Indian Journal of Science and Technology, *Vol 9(16), DOI: 10.17485/ijst/2016/v9i16/88524, April 2016*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Development and Visualization of Domain Specific Ontology using Protege

Usha Yadav[1], Gagandeep Singh Narula[2]*, Neelam Duhan[3], Vishal Jain[4] and B. K.Murthy[1]

[1]Centre for Development of Advanced Computing (CDAC), B-30 Noida - 201307, Uttar Pradesh, India; ushayadav@cdac.in, bkm@cdac.in

[2]Computer Science and Engineering, Centre for Development of Advanced Computing (CDAC), B-30, Noida - 201307, Uttar Pradesh, India; gagan.narula87@gmail.com

[3]Department of Computer Engineering, YMCA University of Science and Technology, Faridabad - 121006, Haryana, India; neelam_duhan@rediffmail.com

[4]Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi -110063, India; vishaljain83@ymail.com

Abstract

Background/Objectives: The research aims to explore differences among various ontology development tools, its languages; finally developed and visualized ontology on specific domain. **Methods/Analysis**: Railway Enquiry System (RES) ontology is being developed with the help of Protege tool and visualized using TGViz tab. It involves creation of various classes and their instances so that a person can find references to its query. **Findings**: The following manuscript makes readers aware of concept of Semantic Web because the search performed by today's search engines is based on keyword extraction technique which leads to irrelevant and incomplete results marked with low precision and high recall. Developed ontology depicts real world scenario of railway reservation system. With this ontology, a person can check its seat availability, train fare details, PNR status and many more. **Improvements/Applications**: The given ontology can be extended to develop railway tracking web based application using Web Ontology Language (OWL) and Semantic Web Rule Language (SWRL).

Keywords: Ontology, Ontology Tools and Languages, Protege, Semantic Web

1. Introduction

World Wide Web (www) is a distributed repository of millions of documents which covers wide range of multidisciplinary information; to extract and retrieve particular information among these documents is a cumbersome job. There are two confusing terms associated with extraction and retrieval. Information Retrieval specifies retrieving information from millions of documents irrespective of documents are relevant or not while Information Extraction specifies extraction of information from relevant documents. WWW is the largest information construct that has gained various advancements ranging from web 1.0 to web 4.0. Web 1.0 is first generation of web that is read only and static web[1]. Web 2.0 is second generation of web and known as Social and Read/Write web[2]. Web 3.0 is considered as third generation of web and is known as Semantic Web (SW)[3]. Till this, machines are not clever as they perform tasks on basis of user input requirements. Web 4.0 is fourth generation of web and is known as Symbiotic Web. It will make machines to think in an intelligent way by reading contents of web and producing that information which loads the website faster[4].

In order to increase degree of relevance, there is need to move towards Semantic Web (web 3.0) and ontology. In broad terms, Semantic Web is known as Global Information Mesh which consists of annotated documents represented in language friendly to humans as well as machines. It curtails the gap between humans and machines. Ontology represents relationship among classes, properties and

*Author for correspondence

instances in hierarchical fashion. Table 1 illustrates the differences among various generations of web. The paper is organized as follows: Section 2 presents brief information about Semantic Web and its layout. Section 3 explicitly defines ontology ranging from its components to development tools and languages. In addition to this, a comparative study has also been described among various development tools and languages. Section 4 presents case study on Railway Enquiry System (RES) and its ontology is being developed with the help of Protege tool.

2. Semantic Web (SW)

The idea of SW was given by the inventor of www-Tim Berners Lee in 1996 that targets to convert present information into machine friendly language[5]. In simple words, it is termed as repository of information and languages involved for presenting such information.

2.1 Architecture

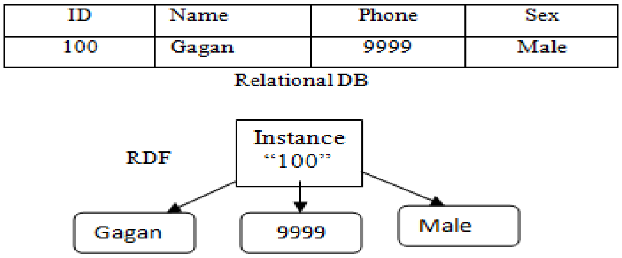
Its layout consists of following components:

- Unicode and URI - Unicode represent each character uniquely and provide intellectual style while URI is Uniform Resource Identifier that represents data in syntactical format.

