Ontology Creation and Development Model

Pallavi Grover, Sonal Chawla

Research Scholar, Department of Computer Science & Applications, Panjab University, Chandigarh, India

Associate. Professor, Department of Computer Science & Applications, Panjab University, Chandigarh, India

ABSTRACT: The continuing need for effective information retrieval, paved way for the inception of semantic web. It's, unquestionably, as extension to the existing web which holds distributed information lacking any logical relationships amongst themselves. This information can be integrated and related to each other using ontologies. The paper intends to bridge this gap by providing a brief idea of ontology creation and how can it help in implementation of semantic web. The paper has threefold objective. Firstly the paper provides multiple definitions of ontology and their relation with in semantic web. Secondly the paper briefly provides a comparison of several ontology development tools justifies the use of a particular ontology tool. Lastly the paper proposes an ontology on data structures.

KEYWORDS: Semi-Formal Ontology, Semantic Web, Ontology, Ontology Tools

I. INTRODUCTION

The idea behind the Semantic Web is to weave a web that not only links documents to each other but also recognizes the meaning of the information in those documents.' (Frauenfelder 2001)[1].This implies that the current web needs to be transformed from being a bank of distributed, non interrelated information to coherent and logically related web. The diagram underneath/aside, as given by Tim Berners Lee, presents a layered structure of such kind of web. Broadly speaking, Uniform Resource Identifier and Unicode form the base layer of Semantic web. Extensible Mark-up Language (XML) provides a platform for data traversal.RDF/RDF-S provides the format in which the data transfer takes place. Above these layers comes Ontology i.e. its web and query language. In the most basic sense, Ontology refers to vocabulary for a language which is RDF complaint and can be used to inter-relate data such that it can exist on Semantic Web.
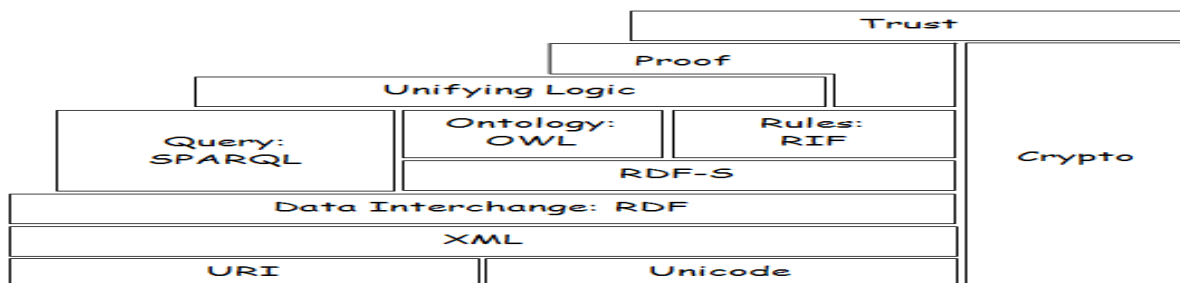

Fig.1 Tim Berners Lee Model for Semantic Web

II. ONTOLOGY

Definitions:

In this section, the most relevant, although quiet different, definitions of the term ontology have been presented. These definitions provide different but complementary point of views of the same reality. The term is borrowed from philosophy, where Ontology is a systematic account of Existence or explanation of being. But in the previous decade and a half this term has become relevant for the knowledge engineering community. Starting from 1991 several definitions have been proposed which have evolved over time. A review of these definitions is followed by explanation of relationships among them.

One of the first definitions for ontology was given by Neches and colleagues (1991), which is as follows: ''an ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary'' [2]. Neches' definition not only provides a description of what

should be done to build ontology but also gives guidelines. It suggests identifying basic terms and relation amongst those terms, providing their definitions as well, there by highlighting the idea of inference of knowledge from it.

A few years later, Gruber (1993) defined an ontology as ''an explicit specification of a conceptualization'' [3]. This definition became the most quoted in literature and most used by ontology community.

In 1995, Guarino and Giaretta collected and analyzed seven definitions and proposed that "A logical theory which gives explicit, partial account of conceptualization" [4] where conceptualization can be thought of as an idea of the world that a person or a group of people can have. It must be noted that this definition highlighted the use of logical theory in building ontology.

