ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

Ontology creation using Noun Phrase Approach

Arti V Waghmare, Ujwala Bharambe

Abstract— **In the last decade Ontologies have begun to play more prominent role in knowledge engineering, information system, domain engineering, artificial intelligence, the semantic web and so on. Semantic web technologies relies on Ontology as a tool for modeling an abstract view of real world and contextual semantic analysis of real documents[1].In order to utilize full knowledge available in modern organizations new approaches to Ontology creation are necessary to bridge the ontology gap and enable domain experts to create formalize knowledge. The construction of new Ontology is normally based on predefined requirements which can either be based on the requirements specified for an application to be developed or requirements specified for specific domain of knowledge. Several research groups have proposed various methodologies for building Ontologies. Ontology creation is also considered as a sense making activity. In our work we reviewed the existing methodologies for Ontology creation and present a new approach for novice developer dealing with Ontology. This helps in improving the traditional life cycle of Ontology development mainly applicable to Domain ontology ecreation activity.**

Index Terms—**Ontology, Semantic web, World Wide Web,OWL.**

I. Introduction

Ontology concept originated from the domain of philosophy. In the field of computer, ontology means a specification of a conceptualization in the context of knowledge sharing, which enables the computer to explain and handle the meaning of information. The target of ontology is to acquire the relevant knowledge, provide the common comprehension, and confirm the common approbation vocabulary in the domain. At the same time, ontology give the explicit meaning for these vocabularies and the relationships among the vocabularies, then eliminate the customer's misunderstandings of the concept and the logic relation, thus attain the purpose of accurately processing information. Generally speaking, two basic characteristics of the ontology are to share and specify. To share means that the ontology expresses the generally accepted knowledge; to specify implies that there are explicit definitions for these concepts and restrictions on the usage of concepts.

Ontology of domain is defined as a group of four members: Dom_Ontology=(C, R, L, f), C,R,L and f denote respectively domain, relation, attribute and value, F is a mapping relation, When c belongs to C and I belongs to L, then f(c)= I. R is a marshal of relations, including different kinds of relationships of concepts in ontology[2].

Development of Ontologies has been moving from the realm of Artificial-Intelligence laboratories to the desktops of domain experts. Ontologies have become common on World Wide Web. The Ontologies on the Web ranges from large taxonomies categorizing Web sites (such as on Yahoo!) to categorizations of products for sale and their features (such as on Amazon.com). The availability and the proliferation of Ontologies are crucial for the success of the Semantic Web.

It is now widely recognized that constructing a domain model, or Ontology, is an important step in the development of knowledge based systems. What is lacking, however, is a clear understanding of how to build Ontologies.As consequence; a great number of researchers are working on method and techniques to build Ontologies through automatic or semi-automatic processes, which perform knowledge acquisition from texts, dictionaries and structured and semi-structured information source.

II. ONTOLOGY DEVELOPMENT AND DOWNSIDES

The general stages in the design and development of ontology are as follows [3]:

- The first step involves determining the domain and source and also purpose and scope of the ontology. Questions that should be addressed at this stage include: what domain will the ontology cover?, what is the purpose of the ontology? and for what sorts of questions should the information in the ontology be able to provide answers?
- The second step is to ascertain if an ontology has been developed previously in the same subject area. If such an ontology exists, it is easier to modify the existing ontology to suit ones needs than to create a new ontology .Reusing existing Ontologies may also be a requirement if the system needs to interact with other applications that have already committed to particular Ontologies.
- The third step is to enumerate important terms in the ontology.
- Steps 4 and 5 are closely intertwined. They entail defining the classes (concepts) and the class hierarchy (Step4), and defining the properties of classes (Step 5).
- Step 4. A number of different approaches can be taken when determining the hierarchy of classes. One could use a top-down approach, which starts with the definition of the most general concepts in a domain and continues with more specialized concepts. Another approach is the bottom-up approach, which starts with the definition of the most specific classes (the leaves of the hierarchy), with subsequent grouping of these classes into more general concepts. From the list of terms drawn up in Step 3, those terms that describe objects that have an

.

independent existence should be extracted as these will form the classes (concepts) of the ontology. To determine the hierarchical organization of the ontology, for each class one should ask if the instances of that class could also be instances of a more general class. If the answer is yes, then this class constitutes a subclass of the other class and, hence, is further from the root concept in the ontology.

- Step 5. Once the classes have been defined, the next step is to describe the internal structures (properties) of the concepts. Again, these should be readily available from the list produced as a result of Step 3.
- Step 6. Involves attaching facets to the properties that is, describing the value type, allowed values, the number of allowed values (cardinality) and other features that are deemed to be necessary. In this way, constraints are placed on the types of data that are allowed.
- Step 7.The final step 7 in the procedure is to create instances of the classes, that is to provide examples of each of the classes.

