Web Ontology Language Editors for Semantic Web - A Survey

N. Madurai Meenachi Scientific Information Resource Division, Resources Management Group, Indira Gandhi Centre for Atomic Research Kalpakkam M. Sai Baba
Scientific Information Resource Division, Resources
Management Group,
Indira Gandhi Centre for Atomic Research
Kalpakkam

ABSTRACT

Ontology is created based on information provided by domain experts and forms the basis for creating metadata. XML, RDF and OWL are some of the languages which are used to represent ontology. The need for knowledge preservation and representation, ontology construction methodology are briefly described. Available tools for ontology development are listed. Integrated development environment (IDE) tools like OntoStudio, TopBraid, SWOOP and Protégé are analyzed and their performance is studied.

General Terms

Ontology, Semantic web, Web Ontology Language

Keywords

OWL, Semantic Web, RDF, Protégé, SWOOP, TopBraid, OntoStudio

1. INTRODUCTION

The Knowledge sharing is a systematic process for creating, acquiring, synthesizing, sharing and using knowledge and experience, to achieve the organizations goal. Sharing the knowledge, would aid in designing a sound knowledge management system for the organization. , This process involves developing tools which can capture the knowledge and making the process amenable for machine processing. Ontology is hierarchically structured and has a set of terms to describe a domain. By using such structure a knowledge base is built [1]. Ontology is created based on information provided by domain experts and forms the basis for creating metadata. XML, RDF and OWL are some of the languages which are used to represent the ontology. At our Centre we have initiated a program to develop ontology in the nuclear energy domain. Having made a survey of development of ontology in various domains [2], we report in this paper, a survey of web ontology language editors utilized for semantic web development. The need for knowledge preservation and representation is briefly described. Ontology construction methodology is described and available tools for ontology development are listed. Four of such tools, namely Protégé, SWOOP, OntoStudio and TopBraid are analysed.

2. NEED FOR KNOWLEDGE PRESERVATION

Knowledge Management [3-5] is emerging with number of application in various domains. Capturing the knowledge contained in each domain and preserving it for future is one of the prime objectives. The system that is developed to capture the knowledge, involves: identifying, creating, representing and finally distributing. The effort is to employ a suitable mechanism for obtaining and integrating the collective

wisdom from all the members in an organization. It facilitates in preserving both tacit and explicit knowledge and involves converting the available data into a reusable knowledge repository for any future services. It involves the creation of knowledge base by collection of all the expertise, experience and knowledge available within any given organization-Explicit knowledge includes reports, which can be easily described and codified for storage and dissemination. The tacit knowledge on the other hand is highly personal, context specific, and can only be manifested through sharing, applying and practicing and cannot be easily codified. Making available tacit knowledge is vital for growth of any organization. It would result in saving time and efforts in solving various types of issues and problems. Tacit knowledge is accumulated through a process of trial and error while practicing or executing any given taks. Knowledge Management tools are used to capture and transfer of knowledge from the experienced members to the new workforce [6].

3. KNOWLEDGE REPRESENTATION

The unique characteristic of initial effort in knowledge representation is the lack of precision [7]. In the present era this is avoided by enabling the machine to process. To make the machine powerful we have to feed our knowledge to the system (machine) which is being achieved by knowledge representing tools like RDF, OWL, XML, F-logic, ObjectLogic etc.

4. ONTOLOGY CONSTRUCTION

Provision to share and re-use knowledge is what is needed for the knowledge based system. For a large scale domain, ontology is very useful and has the ability to store the relations that are connected. The way the information is expressed needs to be changed as it is not an easy task to teach computers to comprehend natural languages, [8]. Ontology is a set of assertions and relations among the objects for specifying the concepts involved in the specific domains [9]. Ontology development enables [10]:

- Sharing common understanding of knowledge among people and machine
- Reuse of the domain and expert knowledge
- Making the domain assumptions explicit
- Increase interoperability among various domains
- Increase the scalability

There are multiple layered concepts in any knowledge system. Domain areas can often be organized in a tree structure composed of a super-concept, sub-conceptand their relevant relationships. The tree structure for the knowledge of ontology-based search engine facilitates the search of web

knowledge simply by adding structure to the largely unstructured web. In addition, the ontology can be indexed to further facilitate efficient searches. Reasoning on general concepts can be performed by matching various concepts in different expression and relevant classes [11].