Borst (1997) modified Gruber's definition slightly and proposed that: ''Ontologies are defined as a formal specification of a shared conceptualization'' [5]

Later in 1998 Studer merged the definitions given Gruber and Borst and stated that: "An Ontology is a formal, explicit conceptualization. Conceptualization refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine readable. Shared reflects the notion that ontology captures consensual knowledge, that is, it is not private of some individual, but accepted by a group"

Lifecycle: Every ontology goes through certain stages which are also referred as it lifecycle.
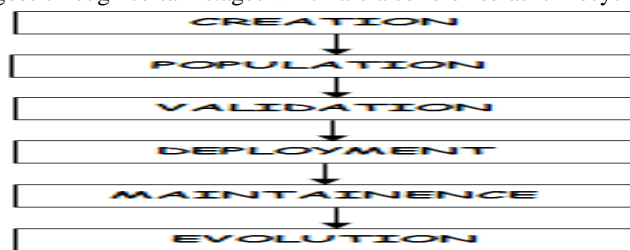

Fig.2 Lifecycle of Ontology Development

III. METHODOLOGY

Several basic questions need to be answered before building ontology. They relate to methodologies tools and languages to be used in the process.

To begin with, one needs to answer if he will be building new ontology altogether or reusing ontologies available on ontology servers. There are several ontology libraries available for reuse which are either associate with ontology language or are attached to ontology tools. One also needs to know the set of activities which will be required to perform and what would be the life-cycle of ontology.

Then one needs to select a tool which allows ontology development, there by taking care of the inference engine, formats of storage, import & export supported by the tool.

The focus now shifts to the ontology language. Its expressiveness, inference mechanisms, compatibility with other languages, the tool which supports that language, need of translators etc are a few parameters that we need to take in account.

Ontology development is a complex and largely domain-oriented process that can be benefited from tool support. As mentioned earlier, there exist a number of tools that for ontology development.

TOOLS AVAILABLE

Ontology development process is very intricate process and principally a domain oriented one.In the recent years, researchers have developed a lot of tools for developing ontology. For example Protégé, SWOOP, Top Braid Composer, OilEd, WebODE, OntoLingua. As defined by N. F. Noy and M. A. Musen [6], ontology development tools allow users to define new concepts, relations and instances. Besides the capability of importing and extending emerging ontologies, development tools may contain some additional features such as graphical browsing, search and constraint checking capabilities [9].There several other feature to every tool that are considered like built-in inference engine, exception handling.

PROPOSED TOOL

There can't be a best tool for ontology development. But definitely there is a tool which is most popular amongst ontology developers. With reference to the survey stated in [9], we conclude that Protégé is one of the most used tools for building ontology despite initially it's a little hard to learn. Out of the 32 ontology developer, 24 claimed to have used Protégé the most with their maximum experience being close to a year.
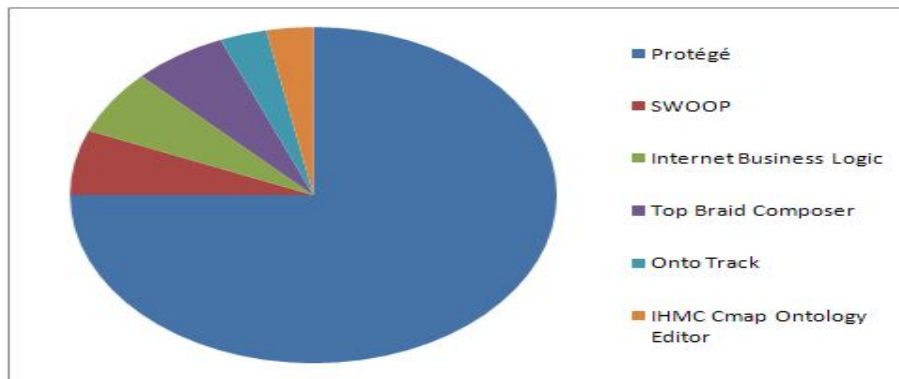

Fig. 3 Most used tool amongst Ontology developers [9]

This is easily available because it is open source tool. There are more than 70,000 registered users of Protégé who are using the system to manage terminologies and ontologies in many different domains.

ONTOLOGY CLASSIFICATION

Ontologies can be classified, as given in the figure 4, based on the strictness of language followed by them.

Highly Informal: As the name suggests, they use language which is least restricted in nature. They're almost expressed in natural language.

Semi-Informal: They are expressed in a restricted and structured form of language.

Semi-Formal: Their expression is more restricted than Highly Informal and Semi-Informal. An artificial and formally defined language is used viz, OntoLingua, OWL etc.

