International Journal of Interactive Digital Media, Vol. 2(1), ISSN 2289-4098, e-ISSN 2289-4101, www.ijidm.org

Ontological engineering languages and semantic web: A short review

Muhammad Waseem Chughtai, Ali. Bin Selamat and Imran Ghani

Abstract— Formal and executable ontologies earn a lot of attention in these days. However, formality of ontology depends on languages, which are being used to present it. This paper presents a short review on ontological engineering languages that have been used for the research community for representing ontologies. The most popular ontology languages such as extensible markup language, resource description framework and its schema, web ontology language, object interaction language and DARPA agent markup language have been reviewed. Furthermore, their working, disorders, and comparison has presented in this paper.

Index Terms— Ontology, Semantic web, W3C, XML, RDF, RDFS, OWL, OIL, DAML.

1 Introduction

WEB itself is an extensive collection of online resources, it is an unprecedented space of heterogeneous and distributed information, which is practically infeasible for the users to combine all possible sources to retrieve optimized information in minimum possible time has thereby turned up as a major challenge. Hereby, the web needed a customized web experience for user's needs and preferences. Tim Berners-Lee (at the mid of 1999) envisioned the semantic concept to accomplish this challenge, his vision includes an extended web that incorporates the machine interpretable information, enabling machines process to the available information and acting on behalf of the users counterparts[1]. In 2001, Tim Berners-Lee officially introduced semantic web with full realization of the actual need of the users and implementation of the semantic web idea [2].

Semantic web is not a separate Web; it is an extension of current web, which aims to optimize the unstructured information with meaning in the sense that a computer program can learn enough about what the data means [2]. Semantic web data can be expressed not only in natural language but also in the form that can be understood, interpreted, better enabling computers and users. Although, with hypertext markup language (HTML) based (audio, video, and image) documents, its performance are not as same as with text document (formalize of meaning from context) [3]. This lack of formation is a leading cause of certain problems based on semantics and the need to adopt an alternative approach that will publish data on the Web, not just in legible format as well as measuring process [2, 4]. To overcome this cause, researchers inherited ontology (structural frameworks for organizing information) that acts on behalf of users. Each user needs to be annotated with ontologies that provide the basic semantic tools to construct the semantic web [5].

- *Muhammad Waseem Chughtai is with the Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia (UTM), Skudai, 81310, Johor Bahru, Johor, Malaysia. Email: mr.chughtai88@gmail.com,*
- *Ali. Bin Selamat is with the Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia (UTM), Skudai, 81310, Johor Bahru, Johor, Malaysia. Email: aselamat@utm.my*
- *Imran Ghani is with the Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia (UTM), Skudai, 81310, Johor Bahru, Johor, Malaysia. Email: imran@utm.my*

2 Ontology

Ontology plays an important role to provide shared knowledge models to semantic-driven applications targeted by semantic web. In the technology stack of the semantic web standards [4], ontology's are called as an explicit layer. Ontology is the vehicle by which we can model and share the knowledge among various applications in a specific domain. In information technology, ontology is the working model of entities and interactions in a particular field of knowledge or practices, such as electronic commerce or "the planning activity." In artificial intelligence (AI), ontology is, according to Tom Gruber, an AI specialist at Stanford University, "the specification of conceptualizations, used to help programs and knowledge sharing in the rights[5]." In this usage, ontology is a set of concepts, such as things, events, and relationships - that are specified in some way (such as specific natural language) to create a vocabulary agreed to exchange information. In IT (Information technology) and CS (Computer Science) the term "Ontology" is an artifact and technical. It is designed for certain area knowledge and real or imaginary modeling.

2.1 Key Terms

Ontology's are part of the W3C standards stack for the semantic web, in which they are used to specify standard conceptual vocabularies in which to exchange data among systems, provide services for answering queries, publish reusable knowledge bases, and offer services to facilitate interoperability across multiple, heterogeneous systems and databases [3]. The key role of ontology's with respect to database systems is to specify a data modeling representation at a level of abstraction above specific database designs (logical or physical), so that data can be exported, translated, queried, and unified across independently developed systems and services. Successful applications to date include database interoperability, cross database search, and the integration of web services [6].