Table 1. A comparison among various generations of Web

S.No	Web 1.0.	Web 2.0.	Web 3.0.	Web 4.0.
1.	Reading	Reading/ Writing	Read-write-execute or portable personal web	Read-write-exec-concurrency
2.	Focus on companies	Focus on communities	Focus on lifestream	Focus on communities and lifestreams.
3.	HTML	XML, RRS, Wikis	RDF, RDFs, OWL	Middleware (WebOS)
4.	Web forms	Web applications	Smart applications	Middleware and parallelized services
5.	Netscape	Google, Wikipedia	Dbpedia	--------
6.	It is like crawling	It is like walking	It is like running	It is running in highly supervised and intelligent way under supervision.

- XML- It stands for Extensible Markup Language that consists of namespaces and schemas to define structure of data on web.
- Resource Description Framework (RDF) - It is used for describing information in form of data models which in turn consists of triples viz. Subject, Predicate and Property. Example of RDF is given in Figure 1
- RDFs - It stands for RDF Schema that acts as vocabulary language to represent and inference RDF data models.
- Ontology - It is defined as set of terms used to describe given domain and derive inferences from it.
- Logic and Proof - In this layer, agents can make inferences in finding requirements of given resources with the help of inference systems[6].
- Trust - It signifies assurance and degree of loyalty to information[7]



ID	Name	Phone	Sex
100	Gagan	9999	Male

Figure 1. Example of RDF.

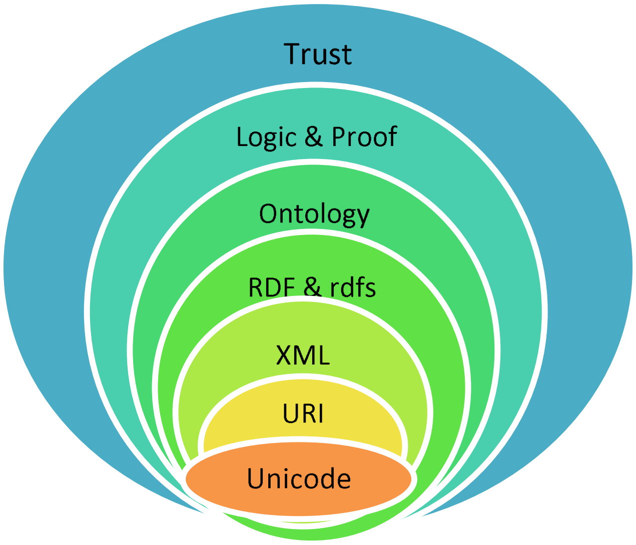


Figure 2. Stack Venn diagram of Semantic Web Architecture[7].

3. Ontology

The word Ontology is derived from two Greek words – onto that means "being" and logia which means "written or spoken discourse". Ontology has wide range of definitions ranging from philosophy to artificial intelligence. Ontology is abbreviated as FESC which means formal, explicit, specification of shared conceptualization[8].

3.1. Components of Ontology

- **A set of concepts**
  These can be the nodes in the representation of ontologies.
- **A set of properties**
  Every node or a concept or a class may or may not have properties related to it, properties can also be summarized as the values of the concepts.
- **A set of relational properties**
  It implies relationship between two or more concepts or nodes. This generally generates a hierarchical way from one concept to another.
- **Hierarchy of concepts**
  Sub concept/super concept relationships.
- **Hierarchy of properties**
  Sub-property/super-property relationship.
- **A subset of symmetric properties**
  It defines set of properties in a concept that have same values and same functionality.
- **Transitive property relation**
  Transitive relation is defined as, if property A is related to property B and property B is related to property C then property A will be necessarily related to property C.
- **Symmetry and Inverse Symmetry relations among properties**
- **Domain values related to properties**
- It defines the class n the level of the properties; concepts that share same property values have same domains.
- **Range values related to properties**
- Range is a characteristic of the concepts, which can be an interval, a list of elements or simply a character.
- **Minimum and Maximum cardinality for each concept-property pair**
- In Set theory cardinality is said to be the number of elements in a set, in this concept cardinality is a positive number that is associated with each concept and showing that how many properties are associated with that concept. Maximum and minimum cardinality is the range, discussed above, of the properties associated with any concepts.

