IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol-6, Special Issue-4, May 2018 E-ISSN: 2347-2693

Analysis of Modelling Frameworks for Knowledge Acquisition

Thirunavukkarasu K.[1*], Abhilash Ashu[2], Digvijay Singh[3], Shahnawaz S. Khan[4]

[1*,2,3]School of Computing Science and Engineering, Galgotias University, Greater Noida, India
[4] Department of Information Technology, University College of Bahrain , Janabiyah, Bahrain

Available online at: www.ijcseonline.org

***Abstract*—** Knowledge acquisition is an elementary stage of knowledge engineering. The process includes extracting the raw data from various sources that are later structured and organized in such a form that contribute in providing knowledge. Usually knowledge can be acquired from sources such as primer, manuals and simulation models but a highly elaborate acquisition comes from human experts. This paper projects the various modelling techniques and framework. The proposed Knowledge Acquisition Framework(KAF)emphasizes on the models like Common KADS, MIKE, PROTEGE and their concept of knowledge acquisition. This technology helps in the process of building and framing various application of knowledge engineering.

***Keywords*—** KADS, MIKE, PROTEGE, KAF, Umbrella Approaches, ESPRIT

I. INTRODUCTION

Knowledge acquisition is the term refers to the knowledge which will be used in an organization and it may obtain from various external sources. External knowledge sources are important and one should therefore take a holistic view of the value chain (Gamble & Blackwell 2001). Knowledge acquisition external sources that include suppliers, competitors, partners, alliances, customers, stake holders and external experts.

Knowledge acquisition is the process of enthralling and storing new information in memory, the success of which is often gauged by how well the information can later be retrieved. This process of storing and retrieving depends on the way we do in representation and organization of the information. How knowledge is being used that depends on structure of information. Knowledge acquisition can be improved by considering the purpose and function of the desired information in an organization.

Knowledge engineering makes use of three types of knowledge i.e. explicit, tacit and embedded knowledge. In real scenario, knowledge does not belong to anyone type but it's a mixture of all the three. However, to understand knowledge as a whole, one must know these each type.

A. Explicit Knowledge

Explicit Knowledge is easily understandable and it is expressed and recorded as numbers, codes, words, mathematical, financial and scientific formulas etc. Explicit knowledge is lucid to store communicate and distribute and also it is the knowledge which is found in books, on the web, and other visual and oral means. Opposite of tacit knowledge. This knowledge is most comfortably handled by Knowledge Management System(KMS), which is very effective at facilitating the storage, retrieval, and modification of documents and texts.

B. Tacit Knowledge

Tacit knowledge tops the hierarchy of the knowledge sources in an organization. It can lead to a revolution in the organization as mentioned by Wellman in 2009. According to Gamble & Blackwell[15], the companies that lose focus on tacit knowledge are reducing their innovative capability and strong competitiveness. Polanyi in 1966 introduced and defined these types of knowledge. Tacit knowledge is personal and context dependent. Sources of tacit knowledge are minds of stakeholders. Knowledge can include personal beliefs, values as well as skills, capabilities, and expertise.

C. Embedded Knowledge

Things like culture, routines, structures etc. contains lots of locked knowledge which is called Embedded Knowledge. It is found in rules, manuals, organizational culture, codes of conduct, products, etc. This embedded knowledge can exist in explicit sources but however, knowledge itself is not explicit. As it is difficult to manage and organize embedded knowledge, firms who able to do it reaps competitive benefits.

II. MODELLING TECHNIQUES

Modelling is an activity, the aim of which is to make a particular part or feature of the world easier to understand, define, quantify, visualize, or simulate by referencing it to existing and usually commonly accepted knowledge. Modelling is the term which refers to the process of generating a model as a conceptual representation of some phenomenon. Typically, a model will refer only to some aspects of the phenomenon in question, and two models of the same phenomenon may be essentially different, that is to say, that the differences between them comprise more than just a simple renaming of components. We are able to model the knowledge acquisition through the following techniques.

