

INTEGRATED APPROACH TO ONTOLOGY DEVELOPMENT METHODOLOGY WITH CASE STUDY

Sandeep Chaware¹, Srikantha Rao²
^{1,2}MPSTME, NMIMS University, Mumbai, INDIA
{¹smchaware@gmail.com, ²dr_s_rao@yahoo.com}

Abstract

Knowledge can be represented by ontology. In an enterprise context, they reflect the relevant knowledge based on enterprise-specific concepts and their relations. In order to develop ontology, there are various methodologies, where each one may have some pitfalls depending on context. In this paper, based on an analysis of existing methodologies, we explore the possibilities by proposing an integrated model for developing ontology, which can be used to build any kind of ontology. Our main intention is to reduce development time and effort. We had proposed the system with respect to shopping mall domain, where dynamically ontologies can be prepared to get the information faster and correct. Further, these ontologies can be used for mapping. We compare our model with the existing developing methodology, and we had tried to remove the possible pitfalls of the existing techniques.

Keywords

Ontology, Skeleton, Sensus, Methontology, Gruinger and Fox, WordNet, Shopping Mall.

1. INTRODUCTION

There are various domains where there is need of an application which will take care of various linguistics people to provide better services. For example, holy places, where millions of multilinguistics devotees come and demands services in their local language. Another example is government services, where each service should be in local language for common people. The popular example is Shopping Mall, where millions of people come with various language speaking communities. In this case, it is very difficult to provide better and faster service, especially when the people demands service in their local language. There are several solutions suggested, the best one is use of ontology. The ontology can be used to represent the information in a better way, which can be used to provide the service to all. Ontology can be defined as “An explicit specification of a conceptualization”. Ontology is arranged in a lattice or taxonomy of concepts in classes and subclasses. Each concept is typically associated with various properties describing its features and attributes as well as various restrictions on them. It is a shared conceptualization of knowledge in a particular domain. The top-level ontologies describe very general concepts like space, time, matter, object, event, action etc. which are independent of a particular problem or domain. Other ontologies are domain and task related to domain or activity. Examples of existing ontologies are: top-level ontologies like SUO (Standard upper Ontology) provides definition for general purpose terms, SENSUS, natural language-based ontology developed by NLG at ISI to provide a broad conceptual structure for working in machine translation, WordNet, which is a large lexical database for English created at Princeton University or IITB for Indian languages and medical ontologies such as gene, Galen and Menelas [1]. These ontology can be used in many applications like Information and Knowledge Management, Military, Education, small-and-large enterprises, industrial risk analysis, medical, communication, construction of emergency plans.

In this paper, we had made survey of various ontology development methodologies and found out some pitfalls. We had proposed an integrated approach to the development of ontology. The case study of Shopping Mall ontology had proved the approach the best one.

2. INTRODUCTION TO BUILDING ONTOLOGIES

The basic steps in building ontology are straightforward. Various methodologies exist to guide the theoretical approach taken and numerous ontology building tools are available. Ontology is typically built in more-or-less the following manner [1]:

1. **Acquire domain knowledge:** Assemble appropriate information resources and expertise that will define, with consensus and consistency, the terms used formally to describe things in the domain of interest. These definitions must be collected so that they can be expressed in a common language selected for the ontology.
2. **Organize the ontology:** Design the overall conceptual structure of the domain. This will likely involve identifying the domain's principal concrete concepts and their properties, identifying the relationships among the concepts, creating abstract concepts as organizing features, referencing or including supporting ontologies, distinguishing which concepts have instances, and applying other guidelines of your chosen methodology.
3. **Flesh out the ontology:** Add concepts, relations, and individuals to the level of detail necessary to satisfy the purposes of the ontology.
4. **Check your work:** Reconcile syntactic, logical, and semantic inconsistencies among the ontology elements. Consistency checking may also involve automatic classification that defines new concepts based on individual properties and class relationships.
5. **Commit the ontology:** Incumbent on any ontology development effort is a final verification of the ontology by domain experts and the subsequent commitment of the ontology by publishing it within its intended deployment environment.