Ontology is created from scratch by extracting information from domain experts and merging already existing ontology into new ontology. It forms the basis for syntactic and semantic metadata which can be used for annotating or tagging content [12]. Many of the currently available ontology supported languages are: XML, RDF and OWL. The available tools are: Protégé, Model Futures OWL Editor, TopBraid Suite, OntoLingua, OntoEdit, WebODE, KAON, ICOM, DOE, WebOnto, Medius Visual Ontology Modeler, LinKFactory Workbench, K-Infinity and OntoStudio. Some of these are available as freeware tools. In this paper, we have chosen, Protégé, SWOOP (open source tools) and OntoStudio, TopBraid (commercially available) for the analysis.

4.1 OntoStudio Tool

OntoStudio 3.1.2 is a commercially available engineering work bench which can be installed easily. It consists of classes, properties, rules, queries and mapping. Provision to define properties, range, minimum and maximum values are available. Apart from this ObjectLogic source, view of ontology shows classes and properties are also available. OWL, RDF and ObjectLogic formats are supported in this editor [13]. While creating a project, users have a choice to select the storage type like file based internal repositories or collaboration server. It has the provision to get data from database and also facilitates with chart and report generation. Extraction of the report data in different formats like xml, html, doc, ppt, pdf, postscript is also possible. It is very user friendly for connecting SQL database with a query builder facility. Query Builder SQL is available so that the queries can be created easily.

OntoBroker tool is an added package available along with OntoStudio. It is used for establishing network connection and for setting up the startup configuration. In the OntoBroker ontology construction tools are available to convert OWL to object logic, RDF/XML, Turtle, N-triples, n3, ANSI thesaurus etc. OntoBroker consists of ObjectLogic query and SPARQL for querying the graph. SPARQL is a W3C (World Wide Web Consortium) standard. It is a query language for RDF (SPARQL: SPARQL Protocol and RDF Query Language). OntoBroker only implements a subset of SPARQL (and SPARQL only supports a subset of the ObjectLogic query features). One important feature of OntoBroker is that it can interface with programming languages like .net, java. It is also a multilingual editor and languages like French, English, and German can be used for defining equivalent synonyms.

4.2 Protégé Tool

Protégé is run on a variety of platforms like Windows (32 and 64 bit), Linux (32 and 64 bit), MacOSX, Sun, Solaris, HPUX, IBM. It allows the users to customize interface extensions, incorporates the Open Knowledge-Base Connectivity knowledge model there by interacts with the standard storage formats for example like relational databases, XML and RDF [14]. OWL extends RDFS to allow for the expression of complex relationships between different RDFS classes and of more precise constraints on specific classes and properties (for example, cardinality relations, equality, enumerated classes, characteristics of properties etc).

This tool allows formats like RDF/XML, OWL/XML, OWL Functional syntax, Manchester OWL syntax, OBO 1.2 flat file, KRSS2 syntax, Latex and Turtle (Terse RDF Triple Language). Protégé uses the reasoners like FACT++, Hermit and RACER to fix the inconsistency. The main services offered by reasoner are to test whether or not one class is a subclass of another class. By performing such tests on all classes, it is possible for a reasoner to compute the inferred ontology class hierarchy. Another reasoning service is "consistency checking" which is to check whether it is possible or not for the class to have any instances. The ontology can be sent to the reasoner automatically to compute the classification hierarchy, and also for checking the logical consistency of the ontology. In Protégé, the manually constructed class hierarchy is called the asserted hierarchy, whereas automatically computed by the reasoner is called the inferred hierarchy. While asserted knowledge inference is derived from original data base, that of inferred knowledge, is through additional data that a reasoner (eg. a reasoner that is plugged-in Protégé) will provide from the original data. By maintaining relational database at the backend, enables Protégé to process ontologies which are too large to reside in the memory. Finally, through Protégé's plug-in mechanism, user can generate their own custom import or export plug-ins to work with custom or specialized formats [15]. OWLGrEd is a unified modeling language style graphical ontology editor for OWL and allow graphical ontology exploration and development including interoperability with Protégé [16].