Rigorously Formal: This is the most restricted in nature. They provide meticulously defined terms with formal semantics, theorems, proofs of properties such as soundness and completeness.
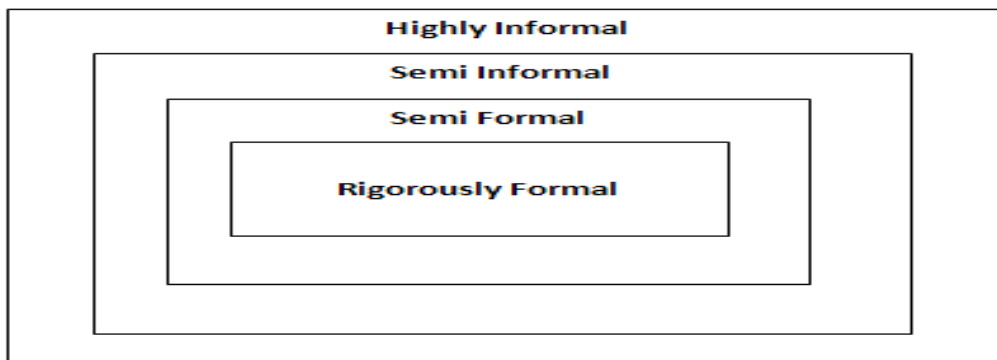

Fig. 4 Classification of Ontologies

ONTOLOGY CREATION

Following the definition of ontology, as given by Neches and colleagues, we need to identify basic terms in ontology and define relations among them. The terms can be related in various ways. The three most commonly used relationships are is-a, synonym, and related-to where the semantics of the first two are well known. ''Related-to'' is used for generic associations between entities.
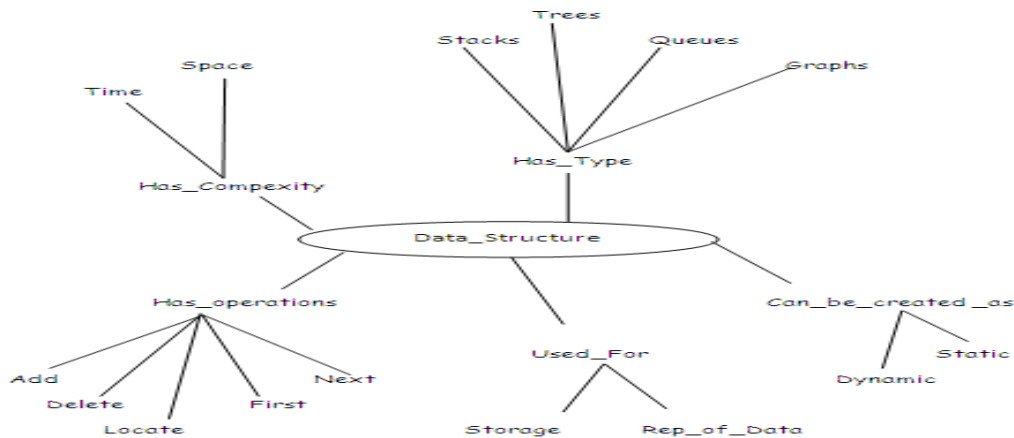
Fig.5 Representation of Highly Informal Data Structure Ontology

The diagram above can be referred as representation of Highly Informal Ontology primarily because this has been expressed in natural language.

The methodology for the creation of domain ontology, as given below, consists of a set of steps.

Step 1: Identification of basic terms

The first step in the development of any ontology is identification of the basic terms and definitions given for them and their properties. Here the main challenge is of completeness which can be addressed by through identification of use cases and by enabling the ontology to evolve.

1.1 Identification of use-cases

These are nothing but those pieces of information or actions which act as a pre requisite for the other action to be performed or for a relationship to exist.

E.g. If I need to represent information through Queues, the pre requisite is existence of information

1.2 Identification of synonyms or related terms

This can either be done online using a thesaurus or manually depending on the domain and type of ontology being built.

 E.g. A tree can be referred to as a Hierarchical Structure.

Tree	Hierarchical Structure

Fig. 6 Representation of synonymous terms

Step2: Identification of relationships

As mentioned earlier, identification of relationship among the basic term in ontology is important. There are three types of relationships.

- Is-a: e.g., Queue "Is-a" Data structure.
- Synonymous: e.g., Tree is "synonymous" to Hierarchical Tree structures.
- Related To: e.g., Storage is "related to" Data Structures

Besides this relationships between ontologies should also be defined especially in case of large domain ontologies which have been divided into sub ontologies.

Step3: Identification of basic constraints

There are four types of constraints that exists

- Pre-Requisite Constraint: One terms existence depends on another. E.g., Data existence is must for the usage of data structures utility.
- Temporal Constraint: A term/relationship must precede the occurrence of other.

- Mutually Inclusive Constraint: One term requires the other term for existence.
- Mutually Exclusive Constraint: Two particular terms can't occur at the same time. E.g., Only one is achievable either Time Complexity or Space Complexity.

Step 4: Identification of higher-level constraints capturing domain knowledge.

Here identification of

- Domain constraints: Constraints on domain terms.
- Domain dependencies: Constraints on multiple terms or relations in a given application. E.g., To perform Depth First Search a Queue can't be used. This is a domain dependency.