3. ONTOLOGY METHODOLOGY: A SURVEY

Basically, a series of approaches have been reported for developing ontologies. In 1990, Lenat and Guha published the general steps and some interesting points about the Cyc development. Initially, the Enterprise Ontology and the TOVE (TOronto Virtual Enterprise) project ontology had been proposed in the domain of enterprise modeling. Bernaras et al. presented a method used to build ontology in the domain of electrical networks as part of the Esprit KACTUS project. The methodology METHONTOLOGY appeared at the same time. In 1997, a new method was proposed for building ontologies based on the SENSUS ontology. Some years later, the On-To-Knowledge methodology appeared as a result of the project with the same name. However, all these methods and methodologies do not consider collaborative and distributed construction of ontologies. In this paper, we are describing some of the methodologies for building ontologies [1].

3.1 Skeleton Methodology

This methodology is based on the experience of developing the Enterprise Ontology, ontology for enterprise modeling processes. A plan or draft for a project along with activities can be represented as ontology. The steps are, first, identify the main purpose of the ontology, second build the ontology, where the key concepts and its relationships can be captured and third, code it with proper language and may be integrated with existing ontology. We can use either top-down or bottom-up approach to represent the ontology [1] [2]. This method is simple to implement but limited to scope.

3.2 Gruninger And Fox Methodology

Gruninger and Fox proposed a methodology that is inspired on the development of knowledge-based systems using first order logic. This methodology has been suggested as TOVE project ontology within domain of business processes and activities modeling. This represents logical model of knowledge. The steps are: 1) Capture the motivating scenarios. 2) Formulation of informal competency questions, where the scope of the ontology can be decided. 3) Formulation

of formal competency questions, which specify the terminology with definition and constraints. 4) Specification of axioms and definition within the formal language. 5) Finally, specify the conditions under which the solutions to the questions are complete. In this methodology, the ontology can be built by using questions and answers for motivating scenarios, which represents main concepts, properties, relations and axioms on the ontology [1][2]. The methodology will extend the scope but the procedure is complex.

3.3 Methontology Methodology

This methodology will give the construction of ontology at knowledge level. The ontology development process is: 1) Determine the tasks to be performed when building ontology i.e. scheduling, control, quality assurance, specification, knowledge acquisition, conceptualization, integration, formalization, implementation, evaluation, maintenance, documentation and configuration management. 2) Determine the life cycle of ontology as number of stages. This represents the activities to be performed in each stage and how the stages are related. 3) Determine the techniques used in each activity, the products that each activity output and how they have to be evaluated [1][2]. These methodologies deal with software engineering concepts. It handles all the activities in details.

3.4 Sensus Methodology

The method based on Sensus is a top-down approach for deriving domain specific ontologies from huge ontologies. The steps are: 1) A series of terms are taken as seed. 2) These seed terms are linked by hand to SENSUS. 3) All the concepts in the path from the seed terms to the root of SENSUS are included. 4) Terms that could be relevant within the domain and have not yet appeared are added. 5) Finally, for those nodes that have a large number of paths through them, the entire subtree under the node is sometimes added, based on the idea that if many of the nodes in a subtree have been found to be relevant, then the other nodes in the subtree are likely to be relevant as well [2]. This methodology uses the existing ontology, where the merging will be complex due to different structures.

3.5 WordNet Methodology

WordNet is a lexical database for the English language. It groups English words into sets of synonyms called *synsets*, provides short, general definitions, and records the various semantic relations between these synonym sets. The purpose is twofold: to produce a combination of dictionary and thesaurus that is more intuitively usable, and to support automatic text analysis and artificial intelligence applications.

The hypernym/hyponym relationships among the noun synsets can be interpreted as specialization relations between conceptual categories. In other words, WordNet can be interpreted and used as a lexical ontology in the computer science sense. However, such ontology should be corrected before being used since it contains hundreds of basic semantic inconsistencies such as (i) the existence of common specializations for exclusive categories and (ii) redundancies in the specialization hierarchy. Furthermore, transforming WordNet into a lexical ontology usable for knowledge representation should normally also involve (i) distinguishing the specialization relations into *subTypeOf* and *instanceOf* relations, and (ii) associating intuitive unique identifiers to each category. WordNet has also been converted to a formal specification by means of a hybrid bottom-up top-down methodology to automatically extract association relations from WordNet, and interpret these associations in terms of a set of conceptual relations, formally defined in the DOLCE foundational ontology [3].