4.3 SWOOP Tool

SWOOP (semantic web ontology editor) is a tool for creating, editing, and debugging OWL and also acts as a web ontology browser and editor. It was developed and released by the MIND lab at University of Maryland, USA in the year 2007, and is currently an open source project with contributors from all over the world [17]. It supports PELLET, RDFS-like reasoner. Advanced features of SWOOP include, possibility of repairing the ontology and re calculate the rank, split the ontology, pellet query, version control and also partitioning the ontology automatically [18].

Currently SWOOP is not active development at Mindswap and is now under the control of SWOOP's Google Code. It is developed using java and can be executed both in Linux and Windows operating system. Multimedia markup plug-ins that facilitates with image and video annotations is supported [19]. While Protégé supports loading of OWL files that are developed in SWOOP, the files developed in Protégé can not be loaded in SWOOP.

4.4 TopBraid Tool

TopBraid Suite is a collection of integrated semantic solution products. All the components of the TopBraid suite, work within an evolving open architecture platform and implemented adhering to W3C semantic web standards. This software was released in 2011 and is available in three forms like free, standard and maestro. It is found that free version has several limitations. Maestro version has many advanced features like its own internal web server for testing application development [20]. It has a flexible and extensible framework with an application program interface for developing semantic client/server or browser-based applications. Some of the concepts are similar to Protégé, like generation of schema based forms for data acquisition. Many features are available like graphical editor which can be used for designing the ontology easily. Moreover cloning of classes and subclasses is also supported.

5. COMPARISON BETWEEN THE INTEGRATED DEVELOPMNET ENVIRONMENT TOOLS

A Comparison of characterestics of IDE tools like Protégé, OntoStudio, TopBraid and SWOOP is summarized in Table 1. All the four IDEs can run both in Windows and Linux operating systems. Protégé runs in Mainframe, Apple and Sun system, indicating its platform independence.

Supporting file format implies that languages in which ontology data are read in and write out. In Protégé read in and write out file formats are starting from flat file, RDF, OWL, Latex, Turtle format whereas in others OWL, RDF, F-logic and objectlogic are supported. Ontology viewed in graphical format to edit and correct the ontology data, OntoGraf, OwlViz plug in are used in Protégé to view the graphical format of the ontology and also to get image of the ontology in jpg, png, gif format. Other integrated development

environment tools use their own built in software for graphical representation. In programs developed using Java provision exists to interface with OntoStudio, TopBraid and SWOOP tools. Protégé is having its own application program interface (API) to integrate with any given application. Multiple language support is available in OntoStudio. Inference Engine is used to recognize any ambiguity about the meaning of the terminology used in the Ontology [21]. Pellet reasoner is supported by all the IDE. Many reasoners use firstorder predicate logic to perform reasoning. One of the difficulties of IDE is that the OWL file created in one editor tool is not compatible with other editor tools even though the file types are the same. Among the four tools discussed in this paper TopBraid is user-friendly. SWOOP and Protégé are open source editors. OntoStudio is commercially available tool with additional features like multi language support, database interface, reports and charts generation and has more flexibility. Protégé tool is also flexible, easy to use, open editor and widely used by the semantic group.

Table -1 Comparison of Integrated development environment (IDE) tools Protégé, OntoStudio, TopBraid and SWOOP

