow the definition of concepts and instances (of these concepts) that are related with each other using semantically qualified links, the extraction of knowledge from a base and reasoning on this knowledge to deduce inferred knowledge. Applying such an approach to information technologies enable machines to understand the actual meaning of data which is formulated using a distributed and evolving vocabulary (Joly, Maret, & Bataille, 2009) .

II. STATEMENT OF PROBLEM

Many learners and scholars would prefer to use the enormous resources available online for research and learning purposes but the problem here is how to go about getting the right materials faster and easier, which will translate to achieving results in a shorter time. It is also discovered that many institution of learning still design their content driven Web-based Apps using a well-known Hypertext Markup Language (HTML). This dynamic language has its limitation of inability for one to build ontology language functionalities that will enhance the structuring of the web pages. Worthy of note is the fact that the teaching of Sematic Web Technologies has not been introduced in our institution of higher learning.

III. OBJECTIVES

Having identified the issues, this paper seeks to provide the solutions or means to solutions to those stated problems. To be specific, our objectives are defined as follows

- To advocate the integration and teaching of semantic web technologies in our higher institutions of learning.
- Employ the concept of semantic web technologies in our future web sites engineering and designing.
- The time and efforts involved in carrying out research and learning would be reduced greatly, which will promote quality and efficiency in research.

IV. RELATED WORK

The increasing and ongoing need for and reliance on software in almost all industries, coupled with increasing code volume and complexity and changing technologies, results in increasing productivity pressure on software engineers to deliver greater software functionality within stringent cost, time, and quality constraints (Oberhauser, 2010). Moreover, according to him, the software maintenance phase is affected by these pressures, and new approaches are also needed for improving the efficiency and effectiveness of corrective and progressive maintenance activities.

The Semantic Web is the emerging landscape of new web technologies aiming at web-based information and services that would be understandable and reusable by both humans and machines (Sampson, etal, 2004). According to them, Ontologies, generally defined as a representation of a shared conceptualization of a particular domain, is a major component of the Semantic Web. It is anticipated that Ontologies and Semantic Web technologies will influence the next generation of e-learning systems and applications(Sampson, etal 2004.).

The central ideal of the Semantic Web is to extend the current human-readable Web by encoding some of the semantics of resources in a machine-processable form (Berners lee, et al 2001). Moving beyond syntax opens the door to more advanced applications and functionality on the Web, Computers will be better able to search, process, integrate and present the content of these resources in a meaningful, intelligent manner (Oberhauser, 2010.).

(Devedzic, 2004) surveys important aspects of Web Intelligence (WI) which explores the fundamental roles as well as practical impacts of Artificial Intelligence (AI) and advanced Information Technology (IT) on the next generation of Web-related products, systems, services, and activities. He argues that the key advantages of applying Web Intelligence techniques to Artificial Intelligence in Education (AIED) are enhanced adaptively and enhanced learner comfort. Automatic discovery, invocation, and composition of educational Web services can free the learner from many time-consuming activities that often disrupt the learning process itself. Finally according to him, ontology supported learning process greatly increases automation of a number of learners', teachers', and authors' activities related to Web-based learning environments.

Dicheva (2004) outlined the state-of-the-art research on Semantic E-learning and suggested a way towards the Educational Semantic Web. He proposed a modular semantic-driven and service-based interoperability framework and

related ontology-driven authoring tools. The challenge of the next generation web-based educational systems is to support user-friendly, structured and automated authoring, balancing between exploiting explicit semantic information for agreement and exchange of educational information, and on the other hand, collecting and maintaining the information semantics.

As to the integration of Semantic Web technologies in the Software Engineering lifecycle, (Oberhauser R, 2007) discussed a holistic approach. With regard to software artifacts, (Bontcheva, et al, 2006) presented an ontology learning approach that exploits a range of information sources associated with software projects. Work utilizing the Semantic Web for automated software engineering purposes includes (Dinger U., 2006).

Semantic web computing, with its formal structureing of information and machine-processable semantics, has the potential to improve Software Engineering automation and information integration (Oberhauser, 2010.). One of the issues facing Semantic Web Computing is the creation and adoption of standardized ontologies in OWL(Web Ontology Langauge)(Horrocks et al., 2004) for the various industry domains to precisely define the semantic meaning of the domain-specific concepts. The additional modeling effort incurred by ontologies must result in savings elsewhere (Oberle, 2006). (Vargas-Vera, 2004) outlined an e-Learning services architecture offering semantic-based services to students and tutors, in particular, ways to browse and obtain information through web services. They present a proposal for a student semantic portal providing semantic services, including a student essay annotation service. They also claim that visualization of the arguments presented in student essays could benefit both tutors and students.