3.2. Basic Steps for Building Ontologies

- Determine Scope:- It includes defining structure and values associated with ontology.
- Consider re-using:- Recent ontologies can be re-used for defining schema of new ontology.
- Enumerate terms:- Clearly specify all the terms that specifies domain and range of ontology in structured list.
- Define taxonomy:- After specifying terms it is necessary to organize them in hierarchical fashion. If A is subclass of B, then every instance of A must be an instance of B.
- Define properties:- It is most important step to organize the properties that link the classes while organizing these classes in a hierarchy.
- Define facets:- The ontology will only require the expressivity provided by RDF Schema and does not use any of the additional primitives in OWL.
- Define instances:- Ontologies are being used to organize sets of instances[9].

3.3 How to use Ontology

Usage of ontologies depends on number of levels assigned.

Level 1: As vocabulary language for interacting among multi agents in distributed scenario.

Level 2: Represented as database schema that holds information about classes, properties and instances in it. Data can be retrieved easily from database by accessing its schema.

Table 2. Steps for construction of ontologies

i. Determine Scope
ii. Consider Reuse
iii. Enumerate Terms
iv. Define Taxonomy
v. Define Properties
vi. Define Facts
vii. Define Instances
viii. Check for Anomalies

Level 3: As knowledge base that is created after deriving inferences rules in given ontology.

Level 4: For handling complex queries and datasets.

Level 5: Standardization

- Standardization of structure of ontology.
- Standardization of concepts hierarchy.
- Standardization of domain ontology components.
- Standardization of tasks performed on ontology.

Level 6: For integration of ontologies to different systems like knowledge management, ERP systems, E-learning and many more.

3.4 Ontology Development Languages

Following are types of ontology languages used in Semantic Web.

- LOOM[10]:- It is one of knowledge representation languages that is based on description logics and rules to build concepts automatically.
- SHOE[11]:- It is used to extract relevant information from web documents. It also combines knowledge representation data and ontological features.
- OML[12]:- It stands for Ontology Markup Language that is treated as extension of SHOE.
- XOL[13]:- It stands for Ontology Exchange Language that is based on XML and used for development of ontologies in any tool.
- DAML+OIL[14]:- DAML stands for DARPA Agent Markup Language and OIL stands for Ontology Interchange Language. It is used for achieving semantic interoperability among various resources.
- CycL[15]: - It is one of formal languages that use predicate logic to define concepts in domain. It comes under category of generic ontologies.

Table 3. A Comparison among ontology development languages

Features	LOOM	SHOE	OML	XOL	DAML+OIL
Concept documentation.	Yes	No	yes	No	yes
Instance attributes	Yes	yes	yes	yes	yes
Class attributes	yes	No	yes	yes	Yes
n-ary relations	yes	Yes	yes	No	No
Cardinality constraints	yes	No	No	No	yes
Concept instances	yes	yes	yes	yes	yes
Rules	yes	yes	yes	no	no

3.5 Ontology Development Tools

In general, ontology development includes phases like specification, design and formalization phases. All these phases are treated as SDLC phases[16]. Table 4 lists differences among various ontology editors[17].

Table 4. A comparison among various ontology editors

Tool	Version	Owner / Developer	Features / Limitation	Primary Language	FOSS (free open source software)
Adaptiva	–	Sheffield University	Knowledge Acquisition	Java	Yes
Semantic-Works 2008	2008 sp1	Altova	OWL+RDFS Editor	Java	No
Conzilla2	2.2	Knowledge Management Research Group	Concept Browser	Java	Yes
HOZO	5.01	Osaka University	Role concept; User-friendly	Java	Yes
OWL Editor	0.2.0.36	Model Futures	Tree-based	Other	Yes
Onto-Track	–	Ulm University	Fast browsing & Easy editing	Java	Yes/No
OWL-S Editor	23	Linkoping University	Semantic Web Services	Java	Yes
Protégé	3.4 beta	Stanford Medical Informatics	Multiple Inheritance	Java	Yes
SWOOP	2.3 beta	MINDSWAP	Web-browser look & feel	Java	Yes
Web Onto	–	Open University	Knowledge Modelling	Java	Yes

Besides this, there are various versions of Protégé like 2000, 3.1, 3.2, 3.4, 3.4 beta, 4.0, 4.0 beta and 5.0 desktop. Table 5 lists differences between most common versions of Protege [18].

4. Case Study

The paper presents Railway Enquiry System (RES) ontology that describes terms involved in a railway reservation system. A person can see the train or can see the seat availability or also can see the fare, but a person can't book the ticket.