Characteristics	Protégé	OntoStudio	TopBraid	SWOOP
Supporting platform	Windows(32 and 64 bit), Linux (32 and 64bit), MacOSX, Sun, Solaris, HPUX, IBM	Windows XP, Vista, Windows 7, WindowsServer2003, SUSE Linux 10.x	Windows, Linux Macintosh (32 and 64 bit)	Windows, Linux
Supporting File Formats	RDF/XML, OWL/XML, OWL Functional syntax, Manchester OWL syntax, OBO 1.2 flat file, KRSS2 syntax, Latex and Turtle (Terse RDF Triple Language)	OWL, RDF, F-logic ObjectLogic	OWL, RDF, turtle, n-triple, xml	OWL, RDF, XML, text, SWOOP ontology object files
Graphical representation	OntoGraf,OwlViz, OWLGrEd for UML	Supported	Supported	Not available
Query Supports	DL query	SPARQL, ObjectLogic query, Query Builder	SPARQL, SPIN are supported	Pellet Query
Reasoner	Fact,Fact++,Hermit ,Pellet	Pellet	Pellet,Jena, Jena built in reasoned,TopSPIN, SwiftOWLIM	Pellet, Racer, Fact
Programming languages interface	PROTÉGÉ API, Protégé Script Tab	Interface with .net, Java program	Adobe Flex, HTML and Javascript: used with SPARQL Web Pages (SWP), Java Web service API is implemented using SPARQLMotion	Interface with Java program
Conversion from one form to another	RDF/XML, OWL/XML	ObjectLogic to F_logic, Ontology to object logic, RDF/XML, Turtle, N-triples, n3, ANSI thesaurus	RDF, spread sheet, share point, Graphs are supported	Not available
Availability	Freeware	Licensed version, license for three months trail is available	Free, standard, Maestro versions are available	Freeware
Multi user support	Supported	Not Supported	Supported	Not Supported
Consistency check	Supported	Supported	Supported	Supported

Characteristics	Protégé	OntoStudio	TopBraid	SWOOP
Language supported to define synonyms	English	English, French and German	English	English
RIF (Rule Interchange Format)	Not Supported	Supported	Supported Jena rule engine	Not Supported
Report generation support	Not Available	Available. Using Business Intelligent Reporting Tool	Available	Not available
Chart Generation support	Not Available	Available	Available	Not available
Graphical editor support	Not Available	Available	Available	Not available
Implemented using	Java	Eclipse IDE	Eclipse IDE, Jena, Apache	Java
Ontology version comparison facility	Available	Not available	Available	Available
Developed and Supported by	Stanford university	OntoPrise	TopQuadrant	MINDSWAP research group
Database support	Not Supported	Microsoft SQLServer 2005 and 2008; Oracle 10g and 11g; DB2 8.1.2, 8.2,9.0 and 9.5; JDBC	Supports Oracle data base	Not Supported
Ontology difference	PromptDiff plugin-to compare the ontology version	Not Available	Not Available	Not Available
Split ontology	Not Available	Available	Available	Available
UML representation	Using OWLGrEd plug-ins	Supported	Supported	Not Supported
Root node	Thing	owl:Thing	owl:Thing	owl:Thing
Last stable version	Protégé 4.1	V3	V3.6.1	V2.2.1
Version control	Not Availbale	Available	Available	Available
Documentation	Very Good	Good	Good	Medium
Dynamic help facility	Not Available	Not Available	Available	Not Available
Syntactic validation	Not Supported	Not Supported	Supported	Not Supported
Cloning of concepts	Not Supported	Not Supported	Supported	Not Supported

6. NUCLEAR REACTOR DOMAIN KNOWLEDGE REPRESENTATION

A sustained controlled nuclear fission reaction is a reactor. The heat generated in nuclear reactor is extracted by a coolant and exchange the same to water for generating steam which would run the turbine for generating electricity. In addition to generating electricity, nuclear reactors are used for producing isotopes for medical purposes. They are also used for carrying out research on materials. Preserving the operational, commissioning and decommissioning experience of nuclear reactor is essential for continued effective use of enormous potential of this form of energy. In our group we have initiated a program of creating knowledge representation in the nuclear reactor domain. Based on the analysis of various tools available we are employing Protégé tool for the ontology development.

7. SUMMARY

Development of knowledge management system requires employing tools which would aid in capturing the knowledge and amenable for machine processing, which requires development of ontology. XML, RDF and OWL are some of the languages which are used to represent ontology. The methodology of construction of ontology is described briefly and available tools for ontology development are listed. Four of the available tools, Protégé, SWOOP, OntoStudio and TopBraid, are analyzed. Protégé and SWOOP are open source tools and OntoStudio and TopBraid are commercially available. We are pursuing Protégé for ontology development in the nuclear energy domain,

8. ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the encouragement of Shri S.C. Chetal, Director, IGCAR. We also acknowledge the colleagues of SIRD, RMG for their participation and support.

