DESIDOC Journal of Library & Information Technology, Vol. 31, No. 2, March 2011, pp. 84-89

Construction and Applications of Ontology: Recent Trends

Neela J Deshpande and Rajendra Kumbhar

DLIS, University of Pune, Maharashtra-411 007
E-mail: njdeshpande @unipune.ac.in; rajendra_kumbhar@unipune.ac.in

ABSTRACT

The purpose of ontology is to achieve inter-operability by providing a common terminology and understanding of a given domain of interest, which in turn allows for the assignment of clear meaning to learning object. Ontologies are used in artificial intelligence, biomedical informatics, library science, chemistry and many other subjects. The purpose of this paper is to make the library professionals alert about the literature published on ontology and applications of ontology.

Keywords: Ontology, construction of ontology, semantic Web, literature review

1. INTRODUCTION

The socio-economic and technological developments have been responsible for the unmanageable growth of information and knowledge. Libraries and other similar organisations have been entrusted to collect, organise and make useful knowledge available without loss of time. To serve this purpose, libraries and the other agencies involved in the knowledge generation, have been developing various knowledge organisation tools. These include classification schemes, lists of subject headings, thesauri, taxonomies, ontologies, etc.

Ontology has been widely applied to information retrieval, artificial intelligence, knowledge network and knowledge management[1]. Ontologies ensure efficient retrieval by enabling inferences based on domain knowledge, which is gathered during the construction of knowledge base[2]. Ontologies specifically play an important role in supporting knowledge-based applications in semantic web[3]. They are the backbone of a semantic Web, a semantic-aware version of the World Wide Web[4]. The success of the semantic Web depends strongly on the proliferation of ontologies[5]. Ontologies represent a key aspect for the integration of information coming from different sources, for supporting collaboration within virtual communities, for improving information retrieval, and more generally, it is important for reasoning on available knowledge[6]. Thus with the growth of semantic web and the knowledge management systems in the corporate world and other organisations, ontologies will be vital tools for knowledge sharing.

2. PURPOSE, SCOPE AND METHODOLOGY

Purpose of this paper is to review the literature dealing with the construction and applications of ontology. The review is restricted to the literature published up to July 2010. The review is based on the literature reported in *LISA*. Since it is based on the abstracts in the *LISA*, it covers the literature published in non-English languages also.

3. ONTOLOGY: A DEFINITIONAL ANALYSIS

Gruber[7] defined ontology as an explicit specification of a conceptualisation. Guarino[8] argued that it is not a conceptualisation but an agreement about a conceptualisation. While further elaborating the meaning of ontology, Guarino[9] stated that, ‘ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e., its ontological commitment to a particular conceptualisation of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. Ontology indirectly reflects this commitment (and the underlying conceptualisation) by approximating these intended models’. The characteristics of ontology are: Sharing: The notion of ontology is to capture commonly agreed knowledge. Conceptualisation: Refers to the mental formulation of a phenomena in the world. It is developed by identifying the related concepts of that phenomena. Formal relationships: Ontology formalises the relationships among the concepts, which makes computer to interpret the semantic relationships among the concepts and infer the implicit knowledge. Polynomial hierarchy: It develops in a polynomial hierarchy instead a rigid monolithic hierarchy structure.

4. CONSTRUCTION OF ONTOLOGY

Ontology construction is a complex task. It involves various methods and use of tools such as vocabulary lists, softwares, Web ontology language (OWL), etc. It also involves use of principles and theories. In the literature reviewed, it was observed that most of the researchers are proposing the use of automatic or semi-automatic methods for the construction of ontology because it is a time-consuming and tedious task[1].

According to De Nicola, Missikoff, and Navigli[4] ontology building exhibits a structural and logical complexity that is comparable to the production of software artifacts. Nevertheless it is useful to have varieties of high quality ontologies for the effective use of knowledge. Such availability depends on effective and usable methodologies aimed at supporting the crucial process of ontology building[4].

