International Journal of Advance Research in Science and Engineering
Vol. No.4, Issue 08, August 2015
www.ijarse.com
IJARSE
ISSN 2319 - 8354

SEMANTIC WEB TECHNOLOGY AND ONTOLOGY DESIGNING FOR TAXONOMY OF LIVING ORGANISMS

Princy Aggarwal[1], Anuj Mehta[2]

[1]M.tech Student, [2]Assistant Professor, Department of Computer Science & Engineering,
Shri Krishan Institute of Engineering &Technology, Kurukshetra University, Kurukshetra, (India)

ABSTRACT

The Semantic Web is expected to be the next generation of the WWW. The increasing interest in the semantic web is producing a growing number of publicly available domain ontologies. Ontologies are mainly used to build smart and rational relationships among the concepts of a specific domain, so that the semantically related information are retrieved and queried whenever they are needed. In this paper we will use the components of the semantic ontology science and we will propose new way of building semantic and accurate relationships that will eventually lead to more precise results. In this paper we will discuss the techniques used for building semantic web applications and then present the development of ontology for taxonomy of living organisms by the use of Protégé which is very user friendly and easy to handle tool.

Keywords: ***Semantic Web, Ontology, Resource Description Framework (RDF), Ontology Web Ontology, Protégé***

I. INTRODUCTION

Biology is a complex and diverse science that is ever evolving. One aspect of the complexity of Biology is the complexity of the living systems themselves that are studied and represented. Biology, as the study of living things, is organized into a structured hierarchical organization which includes concepts such as 'class', 'order', 'genus', and 'species', from which the binomial naming convention of organisms evolved.Today, knowledge representation systems, including ontologies in the information sciences, are very much reminiscent of Linnaean classification and may be considered the modern day continuation of the Linnaean enterprise.

Ontology is a framework to represent relationships between objects or relationships of entities that can be said to exist in nature [1]. The relationships derived between these objects lend a method to classify objects based upon what is similar or different, where the resultant organization then resembles a hierarchy. Similar types of hierarchical organization can be found in the biological sciences in the form of phylogenetic trees and taxonomies like that of the Linnaean classification of living systems; however, there are key differences between ontology and a simple taxonomical hierarchy. In the main, a taxonomical hierarchy categorizes objects based on similarities and differences but does not attempt to capture meaning behind the classification as does an ontology. Ontology strives to define relationships betweenthe objects in an attempt to model a more formal framework for the representation of reality [2].

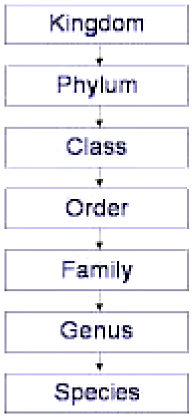

Fig. 1 Linnaean's System Classification of Living Organisms

II. SEMANTIC WEB

The Semantic Web is an extension of the current enabling computers and people to reuse the information. This is realized by marking up Web contents with properties, and relations, in a reasonably expressive markup language with a well-defined semantics [3].In such a context, some languages also known as Semantic Web languages are used to represent information about resources on the Web. This information is not limited to Web resource description, but can be about anything that can be identified. Uniform Resource Identifiers (URIs) are used to uniquely identify entities. Semantic web uses number of techniques like RDF, OWL, XML and SPARQL [4].

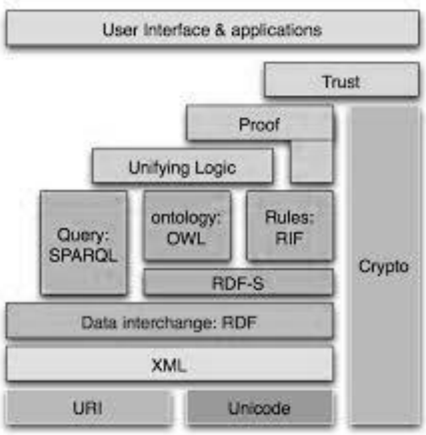

Fig. 2 Semantic Web Architecture [5]

2.1 Rdf

RDF stands for Resource Description Framework. It is a graph model similar to relation data model to organize data in a more meaningful way. The RDF data model is based upon the idea of making statements about resources in the form of subject-predicate-object expressions. These expressions are known as triples in RDF terminology. Triples are statements that contain a subject, a predicate, and an object. RDF can be viewed as an application neutral data model [6]. It is used to describe various attributes of thing like name, designation, salary, address etc. For example, Delhi is the capital of India. There are two thing named Delhi and India which are related to each other by the link "is the capital of".