2.2 Ontology Engineering

The term "engineering" has adopted by artificial intelligence (AI) researchers in the early years. The Ontological Engineering refers to all activities concern the terms and concepts implicated in knowledge engineering, the management of knowledge and smart integration of the information for building ontologies [5]. Instantly, ontology is engineered into three categories .i.e., Natural Language Ontology (NLO), Domain Ontology (DO) and Ontology Instance (OI). NLO is the relationship between generated lexical tokens of statements based on natural language, DO is the knowledge of a particular domain and OI is the automatically generated web page behaves like an object [7].

2.3 Ontology Development

Ontology development in ontological engineering is an iterative process based on seven steps activities as shown in Fig. 1.

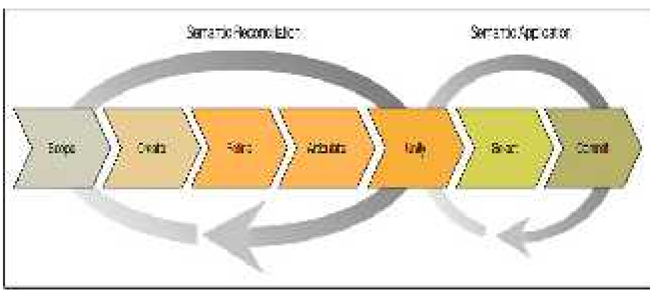

Fig. 1. Ontology development activities [8].

Ontology process of development starts with the determination of the scope by collecting and a list of terms based unorganized information pertinent. Comparison information in the form of classes, such as names, the same properties as verbs containing the values, the constraints, and relationships with other properties based on classes and their subclasses [3]. Conditions are arranged in a top-down or bottom-up class and the hierarchy subclass called taxonomies hierarchy, operates in a manner that the instance of the child class can behave as the instance of the parent class and ownership of the parent class may also be applied to the instance of child class. Size ranges the ontology's with respect to the number of classes and instances, while the number of classes and instances will increase when height of the ontology will rise.

Ontology development process with respect to implementation point of view depends on some currently available ontology supported languages; few of them are discussed in this paper, such as XML, RDF, RDF(S), OWL, OIL and DAML.

3 Ontological Lamnguages

Ontological languages are the formal languages, which are implemented into ontology development or used in particular methodologies. Perhaps, ontology engineering itself is the existence of many methodologies, tools and languages, it is not easy to select an appropriate ontology development technique[9]. Sets of ontology languages, ontology tools, and developing methodologies are presented in Figure 2. One thing need to be notice that not all available methodologies, languages and tools are being selected for ontology development (dark area shows the un-used sources) because some ontology development methodologies were built for specific ontology languages, some ontology development methodologies could be used only with certain tools.

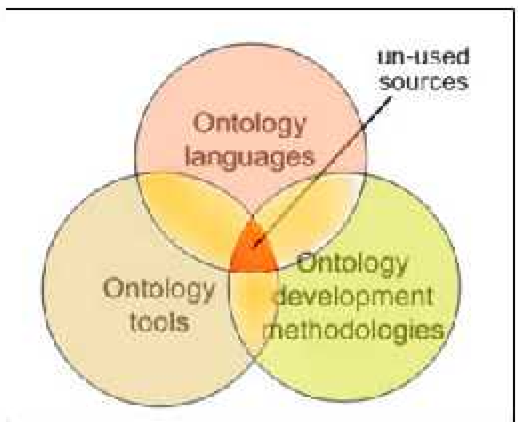

Fig. 2. Ontology development methodologies, tools and languages [9].

Ontology languages in the web environment are called web-based ontology languages, which are based on the particular web standards, for example: OWL pretends to be an extension to RDF for expressing the logical states. This paper also covers some web-based ontology languages (XML, RDF, RDF(S), OWL, OIL, and DAML), its description, and way of working, which are going to compare in the next session of this paper.

