Vol. IX, Issue 1
February 2019

Collaborative Research and the Semantic Web

Authors: **Julius MURUMBA,** Kibabii University, P.O. Box 1699-50200, Bungoma, Kenya**; Elyjoy M MICHENI,** The Technical University of Kenya, P.O. Box 52428, Nairobi, 00200, Keny**a; Alice N WECHULI,** Kibabii University, P.O. Box 1699-50200, Bungoma, Kenya

This paper is a review of the literature on semantic web technologies and states their relevance in collaborative research. The semantic web is comprised of modern technologies that can allow academia in developing countries learn to adopt innovative and collaborative techniques of research.

The objectives guiding this paper were, to examine the relevance of semantic web technologies in collaborative research and to establish the benefits semantic web technologies present to collaborative research. This paper explores the semantic web environment to present an assessment of its effect on collaborative researches.

The study is carried out through a literature review that involved an examination of scientific research papers in journals and conference proceedings, and from online journals. The paper concludes that the semantic web has changed the way we interact, potentially changing collaborative research by adding new levels of functionality on the web by enabling activities which are interactive in nature between people and computers.

These activities include but not limited to interpreting content on the web, update alerting, data sharing and customized responses to web queries.

***Keywords**: Semantic Web, Collaborative Research, Web Services*

Introduction

Research scholars and academicians in Kenya as well as many other developing countries of the world have expectations that the technological advancements in hardware and software in recent years that are leading to continued increase in computing power and the growth of the internet will enable them use these technologies to enhance collaborative research. These technologies include broadband networks, cloud computing, the semantic web and grid computing. The University landscape of many developing nations is being changed by these technologies that have led to use of distributed computing, internetworking and cloud based applications, which are making collaborative research possible [1]. Collaborative research in universities provides academic members of staff with both formal and informal professional learning opportunities, as well as opportunities for the universities to benchmark against best practices in academia. [2] state that collaborative research refers to a situation where, researchers work together in a collaborative manner in order to solve shared challenges or to achieve common goals, and to produce new scientific knowledge. The emergence of the notion of collaborative researches is due to demands for collaboration on research projects among researchers in universities especially in situations where such projects involve research groups that share challenges are geographically distributed. [2] further hold the view that enabling and managing such research collaborations requires proper internetworking infrastructure and extensive distributed computing resources. Information Technologies such as the internet and the World Wide Web offer universities the opportunities of enabling and managing collaborative research by helping researchers manage joint and sometimes complex tasks. Collaborations in universities provide a medium for mutual reflection on peer practices and assists professionals refine knowledge in their areas of specialization, therefore, it provides universities opportunities to mentor researchers and to be elevated to the level of the world's best.

The web has become an important educational and collaborative research medium as a result of the continuing growth of web-based applications anchored on these emergent technologies. The semantic web in particular can promote research activities of academic members of universities since it is capable of enabling access to academic resources. According to [3], such academic resources may include research journals

and repositories accessible online, databases and generally information about conferences. The semantic web further offers researchers and academicians opportunities for sharing experiences in diverse areas of research, in addition, it enables them to exchange research findings throughout the world. The semantic web refers to the automated access and use of information on the web based on semantics of data which is processed by networked computers. At the core of the semantic web are the ontolgy's which are used for management of knowledge in distributed systems [4]. The semantic web was designed as a machine process able web of data; a web where computerized agents collect, integrate, exchange and reason based on large quantities of heterogeneous online content [5]. The semantic web community can therefore exploit theories, methods, and tools from various disciplines and collective intelligence ultimately building a global semantic web comprising of distributed interleaved human-machine computation. Ontology is the core of the semantic web, and is used for management of knowledge that is formalized in distributed systems. Ontology makes information on the web to be understood by the computers and also assists people in processing and interpreting information on the web. Therefore, Ontology's provide semantics of data that are process able by machines and communicable between different software applications and people [4]; [6].

Semantic Web Technology Description

Semantic web technologies are today frequently used by scholars because of their ability to provide easy access to academic resources, and to provide opportunities for sharing information regarding research findings or information about available research opportunities as well as information on conferences and symposiums [3]. [7] Observe that technologies that make up the semantic web have transformed it into an interactive and collaborative medium for information sharing since everyone can create and share a variety of multimedia information. According to [8], the emergence of the semantic has led to web applications that encode different types of resources in XML, and are usually identified using a uniform resource identifier (URI).