Developed ontology is partial (as it only shows the terms used in ontology) that describes real-world phenomena – Railway Enquiry System (RES).

4.1 Screen shots

Tool used: Protege 3.4 beta. It is created at Stanford University[19] and acts as an open-source knowledge requisition system that is written in Java[20].

In Figure 3, Railway Enquiry System is marked as super class and it consists of various sub-classes like Fare Enquiry, Find Your Train, PNR Status and Seat Availability. Fare Enquiry class is futher divided into classes like CLASS,Concession, Train Number etc.

Figure 4 displays slots of one of classes named CLASS under Fare Enquiry of RES ontology. It holds type of values in CLASS whether it is AC Chair, First AC, Second AC etc.

Table 5. Differences between 3.4.1 and 3.4 beta

Features	Protégé 3.4.1	Protégé 3.4 beta
Compression Algorithm for Client server communication	Yes	No
Memory leaks in database mode	Yes	No
Inheritance of browser slot Patterns by subclasses	No	Yes
OWL file to OWL database conversion	Slow	Fast
Debug and performance	No	Yes
Support for Derby Database	No	Yes
Protégé script console support for manipulation of ontologies	No	Yes
Database inclusion	No	Yes

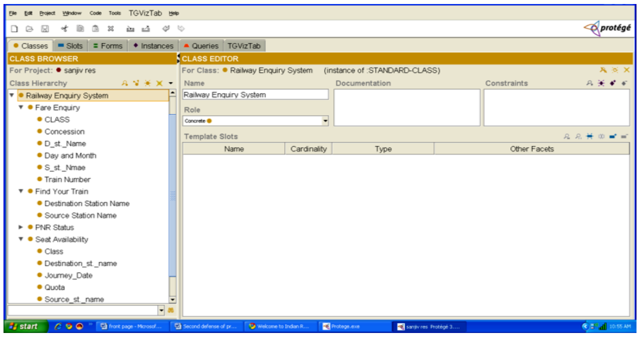
Figure 5 displays references of given ontology like Fare Enquiry class is direct super-class of CLASS which further has instances AC Chair Car, First AC and so on.

Figure 6 displays classes corresponding to RES ontology in form of graph by using TGViz tab. TGViz stands for Touch Graph Visualization tab that visualizes classes and instances in developed ontology.

4.2 Code Snippet

RDF/XML source code

```
<?xml version='1.0' encoding='UTF-8'?>
<!DOCTYPE rdf:RDF [
  <!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-
    syntax-ns#'>
```

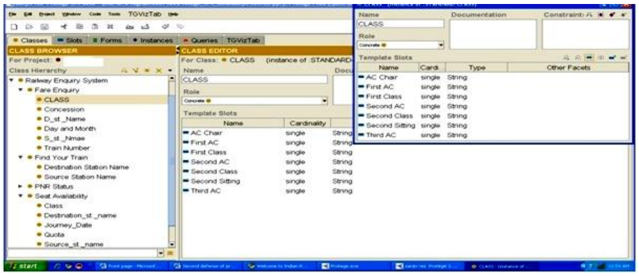


For Class: Railway Enquiry System (instance of :STANDARD-CLASS)

Name: Railway Enquiry System

Role: Concrete

Figure 3. Super class-sub class hierarchy of RES ontology.



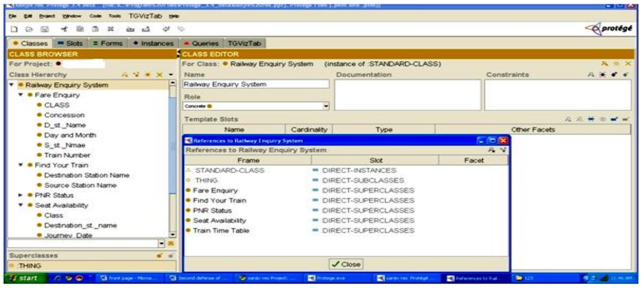
For Class: CLASS (instance of :STANDARD-CLASS)

Name: CLASS

Role: Concrete

Name	Cardinality	Type	Other Facets
AC Chair	single	String	
First AC	single	String	
First Class	single	String	
Second AC	single	String	
Second Class	single	String	
Second Sitting	single	String	
Third AC	single	String	

Figure 4. Slots class "CLASS" of RES ontology.



