

Ontology Development for the Domain of Software Requirement Elicitation Technique

A. Awal, A. Mishra, Gana M. Usman, A. AbdulG

^{1,3,4}Department of Mathematics/Computer science, IBB University

²Baze University

Abstract: - Domain ontology is the ontology of describing domain-specific objects, it provides definitions of the concepts in a specific domain, as well as the relationships between them. 5W1H analysis method is widely utilized to give a comprehensive and delicate analysis to a specific issue. The Hierarchical Modelling Methods which includes: Button-Up, Top-Down and Middle-Out is used to model concepts. These conceptual modelling methods solved the real problem in different application domain, but they did not give an analysis and extraction method for concepts and relations in domain ontology, especially large and complex ones, and not present a general modelling method suitable for all kinds of domain ontology. Domain ontology for software requirement elicitation is developed using 5W1H approach, the ontology specify concepts (classes) and relations between the concepts in software requirement elicitation domain. The 5W1H approach helped in extracting and analyzing concepts and relations within the domain of discourse. In this project, 5W1H conceptual modelling method is used for classifying the concepts in the software requirement elicitation domain, the concepts in the domain were extracted from six interrogative elements of 5W1H, which include: Who, When, Where, What, Why and How, and relations between those concepts is built.

Keywords: 5W1H, Domain Ontology, Requirements Elicitation, Class, Concepts.

INTRODUCTION:

The success of a software project is heavily dependent on the quality of requirements capture and analysis. According to the survey in many failed software projects, the projects that failed because of the bad requirements account for about 45%. Errors made in this process are extremely expensive to correct when they are detected during implementing or testing period. Affected software quality would give a significant impact on society and economy such as increasing costs and business interruption (Fauziah B, et al. 2005). A good requirement specification document can help to avoid the early mistakes, improve efficiency and rate of success of software project development but this can only be achieved if requirements are elicited properly. In the past two decades, many experts in the field of software requirement elicitation had come up with many methods of capturing requirements, however, there is still communication gap between the customer and requirement analyst, one of the reasons is lack of explicit or agreed-upon definition of the key concepts in software requirement elicitation domain. Thus, the need to develop Ontology for the Software requirement elicitation domain. The creation of ontology in the domain of requirement elicitation can

strengthen the communication between customer and requirement analyst which lead to the development of more reliable software. The ontology is a repository of questions, can be expanded by adding more questions. Ontology is a discipline that is part of the knowledge representation field (Sowa, 2000). In ontology, a precise definition is associated with each concept and relationship type that is used.

LITERATURE REVIEW

In the study by (LI *et al*, 2007) a formal approach that uses description logic to describe ontology is suggested, integrity and derivation rules are modelled, these rules restrict business behaviour. Syntax, semantics and visualization are the three aspects in which these rules are represented. The requirements model checking framework is also provided combining domain ontology and domain rules, which makes the requirements elicitation process both guided by domain ontology and restricted by domain rules. Hence, making the acquired requirements comply with both the business rules and domain knowledge. It gives a requirements model with UML class diagram and checks the model using domain ontology and rules.

According to (Yuqin and Wenyun, 2006), domain ontology-based approach was proposed which can be used to elicit and also analyze requirements. The top down method where problem is decomposed into several sub problems is used. when decomposing problem domain, abstract stakeholders other than practical ones where used. Incompleteness and inconsistency of domain requirement is detected using reasoning logic based on ontology relations. The proposed requirements elicitation method using ontology does not give a method to improve the quality of the elicited requirements.

In the study by (Maria and Achilles, 2011), A domain Ontology was developed using use case diagrams. this is done so that creators of educational materials that are related to software engineering course can have common references. Some software engineering experts assessed the ontology and satisfies that all aspects of UCD is adequately covered. the Ontology proposed can be used in various aspects such as in the field of object-oriented analysis, a part of e-learning application etc. The proposed Use Case diagram ontology fulfils the requirements that he had defined, as it can help to the semantic annotation and retrieval of learning resources, such as learning objects and learning goals, and it can answer queries about the field of

Use Case Diagrams. This ontology is only suitable for object-oriented analysis domain.