In other words, the address of a resource on the World Wide Web is identifiable through its URI or through alternative forms of identifier

systems, such as digital object identifier (DOI). The standards and most of the technical recommendations of the semantic web are developed by World Wide Web Consortium (W3C). These standards include standards for ontology, web queries, linked data, Hypertext Markup Language, Hypertext Transfer Protocol and Extensible Markup Language [9]. According to [10], The introduction of Extensible Markup Language, Resource Descriptive Framework, and Web Ontology Language into the semantic web were aimed at addressing the limitations of HTML. Extensible Markup Language (XML) defines portable data in a structured and self-describing format and therefore enables transmission of structured documents. Ontology standards are used in taxonomies or in classification of terminologies that are used in applications or datasets, and to help in integration of different data sets or explain relationships between the terms or datasets [9]. Resource distribution framework provides semantics f data and enables the linking of data together through attributes of data or meta-data on the World Wide Web [8]. The Really Simple Syndication (RSS) web feed that provides to updates for users is a good example of an RDF object.

Technologies supporting the semantic web are best illustrated using a layered architecture [11]; [12]. This is depicted in Figure 2.4 below. The World Wide Web lies at the lowest level, and is made up of various elements like the uniform resource identifier and UNICODE. Other elements in the architecture include XML, RDF, SPARQL and OWL among others all of which support the layered architecture.

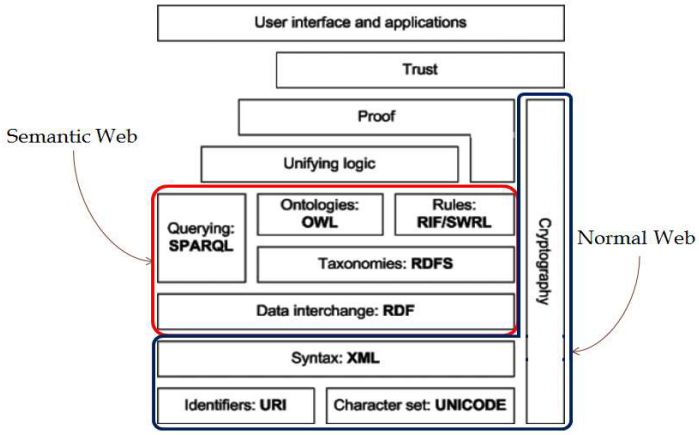

Figure 1: Architecture of the Semantic Web (Adapted from [11]).

These elements have each a specific function in the architecture: (i) The Uniform Resource Identifier is a string that uniquely identifies resources on the world wide web while UNICODE provides standards for exchange of symbols on the web. (ii) XML technique is a communication medium for electronic data exchange which is used to define portable data in a structured and self-describing format. (iii) Resource Description Framework (RDF) is a model that enables data interchange. It is a structure that uses XML schemas to provide methods for expressing semantics, and supports the exchange of information in machine understandable form on the web. The RDF structure is composed of a resource which is identifiable by a URI; as well as properties which are a description of the characteristics of the resource and statement which is a property having a value for that resource to enable it form an RDF statement [9].

Developments in Semantic Web Technology

The semantic web, commonly referred to as web 3.0 is characterized by personalization, for example iGoogle, intelligent search, collaborative use and behavioral advertising etc [13]. The semantic web is a medium for data and this has made it necessary to find ways of attaching meaning to the data or extracting useful and valuable information from this data. Queries are used in a way that is similar to the way relational databases use Structured Query Language (SQL). To achieve this, XQuery is useful when it comes to extracting information from XML data while SPARQL is used in getting information from RDF objects through protocols supporting open access technologies such as HTTP and SOAP [9]; [14]. SPARQL is a query language for semantic web data sources that is be used to query diverse data sources including data stored natively as RDF and data viewed as RDF through a middleware. Ontology is a term derived from philosophy and is usually meant to refer to the systematic study of the nature of existence. Ontology is an explicit and a formal specification of a shared concept of a domain of interest [15]. It can be represented in the form of patterns of knowledge representation that may be done using objects or concepts within a specific domain such that their definitions, properties and their associations are able to clearly model a domain. The Web Ontology Language (OWL) allows users to describe the semantics or meaning of resources and their properties on the web, making it possible for users to add vocabulary in the domains.