A Simple Knowledge-Engineering Methodology [4]:

There is no one correct way or methodology for developing Ontology. Ontology development is an iterative process. With every iteration the developed ontology get revised and refined. The knowledge acquiring process should follow the given steps of Ontology creation:

Step 1: Determine the domain and scope of Ontology.
Step 2: Consider reusing existing Ontology.
Step 3: Enumerate important terms in the Ontology.
Step 4: Define the classes and class hierarchy.
Step 5: Define the properties of classes.
Step 6: Define facets of the slot.
Step 7: Create instances.

All the above methods resembles to the following life-cycle model of Ontology creation shown in Figure 1

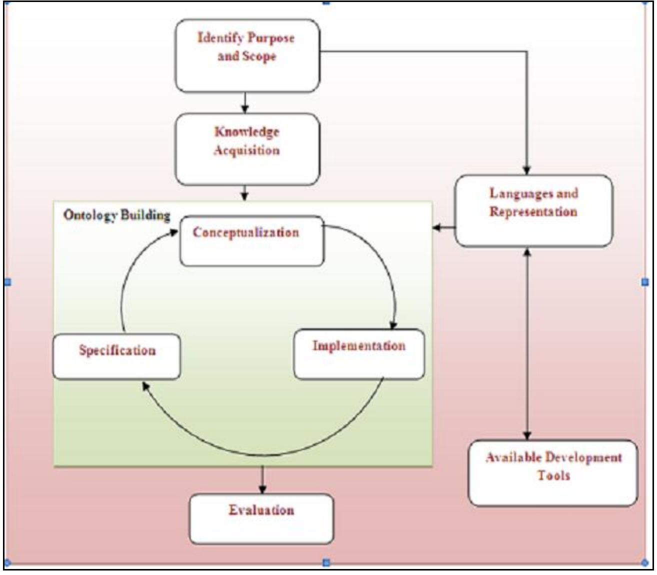


Figure 1

Ontology life-cycle is the specific sequence of activities that Ontology practioners carry out for developing an Ontology. The basic three phases of Ontology development life-cycle as shown above in the figure 1 includes:

Specification phase in order to define why Ontology being built, what its intended uses are.
Conceptualization phase in order to conceptualize the domain knowledge.
Implementation phase in order to transform the conceptual model into a formal computable model.
Followed by Evaluation phase in which the resulting Ontology is assessed.

Downsides:

Creating Ontology is not an easy task, the creator has to familiarize with the context of the ontology and possibly define the users of forthcoming Ontology [5]. Appropriate literature: textbooks, reference books, encyclopedias as well as newspapers, may serve as information sources, when creating the Ontology.

For the purpose or reuse there must be prior Ontologies or thesauri, classification schemes or glossaries dealing with the subject or some sector of the subject, they are worth as well.

Ontology creation is a tedious, time consuming and expensive task due the lack of unified approach for extracting the meaningful data to form the knowledgebase [10].

Ontology development is often error-prone, since there might be many different interpretations of same domain and the resulting model is too subjective.

In addition there is no single correct way to model a domain, there are always variable alternatives. The quality of solution depends on how the subjective domain information can be processed to form the concept hierarchy.

III. STRUCTURE OF ONTOLOGY

Ontologies are the structural framework for organizing information. Ontologies typically have two distinct components: 1.Names for important concepts in the domain; 2 Background knowledge or constrains in the domain. These components are commonly described as:

Individuals: Instances or objects (the basic or “ground level” objects)
Classes: Sets, collections, concepts, classes in programming, types of objects, or kinds of things
Attributes: Aspects, properties, features, characteristics, or parameters that objects (and classes) can have.
Relations: Ways in which classes and individuals can be related to one another.
Function terms: Complex structures formed from certain relations that can be used in place of an individual term in a statement.
Restrictions: Formally stated descriptions of what must be true in order for some assertion to be accepted as input.
Rules: Statements in the form of an if-then (antecedent-consequent) sentence that describe the logical inferences that can be drawn from an assertion in a particular form.
Axioms: Assertions (including rules) in a logical form that together comprise the overall theory that the ontology describes in its domain of application.
Events: The changing of attributes or relations.

IV. ONTOLOGY DEVELOPMENT APPROACHES

Task based Approach [7]:This method focuses on the manual domain ontology creation, by a novice ontology developer, by proposing an approach, which is adaptive with respect to the knowledge level of the developer in a specific domain and the availability of the resources, needed in the ontology engineering. It suggests four alternative scenarios, which include mandatory tasks that must be carried out by an individual, who has no experience in ontology engineering, in order to build his personalized ontology.