- Protocol analysis techniques
- Matrix-based techniques
- Sorting techniques

A. Protocol analysis techniques

Protocol Analysis works by identifying the knowledge-objects within a protocol [8], usually, it's called as transcript. Information in this is acquired via interviews or texts to obtain the type of knowledge, such as relationships, decisions, attributes, and goals. This functions used as a bridge between the use of knowledge modelling techniques and protocol-based techniques. For instance, if the transcript lookup for the task of an airport, then attributes like the type of airport, runway length, hanger capacity and maintenance crew etc. would be used for the analysis. Another example, an interview transcript would be looked up and analysed by emphasising all the attributes that are relevant to the project. This would be iterated for all the relevant relationships, attributes, tasks and values. The attributes are used for the problem-solving.

In some cases, more informative categories are used for the identification based on the requirements of the project. For instance, if the transcript lookup for the task of diagnosis, then such categories like symptoms, diagnostic techniques and hypotheses would be looked for the analysis. Such categories may be taken from problem-solving models and generic ontologies[10].

B. Matrix-based techniques

In this kind of technique, a matrix is first created consisting of things as the type of problem encountered according to its possible solution. This technique has a two-dimensional matrix. Important types require the use of frames for illustrating the properties of concepts and the repertory grid technique used to rate, analyse, elicit and categorise the properties of concepts.
These techniques require the making and filling-in of a 2-dimensional matrix (grid, table). Useful examples are:

- Problems v Solutions
- Concepts v Properties (attributes and values)
- Tasks v Resources
- Hypotheses v Diagnostic techniques

The elements within the matrix can contain:

- Symbols (ticks, crosses, question marks)
- Colours
- Numbers
- Text

C. Sorting techniques

This technique is used for analysing or knowing the pattern that how people made a decision by comparing the various concepts which direct to the generation of knowledge about its features, properties, etc. The simplest form is card sorting. Here a number of cards are given to the expert each representing the name of a concept. The only task an expert has that is to repeatedly sort the cards into the stack such the cards have something in common in each pile. An example can be given that a highly skilled expert of astronomy will get different set of cards showing different planets, names and sizes. The expert gives information on the attributes and values they use to denote the properties of concepts by naming each pile. Different types of sorting work around data or photographs as compared to other cards where descriptions are not easy to use.

Triadic elicitation, a method often used in relation with sorting techniques is (aka 'Three Card Trick'). This technique compels the expert to make new attributes as per the requirement. This technique contains information asking the expert what is same and unlike about these three randomly chosen methods, i.e. in what way are two of the same and unlike from the other. This is a way of triggering attributes that are not immediately and very easily managed by the expert.

III. UMBERALLA APPROACHES FOR KAF

A. Common KADS approach

Common KADS (Knowledge Acquisition Documentation and Structuring) [3] is one of the most prominent models that support well-defined knowledge engineering. Various companies and other organizations have developed and have successfully validated this approach in the context of the European Strategic Program on Research in Information Technology IT Programme (ESPIRIT). Common KADS [9] helps in developing, distribution and applying the knowledge resources of an organization. Different methodologies and techniques provided by this approach give a detailed analysis of whole task and processes from where we can derive knowledge. An important aspect of this approach is its dependency on different models to acknowledge the complexity of a knowledge engineering [1] project.

Common KADS consists of the following model: -

1. *Organizational model*- In this, the structure of the organization is described and the detailed information of all the functions and techniques which are performed are mentioned [1]. It lays emphasis on the analysis of both current as well as future ones, as they may be implied by the proposed solutions.

2. *Task model* – It provides top-down hierarchical description of the task which is performed in an organizational unit. Here the specification of the task of each agent is also mentioned. It lists down all the inputs and outputs, preconditions and performance criteria as well as the resource needed.

3. *Agent model* - This model tells us the role of each agent involved in the functioning of the task. It describes what needs to be done and to whom the task to be assigned. The agent can be a robot, human or any software.

4. *Communication model* – Common KADS approach aims in designing of the multi-agent system. Here in this model interaction among various agents are specified.