The literature reviewed below is evident that the view of De Nicola, *et al.* is supported by other researchers, as number of ontology development methodologies have been evolved and tested successfully. Clustering-based method is another method useful in ontology construction. It is suitable for the construction of cultural ontologies for community-oriented information systems[10]. This semi-automated method merges distributed annotation techniques or subjective assessments of similarities between cultural categories with established clustering methods to produce “cognate” ontologies.

Another semi-automated system for ontology construction is proposed by Lin and Hanqing[5]. It adopted an approach of non-dictionary Chinese words segmentation techniques based on N-Gram to acquire domain candidate concepts. It takes the method based of natural language processing in the recognition of domain concept property relation, extracting subject, predicate and object of sentences. This triangle data can he treated as the triplet of data, object type and property.

Jun and Yuhua[11] introduced an automatic approach for ontology building by integrating traditional knowledge organisation resource. It first builds a primary ontology describing the classes and relationships involved in bibliographic data with OWL, and then fill the primary ontology with instances of classes and their relations extracted from catalogue dataset and thesauri and classification schemes used in cataloguing.

Conversion and or use of existing vocabulary tools is one of the sub-systems adopted in ontology construction. Junzhi and Dandan[12] suggested a method of conversion from law FrameNet database to OWL ontology. Similarly a procedure for modeling the semantics of Library of Congress subject headings into ontology is described by Papadakis[13]. Hua, *et al.* suggested a SUMO-based emergency response preplan ontology model. This model facilitates shared understanding for people from different domains. It uses SUMO which is the most authoritative common ontology as upper ontology. It analysed the preplan-related concepts, properties and relations in depth, extracted domain knowledge, refined norms of professional and business glossary of terms and clearly defined their meaning and relationship among them. Ontology engineering and thesaurus-based ontology construction are the two major methods for the construction of domain ontology.

Ziyu and Lei[15] proposed a construction methodology of domain ontology that is composed of multi-disciplinary approach based on thesaurus and thematic words of high-speed railway that is built by the experts. Ontology and software development share a fare amount of commonality. Considering this feature De Nicola, Missikoff and Navigli proposed an ontology development methodology called UPON, i.e., Unified Process for Ontology[4]. This system is based on the unified software development process or unified process (UP), which is a widely used standard in software engineering.

Guangron[16] developed a course knowledge ontology for an e-learning course in C programming. The ontology is constructed through drawing out the core concepts of the course as well as the relations among the concepts. The ontology consists of 183 concepts, 130 sub-class relations, and 48 properties which are described in standard OWL. Most ontology construction methods focus on concept types.

Nguyen[17] *et al.* presented an ontology construction method that focused on relation types instead of concept types. The authors formalised the concept of “predicate of predicates” as meta-relation type and introduced the new hierarchy of meta-relation types as part of the ontology definition.

The new notion of closure of a relation or meta-relation type is presented as a means to complete that relation or meta-relation type by transferring extra arguments and properties from other related types. The end result is an expanded ontology, called the closure of the original ontology, on which automated inference could be more easily performed.

The domain ontology evolution approach is yet another approach useful in constructing ontology. This approach is adopted in construction of aviation ontology by Yingfang and Lei[18]. The FCA-based ontology construction method is another one. It consists of three steps, i.e. (i) PAT-tree based phrase extraction, (ii) feature selection, and (iii) formal context acquirement. This method greatly improves automatic construction of ontology and the constructed ontology by this model is reliable[1].

Jiang and Tan[3] criticised that the traditional ontology construction systems employ shallow natural language processing techniques and focus only on concept and taxonomic relation extraction. To overcome this lacuna, the authors proposed a new system for mining ontologies automatically from domain-specific documents, known as Concept-Relation-Concept Tuple-based Ontology Learning (CRCTOL).

This system: (i) adopts a full text parsing technique, (ii) employs a combination of statistical and lexico-syntactic methods, including a statistical algorithm that extracts key concepts from a document collection, a word sense disambiguation algorithm that disambiguates words in the key concepts, (iii) uses a rule-based algorithm that extracts relations between the key concepts, and (iv) also adopts a modified generalised association rule mining algorithm that prunes unimportant relations for ontology learning.

