Malaysia Indigenous Herbs Knowledge Representation

Zulazeze Sahri[1], Sharifalillah Nordin[2] and Haryani Harun[3]

[1]*Universiti Teknologi MARA (UiTM), Malaysia, azeze@pahang.uitm.edu.my*
[2]*Universiti Teknologi MARA (UiTM), Malaysia, sharifa@tmsk.uitm.edu.my*
[3]*Universiti Teknologi MARA (UiTM), Malaysia, haryani@tmsk.uitm.edu.my*

ABSTRACT

The values of Malaysia indigenous herbs are not limited to its medicinal benefits but also in local and traditional knowledge of each herb. However, this knowledge is scattered and unorganized in internet and merely own by local people that led to difficulty in designing knowledge based system. This paper proposed a Malaysia indigenous herbs knowledge representation and its implementation in Web Ontology Language (OWL) using Protégé 4.2. The herbs knowledge mainly in preparation methods and plants part uses is collected from literatures and experts in herbal field. Additional interview session is conducted with expert to validate the information gathered. In order to ensure the knowledge model and ontology view is well defined, a prototype of knowledge based system is developed based on the ontology classes or concepts and relationships defined which requires conversion the OWL ontology into the relational database system at first hand.

Keywords: Knowledge Management, Herbs Traditional Knowledge, Herbs Knowledge Based System, Ontology.

I INTRODUCTION

Herbal medicine is a well recognized system of medicine throughout the world. Herbal medicine is used as remedies for illness such as fever, diarrhea, sore throats, sinus problems, respiratory problems and skins condition (Sujaidi, 2009). Almost all civilizations have some way or another, a connection with the herbal world. Herbs also have been recognized as an alternative medicine and economical resources. Nearly 80% of the world populations rely on the use of traditional medicines to meet their health care needs (Sandhya et al., 2006) and up to 90% of developing world relies on the use of medicinal plants (WHO, 2002). In Malaysia, there are more than 2000 plant species that have healing qualities and highly potential to be commercialized (MARDI, 2010). Malaysian Agricultural Research and Development Institute (MARDI) recorded gross profits of more than RM 5.4 billion a year from herbal related products. Therefore, it is important to preserve the value of the medicinal plants as it brings economical importance and vast valuable indigenous knowledge such as preparation methods, plant part uses and other traditional knowledge.

Indigenous Knowledge (IK) is the unique, traditional, local knowledge that exist within and developed around the specific conditions of women and men indigenous to a particular geographical area (Grenier, 1998). It is the basis for local level decision making in agriculture, healthcare, food preparation, education, natural resource management and a host of other activities in rural communities (Warren, 1991). Herbal medicine is a good example of Indigenous Knowledge which has affected the lives of people around the globe (World Bank, 1998). However, as Malaysia moves towards the stream of global modernization, the art of traditional healing using medicinal plants suffers from the global emphasis on modern biomedical healthcare facilities and the disinterest of younger generations in traditional knowledge and practices (Kulip, 2003; Lin, 2005). Thus, the great store of indigenous knowledge held by traditional healers, local people and other practitioners are important to be captured in order to preserve the knowledge. There are many research done in ethnobotany field on the potential medicinal plants in the traditional territories of indigenous people and to preserve the indigenous knowledge of medicinal plants (Ahmad (2008), Ahmad & Ismail (2003), Ong (1999), Kulip (2003), Lin (2005) and Samuel et al. (2010)). However, the knowledge found from research are stored in various forms and scattered in websites, portals, journals and books. It also described in natural language with many local terms with different meaning. Currently, there is no centralized knowledge based system to store and share the indigenous herbs knowledge in Malaysia.

This research aims to capture and represent Malaysian indigenous herbs knowledge focusing on the preparation methods. The knowledge representation scheme can be used to encode and store knowledge in a database or "Knowledge Base" (Rosenberg, 1986). In other words, this research's objective is to structure and formulate the information and natural language into formal language. Twenty-eight (28) herbs are studied from five ethnobotanical studies namely; Ahmad & Ismail (2003), Ong (1999), Kulip (2003), Lin

(2005) and Samuel et al. (2010). In addition, knowledge from expert in herbal field has been collected by interview session mainly to validate and to gain additional indigenous knowledge.

Based on unstructured and descriptive herbs information, the research continues with identifying the suitable knowledge representation techniques that can be applied in representing this data. Further literature review can be summarized in the next section of this paper.