Delhi is the capital of India

Subject predicate Object

Fig. 3 RDF Data Model

2.2 Owl

The Web Ontology Language (OWL) is one of the most expressive standardized Semantic Web languages. It is layered on top of RDF and RDF-S. OWL is a family of knowledge representation languages based on DLs. OWL languages are well-founded, useful and efficient enough for being the basis of knowledge representation for the Semantic Web, and thus for representing ontologies. OWL can be used to define classes and properties as in RDF-S but also provides constructs to create new class descriptions as logical combinations of other classes. OWL has three different levels of expressiveness: OWL-Lite, OWL-DL and OWL-Full. Each of these sublanguages is a syntactic extension of its simpler predecessor [7].

2.3 Sparql

The Simple Protocol and RDF Query Language (SPARQL) is a SQL like RDF query language for databases which is able to retrieve and manipulate for any data store in RDF format.SPARQL thus provides a full set of analytic query operations such as JOIN, SORT and AGGREGATE for data whose schema is intrinsically part of the data rather than requiring a separate schema definition. Schema information (the ontology) is often provided externally to allow different datasets to be joined in an unambiguous manner. In addition, SPARQL provides specific graph traversal syntax for data that can be thought of as a graph [8].

III. PROPOSED WORK

A taxonomical hierarchy categorizes objects based on similarities and differences but does not attempt to capture meaning behind the classification as does an ontology. The proposed ontology aims to define relationships betweenthe objects in an attempt to model a more formal framework for the representation of reality. This mechanism will work on the various techniques of semantic web like Ontology, RDF and XML [9].

IV. ONTOLOGY DEVELOPMENT FOR PROPOSED WORK

This section covers and explains the development of ontology for proposed taxonomy of living organisms. The ontology discussed here is developed in protégé tool [10]. It is freely available, open source, having a great set of plug-ins like OWL VIZ, onto graph, DL QUERY etc.

Fig. 5 shows the ontology building[11] for proposed taxonomy which has life as its top class which has subclass super kingdom which is again subdivided into three classes Archaea, Eubacteria and Eukaryote which are further subdivided till low level class genus comes. Any class can have different individuals which can be instantiated whenever required. Fig. 6 represents the object property assigned to various kingdoms to relate their features.

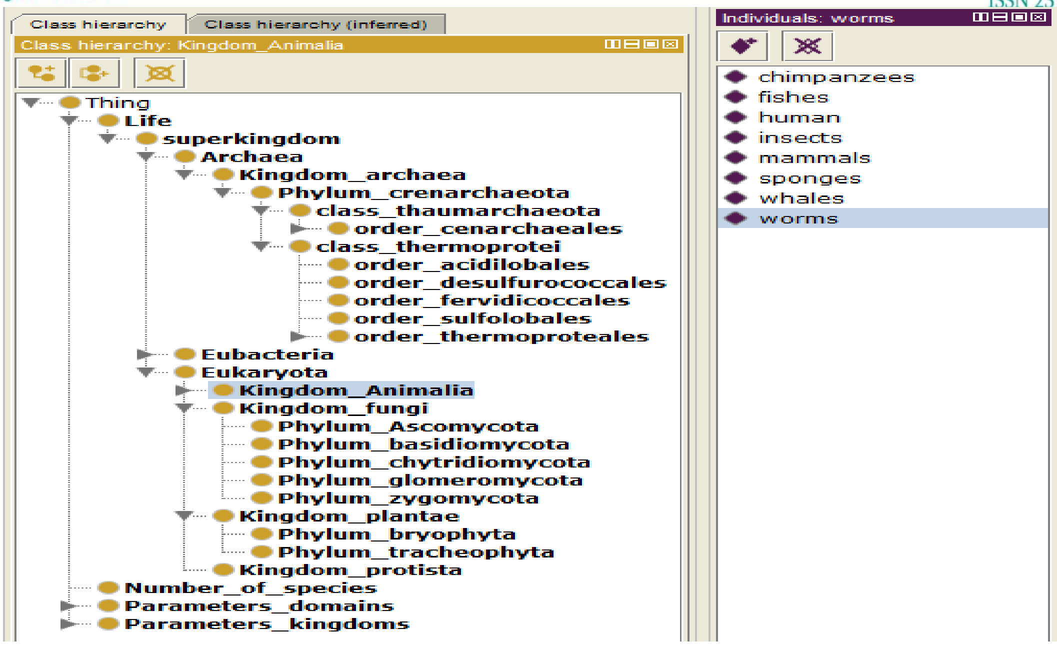

Fig. 5 Class Hierarchy and Instances of Particular Class Kingdom_Animalia of the Proposed Work