Using this method the Jiang and Tan built terrorism domain ontology and sport-event domain ontology.

5. ONTOLOGY MATCHING FOR INTEGRATION

Matching relevant ontology data for integration is vitally important as the amount of ontology data increases along with the evolving semantic Web, in which data are published from different individuals or organisations in a decentralised environment. For any domain that has developed a suitable ontology, its ontology annotated data (or simply ontology data) from different sources often overlaps and needs to be integrated. Considering this requirement Wang, Jie and Zhang[19] proposed an intelligent Web ontology data matching method and framework for data integration. This method is different from existing data matching or merging methods applied to data warehouse in that it employs a machine learning approach and more similarity measurements by exploring ontology features.

To resolve ontological heterogeneities among distributed data sources in an organisational memory and subsequently generate a merged ontology to facilitate resource retrieval from distributed resources for organisational decision making, Kiu and Lee[20] proposed an automated ontology mapping and merging algorithm, namely OntoDNA, which employs data mining techniques (FCA, SOM, K-means).

6. USES OF ONTOLOGY

Ontologies are constructed keeping in mind specific applications. Ontologies are concept specification devices that help in very many areas of work. This is substantiated by the literature reviewed below which deals with diverse areas ranging from library and information science, knowledge management, education, medical science, organisation of information on the Internet and for retrieving the same through search engines to financial institutions. The literature in this section is reviewed in the same sequence as mentioned above.

Ontologies could be used by users to find their actual and potential search requirements step by step which depends on two methods; one is using edit distance-based method to search word in ontology base and the other is using concept space-based method to help user to extend keywords. The benefit of this method is that it utilises the relations between words built by field ontology to find word's actual meanings in its context[21]. Keyword-based search in the traditional system performed poorly when literal term matching was carried out for query processing, due to synonymy and ambivalence of words. According to Saravanan, Ravindran and Raman[2] ontological frameworks can be used to improve this performance. Another application of ontology in library related activity is worked out by Liao[22], *et al.* Liao and his co-authors developed a novel library recommender system for English collections by integrating personal ontology model and collaborative filtering model with domain specification. The personal ontology recommender (PORE) system offers a friendly user interface and provides several personalised services. This system was implemented and tested in the Library of National Chung Hsing University in Taiwan and proved satisfactory.

As a greater volume of information becomes increasingly available across all disciplines, many approaches, such as document clustering and information visualisation, have been proposed to help users manage information easily. However, most of these methods do not directly extract key concepts and their semantic relationships from document corpora, which could help better to illuminate the conceptual structures within given information. To resolve this problem Zheng[23], *et al.* proposed "Clonto" approach which identifies the key concepts, and automatically generates ontologies based on these concepts for the purpose of conceptualisation. LIS professionals have always shown interest in integrating various vocabulary control tools. As such integration brings many advantages. Zhonghua[24], *et al.* proposed integration of ontology into thesauri so as to build 'thesauri-ontology', as a solution to building multilingual thesauri. The MatSeek database which enables federated search interface to key material science databases and analytical tools uses MatOnto ontology to integrate data across disparate databases Cheung, Hunter and Drennan[25].

Chen[26], *et al.* developed information needs ontology based on a comprehensive understanding of personalised user information needs. For this purpose the authors first formalised an ontology model applicable to the ubiquitous computing environment to describe user information needs. Second, in light of the user's explicit information needs, the digital library finds the hidden information needs from the ontology knowledge base, and also extracts the information resources required by the user from the corresponding database and then pushes them to the user by means of RSS feeds. Finally, users can browse through a man-machine interface to access the information.

Ontology with its inference mechanism is the key technology of semantic Web. With the development of semantic Web, ontology is widely used in many domains. Ontology provides a brand new model of knowledge management, which solves the problems of knowledge organisation, knowledge retrieval and so on[27]. Most users of knowledge management systems are not knowledge management professionals as such they lack the basic understanding of knowledge organization. Visualisation is a better technique to show the knowledge structure to such users.