3.6 Pitfalls of Existing Ontology Development Methodologies

- 1) Some of the methodologies are too formal and only useful for small-scale applications or contexts.
- 2) Some methodologies like Methontology, is more mature and detailed where as some steps can be either integrated or rejected depending on context [5].
- 3) Integration of existing ontologies may be difficult due to change in structure or plan.
- 4) For each scenario, we can not decide the competency questions, which will represents the definition and constraints of terms used in ontology.
- 5) Some of the methodologies are complex to build and takes long time and utilize large resources.

4. PROPOSED INTEGRATED ONTOLOGY DEVELOPMENT METHODOLOGY

When ontology technologies emerged in the 1990s, the focus on knowledge acquisition influenced the way new capabilities were put to use in the field. Early ontology methodologies adopted the method for developing knowledge bases. This orientation is not as evident in today's tools. There is also increasing support for common upper level ontologies like WordNet, Cyc, and others.

Figure 1 shows the proposed integrated ontology development methodology. Here, we can integrate the existing ontology methodologies in order to remove the pitfalls, and hence to improve the overall procedure to build ontology. The modules are motivating user scenarios module, formal/informal questions and answer generation module, extraction of terms and constraints module and build ontology module. Each module is described below in brief.

- 1) **Motivating User Scenarios Module/Keyword:** This module is responsible for capturing the motivating user scenarios for particular domain. In this module, a keyword can be entered and processed to extract the abstract concept. This module can be manually maintained or we can use UML diagram to represent them. With these scenarios, we can formulate the exact purpose and need of ontology for the domain.
- 2) **Formulation of Formal/Informal Questions and Answer Module:** Within this module, the possible informal and formal questions and answers can be generated for the motivating scenarios. These questions and answers can be generated either manually or by the system according to scenario or from abstract concept of entered keyword. This module will determine the scope of the ontology. These questions and answers may be different for different users and scenarios. No single ontology structure will satisfy the need of user.
- 3) **Extraction of Terms and Constraints Module:** Once you know the scope of the ontology, the terms and constraints can be extracted from them to know the concepts and their relationships for the domain. This step can be carried out manually or by parsing of the keywords from the questions/answers.
- 4) **Build Ontology Module:** Finally, we can build the ontology by looking at these concepts and their relationships. We can use any approach, but the top-down approach is better since it extract the terms from abstract to specific concepts.