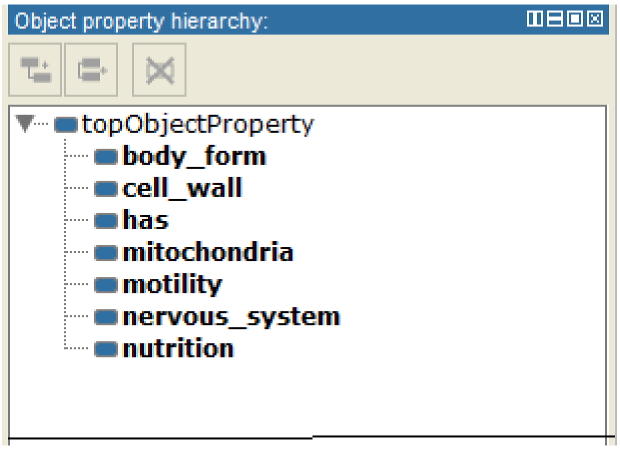

Fig. 6 Object Properties

ext comes visual representation of different classes created in this proposed work. Different plug-ins are available for this purpose in protégé like Onto graph, OWL VIZ etc []12. Fig. 7 represents ontograph for the class life. Fig. 8 represents ontograph for the class Kingdom_archaea. Fig. 9represents OWL VIZ for the class Phylum_chordata