Hui[28] bring out the difference between visualisation in knowledge management system based on ontology molecule and in the ordinary visualisation system and puts forward the principle of visualisation designation in knowledge management systems based on ontology molecule. A case study of application of ontologies in knowledge management in Brazilian energy utility is presented by Almeida and Barbosa[29].

Benson[30] explored potential of ontological principles in the formal description of archival photographs. This exploration was for knowing whether the photograph archivists need more formalised system of representation or existing schemes are satisfactory. Based on the exploration the author has proposed a new semantic archives model.

Guo and Zhang[31] used ontology to represent domain knowledge for the purpose of a question-answering system. Ubiquitous learning grid uses ubiquitous computing environment to infer and determine the most adaptive learning contents and procedures at anytime, anyplace and with any device. Liao[32], *et al.* proposed an ontology-based ubiquitous learning grid (OULG). It helps to resolve the difficulties concerning how to adapt learning environment for different learners, devices, places. OULG identifies and adapts the aspects of domain, task, devices, and background information awareness, so that the adaptive learning content could be delivered. Ontologies could be used in the e-learning field to model educational domains and to build, organise and update specific learning resources[33]. By using ontologies it is possible to define a description base for scholarly events to enable software agents to crawl and extract scholarly event data, and to facilitate unified access to this data. The collected data may then be mined for non-obvious knowledge. Jeong and Kim[34] developed and implemented the SEDE, i.e., Scholarly Event DEscription ontology for this purpose.

Chi and Chen[35] demonstrated how the semantic rules in conjunction with ontology can be applied for inferring new facts to dispatch news into corresponding departments. The system comprises finding a glossary from electronic resources, gathering organisation functions as controlled vocabularies, and linking relationships between the glossary and controlled vocabularies. Web ontology language is employed to represent this knowledge as ontology, and semantic Web rule language is utilised to infer implicit facts among instances. Use of ontology in financial institutions is demonstrated by Ye[36], *et al.* The authors first developed static ontology, dynamic ontology, and social ontology and then they integrated these in the OWL and the Semantic Web Rules Languages (SWIRL) framework, both of which are machine readable.

E-science is increasingly being used to address scientific problems that require cross-disciplinary knowledge, such as climate change, natural disasters, and environmental health. Brodaric and Probst[37] integrated two ontologies so as to derive benefits from unified disciplinary ontologies for retrieving information on e-science. Ontologies could be used for seamless integration of medical images and different user applications by providing direct access to image semantics. Semantic image retrieval provides the basis for clinical decision support and computer-aided diagnosis[38]. Ontologies could also be used to develop first-aid knowledge question answering system[39].

Ontology's role in organising the vast stock of knowledge on the Internet is highlighted by Gokhale[40]. Traditional search engines work on the basis of word-match technology which results in low efficiency. This problem could be overcome by using an intelligent search method based on domain ontology[41]. Further Aguilar-Lopez, Lopez-Arevalo and Sosa-Sosa[42] suggested that the relevance of search engine results could be enhanced by filtering the contents through the integration of domain ontologies, the WordNet thesaurus, and a hierarchical similarity measure.

7. CONCLUSION

It could be seen from the above that a large number of ontologies have been developed for diverse applications. Many applications benefit from the use of a suitable ontology but it could be difficult to determine which ontology is best suited to a particular application. To be able to select a suitable ontology, it is necessary to evaluate them.

Though there are many criteria available such as application, methodologies, etc. Yu, Thom and Tam[43] believe the requirement should be the most important criteria for ontology evaluation. So the authors presented a new method named ROMEO, i.e., Requirements-Oriented Methodology for Evaluating Ontologies. It identifies requirements that the ontology must satisfy and map these requirements to evaluation measures.

REFERENCES

1. Lei, Wang; Kuanjiu, Zhou & Peng, Qiu. Automatic domain ontology construction. *J. China Soc. Sci. Tech. Inf.*, 2010, **29**(1), 45-52.