V. METHODOLOGY

Many models and frameworks exist as ontology languages, which are classified into traditional syntax ontology language, markup ontology language, description logic-based e.t.c. Their sub categories ranges from Open Biomedical Ontologies (OBO) for modeling life sciences, Web Ontology Language (OWL), Ontology Interference Layer (OIL) to Resource Description Framework (RDF) for modelingMarkup Languages and more. In this paper, we employ a popular RDF as the basis for our discussion and implementation of SemWeb since it has syntax that is compactable with Extensible Markup Language (XML). We refer to the work of Cyganiak, Wood, & Lanthaler (2014) for detials about RDF framework.

A. Resource Description Framwork Application in Semantic Web

The Resource Description Framework (RDF) is a general framework for how to describe any Internet resource such as a Web site and its content. An RDF description (such descriptions are often referred to as metadata, or "data about data") can include the authors of the resource, date of creation or updating, the organization of the pages on a site (the sitemap), information that describes content in terms of audience or content rating, key words for search engine data collection, subject categories, and so forth. The Resource Description Framework will make it possible for everyone to share Web site and other descriptions more easily and for software developers to build products that can use the metadata to provide better search engines and directories, to act as intelligent agents, and to give Web users more control of what they're viewing. The RDF is an application of another technology, the Extensible Markup Language (XML), and is being developed under the auspices of the World Wide Consortium (W3C). Figure 4.1 describe the application of RDF as the building block or framework of SemWeb with XML as its baseline tool where all the RDF syntax are declared and encoded in order to communicate with other websites and or search engines.

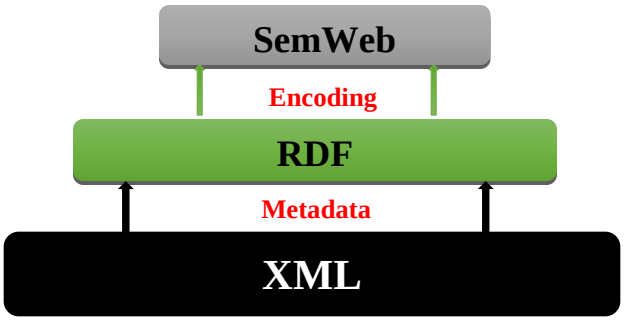

Fig 4.1: Application of RDF, XML in SemWeb Design

B. Definition and Declaration of RDF/XML

All RDF syntax is embedded in the XML code with its metadata usually declared at the top of XML file. RDF uses Universal Resource Identifiers (URIs) to identify resources on the web pages and its syntax is specified to describe the web resources using properties and its values. The combination of resource, property and its value is referred to as RDF statement. Logically, the resource is called subject, property is called predicate and the value is called an object. Figure 4.2 gives an example of RDF statement where *http://www.agfritech.com/rdf*on line 3 is the subject, <author> and <homepage> on line 4 & 5 are the predicates while their values *Agbo& Kissinger* and*http://www.agfritech.com* are the objects of the statement. The fundamental knowledge required

before using RDF syntax are the HTML, XML and Namespace.

```
<?xml version="1.0" encoding="utf-8"?>
<RDF>
  <Description about="http://www.agfritech.com/rdf">
    <author> Agbo & Kissinger </author>
    <homepage>http://www.agfritech.com</homepage>
  </Description>
</RDF>
```

Fig: 4.2 Example of RDF/XML Definition and Declaration of Web Metadata

Table 4.1: Three records from a Journal List

Publisher	Author	Title of Paper	Year	Pages
International Journal of Plant Sc.	Agbo Fred	Alkaline level of unripe orange Juice	2013	3
Unilorin Postgraduate Journals	Kissinger K. and Joseph A.	Effect of Afforestation in the North East of Nigeria	2009	8
Online Journal of Multidisciplinary Studies	AminuMuktar	Semantic Web in 21st Centuries	2014	5

The description of the records given in table 4.1 using a SemWeb with RDF/XML is given below.

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:cd="http://www.agfritech.com/journal#">

<rdf:Description
rdf:about="http://www.agfritech.com/journal/International Journal of Plant Sc.">
  <journal:author>Agbo Fred</journal:author>
  <journal:paper>Alkaline level of unripe  orange Juice</journal:paper>
  <journal:year>2013</journal:year>
  <journal:pages>3</journal:pages>
</rdf:Description>

<rdf:Description
rdf:about="http://www.agfritech.com/journal/Unilorin Postgraguate Journals">
  <journal:author>Kissinger K. and Joseph A.</journal:author>
  <journal:paper>Effect of Afforestation in the North East of Nigeria</journal:paper>
  <journal:year>2009</journal:year>
  <journal:pages>8</journal:pages>
</rdf:Description>