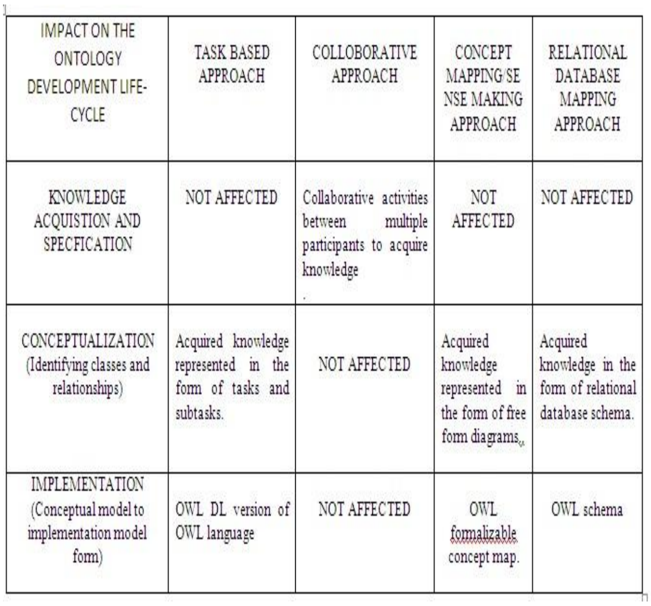
Collaborative Approach [9]: This approach presents an online system that supports initial collaborative activities between multiple participants in the knowledge acquisition phase of ontology development. The main objective of having such a system is mainly to accelerate the initial stage of ontology development where the presence of domain experts is not required. Other available participants can kick off the ontology development by acquiring and gathering the relevant terms and relations from readily available resources from the Internet, e.g. web directories and indexes. The tool specifically covers the phase prior to the formal ontology construction, where the acquired data are deposited into a system for automatic consolidation and structured into semi-formal semantic objects that can easily be converted into ontology elements, e.g. class, instances and properties. This means that the form of collaboration work described in takes place without having to construct any ontology yet. With the easy to follow steps, such knowledge gathering can be performed by anyone, not only limited to those that have prior knowledge in ontology construction.

Concept mapping/sense making Approach [6]: This approach presents the idea of diagramming to support simultaneous sense making and ontology development. Initial acquired knowledge represented in the free form diagrams showing node-link-node triples are converted into the description logic concept maps that combine human readability and understanding with local formalism and machine understandable format.

Relational Database Mapping Approach [8]: Mapping relational data into Ontology or filling Ontology with data from relational database respectively plays an important role during the creation and updating of ontology. If we consider the quantity of data found in relational databases and potential of their joint use in various applications , the question of finding the method for their use becomes obvious. This approach focuses on the principles of automatic conversion of constructs of relational data model to the constructs of OWL (Web ontology language) Ontology.

The comparative summary of the approaches affecting the specific phase of general ontology creation life cycle is shown in the table 1:

Table 1



IMPACT ON THE ONTOLOGY DEVELOPMENT LIFE-CYCLE	TASK BASED APPROACH	COLLOBORATIVE APPROACH	CONCEPT MAPPING/SENSE MAKING APPROACH	RELATIONAL DATABASE MAPPING APPROACH
KNOWLEDGE ACQUISTION AND SPECFICATION	NOT AFFECTED	Collaborative activities between multiple participants to acquire knowledge .	NOT AFFECTED	NOT AFFECTED
CONCEPTUALIZATION (Identifying classes and relationships)	Acquired knowledge represented in the form of tasks and subtasks.	NOT AFFECTED	Acquired knowledge represented in the form of free form diagrams.	Acquired knowledge in the form of relational database schema.
IMPLEMENTATION (Conceptual model to implementation model form)	OWL DL version of OWL language	NOT AFFECTED	OWL formalizable concept map.	OWL schema

V. NOUN PHRASE APPROACH

We restrict the kind of Ontology to be built is to a Domain Ontology. For knowledge to function as an asset in a corporation, it must be explicit and machine-processible. “Explicit” means

that it is “written down,” as opposed to carried in people’s heads or derivable via data analysis. Knowledge must be machine-processible because the volume of data makes human processing impossible. It is commonly said that 80% of corporate data exists as text: program notebooks; problem summaries in warranty records, technical-assistance center logs, customer surveys, and various other archives of records, logs, and diaries.

More or less all the approaches studied so far mainly focuses on capturing domain knowledge and then processing further for relations, facets, slots, instances etc. For novice user in the field of Ontology, it is difficult to frame the concept hierarchy with only the knowledge of domain present in textual form. Here what needed is the systematic approach where domain knowledge can be utilized in well disciplined way to form Ontology. According to the noun phrase approach nouns in the textual description are considered to be concepts and verbs helps us to relate the concepts. The relations identified can constitute aspects, properties, features or attributes that class can have. The conditional statements impose restrictions in relating the concepts.