3.1 Extensible Markup Language

XML (eXtensible Markup Language) is one of the essential contributions towards middleware technologies [10]. It's a meta markup language that allows the sharing of information among various applications throughout the markup, structure and processing. As the important contribution towards the semantic web supplies XML serialization syntax and the abbreviations for information modeling using data type definitions [3], but XML schema is limited and can be used for structured documents, as it does not provide semantics, names and arbitrary executions structuring the elements.

XML provides the required tags created for a particular type of document. The only constraint is that a well-formed and valid XML document conforms to a schema i.e. it requires the XML file adhere to a pre-defined grammar [11]. The XML schema provides a syntactic description of data, its schemas are used a restricted tag set to describe the structure of an XML document and are syntactically constrained to form a balanced tree (i.e., every starting tag must have a corresponding closing tag). Perhaps, RDF can represent semantics more efficiently than XML [12].

3.2 Resource Description Framework

The Resource Description Framework (RDF) [13] is a W3C standard for describing Web resources, such as the title, author, modification date, content, and copyright information of a Web page. It is a flexible, extensible architecture to represent metadata information about World Wide Web resources [14]. RDF provides a common framework for expressing web resource metadata so that it can be automatically exchanged between computers without loss of meaning [15]. It was originally created in 1999 as a standard on top of XML for encoding metadata. The most exciting uses of RDF are not in encoding of information about web resources, but information about things and relations between them in the real world like people, places, concepts, etc. This mechanism for describing resources is a major component that has been proposed in the W3C's semantic web activity. This activity is an evolutionary stage of the world wide web (W3) in which automated software can store, exchange, and use machine-readable information distributed throughout the Web, in turn enabling users to deal with the information with greater efficiency and certainty.

RDF contains a model and XML syntax for representing information in a way that allows programs to understand the intended meaning. RDF's simple data model and ability to model disparate and abstract concepts has also led to its increasing use in knowledge management applications unrelated to Semantic Web activity [11].

3.3 Resource Description Framework Schema

RDFS has developed by the W3C. It is an extensive term of RDF that can be used to formally describe an ontology schema as a set of classes (resource types) and their properties; to specify the semantics of these classes and properties; to establish relationships between classes, between properties and between classes and properties; and to specify constraints on properties. In RDF(S) ontology, relationships between classes and properties are created by specifying the domain of a property. Thereby constraining the class or set of classes to which a given property may be applied. However, the structure of an RDF(S) ontology consists of a system of elements or slots whose possible range of values may or may not be established by the ontology. RDF(S) does provide for establishment of a controlled vocabulary (or vocabularies) within the structure of the ontology: specifying the range of a property stipulates that any value of that property must be an instance of a particular class of resources (e.g., the class Literal).

An RDF(S) ontology is further distinguished from a traditional thesaurus in that it does not incorporate a lead-in vocabulary. While it is possible to map natural language synonyms to the appropriate classes or properties in the ontology, this must be accomplished through a domain lexicon that is external to the ontology itself.

3.4 Web Ontology Language

OWL (Web Ontology Language) [16] derived from American DARPA Agent Markup Language (DAML) and based on ontology. It is a standard ontology language for the semantic web and an extension of RDF Schema [13]. OWL is compatible with early ontology languages and provides the engineer more power to express semantics [9]. OWL describes knowledge in terms of classes and properties; it has based on description logics and decidable reasoning. There are three different categories or sub-languages of OWL with varying expressively: OWL-Lite, OWL-DL and OWL-Full [16]. TABLE I gives a short description of OWL three sub-languages.

TABLE I. SHORT DESCRIPTION OF OWL SUB-LANGUAGES (OWL-LITE, OWL-DL, AND OWL-FULL)

Sub-Language	**Description**
OWL-Lite	It is the simplest sub-language that intended to be used in those situations where only a simple class hierarchy and simple constraints are needed [16].
OWL-DL	It is much more expressive than OWL-Lite, it is based on description logics that are decidable fragment of First Order Logic, amenable to automated reasoning and increase the possibility rates to compute the classification hierarchy automatically [16].
OWL-Full	It is the most expressive OWL sub-language. It is used in those situations where very high expressiveness is more important than being able to guarantee the decidability or computational completeness of the language [16]. That is why it is not possible to perform automated reasoning on OWL-Full ontologies.