IJARSE
ISSN 2319 - 8354

Fig. 7 OWL VIZ for the Class Life

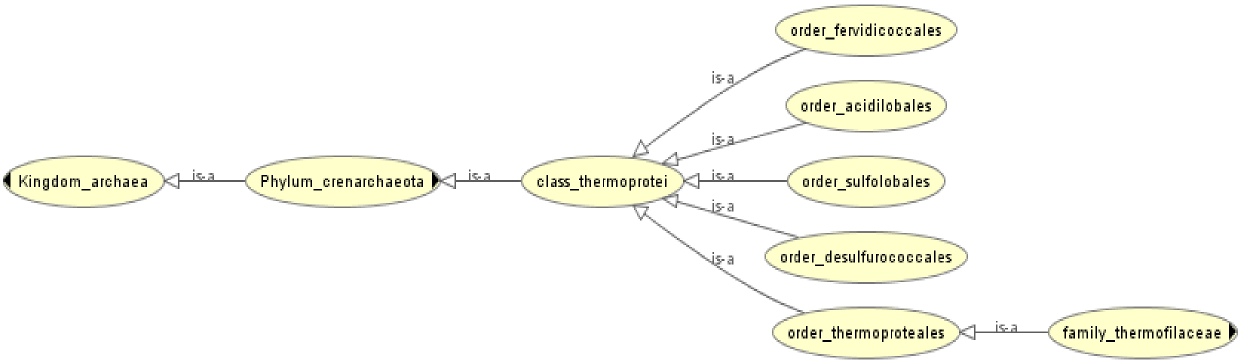

Fig. 8 OWL VIZ for the Class Kingdom_Archaea

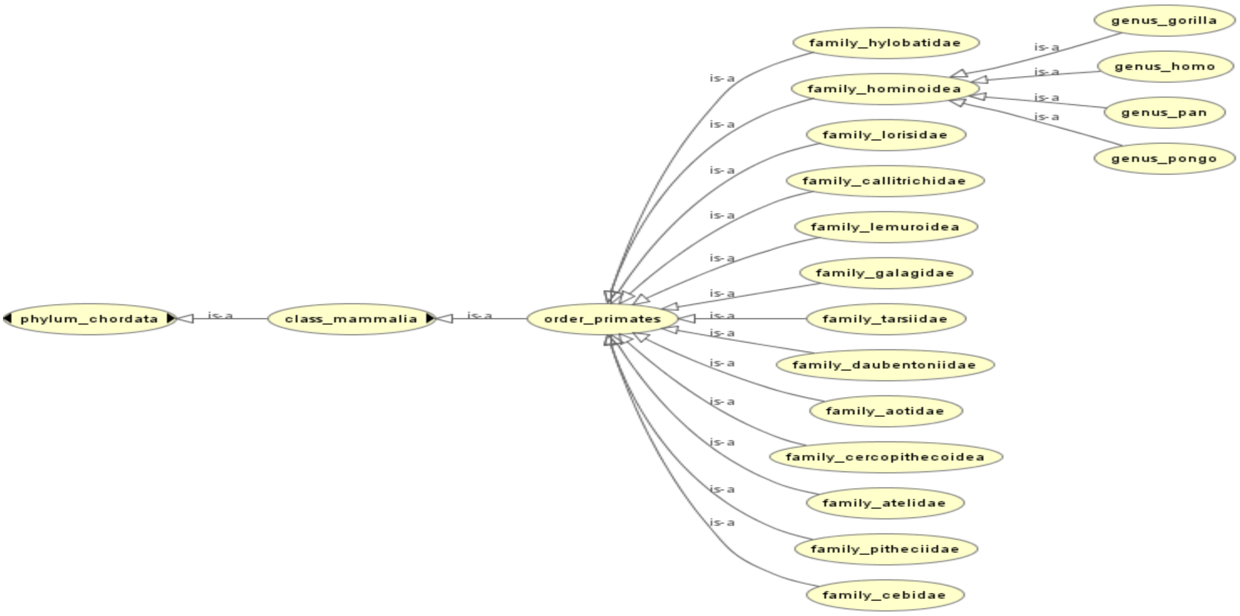

Fig. 9 OWL VIZ for the Class Phylum_Chordata

Onto graph is another plug-in for visual representation. Fig. 10 represents different sub classes in the class kindom_animalia. Fig 11 represents all the levels of hierarchy from superkingdomarchaea to genus thermofilum.