[16] noted that OWL provides additional vocabulary together with a formal semantics which make it possible for machines to interpret web content that is not supported by XML and RDF. Applications developed using OWL are capable of working at the human conceptual level while enabling integration of information and searches through communities.

A web service is software function enabling machine to machine interoperation and is usually hosted at a network addressable location. The interface of the web service serves the purpose of hiding the implementation details of the Application Programming Interface (API), therefore enabling the web service to be used independently of the programming language used to write it and the hardware as well as software platform on which it is implemented. Because of this independence web service based applications can be component-oriented and can be loosely coupled, and can support cross-technology implementations. According to [17], Web Services are deemed an integration middleware for web based applications because of their viability in coordination and transaction capabilities for web applications. Middleware is essentially a layer of software running between client and server processes in a distributed computing environment and is defined as a distributed software functionality upon which applications and system s can be constructed, for example web-enabled access to databases[18]. Web services middleware therefore plays the role of a middleman in the overall software architecture by addressing cross-platform communications and things like security. Web Services make full use of open standards including HTTP, XML, SOAP, WSDL, and UDDI [19]; [18]. [20] Observed that the objective of Semantic Grid computing is to incorporate the approach of Semantic Web into Grid computing. The semantic web infrastructure is therefore not orthogonal to grid computing but rather an infrastructure that can be amalgamated with grid computing applications.

Collaborative Research and the Semantic Web

The modern technologies of the semantic web can allow researchers from developing countries to learn to adopt innovative techniques of research. [27] state that the distributed, interlinked and highly structured information that is found on the semantic web allows researchers to share their knowledge, research activities, interrelations and also their publications.

These technologies are being applied in many applications by conceptualizing and formalizing domain specific knowledge. This is being achieved using ontology to offer support for diversified processing and some form of reasoning operated by machines. Many semantic web applications are concerned with search, which typically requires that ontological descriptions be used in the searches in order to improve the aspect of locating and retrieving of relevant information. Semantic web applications also focus on annotations which in turn provide a foundation for display of resources on the web. Annotations improve the way information about known concepts that are appearing in ontology is displayed [21], and are important in achieving interoperability across organizations. Ontology's enable software agents and people to have a common understanding of the structure of information on the web and also enable a common analysis and re use of domain specific knowledge [22]; [15].

The resource description framework organizes information on the web by considering everything as a resource or named things uniquely identifiable by a URI. Data linking problem on the web is therefore resolved through this form of organization. The work of [9] indicated that the semantic web revolved around using linked data to create typed links between different sources of data that conform to the semantic web standards. Examples of the linked data are the RDF, OWL and XML all of which have provisions for descriptions that are machine readable and therefore append meanings to content on the web. Semantic web technology has previously been used in collaborative researches, for instance in clinical research and bioinformatics [23] where the technologies were used to remove barriers that existed to patient oriented research. This was achieved by combining mechanisms used for addressing and naming that are highly distributable and use RDF and OWL in formal representation of knowledge. This mechanism for rendering documents is known as the Gleaning Resource Descriptions from Dialects of Languages (GRDDL). GRDDL is used to obtain data in RDF format from XML and XHTML documents using SPARQL. Similarly, Semantic wikis are commonly used to allow for explicit annotation of content and therefore makes such content more readable for machines. The semantic wikis also allow for use of Meta data, embedded queries and semantic searches. Examples of semantic wikis include Semantic MediaWiki, AceWiki4, OntoWiki7, among others [24].

Semantic Web Benefits

One of the advantages of the semantic web is its ability to facilitate the management of heterogeneous information. It enables machine to machine communication through use of semantic documents and data rather than human readable writings or speech. Semantic web based applications are generally developed to enable different individuals and institutional structures exchange information in machine readable formats, and offers diverse search opportunities for different search engines. Searches based on ontology use axioms of concepts and their relationships and therefore perform fine grained filtering of content to give results. The work of [9] shows us that semantic search engines are divided into human oriented search engines and application oriented search engines. The human oriented search engines are similar to the Google search engine with keyword based search.