<rdf:Description
rdf:about="http://www.agfritech.com/journal/Online Journal of Multidisciplinary Studies ">
  <journal:author>Aminu Muktar</journal:author>
  <journal:paper>Semantic Web in 21st Centuries</journal:paper>
  <journal:year>2014</journal:year>
  <journal:pages>5</journal:pages>
</rdf:Description>
.
.
.
</rdf:RDF>
```

Fig 4.3: Screenshot of drescribing table 4.1 in a SemWeb using RDF/XML

The first line of the RDF document is the XML declaration. The XML declaration is followed by the root element of RDF A fundamental specification that is used to describe SemWeb is the namespace. We use the next figure below to demonstrate how RDF can be used to describe list of research papers in plant sciences published by different online publishing companies with semantic websites. The description of the journal list is presented in table 4.1 for better understanding.
documents: **<rdf:RDF>**.What followed is called the XML Namespace. An XML namespace is basically a document that tells applications the meaning of all the tags in another document. The creator of an XML document declares the namespace at the beginning of the document with a line of code. In our example, our namespace declaration is found on line 3 of figure 4.3 above, which specifies that elements with the rdf prefix are from the namespace "http://www.w3.org/1999/02/22-rdf-syntax-ns#".The **<rdf:Description>** element contains the description of the resource identified by the **rdf:about** attribute. While the elements: **<journal:author>, <journal:paper>, <journal:year>,** etc. are properties of the resource.

VII. CONCLUDING REMARKS

Most popular search engines of the internet are keyword based and thus, base their queries on the occurrence of given words in the textual content of pages. The drawback of this approach is that words can have very different meanings (interpretations), and also a concept can be formulated with many different words; therefore if I would like to eat an apple and search for an <<apple store>> the search engine will probably propose a store vending computers and mp3 players! On the other hand, a semantic search engine would ask you if you are looking for fruits or electronic devices to return relevant results. If we must reason on meaningful knowledge, semantic technologies are identified as the most promising candidate to enable this.
In order to achieve this, in this paper we started by presenting a survey of Semantic Computing technologies as applied to eLearning. We also demonstrated how the RDF technology can be used for the development of the SemWeb. In particular, we have provided an example of how to populate this ontology using an RDF based approach. However, the proposed model is generic, and any topic model can be used. This way, we can leverage the techniques that can be applied to enhance learning and research.

VIII. CONTRIBUTIONS TO KNOWLEDGE

This research work has been able to point out that employing the concept of semantic web technologies makes learning and research easier and faster considering the time and speed involved in research. It is also noted that shifting the website design from human readable domain to machine-readable domain will help search engines otherwise known as agents to return a more concise and accurate results when a user sends requests.

REFERENCES

- Berners lee, T.; Hendler, J. ; Lassila O. (2001). The Semantic Web, Scientific American. 28-37.
- Bontcheva, K., & Sabou, M. (2006). Learning Ontologies from Software Artifacts: Exploring and Combining Multiple Sources, In: Proceedings of 2nd International Workshop on Semantic Web Enabled Software Engineering.
- Cyganiak, R., Wood, D., & Lanthaler, M. (2014, February 25). *W3C Recommendation* . Retrieved January 30, 2015, from RDF 1.1 Concepts and Abstract Syntax: http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
- Devedzic, V. (2004). Web Intelligence and AIED.
- Dicheva, L., & Aroyo, D. (2004). The New Challenges for E-learning: The Educational Semantic Web.
- Dinger, U., Oberhauser, R., & Reichel, C. (2006). Leveraging Web Service-based Software Engineering, In: Proceedings of the International Conference on Software Engineering Aadvances (ICSEA'06), IEEE Computer Society Press.
- Gómez-Pérez, A., & Corcho, O. (2002). *Ontology Languages for the Semantic Web.* IEEE Intelligent Systems.
- Joly, A., Maret, P., & Bataille, F. (2009). Leveraging Semantic Technologies Towards Social Ambient Intelligence.
- Lucent, A., Bataille, F., & Joly, A. (2004). Leveraging Semantic Technologies towards Social Ambient Intelligence.
- Oberhauser, R. (2010). Leveraging Semantic Web Computing for Context-Aware Software Engineering Environments, Semantic Web, Gang Wu (Ed.), . *Open Science.*
- Oberhauser, R., & Schmidt, R. (2007). Towards a holistic integration of software lifecyclem Processes using the Semantic Web In: Proceedings of the 2nd International Conference on Software and Data Technologies. 133-144.
- Sampson, D. G., Lytras, M., Wagner, G., & Diaz, P. (2004). Ontologies and the Semantic Web for E-learning, Educational Technology & Society. 26-28.

- Sebastin, A., Rios, F. A., & Bustos, F. (2002). Leveraging Social Network Analysis with Topic Models and the Semantic Web.
- Strassner, J., O'Sullivan, D., & Lewis, D. (2007). Ontologies in the engineering of management and autonomic systems: a reality check. *Journal of Network and Systems Management*.
- Vargas-Vera, Moreale, E., & Maria. (2004). Semantic Services in e-Learning: an Argumentation Case Study. Educational Technology & Society. 112-128.
- w3schools.com. (2004, February 10). *RDF Rules*. Retrieved January 31, 2015, from w3schools.com: http://www.w3schools.com/webservices/ws_rdf_rules.asp