METHODOLOGY

In this study, the methodology of (Uschold, M. King, M. 1995) is employed. This is based on the experience of developing the *Enterprise Ontology*, which is a collection of

terms and definitions relevant to business enterprises (Grüniger, M., Fox, M, 1995). The stages are as follows: Identifying the domain, Concepts Identification, Concepts Organization, Ontology representation and finally, Evaluation. It is worth mention that, for the Concept Organization, the “5W1H” conceptual modelling method is employed and the Ontology representation is done using Protégé. Fig 1 below represents the methodology

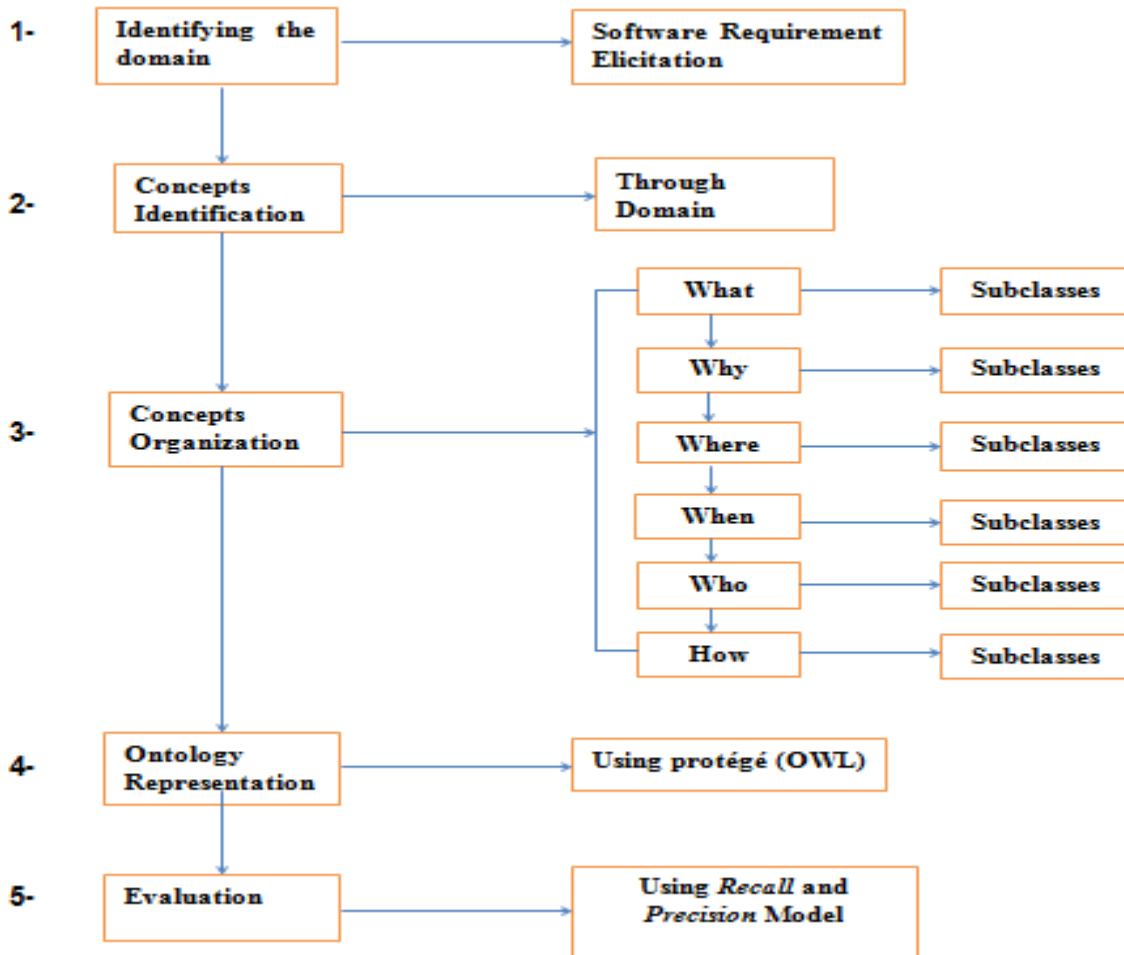


Figure 1: Research Methodology

3 Identifying the domain: This is the first stage. This is where we identify the domain to build the ontology for. In our case, the domain is Software Requirement Elicitation.