The human oriented semantic search provides a more detailed interface where the user can exploit the underlying structure of the data instead of simply providing links to the pages the user searches, and also provides options for searching documents, objects and concepts with each option being capable of yielding different results. On the other hand, the application oriented search serves the needs of applications built on top of linked data. Application oriented searches provide application programming Interfaces through which linked data applications discover RDF documents when searching on the web. Benefits of the semantic web to researchers according to [25] include; programs and web applications can exchange information seamlessly, results of searches of search engines display information that seems more relevant and data mashers are used to find new things as a result of combining data from different datasets.

As a consequence of the semantic web, researchers can make sense of large amounts of data because of the help offered by the semantic web in enabling machines understand a variety of information web pages and the relationship between different pieces of information. Institutions of higher learning can leverage on semantic technologies to build large and flexible web applications for diverse purposes. The semantic web technologies and standards have led to several software advantages [26]. These include 1) Formation of virtual research communities. This is achieved when Institutions with common interests become tightly connected on the web as

a result of preferences that are mutually beneficial, that are then defined in terms of ontology to form a virtual community. 2) Filtering of information which occurs during the processes of information sending, receiving and storing. The web services invocation and information filtering sends information selectively to the rightful clients.

Conclusions

This paper has discussed the semantic web and its relevance in collaborative research initiatives in institutions of higher learning in the perspective of developing countries. This paper has provided a technological description and examined the role of the semantic web technologies in collaborative research and these technologies offer to researchers. The semantic web has changed the way we interact, and has the potential to enhance collaborative research. It adds new levels of functionality on the web and enables numerous interactive activities between computers and among users including but not limited to update alerts, interpreting web content, intelligent searches and customized responses to queries. This paper concludes that the semantic web provides the necessary infrastructure for collaborative research and platforms for data sharing, and is therefore a compelling reason for institutions of higher learning in many African countries to build semantic web applications that would enable them engage in collaborative research and gain value from their web applications.

References

[1]. Olson, J., Codde, J., deMaagd, K., Tarkleson, E., Sinclair, J., Yook, S., et al. (2011). An Analysis of e-Learning Impacts & Best Practices in Developing Countries: With Reference to Secondary School Education in Tanzania. Michigan, USA: The ICT4D Program, 209 Communication Arts & Sciences Building- Michigan State University.

[2]. Keraminiyage, K., Amaratunga, D., & Haigh, R. (2009). Achieving Success in Collaborative Research: The Role of Virtual Research Environments. Journal of Information Technology in Construction, 14, 59.

[3]. Hussain, I., & Durrani, M. I. (2012). A Study On the Role Of Web Technology In Enhancing Research Pursuance Among University Academia. i-manager's Journal of Educational Technology, 9 (3), 32-40.

[4]. Lausen, H., Ding, Y., Stollberg, M., Fensel, D., Herna´ndez, R. L., & Han, S.-K. (2005). Semantic web portals: state-of-the-art Survey. Journal of Knowledge Management, 9 (5), 40-49.

[5]. Sarasua, C., Simperl, E., Noy, N., Bernstein, A., & Leimeister, J. M. (2015). Crowdsourcing and the SemanticWeb: A Research Manifesto. Human Computation, 2 (1), 3-17.

[6]. Hall, W., De Roure, D., & Shadbolt, N. (2008). The Evolution of the Web and Implications for e-Research. Southampton: University of Southampton.

[7]. Saeed, N., Yang, Y., & Sinnappan, S. (2009). Emerging Web Technologies in Higher Education: A Case of Incorporating Blogs, Podcasts and Social Bookmarks in a Web Programming Course based on Students' Learning Styles and Technology Preferences. Educational Technology & Society, 12 (4), 98–109.

[8]. Downes, S. (2005). Semantic networks and social networks. The Learning Organization, 12 (5), 411-417.

[9]. Gylfason, B. P. (2010, March 25). The Future of the Web. The Semantic Web. Reykjavik University: Unpublished.

[10]. Alani, H., Hall, W., O'Hara, K., Shadbolt, N., & Szomszor, M. (2008). Building a Pragmatic Semantic Web. IEEE Computer Society, 61-68.