2. Saravanan, Ravindran & Raman. Improving legal information retrieval using an ontological framework. *Arti. Intell. Law*, 2009, **17**(2), 101-24.

3. Jiang & Tan. CRCTOL: A semantic-based domain ontology learning system. *J. Amer. Soc. Inf. Sci. Technol.*, **61**(1), 150-68.

4. De Nicola, Missikoff & Navigli. A software engineering approach to ontology building. *Information Systems,* 2009, **34**(2), 258-75.

5. Lin & Hanqing. Research on semi-automatic construction of domain ontology based on statistical NLP technique. *J. China Soc. Sci. Tech. Inf.,* 2009, **28**(2), 201-07.

6. Gaeta, Orciuoli & Ritrovato. Advanced ontology management system for personalised e-learning. *knowledge-based Systems*, 2009, **22**(4), 292-301.

7. Gruber, T.R. A translation approach to portable ontologies. *Knowledge Acquisition,* 1993, **5**(2), 199-220.

8. Guarino, N. Understanding, building, and using ontologies. *Inter. J. Human-Computer Stud.,* 1997, **46**, 293-310.

9. Guarino, N. Formal ontology and information systems. *In* Formal ontology in information systems, edited by N. Gurino. IOS Press, Amsterdam. 1998. pp. 3-15.

10. Srinivasan, Pepe & Rodriguez. A clustering-based semi-automated technique to build cultural ontologies. *J. Amer. Soc. Inf. Sci. Technol.*, **60**(3), 608-20.

11. Jun & Yuhua. An automatic approach to ontology building by integrating traditional knowledge organisation resources. *J. China Soc. Sci. Tech. Inf.*, 2009, **28**(5), 651-57.

12. Junzhi & Dandan. Conversion from law framenet database to OWL ontology. *J. China Soc. Sci. Tech. Inf.*, 2009, **28**(6), 851-56.

13. Papadakis. Subject-based information retrieval within digital libraries employing LCSHs. *D-Lib Magazine,* 2009, **15**(9-10).

14. Hua, *et al*. Research on SUMO-based emergency response preplan ontology model. *J. China Soc. Sci. Tech. Inf.*, 2009, **28**(3), 331-38.

15. Ziyu & Lei. Research on construction methodology of domain ontology for high, speed railway. *J. China Soc. Sci. Tech. Inf.*, 2009, **28**(2), 195-200.

16. Guangrong. Construction and realiation of course knowledge ontology in e-learning system. *J. China Soc. Sci. Tech. Inf.*, 2009, **28**(4), 499-508.

17. Nguyen, *et al*. Meta-relation and ontology closure in conceptual structure theory. *Arti. Intelli. Law*, 2009, **17**(4), 291-320.

18. Yingfang & Lei. Study on construction of aviation ontology. *J. China Soc. Sci. Tech. Inf.*, 2010, **29(**2), 223-31.

19. Wang, Jie & Zhang. Web ontology data matching for integration: Method and framework. *Int. J. Web Inf. Sys.,* 2009, **5**(2), 220-38.

20. Kiu & Lee. Interoperating ontologies of organisational memory through hybrid unsupervised data mining. *VINE: J. Inf. Know. Manage. Sys.,* 2009, **39**(4), 353-75.

21. Mingbao, Jing & Xiuli. Research on field ontology applications in information retrieval. *J. China Soc. Sci. Tech. Inf.,* 2010, **29**(2), 215-22.

22. Liao, *et al.* A library recommender system based on a personal ontology model and collaborative filtering technique for English collections. *The Electronic Library,* 2010, **28**(3), 386-400.

23. Zheng, *et al.* Exploiting corpus-related ontologies for conceptualising document corpora. *J. Amer. Soc. Inf. Sci. Technol.,* 2009, **60**(11), 2287-99.

24. Zhonghua, *et al.* Researches on multilingual thesauri-ontology. *J. China Soc. Sci. Tech. Inf.*, 2009, **28**(5), 664-71.