(a) Identification of the key concepts and relationships: The second task towards generating the ontology was to compile a list of possible concepts to include in the ontology. The domain experts help in providing the concepts for the ontology, they help in providing questions and concepts that are likely to come under each element of the 5W1H. Minimal set of relationships was created that would be sufficient to represent and link the concepts within the Ontology which include: the ‘HasSubtype’, ‘HasAsPart,’ ‘HasAsExample’ and ‘HasAsSubquestion’ relationships.

(b) Concepts Organization (Hierarchy): There are several possible approaches in developing a class hierarchy (Uschold and Gruninger 1996). In this study, we use the top-down approach, which begins with the definition of the most general concepts and subsequent specialization of the concepts. The 5W1H conceptual modelling method is a modelling method that is made up of six interrogative element which are *What, Why, Where, When, Who* and *How*. Through 5W1H approach, concepts of domain ontology is extracted from six aspects and relations between those concepts is built. 5W1H method is originally used to completely report a piece of news story by answering six questions as follows: *What, Why, Where, When, Who* and *How* (Carmagnola, F. 2008). The 5W1H method is widely utilized to give a comprehensive and delicate analysis to a specific issue or knowledge, such as in production,

management, marketing and so on. Similarly, domain ontology covers distinct knowledge among certain domains. To construct domain ontology, the first important step is to completely and accurately extract knowledge as a set of concepts and relations, and their domain as the same (Stader, J. 1996).

In the ontology, the most general classes of questions were created first: *What, Where, Why, When, Who, and How*, all as subclasses of the default root class *Thing*. Each of these classes is followed by subclasses which are less general than the upper classes, and subsequent specialization of these classes. The figures below diagrammatically depict how the ontology hierarchy has been designed:

(c) To represent this ontology, Protégé is used. This is an ontology editing tool, made by Stanford University. its user Interface is highly usable. Protégé is highly recognised in terms of its scalability and extensibility. Another advantage of Protégé is that it has component based architecture as such the builder can add new, create appropriate plug-ins and add new functionalities. The system is also constructed in an open modular fashion (OntoWeb, 2002). It is open-source software, and has both application type of GUI and API interface, which greatly enhances its flexibility. Protégé is one of the most commonly tools for ontology editing. OWL/XML was chosen as an underlying representation language. Protégé has classes, individuals and

properties as the building blocks for representing knowledge. The Classes Tab is an ontology editor, which designs classes in flexible style and organizes classes as hierarchy.

(d) Evaluation: there are many categories of evaluation. For this study, the category of evaluation used is where evaluation is done by humans who try to assess how well the ontology meets a set of predefined criteria, standards, requirements, etc. (Gómez-Pérez, A., et al, 2004). This Ontology is evaluated using Recall and Precision method. The Ontology is assessed by some set of people using certain predefined criteria A number of agreed-upon subclasses under each element of the 5W1H are considered relevant or irrelevant being under the superclass (*what, where, why, who, when, or how*) and to the topic of the ontology. *Recall*: is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. While, *precision*: is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. They are both usually expressed as a percentage. *Recall* of each element of the 5W1H will be $100\% : (\frac{A}{A} \times 100)$, that is, No. of relevant subclasses assess over total No. of relevant subclasses under each of the elements, only the *Precision* will vary. *Precision* is $(\frac{A}{A+C} \times 100)$, where A= No. of relevant subclasses, and C= No. of irrelevant subclasses. See the table below:

Classes	No. of Relevant (A)	No. of Irrelevant(C)	Recall (in per cent)	Precision (in per cent) ($\frac{A}{A+C} \times 100$)
<i>What</i> : subclasses	43	2	100	95.5
<i>Where</i> : subclasses	3	0	100	100
<i>Why</i> : subclasses	5	1	100	83.3
<i>When</i> : subclasses	8	1	100	88.8
<i>Who</i> : subclasses	4	0	100	100
<i>How</i> : subclasses	3	0	100	100