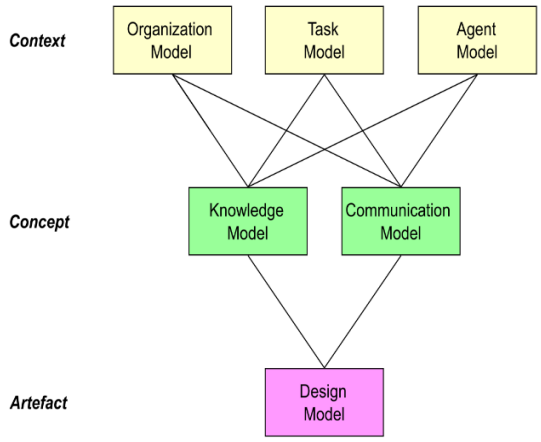

Fig.1: Common KAD model set

5. *Knowledge model* –Knowledge modelling is the specialized form of requirements specification. The first step for this model is to identify the knowledge by exploring and structuring the key information sources and studying the nature of the task in more details. While building the knowledge model, the main problem is to find the balance between learning about the domain without becoming a full. The second step is knowledge specification. The primary goal of the second step is to develop complete specifications of knowledge except for the contents of domain models [13].

6. *Design Model* -Design model process the problem-solving requirements from knowledge model, interaction requirements from communication model and non-functional requirements from other models. After processing these requirements, design model outputs specification of a software architecture and design of the application within this architecture [13].

B. MIKE approach

MIKE(Model-based and Incremental Knowledge Engineering)approach is derived from conventional software engineering [1].It is based on Boehm spiral model approach and different prototyping approaches. Its main principles are:

- The whole process of this approach is sub-divided into different phases, each phase covering a particular aspect of the development process. The main advantage of such modularity is that it gives a clear focus and reduces complexity. Different phases of the process model of MIKE include elicitation, interpretation, formalization, design, implementation and evaluation.

- Evaluation is the most important aspect of each phase [7]. This phase makes use of prototyping approaches to evaluate the operation description result obtained after every other phase.

The MIKE's process model [6] tries to integrate the merits of well-structured process models into an incremental system development and prototyping. Looking in greater detail at its process model, the big modeling gap in between informal descriptions of the expertise which are acquired from the expert via making use of knowledge-acquisition methods, and the expert system's final realization is bridged with several intermediate models. Decomposing this gap into smaller ones minimizes the complexity of the entire modeling process. Since in each step particular aspects can be considered independently of other aspects.

Firstly, task and knowledge are elaborated in natural-language documents. These documents can be obtained from observations or interviews or may pre-exist as books or manuals etc. All these documents and files can be structured and represented in a protocol model with a hypermedia representation.

Secondly, these structure model are those informal descriptions which are transformed into a representation in a semiformal way. For the making of the protocol model with

the structure model the hypermedia tool Mediating Model Construction Kit (MeMo-Kit) can be used [Neu93]. As a result, the knowledge and the task are described along the lines of a model of expertise as defined in KADS [12]. This description of knowledge is organized in numerous layers via the help of appropriate primitives, which are also related with a suitable illustrious representation. The natural language is still defined in semantics of elementary knowledge pieces. Such a mediating description possess some of the merits like the structuring process which creates the mediating representation gives early feedback for the knowledge engineer and the expert. Representation in a semi-formal manner of the expertise gives a useful basis for communicating with the expert; the contents of the model can be used for the explanation facility of the final resultant system; and the model documents models decisions and thus may be used for the maintenance of the final system.