More formally the phases of noun phrase approach for acquiring the domain knowledge to form Ontology are stated below:

Information/domain knowledge gathering: Collect the domain information from various available sources.

Cleaning and refining the information/domain knowledge: Remove the redundant, fuzzy and obsolete information.

Identify Noun phrases and nouns: List all nouns and noun phrases. Nouns can be taken as concepts or classes and phrases identify relations.

Eliminating irrelevant classes and adding new: All classes must make sense in the application domain. Remove such fuzzy or not useful classes. Few classes which are not explicitly mentioned can be derived from the other identified classed and can be added.

Iterate and refine: Iterate and refine the concepts and relations identified to form Ontology. Repeat above steps during the course of every iteration.

Classes and relations being the basic building blocks of concept-hierarchy, approach only look for nouns and phrases. During the successive iterations as a part of refinement, adjectives can be considered either as class or subclass depending on the type of application.

VI. ONTOLOGY EDITORS

Ontology editors are applications designed to assist in the creation or manipulation of Ontologies after capturing the domain concepts and relations using any feasible approach. Few of these editors are discussed below which assist developer to take the support of automatic tools and the application programming interfaces (APIs) supported by these tools to incorporate Ontology into a programming application:

Protégé

Protege is a free, open source ontology editor and a knowledge acquisition system. Protégé is being developed at Stanford University in collaboration with the University of Manchester. This application is written in Java and heavily uses Swing to create the rather complex user interface.

DOME

The DERI Ontology Management Environment (DOME) is developed by the Ontology Management Working Group (OMWG). The mission is to create a suite for the efficient and effective management of ontologies.

Knoodl

Knoodl facilitates community-oriented development of OWL-based ontologies and RDF knowledge bases. It also serves as a semantic technology platform, offering a Java service-based interface or a SPARQL-based interface so that communities can build their own semantic applications using their ontologies and knowledge bases.

Onto Edit

Onto Edit is an ontology editor developed by the On-To-Knowledge project. It implements an ontology constructions process. The Editor stores the ontology conceptual model in the SESAME repository and produces RDF concrete representations.

VII. CONCLUSION

Current state-of-the art research in Information Systems has focused on the use of Ontologies. However, there remain many obstacles for the practical and commercial use of Ontologies for Information Systems. One such obstacle is that current Information System designers lack the know-how to successfully design Ontology. Current ontology design methodologies are difficult to use by Information Systems designers having little theoretical knowledge of Ontology. Ontologies have become core components of many large applications yet the training material has not kept pace with the growing interest. This paper presents new approach of creating Ontology which benefits domain analyst to put the domain knowledge in the form which can be directly used to build Ontology.

REFERENCES

[1] Kaushal Giri,"Role of Ontology in semantic web",DESIDOC Journal of Library and Information Technology , Vol. 31, No. 2, March 2011, pp 116-120.

[2] Guilai Han, Anquan Jiao, "Personalized Data mining based on Ontology and Agent", 2010 WASE Interantional Conference on Information Engineering, vol.-02, pp. 55-56, August 2010.

[3] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, "Acquiring owl Ontologies from data intensive websites," International Conference on Web Engineering, pp. 11-14, July 2006.

[4] Trinity College Dublin, Department of Biochemistry, Dublin, Ireland , Dublin city University, School of computing ,Dublin,"Developing domain ontologies for Course Content, International Forum of Educational Technology and Society,1176-364, 2007.

[5] Mizoguchi, R. (2003). Tutorial on ontological engineering – Part 1: Introduction to ontological engineering. New Generation Computing, 21(4), 365-384.

[6] Eskridge,Thomas C., "Ontology creation as a sense making activity',IEEE Journal on Computing and processing, vol. 27, pp. 58-65, 2012..

[7] Lambrini Seremati,Achilles Kameas "Task based Ontology Engineering Approach," IEEE Conference in Informatics, 2009.

[8] Lambrini Seremati,Achilles Kameas, "Relational Database as a source of Ontology creation" ,IEEE proceedings of International Multiconference on Computer Science and Information Technology, pp. 135-139, 2010.

[9] Norliza Zaini,Hasmila Omar,"An Online System to support Collaborative Knowledge Acquisition for Ontology Development", International Conference on Computer Applications and Industrial Electronics, 2011.

[10] Nikolai Dahlem, Jianfeng Guo," Towards an User friendly Ontology design methodology",IEEE International Conference on Interoperability for Enerprise Software and Applications,2009.