Ontology Engineering Methodologies: An Analytical Study

Syamili C ***Rekha R V***

Abstract

Ontology is a broad term including a wide range of activities. Ontology can provide semantics for the next generation of World Wide Web. Recently ontology has become a major research issues in several disciplines of Computer Science and Information Systems. There are a range of domain ontologies on the semantic web such as Gene Ontology, Biological science ontology, CIDOC-CRM ontology of culture heritage documentation, FRBR in Bibliographic and NCI cancer ontology. When a new methodology is going to build, several basic questions arise related to the methodology, tools and techniques which are to be used in the development process. The study discusses various ontology engineering methodologies and compares major existing domain ontologies with these ontology engineering methodologies. The study critically analyses the problems of existing engineering methodologies and put forward some suggestions to overcome these limitations. The study will act as a guide to ontology developers and ontology experts.

Keywords: Ontology, Ontology Engineering, Semantic web

1. Introduction

Back to the history from a philosophical perspective, Aristotle (384-322BC) invented ontology as a study of the ways that the universe is organised into categories. Recently, the development of domain ontologies has become increasingly important for knowledge level interoperation and information integration. There are a range of domain ontologies on the semantic web such as Gene Ontology (GeneOntology, 2009), Biological science ontology (save 2005), CIDOC-CRM ontology of culture, heritage documentation, FRBR in Bibliographic and NCI cancer ontology (Golbeck et al., 2008). This is a significant problem for scholars and researchers who need to be able to access information within their interest area.

2. Objective of the Study

When a new methodology is going to build, several basic questions arise related to the methodology, tools and techniques which are to be used in the development process. The major questions related to methodology are: i) Which methodology can I use for building ontology? ii) Whether it should start from scratch or using other ontologies available on the server? iii) Which is the life cycle of an ontology that is developed with a specific methodology? Keeping these issues in mind, the study attempts to critically evaluate the existing

ontology engineering methodologies. The major objectives of the study are:

- ❖ Discusses various ontology engineering methodologies
- ❖ Compare major existing domain ontologies with these ontology engineering methodologies.
- ❖ Critically analyses the problems of existing engineering methodologies and put forward some suggestions to overcome these limitations.

3. Methodology of the study

The study includes two phases. In first phase an exhaustive literature search has been done to identify the existing ontology engineering methodologies. These methodologies were critically analyzed and discussed in the first section. Many individual projects also describe their specific methodologies; these are purposefully not included. In the second phase of the study, various documented ontologies of different domain has been compared with respect to their methodologies. For these various ontology libraries has been searched. This include

Ontology Library	URL
Swoogle	http://swoogle.umbc.edu/
TONES ontology repository	http://www.inf.unibz.it/tones/
Protege ontology library	http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library
LOV	https://lov.okfn.org/dataset/lov/
Ontolingua ontology library	http://www.ksl.stanford.edu/software /ontolingua/
DMOZ	www.dmoz.org
DAML ontology library	http://www.daml.org/ontologies/
UNSPSC	www.unspsc.org
RosettaNet	www.rosettanet.org

4. Ontology Engineering Methodology

A series of approaches for developing ontologies have been reported. The methods and methodologies reported for ontology engineering is discussed in this section.

4.1. Cyc

The Cys project started by D.B. Lenat at MCC (Micro Electronics and Computer Technology Corporation) in 1984 as a knowledge representation project. The report of the project published by D.B. Lenat and R.V. Guha in 1989. The Cyc methodology came from the experience of developing a Knowledge Base. Cyc uses a hybrid language called Cyc Language. The building Cyc ontology consists of three steps.

- 1st phase: In this step implicit and explicit knowledge will be coded without the help of natural language systems and machines. It includes the manual codification of articles and knowledge which contain common sense knowledge. This phase is carried out by hand.
- 2nd phase: This phase includes the computer aided extraction of common sense knowledge by humans.
- 3rd phase: This phase includes computer aided extraction of common sense knowledge, though no humans are involved in this process.

This methodology is only used for Cyc Knowledge Base. Parts of this project are published as Open Cyc and its latest version released on June 2012.

4.2. Uschold and King 1995

The first methodology for developing ontology was put forward by Uschold and King in 1995. It is a four step process. This methodology is also based on the experience of developing an ontology for enterprise modeling processes. The four steps are

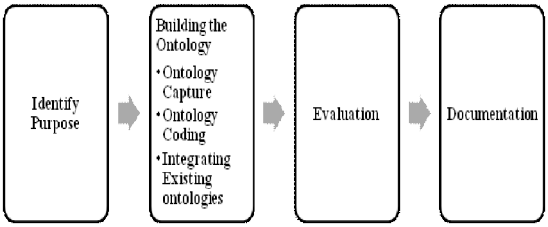
Step 1: The first step includes the identification of purpose, use, and domain and the intended users of ontologies being built.