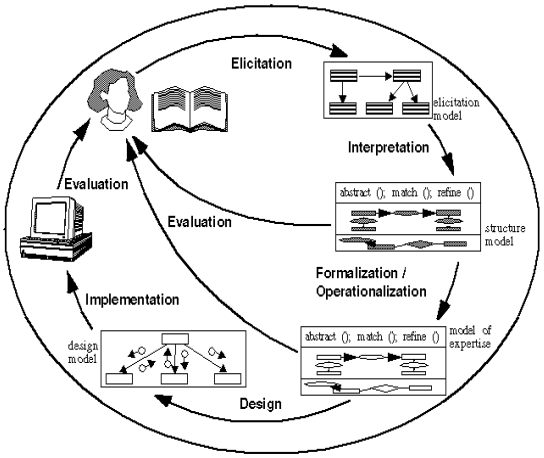

Fig.2: Visual Representation of MIKE development

C. The PROTEGE-2000 approach

Knowledge-based [2] systems are developed in an environment which is called Protege system[14], evolving from more than a decade. This system was started as a micro application for the purpose of the medical domain (protocol-based therapy planning), but it was built as a general-purpose set of tools. The main purpose of this system is to reduce the bottle-neck situation of knowledge-acquisition by minimizing the use of knowledge engineering in the creation of knowledge bases. Recently, this system has developed a community of users throughout the worldwide, who updates new features and themselves adding it with new capabilities for future evolution. In order to do this, Musen[12]stated that knowledge-acquisition functions in a well-defined layer and the knowledge which is collected in one layer can be used for generating and customizing knowledge acquisition [5] tools for another subsequent layer. Thus, the original version of the Protégé software was an application that searches for the usefulness of structured information to simplify the knowledge-acquisition process.

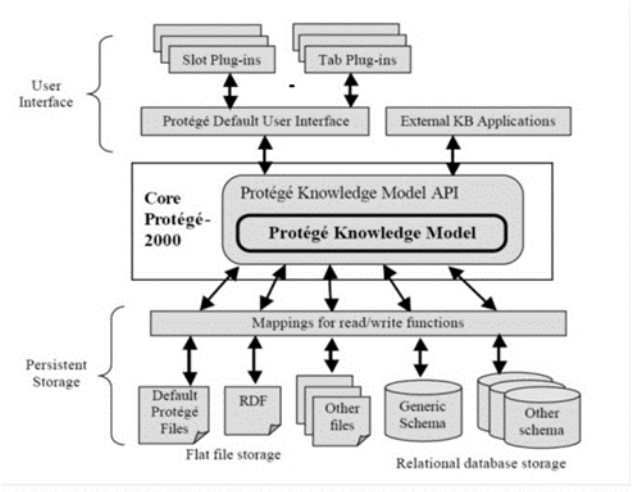

Fig.3: The protégé-2000 architecture

IV. RESULTS AND DISCUSSION

Knowledge acquisition is the most important stage of knowledge engineering. In order to get a well-defined knowledge, a proper acquisition is very much needed. This paper projects the various modeling techniques and framework for knowledge acquisition. These techniques follow different approaches in order to acquire the knowledge. It forms a basis for building the various application of knowledge engineering. Various companies and organization develop these techniques and frameworks in order to derive knowledge from the raw data by proper structuring and organizing the data which is meaningful for the organization. Knowledge which have been build using KAF can be standardized and shared among knowledge workers of global. These umbrella approaches helps in text generation, discovery of linguistic knowledge, re-use based approach, expert system prototyping and bioinformatics for further research.

REFERENCES

[1] Rudi Studer, V. Richard Benjamins, and Dieter Fensel, "*Knowledge Engineering: Principles and Methods*" Data & Knowledge Engineering 25 (1998) 161-197, Elsevier.

[2] John H. Gennari, Mark A. Musen, Ray W. Fergerson, William E. Grosso, Monica Crubézy, Henrik Eriksson, Natalya F. Noy and Samson W. Tu, "*The Evolution of Protégé: An Environment for Knowledge-Based Systems Development*", International Journal of Human-Computer Studies archive,Volume 58 Issue 1, January 2003 Pages 89 – 123.

[3] A. Abecker, S. Decker, K. Hinkelmann, and U. Reimer, Proc. "*Workshop Knowledge-based Systems for Knowledge Management in Enterprises*", 21st Annual German Conference on AI (KI'97), Freiburg, 1997.

[4] Frank van Harmelen, "*Formal Methods in Knowledge Engineering*", PhD Thesis, University of Amsterdam, 1995. The Knowledge Engineering Review, Vol. 10, No. 4, pp. 345-360, 1995

[5] J. Angele, S. Decker, R. Perkuhn, and R. Studer, "*Modeling Problem-Solving Methods in NewKARL*", Proceedings of the 10th Knowledge

Acquisition for Knowledge - Based Systems Workshop (KAW'96), Banff, Canada, November. Karlsruhe: 1996. 18 pp.