25. Cheung, Hunter & Drennan. MatSeek: An ontology-based federated search interface for materials scientists. *IEEE Intell. Syst.,* 2009, **24**(1), 47-56.

26. Chen, *et al.* Dynamic RSS information push service mechanism based on ontology of user information needs. *The Electronic Library*, 2009, **27**(2), 222-35.

27. Hui, *et al*. Research on digital library based ontology evolution and knowledge management (I) ontology molecule theory. *J China Soc. Sci Tech. Inf.,* 2009, **28**(3), 323-30.

28. Hui, *et al*. Research on digital library based ontology evolution and knowledge management (III): The description of dynamic knowledge. *J. China Soc. Sci. Tech. Inf.,* 2009, **28**(5), 643-50

29. Almeida & Barbosa. Ontologies in knowledge management support: A case study. *J. Amer. Soc. Inf. Sci. Technol.,* **60**(10), 2032-47.

30. Benson. The archival photograph and its meaning: Formalisms for modeling images. *J. Archival Org.,* 2009, **7**(4), 148-87.

31. Guo & Zhang. Question answering based on pervasive agent ontology and semantic Web. *Knowledge-based Systems*, 2009, **22**(6), 443-48.

32. Liao, *et al.* The construction of an ontology-based ubiquitous learning grid. *Inter. J. Dist. Edu. Technol*., 2009, **7**(3), 1-25.

33. Jeong & Kim. SEDE: An ontology for scholarly event description. *J. Inf. Sci.,* 2010, **36**(2), 209-27.

34. Chi & Chen. Ontology and semantic rules in document dispatching. *The Electronic Library,* 2009, **27**(4), 694-707.

35. Ye, *et al.* Ontologies for crisis contagion management in financial institutions. *J. Inf. Sci.,* 2009, **35**(5), 548-62.

36. Brodaric & Probst. Enabling cross-disciplinary e-science by integrating geoscience ontologies with Dolce. *IEEE Intell. Syst.,* **24**(1), 66-77.

37. Sonntag, *et al*. Pillars of ontology treatment in the medical domain. *J. Cases Inf. Technol*., 2009, **11**(4), 47-73.

38. Kuanjiu, Yupeng & Huazhen. Ontology based first-aid knowledge mobile question answering system. *J. China Soc. Sci. Tech. Inf*., 2009, **28**(1), 121-27.

39. Gokhale, P.A. Ontology: A tool for organisation of knowledge. *Information Studies,* 2009, **15**(4), 233-42.

40. Hengmin, *et al.* Study on method of the global web intelligent search based on domain ontology. *J. China Soc. Sci. Tech. Infor.*, 2010, **29**(1), 9-15.

41. Aguilar-Lopez, Lopez-Arevalo & Sosa-Sosa. Use of ontologies to enhance the results of web search engines. *Profesional de la Informacion*, 2009, **18**(1), 34-40.

42. Yu, Thom & Tam. Requirements-oriented methodology for evaluating ontologies. *Information Systems*. 2009, **34**(8), 766-91.

About the Authors

Dr (Mrs) Neela J Deshpande is Professor in the Department of Library and Information Science, University of Pune. She obtained her Master's degree in Analytical Chemistry from Nagpur University and MLibSc from University of Pune. She is a recognised guide for PhD of University of Pune, Shivaji University of Kolhapur, etc. She has more than 70 articles and seven books to her credit. Her areas of interest are LIS research, digital libraries, knowledge management systems, information technology applications, and document descriptions.

Dr Rajendra Kumbhar obtained his Masters degree in LIS from Pune University and doctoral degree from Dr B.A.M. University, Aurangabad. He is presently working as Associate Professor, in the Department of Library and Information Science, University of Pune. He has more than 40 articles and 15 books to his credit. His areas of interest are LIS research, IR, thesaurus, knowledge organisation systems. He did research on thesaurus and depth classification and has published a book titled Thesaurus of LIS Terms.