Step 2: This covers building the ontology. This step broken into three parts and they are

1. Ontology capture-this involves the capture of key concepts and ideas from the domain of interest.
2. Coding- Explicit representation of knowledge acquired in the previous step
3. Integrating existing ontologies-It refers to usage of existing ontologies

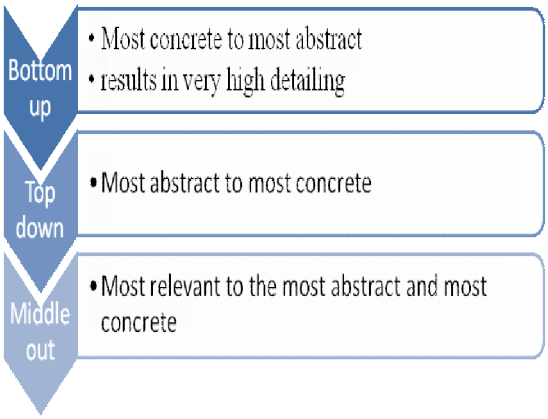
Step 3: Evaluation of the ontology

Step4: Documentation of the ontology



The authors also proposed three strategies for identifying key concepts. They are

1. Bottom up -most concrete to most abstract
2. Top down- Most abstract to most concrete
3. Middle out - most relevant to the most abstract and most concrete

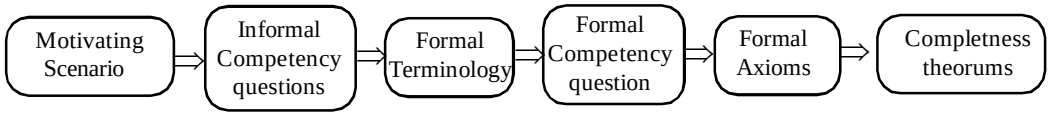


This method gives an argument on the use of middle out instead of bottom up and top down. The method extended in 1996 by Uschold and Gruninger

4.3. Gruninger and Fox

In 1995 Gruninger and Fox designed an approach for developing ontology. This methodology is based on the experience of developing TOVE (TOronto Virtual Enterprise) ontology(Gruninger, M., and Fox, 1995). This method is also called TOVE Methodology. The overall process can be represented using the diagram below. Motivating scenario may be an idea or proposal for new ontology which does not exist as of then. Next is to develop some competency questions. A set of natural language questions used as competency questions in measuring the scope of the ontology. Next is the coding stage and this including specifying the informal questions in formal language. These questions will help in evaluating the ontology. Then specify the axioms using first order logic and define the conditions under which

the solutions to the competency questions are complete.



4.4. Sensus

This is a method for domain specific ontology from large ontologies (Swartout, Knight, Russ, & Rey, 1997). It is developed by Information Sciences Institute Natural Language Group. Sensus include more than 50,000 concepts hierarchically arranged according to their abstraction. Ontosaurus is the software used for building this ontology (Swartout et al., 1997). This web based ontology tool is also developed by ISI. This approach increase the sharability of knowledge as it share same base ontology(Corcho, Fern, Mariano, & Gómez-pérez, 2003). This method using top - down approach. The main process include the following steps

1. Identification of key terms. This key term is called seed term.
2. Linking manually the terms using Ontosarus
3. All concepts from the seed term to root are included.
4. Adding of new term which are relevant and not yet included.
5. Adding of sub trees to the final ontology

4.5. Methontology

It has been created in the Artificial Intelligence Lab of Polytechnic University of Madrid. It is based on IEEE standards for developing software 1074-1995. This method can be used for developing ontologies from scratch, reusing and reengineering. Tools like WebODE support methontology. Other tools are also compatible with methontology. It differs from Uschold and King Gruninger and Fox (Gruninger, M., and Fox, 1995) method as it focusing on evaluation and documentation. Middle out approach is suggested as a strategy in identifying concepts. It helps in identifying primary concept first. Later one can move on either to specialization or to generalization.