[6] J. Angele, D. Fensel, D. Landes and R. Studer, Developing "*Knowledge-Based Systems with MIKE*", Journal of Automated Software Engineering, October 1998, Volume 5, Issue 4, pp 389–418

[7] Juergen Angele, Dieter Fensel, Rudi Studer, "*Domain and Task Modeling in MIKE"*, IS&O 1996: Domain Knowledge for Interactive System Design pp 149-163, Springer

[8] J. McDermott, "*Preliminary steps toward a taxonomy of problem-solving methods*", pp.225-256 In Automating Knowledge Acquisition for Expert Systems. S. Marcus, (Ed.), Kluwer, Academic Publishers, Boston,1998.

[9] Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Robert De Hoog, Nigel R. Shadbolt, Walter Van de Velde and B J. Wielinga, "*Knowledge Engineering and Management: The CommonKADS Methodology*", MIT Press, 1999

[10] Edward H. Shortliffe, A. Carlisl e Scott, Miriam B. Bischoff, A. Bruce Campbell, Willia m Van Melle, Charlott e D. Jacobs, "*ONCOCIN: An expert system for oncology protocol management*". International Joint Conference on Artificial Intelligence (IJCAI '81), Vancouver, CA, 876-881.

[11] B.J.Wielinga, A. Th. Schreiber, J. A. Breuker, "*KADS: A modeling approach to knowledge engineering*", Knowledge Acquisition, Volume 4, Issue 1, March 1992, Pages 5-53

[12] Noy, N. F., Fergerson, R. W., and Musen, M. A., "*The knowledge model of Protégé-2000: Combining interoperability and flexibility*", Second International Conference on Knowledge Engineer- ing and Knowledge Management (EKAW'2000), Juan-les-Pins, France, (2000).

[13] Rose Dieng, Olivier Corby, Sofiane Labidi, "*Agent-based knowledge acquisition",* International Conference on Knowledge Engineering and Knowledge Management EKAW 1994: A Future for Knowledge Acquisition pp 63-82

[14] R. Dieng, O. Corby, and S. Labidi., "*Expertise conflicts in knowledge acquisition*", In Proc. of the 8th KAW, vol. 2, pages 23.1–23.19, Banff, Canada, 1994.

[15] Gamble, P.R., Blackwell, J., "*Knowledge Management: A State of the Art Guide*", Kogan Page, London, 2001

Authors Profile

Mr. Thirunavukkarasu K. is Assistant Professor at Galgotias University, Greater Noida, Delhi-NCR. He is pursuing PhD in CSE in the research area of Spatial Database. He has been a student of Madras University, Bharathiar University, and Anna University, Chennai, India. He has more than 17 years of experience in Teaching and 3 years in software industry. He has taught for APIIT(affiliated to Staffordshire University, UK) at Panipat, various colleges affiliated to Bangalore University, Bangalore and worked as Software Engineer for I2 Technology, UK at Bangalore. He has involved in various academic activities like members of BoE, NAAC, NBA, OBE and Assistant Custodian for PG-Unit, Bangalore University, Bangalore. He has wide research interests that include Knowledge Engineering, Data Mining, Cloud computing and Big Data Analytics. He is a Member of IEEE, LCSI and LSTE. He has 13 SME certificates from IBM and trained students on various domains as SME. He was the organizing Secretary for an International Conference ICACCT^{TM} 2010. He has conducted various conferences, workshops, short term and summer courses. He has published 18 papers in international and 4 papers in national level.

Abhilash Ashu is pursuing Bachelor of Technology in Computer Science from Galgotias University. His areas of interest are Cryptography Algorithms, Network Security, Cloud Security and Privacy, Big Data Analytics, Data Mining and IoT.

Digvijay Singh is pursuing Bachelor of Technology in Computer Science from Galgotias University. His areas of interest are Data Analytics, Network Security, Cloud Computing and Load Balancing, Big Data Analytics, Data Mining, IoT and Wireless communications.

Dr. Shahnawaz Khan received his Ph.D in Computer Science from Indian Institute of Technology (Banaras Hindu University), Varanasi, India. His areas of specialization are machine learning, natural language processing and data science. He was a key player during the accreditation process of NCAAA and NACC at various places. He has always emphasized on bridging the gap between academics and industry keeping in mind the growing IT industry in terms of futuristic technologies. He has been Head/Member of many committees including R&D, faculty and staff development committee, Academic Council, Advisory Committee, Monitoring and Planning Board, Research Advisory Committee, Accreditation committee etc. in Saudi Electronic University and Galgotias University.