3.5 Object Interaction Language

The Ontology Interchange Language (OIL) is a full-fledged web-based ontology language based on RDFS [17]. The OIL has created according to accomplished the requirements [18], which are as follows: (a) be highly intuitive to the human user; (b) have a well-defined formal semantics with established reasoning properties to ensure completeness, correctness, and efficiency; (c) have a proper link with existing web languages such as XML and RDF to ensure interoperability.

OIL is based on the notion of a class and its super-classes and attributes. The relations between super-classes have been defined as independent entities. The meanings of any expression have been described in a mathematic precise way, which enables reasoning with concept description and the automatic derivation of classification taxonomies. OIL has a well-defined syntax in XML; also, it has defined as an extension of the RDF and its extension schema (RDFS). It offers four levels of complexity layers: Core OIL, Standard OIL, Instance OIL (Standard OIL+RDFS) and Heavy OIL [18].

3.6 DARPA Agent Markup Language

Defense Advanced Research Projects Agency (DARPA) started in 1999 for the creation of machine-readable representations for the Web. The effort worked to create technologies and demonstrations for what is now called the Semantic Web. DARPA Agent Markup Language (DAML) is a primary outcome of this workout [19] with aimed to develop a language and tools to facilitate the concept of the Semantic Web based on RDF. DAML consists of two portions, the ontology language, and a language for expressing constraints and adding inference rules. It also includes mappings to other semantic web languages such as OIL, XML, and RDF [20].

The DAML language is an extension of XML and RDF, in latest release of the language (DAML+OIL) provides a rich set of constructs with which to create machine readable and understandable ontologies and to mark-up information. The ontology language (DAML+OIL) has a well-defined model-theoretic semantics as well as an axiomatic specification that determines the language's intended interpretations [21]. This makes it an unequivocally computer-interpretable language. The DAML Inference Language is a logical language with a well-defined semantics and the ability to express constraints and rules for reasoning.

4 Comparison of Ontological Languages

The related work selection criteria for the comparison are depending upon standards, popularity and specification perspectives. This section presents the comparison of ontological languages that has defined and discussed in this paper. TABLE II demonstrates the comparison as follows:

TABLE II. ONTOLOGICAL LANGUAGES COMPARISON.

Source	Languages	Comparison Points	Criteria
[5]	XML, RDF(S),OIL, DAML, OWL	Authors [5] presents an overview of web-based ontological languages	Authors [5] present the following information about languages: Developers; A particular features; Main purpose of a language; Formality;
[22]	RDF(S), OIL, DAML+OIL	[22] Conceptual basis of the language and external representation of the language	[22] Underlying conceptual basis of the language and External representation of the language
[23]	RDF/XML and Formalized-English	[23] Shows how RDF/XML, Formalized English can be used for knowledge representation cases.	[23] Collections and Quantifier precedence; function calls; deceleration and definition.

5 Conclusion

A number of recent research efforts on service matching were discussed in the previous section; these provide important directions for the semantic matching of services in over-coming the limitations of traditional syntactic approaches to service discovery. A number of languages have been developed based on a particular standard. For example, RDF is used for developing OWL, XML is used for developing RDF, etc. The main advantage of using a layered approach for creating a particular language is improving and refining existing language.

RDF is improved and extended to create OWL. However; each language have its advantages and disadvantages and thus the choice of ontological web-based language should be driven by the purpose what for it will be used. For example; XML encloses data in tags, these tags are relate to the meaning of the enclosed text. The tag vocabulary of XML is not restricted and it enables users to create arbitrary tags that have arbitrary meanings. Because XML tags indicate the content and structure of the data they enclose, they make it possible to do things like archiving and searching. XML tags are extensible and allow writing own XML tags to describe the content.

OWL itself have a lack of supporting tools, instead of it; the choice between OWL-Lite, OWL-DL and OWL-Full may be based upon whether it is important to be able to carry out automated reasoning on the ontology or whether it is important to be able to use highly expressive and powerful modeling facilities such as meta-classes (classes of classes). The main advantage of these sub-languages is that an existing language has been extended by creating a sub-language without necessity to define a new language with a set of new constructs. OIL is impossible to define the default-value, to provide the meta-class and to support the concrete domain. Translating from OIL to RDF and back is no longer guarantee to give an identical ontology from a modeling perspective (though semantic equivalence is still preserved).