For Class: Railway Enquiry System (instance of :STANDARD-CLASS)

Name: Railway Enquiry System

Role: Concrete

References to Railway Enquiry System

Frame	Slot	Facet
:STANDARD-CLASS	DIRECT-INSTANCES	
:THING	DIRECT-SUBCLASSES	
Fare Enquiry	DIRECT-SUPERCLASSES	
Find Your Train	DIRECT-SUPERCLASSES	
PNR Status	DIRECT-SUPERCLASSES	
Seat Availability	DIRECT-SUPERCLASSES	
Train Time Table	DIRECT-SUPERCLASSES	

Figure 5. References of RES ontology.

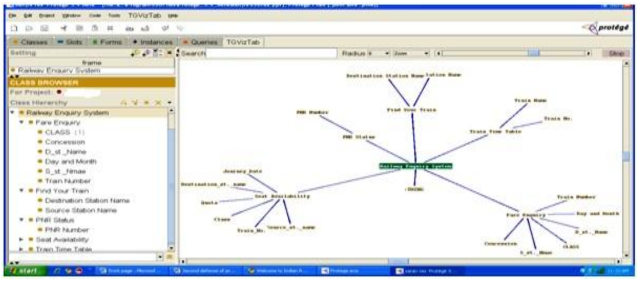


Figure 6. Graph corresponding to RES Ontology using TGViz Tab.

```
    <!ENTITY a 'http://protege.stanford.edu/system#'>
    <!ENTITY rdf_ 'http://protege.stanford.edu/rdf'>
    <!ENTITY rdfs 'http://www.w3.org/2000/01/rdf-schema#'>]>
<rdf:RDF xmlns:rdf="&rdf;"
     xmlns:rdf_="&rdf_;"
     xmlns:a="&a;"
     xmlns:rdfs="&rdfs;">
<rdf:Property rdf:about="&rdf_;AC_Chair"
     rdfs:label="AC Chair">
    <rdfs:domain rdf:resource="&rdf_;CLASS"/>
    <rdfs:domain rdf:resource="&rdf_;Class"/>
    <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&rdf_;Agra"
     rdfs:label="Agra">
    <rdfs:domain rdf:resource="&rdf_;D_st._Name"/>
    <rdfs:domain rdf:resource="&rdf_;Destination_Station_Name"/>
    <rdfs:domain rdf:resource="&rdf_;Destination_st._name"/>
    <rdfs:domain rdf:resource="&rdf_;S_st._Nmae"/>
    <rdfs:domain rdf:resource="&rdf_;Source_Station_Name"/>
<rdfs:domain rdf:resource="&rdf_;Source_st._name"/>
    <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&rdf_;Ashram_Exp."
rdfs:label="Ashram Exp.">
    <rdfs:domain rdf:resource="&rdf_;Train_Name"/>
    <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&rdf_;Bharat_Scout_Guide"
     rdfs:label="Bharat Scout Guide">
    <rdfs:domain rdf:resource="&rdf_;Concession"/>
    <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdf:Property rdf:about="&rdf_;Blind_Concession"
    rdfs:label="Blind Concession">
    <rdfs:domain rdf:resource="&rdf_;Concession"/>
    <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
<rdfs:Class rdf:about="&rdf_;CLASS"
     rdfs:label="CLASS">
    <rdfs:subClassOf rdf:resource="&rdf_;Fare_Enquiry"/>
</rdfs:Class>
<rdf:Property rdf:about="&rdf_;MD3"
     rdfs:label="MD3">
    <rdfs:domain rdf:resource="&rdf_;Train_No."/>
    <rdfs:domain rdf:resource="&rdf_;Train_No._"/>
    <rdfs:domain rdf:resource="&rdf_;Train_Number"/>
    <rdfs:range rdf:resource="&rdfs;Literal"/>
</rdf:Property>
```

5. Conclusion and Future Scope

Ontology is treated as main constituent of Semantic Web that allows explicit well defined understanding of concepts among agents and analyzes domain knowledge. The paper firstly describes evolution of www from web 1.0 to web 4.0. Concept of Semantic Web and ontology is being described. In addition to this, differences among various ontology development tools and languages are listed. Lastly, the paper presents case study on Railway Enquiry System (RES), defines its classes, properties and instances by developing ontology on Protege 3.4 beta and visualizing it using TGViz tab.