9. REFERENCES

- [1] Swartout Patil, R., Knight, K., and Russ, T. 1996. Toward distributed use of large-scale ontologies, Proceedings of the Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop.
- [2] Madurai Meenachi, N. and Sai Baba, M. "A Survey on usage of ontology in various domains", International Journal of Applied Information and Science (IJAIS), 2012, in press.
- [3] Aegis corporation.1995. Introduction to Knowledge Management and Knowledge Base.
- [4] Leonard-Barton, D. and Sensiper. S.1998. "The role of tacit knowledge in group innovation, California management review, 112-125.
- [5] IAEA., 2006. Knowledge Management for Nuclear Industry Operating Organizations, IAEA-TECDOC-1510, 1-185.
- [6] William Hsu, Ricky K. Taira, Suzie El-Saden, Hooshang Kangarloo,and Alex A. T. Bui, "Context-Based Electronic Health Record: Toward Patient Specific Healthcare", IEEE transactions on information technology in biomedicine, 228-234.
- [7] Gordon McCalla. and Nike Cercone .1983. "Introduction: Approaches to knowledge representation". IEEE computer, 12-18.
- [8] Berners Lee, T. and Berners Lee, R.M. 1998. Semantic Web Road map.URL: http://www.w3.org/DesignIssues/ Semantic.html.
- [9] Tom Heath and Enrico Motta. 2007. "Ease of interaction plus ease of integration: Combing Web2.0 and the Semantic Web in a reviewing site", Web semantic: science, device and agents on the World Wide Web, 76-83.
- [10] Bilab K.Sarker., Peter Wallace., and Will Gill. 2008. Some observation on MindMap and ontology Building Tools for Knowledge Management.
- [11] Lizhen Li and Zhifeng Dong. 2010. Ontology of general concept for Semantic Searching. In Proceedings of the Second International Conference on Computer Modeling and Simulation.

- [12] Sheth, A., Bertram, C., Avant, D., Hammond, B., Kochut, K., and Warke, Y. 2002. "Managing semantic contents for the web", IEEE Internet computing, 80-87.
- [13] Moritz Weiten. 2009. OntoStudio as a Ontology Engineering Environment, Semantic Knowledge Management, 51-60, Springer.
- [14] John H. Gennaria, Mark A. Musenb, Ray W. Fergersonb, William E. Grossod, Monica Crubezyb, Henrik Erikssonc, Natalya F. Noyb, and W. Samson. 2003."The evolution of Protégé: an environment for knowledgebased systems development", International Journal of Human-Computer Studies,89-123.
- [15] Daniel L. Rubin., Natalya F. Noy., and Mark A. Musen.2007."Protégé A Tool for Managing and Using Terminology in Radiology Applications", Journal of Digital Imaging, 113-117.
- [16] Janis Barzdins., Guntis Barzdins., Karlis Cerans., Renars Liepins., and Arturs Sprogis.2010. Ontology Editor for Compact UML-style OWL Graphic Notation, http://owlgred.lumii.lv/.
- [17] Semantic Web Ontology Editor. 2007. http://code.google.com/p/swoop/.
- [18] Aditya Kalyanpur., Bijan Parsia., Evren Sirin., Bernardo Cuenca Grau., and James Hendler.2006." Swoop: A Web Ontology Editing Browser", Web Semantics: Science, Services and Agents on the World Wide Web", 144–153.
- [19] Aditya Kalyanpur., Evren Sirin., Bijan Parsia., and James Hendler.2004."Hypermedia Inspired Ontology Engineering Environment: SWOOP".In Proceedings of Third International Semantic Web Conference.
- [20] TopBraid Composer Getting Started.2011. http://www.topquadrant.com/docs/marcom/TBC-Getting-Started-Guide.pdf
- [21] Akbar Siami Namin and Weiming Shen.2005. "Web Services Agent-Based Model for Inter-Enterprise Collaboration", International Federation for Information Processing Digital Library, Emerging Solutions for Future Manufacturing Systems, 231-240.