[11]. Singh, K., Kumar, A., & Yadav, A. K. (2016). Semantic Web Mining: Issues and Challenges. International Journal of Engineering Sciences & Research Technology, 5 (8), 637-664.

[12]. Zeeshan, A., Chimay, A., Darshan, R., Patricia, C., & Dino, B. (2004). Semantic Web Based Services for Intelligent Mobile Construction Collaboration. ITcon, 9, 367-379.

[13]. Farah, J. (2012). Predicting the Intelligence of Web 3.0 Search Engines. International Journal of Computer Theory and Engineering, 4 (3).

[14]. Ma, L., Mei, J., Pan, Y., Kulkarni, K., Fokoue, A., & Ranganathan, A. (2007). Semantic WebTechnologies and Data Management. Retrieved October 20, 2017, from www.w3.org: https://www.w3.org/2007/03/RdfRDB/papers/ma.pdf

[15]. Patil, V. M. (2014). Semantic Web and Ontologies: A Survey. International Journal of Computer Application.

[16]. Varlan, S. E. (2010). Advantages of Semantic Web Technologies in The Knowledge Based Society. Iasi, Romania.

[17]. Karastoyanova, D., & Buchmann, A. (2003). Components, Middleware and Web Services. Darmstadt, Germany: Unpublished.

[18]. Techopedia. (2017). Retrieved July 15, 2017, from Techopedia Inc.: https://www.techopedia.com/definition/30632/web-services-middleware

[19]. Samtani, G., & Sadhwani, D. (2004). Web Services and Application Frameworks (.NET and J2EE). Chicago: Tect.

[20]. Cannataro, M., & Talia, D. (2004). Semantics and Knowledge Grids: Building the Next- Generation Grid. E - S c i e n c e, 56-63.

[21]. MilesBoard, T., Carr, L., Woukeu, A., Wills, G., & Hall, W. (2005). Towards Semantic Web Support for Writing Project Proposals.

[22]. Gontier, E. M. (2015). Web Semantic and Ontology. Advances in Internet of Things, 5, 15-20.

[23]. Ogbuji, C., Blackstone, E., & Pier, C. (2007). Case Study: A Semantic Web Content Repository for Clinical Research. Retrieved October 22, 2017, from Semantic Web Use Cases and CaseStudies: https://www.w3.org/2001/sw/sweo/public/UseCases/ClevelandClinic/

[24]. Zaidan, F. H., & Bax, M. P. (2011). Semantic wikis and the collaborative construction of ontologies: a case study. Journal of Information Systems and Technology Management,8 (3).

[25]. Shazmeen, S. F., & Ramyasree, E. (2012). Semantic Web Mining: Benefits, Challenges and Opportunities. International Journal of Advanced Computer Research, 2 (4), 25-28.

[26]. Zhao, Y., & Sandahl, K. (2003). Potential Advantages of Semantic Web for Internet Commerce. Linköping, Sweden: Unpublished.

[27]. Sure, Y., Bloehdorn, S., Haase, P., Hartmann, J., & Oberle, D. (2005). The SWRC Ontology: Semantic Web for Research Communities. Unpublished.

	**Julius Murumba** is a Doctoral student of Information Technology at Kibabii University, Kenya. He holds a Bsc and MSc (Information System), and is currently a Lecturer at the Technical University of Kenya. His research interests include Semantic web and grid computing technologies
	**Dr. Elyjoy M. Micheni** is a Senior Lecturer in Information Systems and the Chairperson, Department of Management Science and Technology at The Technical University of Kenya. She holds a PhD (Information Technology) from Masinde Muliro University of Science and Technology, Master of Science (Computer Based Information Systems) from Sunderland University, (UK); Bachelor of Education from Kenyatta University; Post Graduate Diploma in Project Management from Kenya Institute of Management. .She has taught Management Information System courses for many years at University level. She has presented papers in scientific conferences and has many publications in referred journals. She has also co-authored a book for Middle level colleges entitled: —Computerized Document Processing‖ Her career objective is to tap computer based knowledge as a tool to advance business activities, promote research in ICT and enhance quality service
	**Dr Alice Nambiro Wechuli** is a Lecturer in the School of Computing and Informatics, Kibabii University