SURVEY ON ONTOLOGY CONSTRUCTION METHODOLOGIES

K.Kalaiselvi[1], K.Thilagam[2]

ABSTRACT

Retrieval of relevant information from the large repositories of information system is a challenging task in today's scenario. Information repositories are much voluminous and stores data in various formats[1]. Most of the information retrieval systems are mainly based on keyword based text matching and retrieve unlimited volume of relevant and irrelevant information. In order to organize the data, a common terminology needs to be adopted. In this regard Ontology plays a major role which is used to capture the domain knowledge and it represents the knowledge Base. This paper discuss about the various methods used for the construction of the Ontology.

Keywords—*Ontology, OWL, RDF*

I. INTRODUCTION

Ontology is expressed as a formal representation of knowledge by a set of concepts within a domain and the relations between those concepts. Ontology necessarily entails or embodies some sort of world view with respect to a given domain[2]. Ontology is used in artificial intelligence, semantic web, software engineering and information architecture, natural language processing etc, as a form of knowledge representation about the world or some part of it. Ontologies are shared conceptualization of a domain [2]. Ontology ensures efficient information retrieval by enabling inferences based on domain knowledge. Reason behind this is shared and common understanding of some domain that can be communicated across people and computers[3]. Ontology enable knowledge reuse and sharing across applications. Ontologies capture the concepts and relationships between the concepts.

II. ONTOLOGY

Gruber[4] defines ontology "as a formal and explicit specification of a shared conceptualization". Each of the terms used in the definition should be interpreted as follows [3]

Conceptualization refers to an abstract model of a phenomenon in the world based on concepts identifying this phenomenon,

Explicit means that the types of concepts and their restrictions are explicitly defined,

Formal refers to the fact the ontology should be unambiguous and machine readable,

[1]Department of CSE, Karpagam University, Coimbatore

[2]Department of CA, Karpagam University, Coimbatore

- Shared reflects the viewpoint that the ontology should cover general knowledge of a domain knowledge that is not recognized by a single individual, but it is agreed upon a group of users

Ontology is mainly used to structure a knowledge Base. Developing an ontology includes

a) classes representing the concepts related to a specific domain of knowledge

b) properties expressing types of interactions among the domain concepts and further divided into object properties and datatype properties,

c) instances(individuals) representing specific entities that are members of a class and

d) axioms that express true facts about the ontology entities.

Ontology capture domain ontology and it is used for organizing and structuring the knowledge in repository[3]. Ontology keeps the knowledge base upto date when the world undergoes changes[5].

Different type of ontologies exists, they are [6]

1. Domain Ontologies - represent knowledge of a certain domain type (example electronic, medical, mechanical [6] etc.)
2. Generic Ontologies - can be applied to a variety of domain types (core ontologies, super theory).
3. Representational Ontologies - formulate general representation entities without defining what should be rpresented. E.g. Frame Ontology[7].
4. Task Ontologies - Provide terms specific for a particular Task.
5. Method Ontologies - Provide terms specific to a particular Problem Solving method.

Generally, construction of ontology could be done in three ways Manual - Ontology is constructed manually Semi Automatic- Human intervention is needed during ontology process

Fully Automatic - The system takes care of the complete construciton

III. Ontology Development Methodologies:

A. Ushold and King: based on their experience in building the Enterprise Ontology they developed the methods as follows [8]. They are

1. Purpose identification
2. Building the ontology
3. Evaluation
4. Documentation

It is also used to build other domain ontologies.

B. Tove Methodology: It was developed to help enterprise process modelling at Toronto University[9]. It follows as

1. Formalize the requirement specification of the ontology by motivating scenarios
2. Using the scenarios formulate competency questions. Ontology should provide vocabulary for expressing theses questions. Questions play the role of constraints and used to evaluate the resulting ontology.

3. Extract a set of terms from the informal competency questions and terms are formalized in a formal language to put into the ontology

4. Formalize the competency questions by defining the terms and writing axioms for interpretation of the terms.

5. Establish conditions for characterizing the completeness of the ontology.

The competency questions strategy is well-accepted and used in On-To-Knowledge methodology.

C. Methontology : developed at Polytechnic University of Madrid and is based on IEEE standards for developing software life cycle processes[10]. It includes

1. Project management process – planning, project control, quality control etc.,

2. Ontology development process

3. Guidelines for knowledge acquisition, evaluation, ontology integration, documentation, version management etc.,

This approach build ontology from scratch, reuse other ontologies .