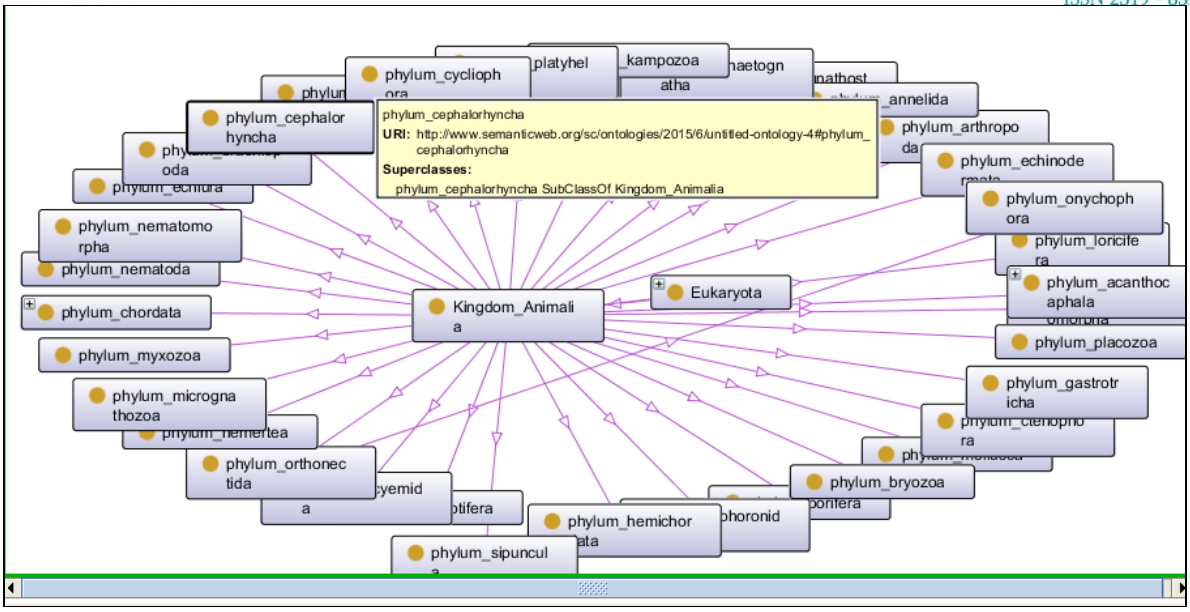

Fig. 10 Sub Classes of the Class Kindom_Animalia

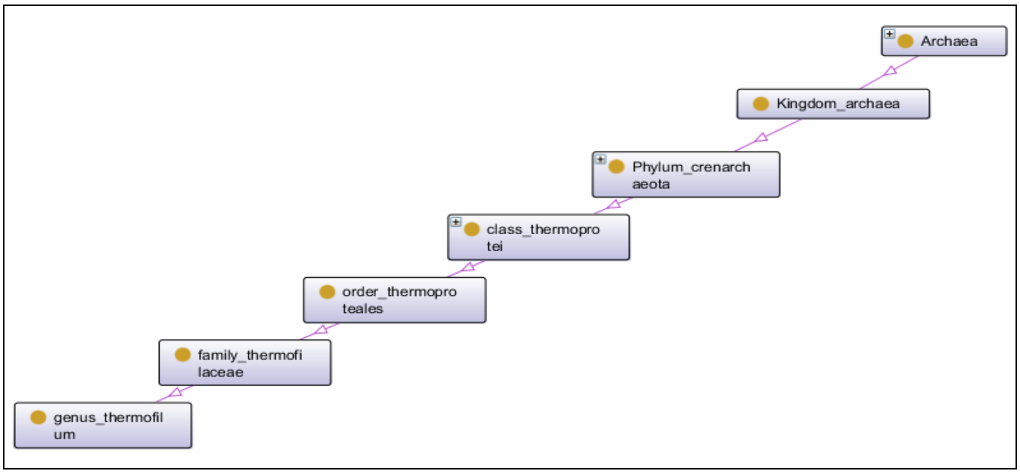

Fig.11 Levels of Hierarchy from Class Super Kingdom Archaea to Class Genus_ Thermofilum.

V. CONCLUSION

The proposed work not onlycategorizes living organisms based on similarities and differences but also attempts to capture meaning behind the classification. This knowledge base is useful for anyone who is interested in knowing different forms of plants, animals, insects, bacteria existing on earth. It will provide fast access to such information in a more meaningful way because in this taxonomy ontology all the resources are related with each other in the sense of some relation in structured manner. This taxonomy ontology can be reused in any of the future ontology which will be an enhanced version of this ontology like adding information regarding diseases and corresponding medicine to different forms of life.

REFERENCES

[1] Gloria L. Zuniga "Ontology: its transformation from philosophy to information systems" Proceedings of the international conference on Formal Ontology in Information Systems - Volume 2001, October 2001.

[2] M. Uschold and M. Grüninger, "Ontologies: Principles, methods and applications". The Knowledge Engineering Review 11 2 (1996), pp. 93–155

[3] Olena Parkhomenko, Yugyung Lee, E. K. Park. "Ontology – driven peer profiling in peer-to-peer enabled semantic web" Proceedings of the twelfth international conference on Information and knowledge management. November 2003.

[4] Chintan Patel, Kaustubh Supekar, Yugyung Lee, E. K. Park. "Ontokhoj: a semantic web portal for ontology searching, ranking and classification" Proceedings of the 5th ACM international workshop on Web information and data management. November 2003.

[5] Bernardo Cuenca Grau, "A possible simplification of the semantic web architecture" Proceedings of the 13th international conference on World Wide Web, May 2004.

[6] Liang Bangyong, Tang Jie, Li Juanzi. "Association search in semantic web: Search + Inference" Special interest tracks and posters of the 14th international conference on World Wide Web. May 2005

[7] Julian Seidenberg, Alan Rector. "Web ontology segmentation: analysis, classification and use" Proceedings of the 15th international conference on World Wide Web, May 2006.

[8]] Xia Wang, Manfred Hauswirth, Tomas Vitvar, MaciejZaremba. "Semantic web services selection improved by application ontology with multiple concept relations"Proceedings of the 2008 ACM symposium on applied computing. March 2008.

[9] Zhang "Ontology and the semantic web" Proceedings of the North American Symposium on Knowledge Organization, Vol.1, 2007.

[10] http://protege.stanford.edu

[11] Sinead Boyce, Claus Pahl. "Developing Domain Ontologies for Course Content" Educational Technology & Society, 2007

[12] M. Uschold, M. King, "Towards a Methodology for Building Ontologies", IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, 1995