Table 1: Evaluation table showing Recall and Precision of the ontology

CONCLUSION

The presence of the domain knowledge for software requirement elicitation will have a big influence on efficiency and the completeness of the requirements gathered in the requirements elicitation process for developing information systems. Acquisition of the domain knowledge is especially important for the requirements elicitation by novices in the task. In this project, a domain ontology was developed for requirement elicitation, which defines concepts in the domain. The effectiveness of the ontology was confirmed by the evaluation carried out using *recall and precision* method which determine the relevancy and irrelevancy of the concepts defined in the domain ontology.

REFERENCES:

[1] Bagiampou, M., & Kameas, A. (2012, September). A Use Case Diagrams ontology that can be used as common reference for Software Engineering education. In 2012 6th IEEE International Conference Intelligent Systems (pp. 035-040).

[2] Carmagnola, F, (2008)“The five Ws in user model interoperability”, in Proc. 2008 UbiqUM.,

[3] Fauziah Baharom, A. D., Abdul Razak Hamdan (2005). "A Survey on the Current Practices of Software Development Process in Malaysia" Journal of ICT 4: 57 76.

[4] Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004): Ontological Engineering. Springer, New York. USA.

[5] Grüniger, M., Fox, M. (1995). Methodology for the design and evaluation of ontologies. In: Proceedings of the Workshop on Basic Ontological in Knowledge Sharing.

[6] Lee, Y., & Zhao, W. (2006, May). Domain requirements elicitation and analysis-an ontology-based approach. In International Conference on Computational Science (pp. 805-813). Springer Berlin Heidelberg.

[7] Li, Z. Y., Wang, Z. X., Zhang, A. H., & Xu, Y. (2007). The domain ontology and domain rule-based requirements model checking. International Journal of Software Engineering and Its Applications, 1(1).

[8] Sowa, J. F. (2000). Knowledge Representation – Logical, Philosophical, and Computational Foundations, Pacific Grove, CA, USA: Brooks/Cole.

- [9] Stader, J., (1996) "Results of the Enterprise Project", in Proc. of the 16th Int. Conference of the British Computer Society Specialist Group on Expert Systems.
- [10] Uschold, M. (1996). Building ontologies: Towards a unified methodology. Technical Report-University of Edinburgh Artificial Intelligence Applications Institute Aiai Tr.
- [11] Uschold, M. Grüniger, M. (1996), Ontologies: principles, methods and applications Knowledge Eng. Rev., 11 pp. 93–155.
- [12] Uschold, M. King, M. (1995) Towards a Methodology for Building Ontologies. Workshop on Basic Ontological Issues in Knowledge Sharing.
- [13] OntoWeb, (2002, May). Ontology-based information exchange for knowledge management and electronic commerce IST-2000-29243

AUTHOR PROFILE:

¹Mrs Aisha Awal (M.Sc.) is a lecturer in the Department of Mathematics/Computer Science, IBB University, Lapai, Nigeria.,

Her current research interests include: Requirements analysis, software evaluation

²Dr. Amit Mishra is a senior lecturer in the Department of Computer Science, Baze University, Abuja, Nigeria. His current research interests include: Network security, Cryptology, Machine learning, Software Engineering, Reverse Engineering. He has been publishing in the above-mentioned fields in various national and international journals. He is also Microsoft Certified professional in Win2K and Exchange Server.

³Mr. Gana Mohammed Usman (M.Sc) is a lecturer in the Department of Mathematics/Computer Science, IBB University, Lapai, Nigeria. Usman M. Gana His current research interests include: Computer networks, Server Security, Green Computing.

⁴Mr. Abdulrahman AbdulGaniyu (M.Sc) is a lecturer in the Department of Maths/Computer Science, IBB University, Lapai, Niger state, Nigeria. His scholarly publication has appeared and has been quoted in reputable peer-refereed Journals, Conference proceedings, Newsletters and edited books. He is actively researching into Big Data Security on Cloud Computing, internet security, Updating Web Search Algorithm.