D. On-To-Knowledge: Developed at Karlsrube University based on a two-loop architecture[11]. They are Knowledge process and Knowledge meta process for introducing and maintaining ontology based knowledge management. Knowledge process is a knowledge use and evolution process and knowledge meta process is a methodology of ontology development and is composed of five major steps. They are

1. Feasibility study
2. Kickoff
3. Refinement
4. Evaluation
5. Application & evaluation

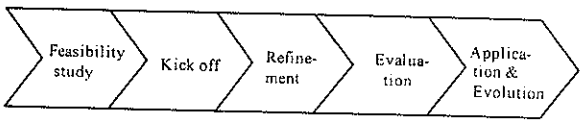

Figure 1 : Steps in On-To-Knowledge Methodology

The above figure-1 summaries the following activities in the construction of ontology. Feasibility study identifies problems and oppurtinities, and reuse of KM application.Kick off capture requirement specification in ORSD and create semi-formal ontology. Refinement refine semi-formal ontology description and formalize into large ontology and create prototype. Evaluation focus on technology and ontology evaluation. Apply ontology and maintenance is carried in the last stage.

None of these methodologies have no technological support and therefore they cannot be easily applied in ontology construction task.[12]

IV. Ontology Representation Languages:

A. Ontolingua: developed by KSL(knowledge Systems Lab) at Stanford University for ontology representation and sharing. It does not have inference functionality. It provides a set of ontology development functions(browse, create, edit, modify) and ontology reuse.

B. RDF(s)(12): Resource Description Framework for metadata description developed by W3C. It includes

Triplet model<object,attribute,value>, object is resource representing a web page, value can take string or resource, attribute link between nodes. A triplet can be a object and a value.

C. **OWL:** Web ontology language developed by W3C. OWL is designed to make it a common language for ontology representation and is based on DAML+OIL. OWL is an extension of RDF Schema and also employs a triplet model. Its principles includes a standard language for ontology representation and it gains the highest priority in extensibility, modifiability and interoperability. First order predicate logic is replaced by rule layer in OWL and is next to the ontology layer. OWL is divided into three syntax-classes, they are

OWL-Lite: provides primary classification hierarchy and simple constraints, quick migration path for thesauri and other taxonomies and also has lower formal complexity than OWL DL

OWL-DL:supports maximum expressiveness, retaining computational completeness and decidability.

completeness - conclusions are guaranteed to be computable decidability - computations will finish in finite time

OWL-Full: meant for users who want maximum expressiveness and syntactic freedom of RDF with no computational guarantees.

Table 1 gives the comparative analysis of ontology construction methods. It also summaries some of the important features for the construction of ontology. Project initiation and ontology quality management has not been proposed in Uschold and king, Methontology and Tove methods. Design, requirements and implementation are proposed in On-To-Knowledge. Knowledge acquisition, verification and validation, documentation are also proposed in On-To-Knowledge. From this On-To-Knowledge is best suited for the construction of ontology.

Table 1 : Comparative Analysis of Ontology Construction Methods[13]

Feature	Uschold & King	Methontology	On-To Knowledge	Tove
Project initiation	Not proposed	Not proposed	Proposed	Not proposed
Ontology quality management	Not Proposed	Not proposed	Proposed	Not proposed
Design	Not Proposed	Described in detail	Proposed	Not proposed
Requirements	Proposed	Proposed	Proposed	proposed
Implementation	Proposed	Described in detail	Proposed	proposed
Maintenance	Not Proposed	Proposed	Proposed	proposed
Knowledge Acquisition	Proposed	Described in detail	Proposed	proposed
Verification and Validation	Proposed	Described in detail	Proposed	proposed
Documentation	Proposed	Proposed	Proposed	proposed

For developing large-scale ontology , Methontology and On-To-Knowledge are very helpful. At the early stage of development to obtain an informal ontology Uschold and kings methodology is useful. Tove is used for enumerating the competency question. On-To-Knowledge is used for knowledge management applications. Users can adopt the features of all the methods in their ontology building processes.

On-To-knowledge methodology are application dependent, since the ontology is built on the basis of a given application. On-To-Knowledge Methodology used in OntoEdit[14] tool for the construction of ontology. OntoEditSupports langauages like XML, DAML+OIL, FLogic, RDF(s). Uschold and King methods, and the methodology METHONTOLOGY are application-independent, since the ontology development process is totally independent of the uses of the ontology. The Tool WebODE[15] uses METHONTOLOGY methodology and give support to languages like DAML+OIL ,OIL,RDF(s), and XML.