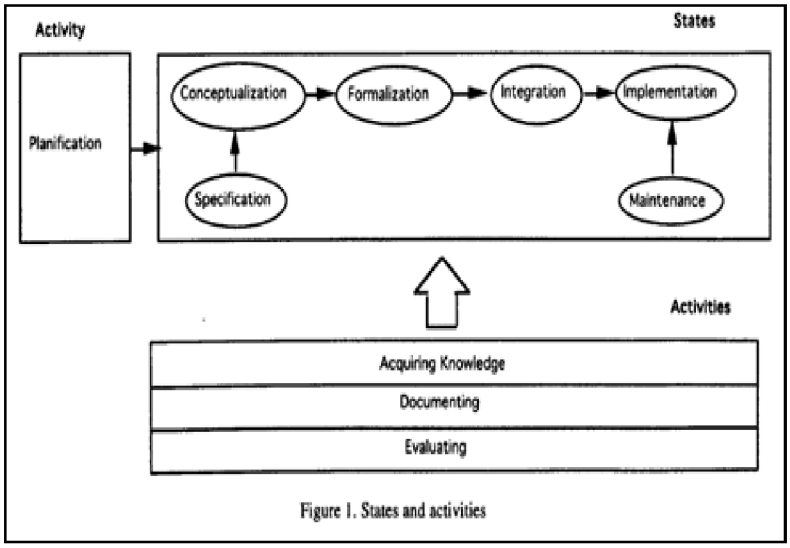
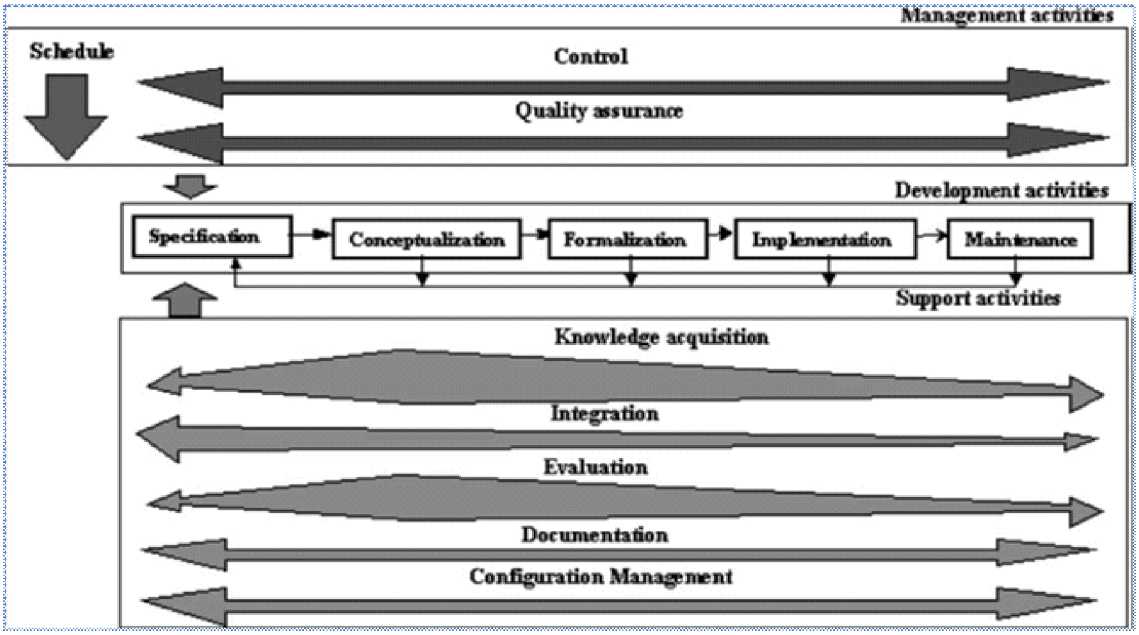


Figure 1. States and activities

4.6. On-To-Knowledge Methodology

It includes identification of goals that is to be achieved by knowledge management tools and is based on an analysis of usage scenarios. The figure shows the steps followed in this method.

Apart from these methodologies there exist some other methodologies too. They are Activity First Method AFM, KACTUS, C04, YAMO. Many individual projects also describe their specific methodologies; these are purposefully not included.



5. Comparison of different ontologies and their ontology engineering methodologies

This section compares the major domain ontologies and their methodologies. It includes only those ontologies which are documented. We attempted to include ontologies of different periods and different domains.

No	Name of the ontology	Author	Year	Methodology used
1	TOronto Virtual Enterprise ontology (TOVE)	(Gruninger, M., and Fox, 1995)	1995	Gruninger & Fox Methodology
2	The Reference Ontology	(Arpírez, Gómez-Pérez, Lozano-Tello, & Pinto, 2000)	(1998)	Methontology
3	Knowledge Acquisition Ontology	Blázquez, J. Fernandez,	(1998)	Methontology
		M. García-Pinar, J. & Gómez-Pérez, A.		
4	Chemical ontology	Fernadez-Lopez, M., et. al	(1999)	Methontology