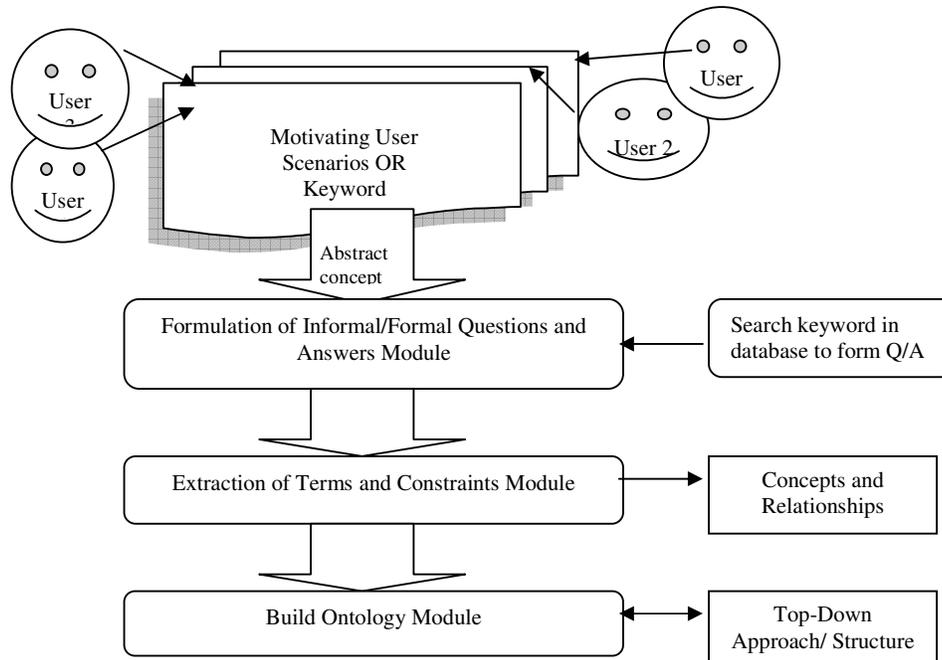


Figure: 1 proposed Integrated Ontology Development

5. THE DEVELOPMENT OF SHOPPING MALL ONTOLOGY: A CASE STUDY

5.1 Purpose and Scope

The purpose is to design and develop ontology in an area where millions of users are visiting the shopping mall every day, with multilingual background. The scope is limited to a number of areas. The visitors are looking for information about a shopping mall and its shops. Each one will look for the details of shops, available products, various schemes and services from shops within shopping mall. The ontology will play an important role in order to serve all these information to the visitors. The entire ontology will provide all the relevant information irrespective of languages [4].

5.2 Domain and Source

Before addressing design issues, the first task was to decide upon an area to investigate as the domain of interest. Shopping Mall database were chosen as the domain. It is a broad subject area that was likely to yield a large number of concepts and associated relationships. These could be used to test the initial hypothesis that the ‘is-a’ relationship is sufficient to express the semantics. It is a mature discipline within computing with an agreed body of core knowledge that is readily available. A Shopping Mall was used as the source – “Raghulila – The Mall, Kandivali (West)”. There are advantages to using a shopping mall as the source of ontology concepts. First, coverage of the domain of interest is extensive as the purpose of visitors is to provide a good grounding in the subject. Second, when each new shop or product is introduced, new terms are explained, thus providing the basis for concept definitions.

5.3 A Systematic Approach to Ontology Modeling

We are proposing the following steps to build ontology for shopping mall, according to the proposed integrated version of methodology.

5.3.1 Step1=> Motivating User Scenarios/Keyword

We can capture various motivating user scenarios, such as first, getting the details about Raghulila shopping mall, i.e. its location, address, phone numbers etc. Second, this shopping mall is running which movies and its details, third, a user wishes to buy jeans of particular brand, he wanted to know its details etc. A keyword can be entered to get the information as service. For Example, Raghulila Shopping Mall, with this the abstract concept has been generated as Raghulila Shopping Mall, further more information can be served from possible questions and answers.

5.3.2 Step2=> Formulation of Formal/Informal Questions and Answers

In this step, formal and informal questions and answers for the various scenarios can be formulated, such as, what is the location of shopping mall you are looking for? What details you want? Which movie, rate of tickets for movie, availability of tickets, date & time? Which brand for jeans? What is the price range? Etc. These questions and answers can be formulated either manually or automatically as per keyword entered.

5.3.3 Step3=> Extraction of Terms and Constraints

The various terms and its constraints can be extracted from the answers from step 2. The terms can be shopping mall name, address, phone numbers, and location. Name of the movie, availability of tickets, date and time of movie etc. Availability of jeans of particular brand along with name of the shop and its details. These terms leads to the concepts and its relationships. For example, concept may be movie 'Ravan'. Its attributes will be date & time of show, price of tickets, availability etc. The relationships can be is-a or has etc.

5.3.4 Step4=> Build Ontology

For building ontology, we will use top-down approach, since we may know the abstract of the shopping mall, and further derive the specification and gene rationalization about the concepts and its relationships. To build the ontology, we can use the tree or graph like structure.