V. Conclusion

This paper analysis the various methodologies for the construction of ontology. Existing methods and patterns for information retrieval do not allow us to accurately retrieve information from the legacy systems. So Ontology based knowledge repository is needed for the efficient information retrieval. All the Knowledge representation languages are within the paradigm of Knowledge representation community. RDF(s) is a kind of semantic network. OWL is the same as RDF(s) in its data model and in top-level ontology. RDF(s) does not distinguish between relations, attributes and features. OWL does not provide user with adequate modelling facility for representing an ontology, but it is appropriate for ontology interchange and sharing. But OWL describes more vocabulary and more effective relationship of any particular domain. From the analysis it is understood that none of the approaches presented is fully mature if we compare them with software engineering and knowledge engineering methodologies. METHONTOLOGY has been recommended by FIPA for the ontology construction task. So it is conclude that On-To-Knowledge and METHONTOLOGY methodology are best suited for the construction of any domain ontology.

References

1. Matteo Gaeta, Francesco Orciuoli, Stefeno Paolozzi and Saverio Salerno *"Ontology Extraction for Knowledge Reuse: The e-learning Perspective"*, IEEE Transaction on Systems, Man and Cybernetics-Systems and Humans 41(4) july 2011

2. Mike.Uschold & Michael Gruninger, *"Ontologies Principles Methods and Applications"*, The University of Edinburgh 1996, Knowledge Engineering Review 11(2) June 1996

3. Rudi Studer, V.Richard Benjamins, Dieter Fensel, *"Knowledge Engineering –Principles and Methods"*

4. Thomas R.Gruber, 1995, Toward principles for the design of Ontologies used for knowledge sharing. International Journal of Human Computer Studies:43(5/6) nov/dec 1995

5. Nicola Guarino *"Formal Ontology, Conceptual Analysis and Knowledge Representation"*, International Journal of Human Computer Studeis 43 625-640 1995

6. G.Van Heijst A.Th. Schreiber and B.J.Wielings *"using Explicit Ontologies in KBS Development"*, International Journal of Human – Computer Studies 46(2/3) 183-292 1997

7. Thomas R.Gruber, *"A Translation Approach to Portable Ontology Specification"*, Knowledge Acquisition 1993

8. http://www.aiai.ed.ac.uk/project/enterprise/enterprise/ontology.html

9. Gruninger.R and Fox.M *"The design and evaluation of ontologies for enterprise engineering"* in proc of Comparison of Implemented Ontologies ECAI'94 Workshop W13 ,105-128 1994

10. Fernandes-Lopes, M., Gomes-Perez, A. and Pazos Sierra, J., *"Building a Chemical Ontology using Methontology and Ontology Design Environment,"* IEEE Intelligent Systems, 14,(1), 37-46, 1999

11. Staab, S.H., Schunurr, P., Studer, R and Sure, Y., *"Knowledge Processes and Ontologies"* IEE Intelligent Systems, Special Issue on Knowledge Management 16(1) 1 26-34 2001 .

12. http://www.w3schools.com/webservices/ws_rdf_example.asp

13. Oscar Corcho, Mariano Fernandez, Lopez Asuncion Gomez Peraz, *"Methodologies Tools and Languages for Building Ontologies, Where their meeting point?"*, Data and Knowledge Engineering 46, 41-64, 2003

14. Y.Sure J.Angele and S.Staab *"OntoEdit Guiding Ontology Developemnt by Methodology and Inferencing"* Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics 2519 1205-1222 2002

15. A.Gomez-Perez. M.Fernandez-Lopez, O.Corcho, and J.aspirez *"WebODE a scalable ontological engineering workbench"* First International Conference on Knowledge Capture 2001

Author's Biography

Mrs. K. Kalaiselvi, is working as an Assistant Professor in the department of Computer Applications. She has got 8 years of Teaching Experience and 11 years of experience in the Industry. Her research interest includes Semantic Web, Data Mining and Artificial Intelligence.

Dr. K.Thilagam, received her Doctoral degree from Karpagam University and is working there as a Professor. She has got 10+ years of teaching experience and 3½ years of Industry experience. Her research interests are Image Processing, Data Mining and Networking.