No	Name of the ontology	Author	Year	Methodology used
5	Environmental pollutants ontology	(Asunción Gómez-Pérez and Dolores Rojas-Amaya, 1999)	1999	Methontology
6	Legal Ontology	(CORCHO, O., 2002)	2002	Methontology
7	Pizza ontology	(Drummond, N., Horridge, M., Stevens, R., Wroe, C. and Sampaio, 2005)	2005	No clue on methodology
8	Wine Ontology	(Graca, J., Mourao, M., Anunciacao, O., Monteiro, P., Pinto, H.S. and Loureiro, 2005)	2005	Enterprise Ontology+ Methontology
9	Information Science Ontology	Sawsaa, A. & Lu, J	(2010)	Methontology
10	Beer Ontology	(Heflin, 2012)	2012	No clue on methodology
11	Quran ontology for Juz' Amma	(Iqbal, Mustapha, & Yusoff, 2013)	2013	ontology merging approach
12	Textile chemical ontology	(Ferrero & Lloret, 2014)	2014	Methontology
13	Food Ontology	(Dutta, Chatterjee, & Madalli, 2015)	2015	YAMO

6. Summary of Observations

By Comparing different domain ontologies and different ontology engineering methodologies the following observations were found.

- ❖ Each of these methodologies following different approaches.
- ❖ There is no correspondence between ontology building methodologies and tools, except for METHONTOLOGY and WebODE, and On-To-Knowledge and OntoEdit.
- ❖ Since there is no technological support for most of the existing methodologies, they cannot be easily applied in the ontology construction task.
- ❖ In fact, most of the tools just focus on few activities of the ontology lifecycle: design and implementation.
- ❖ METHONTOLOGY is most widely used methodology in developing ontology irrespective of the time

Reference

1. Arpírez, J. C., Gómez-Pérez, A., Lozano-Tello, A., & Pinto, H. S. A. N. P. (2000). Reference Ontology and (ONTO)2 Agent: The Ontology Yellow Pages. In Knowledge and Information Systems (Vol. 2, pp. 387–412). http://doi.org/10.1007/PL00011649
2. Asunción Gómez-Pérez and Dolores Rojas-Amaya. (1999). Ontological Reengineering and

reuse. In Dieter Fensel and Rudi Studer (Ed.), Proceedings of the 11th European Workshop on Knowledge Acquisition, Modeling and Management (EKAW '99) (pp. 139–156). London: Springer-Verlag.

3. CORCHO, O., et. al. (2002). Building legal ontologies with METHONTOLOGY and WebODE. Retrieved November 3, 2016, from http://academic.research.microsoft.com/Paper/1856108.aspx

4. Corcho, O., Fern, Mariano, & Gómez-pérez, A. (2003). Methodologies , tools and languages for building ontologies . Where is their meeting point?? Data & Knowledge Engineering, 46, 41–64. http://doi.org/10.1016/S0169-023X(02)00195-7

5. Drummond, N., Horridge, M., Stevens, R., Wroe, C. and Sampaio, S. (2005). Pizza Ontology. Manchester.: The University of Manchester.

6. Dutta, B., Chatterjee, U., & Madalli, D. P. (2015). YAMO?: Yet Another Methodology for large-scale faceted Ontology construction. Journal of Knowledge Management, 19(1), 6–24. http://doi.org/10.1108/JKM-10-2014-0439

7. Ferrero, C. P., & Lloret, E. (2014). Towards the Design of a Textile Chemical Ontology. International Journal of Advanced Computer Science and Applications, Special Issue on Natural Language Processing 2014, 35–41. Retrieved from www.ijacsa.thesai.org

8. Graca, J., Mourao, M., Anunciacao, O., Monteiro, P., Pinto, H.S. and Loureiro, V. (2005). Ontology building process: the wine domain. In Proceedings of the 5th Conference of EFIT, Vila Real. Vila Real.

9. Gruninger, M., and Fox, M. S. (1995). "Methodology for the Design and Evaluation of Ontologies",. In Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95,Montreal.

10. Heflin, J. (2012). Beer Ontology 1.0 (Draft). Retrieved November 3, 2016, from www.cs.umd.edu/projects/plus/SHOE/onts/%0Abeer1.0.html

11. Iqbal, R., Mustapha, A., & Yusoff, Z. M. (2013). An experience of developing Quran ontology with contextual information support. Multicultural Education & Technology Journal, 7(4), 333–343. http://doi.org/http://dx.doi.org/10.1108/METJ-03-2013-0009

12. Matteo, S. (2014). 14th EAHIL 2014 Conference 11-13 June , Rome , Italy What is Altmetrics?? The impact of Altmetrics on San Matteo researchers and on librarians ', 1–38.

13. Swartout, B., Knight, K., Russ, T., & Rey, M. (1997). Toward Distributed Use of Large-Scale Ontologies. AAAI Technical Report SS-97-06., 138–148.

About Authors

Syamili C., Research Scholar, Dept. of LIS, Pondicherry University
Email:csyamili@gmail.com

Dr. Rekha R. V., Assistant Professor, Dept. of LIS, Pondicherry University
Email:rekhatvm@gmail.com