In future, the presented research can be refined and extended according to new ontology languages and criteria of comparison.

Acknowledgment

We would like to express our gratitude to Ministry of Science, Technology and Innovation (MOSTI), Malaysia for funding this research under project vote no: 4S028.

References

[1] T. Berners-Lee, "Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor," New York: Harper San Francisco1999.

[2] T. Berners-Lee, et al., "The semantic web," Scientific American, vol. 284, pp. 28-37, 2001.

[3] K. Tochtermann and H. Maurer, "Semantic technologies–An Introduction," Semantic Technologies Showcase the Austrian Situation, pp. 15-20, 2006.

[4] S. B. Palmer, "The Semantic Web: An Introduction," 2007.

[5] J. Pulido, et al., "Ontology languages for the semantic web: A never completely updated review," Knowledge-Based Systems, vol. 19, pp. 489-497, 2006.

[6] D. Fensel, et al., "Ontologies and electronic commerce," Intelligent Systems, IEEE, vol. 16, pp. 8-14, 2001.

[7] B. Omelayenko, "Learning of ontologies for the web: the analysis of existent approaches," in First International Workshop on Web Dynamics in Conjunction with the Eighth International Conference on Database Theory London, UK, 2001, p. 16.

[8] H. Knublauch, "Ontology-driven software development in the context of the semantic web: An example scenario with Protege/OWL," in Proceedings of 1st International Workshop on the Model-Driven Semantic Web, 2004.

[9] D. Kalibatiene and O. Vasilecas, "Survey on Ontology Languages," in Perspectives in Business Informatics Research, ed: Springer, 2011, pp. 124-141.

[10] H. Panetto, "Towards a classification framework for interoperability of enterprise applications," International Journal of Computer Integrated Manufacturing, vol. 20, pp. 727-740, 2007.

[11] X. Wang, et al., "From XML to RDF: how semantic web technologies will change the design of'omic'standards," Nature biotechnology, vol. 23, pp. 1099-1103, 2005.

[12] S. Decker, et al., "The semantic web: The roles of XML and RDF," Internet Computing, IEEE, vol. 4, pp. 63-73, 2000.

[13] G. Klyne, et al., "Resource description framework (RDF): Concepts and abstract syntax," W3C recommendation, vol. 10, 2004.

[14] K. Hutt, "A Comparison of RDF Query Languages," in Proc. of 21th Computer Science Seminar, Hartfort, Connecticut, 2005, pp. 1-7.

[15] F. Manola, et al., "RDF primer," W3C recommendation, vol. 10, pp. 1-107, 2004.

[16] D. L. McGuinness and F. Van Harmelen, "OWL web ontology language overview," W3C recommendation, vol. 10, p. 10, 2004.

[17] J. Broekstra, et al., "Enabling knowledge representation on the web by extending RDF schema," Computer networks, vol. 39, pp. 609-634, 2002.

[18] D. Fensel, et al., "OIL: An ontology infrastructure for the semantic web," Intelligent Systems, IEEE, vol. 16, pp. 38-45, 2001.

[19] N. Suri, et al., "DAML-based policy enforcement for semantic data transformation and filtering in multi-agent systems," in Multi-Agent Systems and Applications III, ed: Springer, 2003, pp. 122-136.

[20] J. R. Hobbs, et al., "A daml ontology of time," online: http://www. cs. rochester. edu/_ferguson/daml/daml-time-20020830. txt, 2002.

[21] I. Horrocks, "DAML+OIL: a Description Logic for the Semantic Web," IEEE Data Eng. Bull., vol. 25, pp. 4-9, 2002.

[22] X. Su and L. Ilebrekke, "A comparative study of ontology languages and tools," in Advanced Information Systems Engineering, 2006, pp. 761-765.

[23] P. Martin, "Knowledge representation in CGLF, CGIF, KIF, frame-CG and formalized-english," in Conceptual Structures: Integration and Interfaces, ed: Springer, 2002, pp. 77-91.