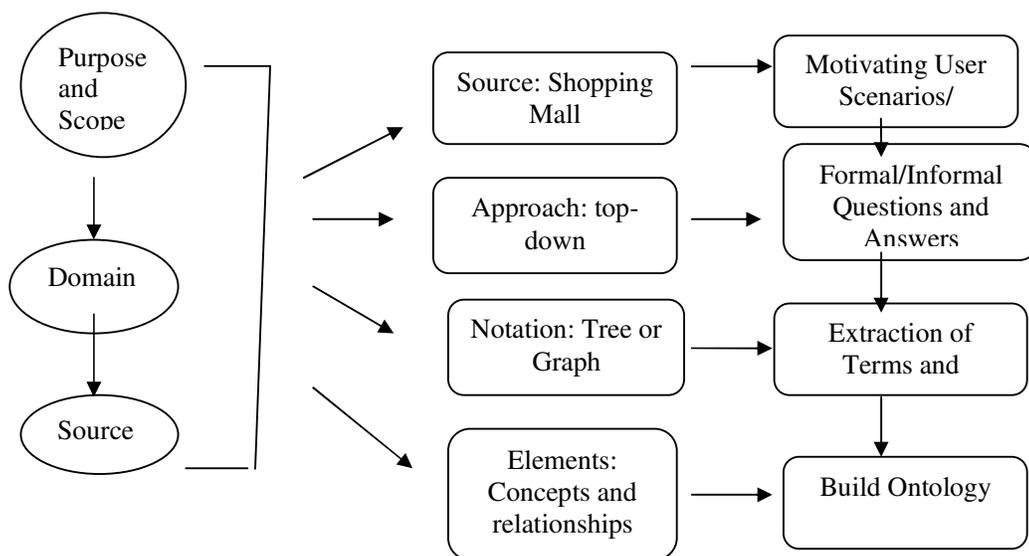


Figure 2: Summary of Ontology Development for Shopping Mall

5.4 Proposed Algorithm for Ontology Development

We are proposing an algorithm for the development of ontology for domain Shopping Mall. The figure 3 shows the steps.

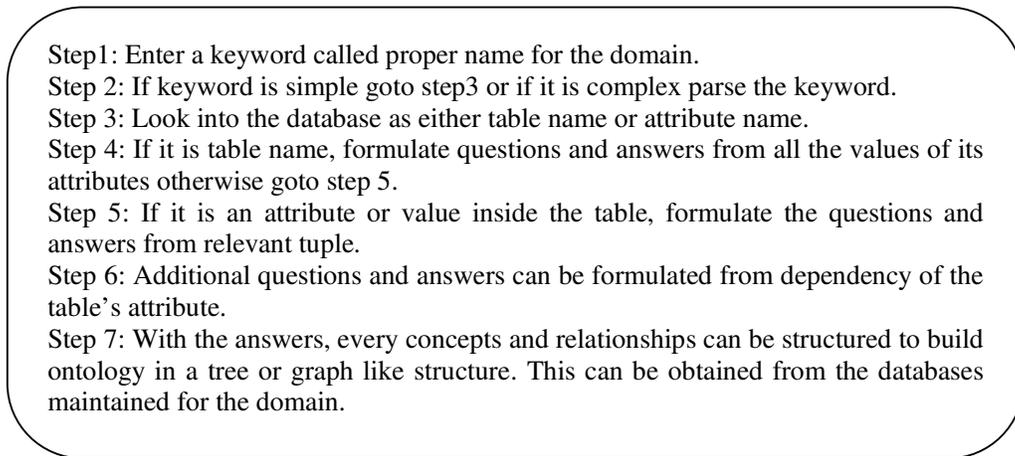


Figure 3: Proposed Algorithm for Building Ontology for Shopping Mall

6. CONCLUSIONS

There are many ontology building methodologies suggested for various domains. With this proposed methodology, we can effectively build the ontology with all possible user scenarios or simple and complex keyword. We are developing the concepts and its relationships dynamically. Also, this methodology is faster than the earlier one, since we are using top-down approach to build the ontology.

REFERENCES

- [1] Fernandez Lopez, M., Overview of Methodologies for building ontologies.
- [2] Oscar Carcho et al., Methodologies, Tools and Languages for Building ontologies. Where is their meeting point?, Data and Knowledge Engineering 46 (2003), 41-64.
- [3] WordNet: Wikipedia, the free Encyclopedia.
- [4] Sinead Boyce and Claus Pahl, Developing Domain Ontologies for Course Content, Educational Technology and Society, 10(3), 275-288, 2007.
- [5] Annika ohgren, kurt Sandkuhl, Towards a Methodology for Ontology Development in Small and Medium-sized Enterprises, IADIA International conference on Applied Computing 2005.

Authors

¹He is working as Asst. Prof. at D.J.Sanghvi College of Engineering, Mumbai. He is pursuing PhD. (Engineering) from NMIMS University, Mumbai, INDIA.

²He is Director at Late Hiray College of Master in Computer Applications, Bandra (E) Mumbai. He is also guiding PhD students at MPSTME, NMIMS University, Mumbai, INDIA.