Considering all four steps involved a more relevant and refined version of the same representation (Figure 6) can be as given in Figure 7, although that would not be a complete representation of ontology.
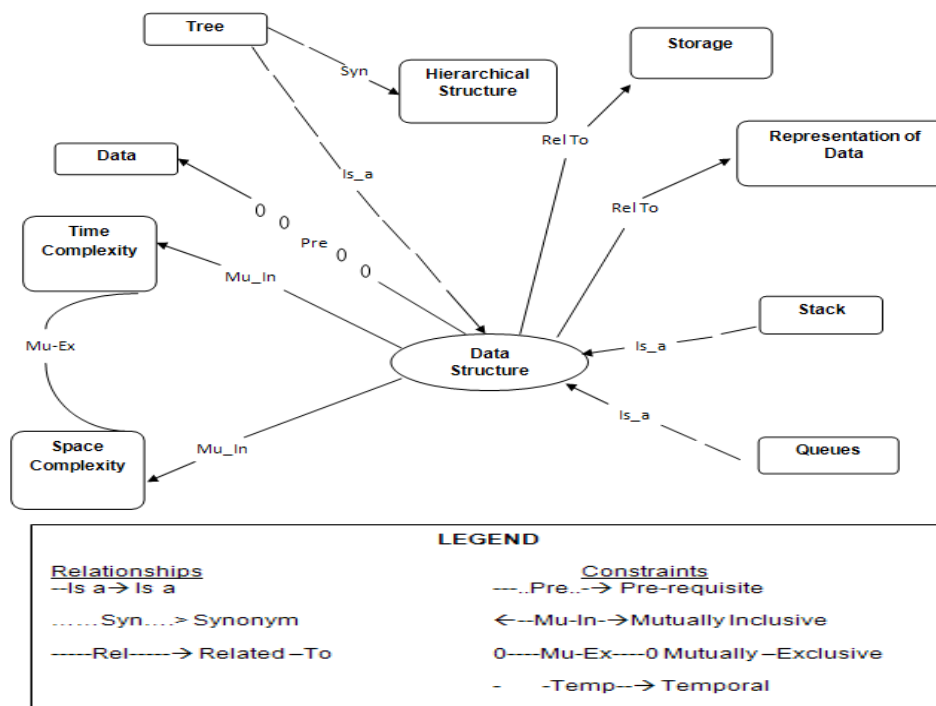

Fig. 7 Partial Representation of Data Structure Ontology

IV. CONCLUSION

Developing ontologies is an important aspect of the Semantic Web. However, to be useful, ontologies must be shared so that there is common understanding among learning object producers about what the terms mean. Since it is likely that different groups of people will use different ontologies for learning objects, mappings between these ontologies are also an important requirement. The Semantic Web opens up a wide range of possibilities for intelligent discovery and reuse of learning objects. By using shared ontologies, it is possible for software agents to perform most of the processing required in discovering and assembling learning objects..

REFERENCES

1. G Kück, "Tim Berners-Lee's Semantic Web" South African Journal of Information Management, vol6 (1) 2004.
2. R. Neches, R.E. Fikes, T. Finin, T.R. Gruber, T. Senator, W.R. Swartout, Enabling technology for knowledge sharing, AI Magazine 12 (3) (1991) 36–56.
3. T.R. Gruber, A translation approach to portable ontology specification, Knowledge Acquisition 5 (1993) 199–220.
4. W.N. Borst, Construction of Engineering Ontologies, PhD Thesis, University of Tweenty, Enschede, NL—Centre for Telematica and Information Technology, 1997.

5. N. Guarino, M. Carrara, P. Giaretta, Ontologies and knowledge bases: towards a terminological clarification, in:N. Mars (Ed.), Towards Very Large Knowledge Bases, Knowledge Building and Knowledge Sharing, IOS Press, Amsterdam, 1995, pp. 25–32.
6. http://en.wikipedia.org/wiki/Ontology
7. Permanand Mohan, Christopher Brooks" Learning Objects on the Semantic Web" Proceedings of The 3rd IEEE International Conference on Advanced Learning Technologies (ICALT'03
8.] Noy, N.F., Musen, M.A.: Evaluating ontology-mapping tools: Requirements and experience In: Proceeding of OntoWeb-SIG3 Workshop, pages 1-14 (2002)
9.] M. Rahamatullah Khondoker, Paul Mueller, "Comparing Ontology Development Tools Based on an Online Survey" Proceedings of the World Congress on Engineering Vol I (2010).
10. http://lstc.ieee.org/wg12/
11. T.R. Gruber, "A translation approach to portable ontologies," Knowledge Acquisition, vol.5, no. 2, pp.1999-220, 1993