As a future work, knowledge can be extracted from developed ontology by importing in any IDE like Eclipse, NetBeans and IntelliJ etc. with the help of some open source framework like Jena and Sesame. A user GUI can be designed which helps in document classification[21] as well as promoting E-learning with the help of Semantic Web technologies[22].

6. References

1. Getting B. Basic Definitions: Web 1.0, Web.2.0,Web 3.0". 2007. Available from: http://www.practicalecommerce.com/articles/464-Basic-Definitions-Web-1-0-Web-2-0-Web-3-0
2. Murugesan S. Understanding Web 2.0. Journal IT Professional. 2007 Jul-Aug; 9(4):34–41.

3. Palmer SB. The Semantic Web: An Introduction. 2001 Sep. Available from: http://infomesh.net/2001/swintro
4. Hemnath. Web 4.0 - A New Web Technology. 2010 Dec. Available from: http://websitequality. blogspot. com/2010/01/web 40-new-web-technology.html
5. Lee TB. The Semantic Web. Scientific American; 2007 May.
6. Greenberg J, Sutton S, Campbell DG. Metadata: A fundamental component of the Semantic Web. Bulletin of the American Society for Information Science and Technology.2003 Apr/May; 29(4):16–8.
7. Lee B, Lassila J. Ontologies in Semantic Web. Scientific American; 2001 May. p. 34–43.
8. Singh S, Jain V. Information Retrieval (IR) through Semantic Web (SW): An overview. Proceedings of CONFLUENCE 2012 - The Next Generation Information Technology Summit at Amity School of Engineering and Technology; 2012 Sep. p. 23–7.
9. Antoniou G, Von F. A Semantic Web Primer. London, England: The MIT Press Cambridge; 2000.
10. MacGregor R. Inside the LOOM classifier. SIGART bulletin. 1991; 2(3):70–6.
11. Luke S, Heflin J. SHOE 1.01 Proposed Specification. Parallel Understanding Systems Group. 2000 Feb. Available from: http://www.cs.umd.edu/projects/plus/SHOE/ spec1.01. html
12. Kent RE. Conceptual Knowledge Markup Language. Proceeding of the Twelfth Workshop on Knowledge Acquisition, Modelling and Management; Banff, Alberta, Canada.1999 Oct.
13. Karp R, Chaudhri V, Thomere J. XOL. An XML-based Ontology Exchange Language (version 0.4). Artificial Intelligence Center. SRI International: 1999 Aug.
14. Horrocks F, Van H. Reference description of the DAMLOIL (March 2001) Ontology Markup Language. 2001. Available from: http://www.daml.org/2001/03/ reference.html
15. Corcho O, Fernandez-Lopez M, Gomez-Perez A. Methodologies, tools and languages for building ontologies. Where Is Their Meeting Point? Data and Knowledge Engineering. 2003 Jul; 46(1):41–64.
16. Singh G, Jain V, Singh M. Ontology development using Hozo and semantic analysis for information retrieval in Semantic Web. ICIIP-2013 IEEE Second International Conference on Image Information Processing; Jaypee Univ. Shimla. 2013 Dec. p. 9–11.
17. Suresh K, Malik SK, Rizvi SAM. A case study on role of ontology editors in National Conference on Advancements in Information and Communication Technology (NCAICT). CSI, Allahabad. 2008 May 15-16.
18. Sivakumar R, Arivoli PV. Ontology visualization protege tools - A Review. IJAIT. 2011 Aug; 1(4):1–11.
19. Slimani T. Ontology development: A comparing study on tools, languages and formalisms. IJST. 2015 Sep; 8(24):1–12. Doi: 10.17485/ijst/2015/v8i34/54249. ISSN 0974-5645.
20. Jain V, Singh M. Ontology development and query retrieval using protege tool. IJISA. Hong Kong. 2013 Aug; 5(9):67–75. ISSN No. 2074-9058. Doi: 10.5815/ijisa.2013.09.08.
21. Ramana AV, Reddy EK. OCCSR: Document classification by order of context, concept and semantic relations. Indian Journal of Science and Technology. 2015 Nov; 8(30):1–8. Doi no:10.17485/ijst/2015/v8i1/75398.
22. Manickasankari N, Arivazhagan D, Vennila G. Ontology based Semantic Web technologies in E-learning environment using protege. Indian Journal of Science and Technology. 2014 Oct; 7(S6):64–7. Doi no: 10.17485/ ijst/2014/v7iS6/54591.