II LITERATURE REVIEW

A. Knowledge Representation Techniques

Knowledge Representation (KR) is the area of Artificial Intelligent (AI) which covers the method of how knowledge can be best represented symbolically and manipulated in an automated way by reasoning program (Brachman & Levesque, 2003). From the—knowledge management perspective, knowledge representation is required to convert tacit knowledge to explicit knowledge, and to represent the explicit knowledge in suitable form to be modeled and applied in knowledge sharing system such as database, repository or library (Obamsawim, (2002); Grenier (1998); Mondo et al. (2007). The most popular and commonly used knowledge representations are *Logic* which related to the truth of statements about the world (Chakraborty, 2010). It is a formal system in which the formulas or sentences have TRUE or FALSE values only. Secondly is *Production Rules* knowledge representation technique that relies on IF-THEN rules which be able to provide the flexibility of combining declarative and procedural representation for using them in a unified form. It is popular as a knowledge representation mechanism used in design of many "Rule-Based System" or "Production Systems" (Chakraborty, 2010). *Semantic Nets* represented knowledge as concept nodes related by directional relationship links, representing the world as a directed graph (Quillian, 1968). Semantic network commonly used to represent the inheritable knowledge by organized into classes and classes must be arranged in a generalization hierarchy (Poonam et al. 2010). *Frame* is collection of attributes or slots and their associated values which describe the real world entity. The frame used to represent class (set) and instances (class' elements). It components consists of frame name, attributes (slots) and values. Finally, Yuanyuan et. al (2010) suggest that the *Ontology* can be used to develop the knowledge based by constructing the ontology model through the relevant concepts and their relationship. Therefore, the formalized ontology knowledge representation makes knowledge sharing and reuse possible through the knowledge based system.

B. Ontology as Knowledge Representation

Based on the characteristic from each of the knowledge representation technique discussed, the research use ontology as the knowledge representation technique because of its ability to identify class or concept, property and relationship within a domain of discourse. In addition, ontology is an important tool for the organization and contextualization of knowledge, particularly in well-bounded context such as in research findings data (Thunkijjanujkij, 2009). Therefore, it is appropriate for this research which focuses on the conversion of unstructured herbs data from research findings and other sources.

Studer et al. (1998) defined ontology as a formal, explicit specification of a shared conceptualization. The *"formal"* can be described as to make the ontology model machine-readable, *"explicit specification"* describes the explicitly defined the concepts, properties, relations, functions, constraints and axioms. *"Shared"* means consensual knowledge whereby the need to describe the common understanding of the structure of information among people in the domain and finally *"conceptualize"* is the abstract model of some phenomenon in the world.

Meanwhile, Noy & McGuinness (2001) defined an ontology as a formal explicit description of concept in a domain of discourse (class or concept), properties that describe the characteristic of the concept (slots or roles), and restriction on slots (facet or role restriction). A complete Ontology with a set of individual instances of classes is able to form a knowledge base. Thunkijjanujkij (2009) claimed that Ontology can provide an organization framework about a concept which organized in a system of hierarchy and associative relations that allows reasoning about the knowledge. Generally, ontology as a graph/network structure consisting of;

- A set of concepts (vertices edges in a graph)
- A set of relationship connecting concepts (directed edges in a graph)
- One set of instances assigned to a particular concepts (data records assigned to concepts or relation)

C. Ontology Development

The process of ontology development involves four main steps as suggested by Noy & McGuinnes (2001), which are enumerating all important terms in the domain, identifying terms' definition and related meaning, defining classes and class hierarchy and finally identifying relationship between classes. Uschold & Gruninger (1996) suggest three processes in identifying classes and class hierarchy such as the top-down, bottom-up and combination development process.

III METHODOLOGY

This section describes the methodology in designing the Malaysia indigenous herbs knowledge representation using ontology approach. In this case, the research needs to identify the concepts available in the domain of herbs, its relationship between concepts and instances. The scope of this research is the indigenous preparation methods and plant part uses. The methodology adopted four main steps in ontology development steps as suggested by Noy & McGuinnes (2001) as discussed in section 3. The steps are;

- Enumerate all important terms in the domain
- Identify terms' definition and related meaning
- Define classes and class hierarchy
- Identify relationship between classes

Term enumeration process is performed to understand what is the commonly shared knowledge used in the herbs preparation methods and plant part uses from all data sources. The class is identified by referring to the meaning of all terms enumerated. The most generalize term that has meaning and purpose which can represent the preparation methods category is chosen as the class name. Then, the research classifies and verify the herbs preparation methods and plant part uses into it classes or concepts identified previously by consulting with the expert. Thesaurus and dictionary also was used to understand the meaning of each term and its concepts. This research applied the top-down process where class identification starts from general to specialize class. The relationship between classes was adapted from existing relationship developed by the reliable authoritative from this field mainly in agriculture and biotechnology such as Soergel et al. (2004), Schmitz-Esser (1999) and FAO (2011). Based on the identified concepts and relationship, the overall concepts connection with its relationship will convert to a knowledge model that represents the Malaysia Indigenous Herbs Knowledge. The next step is the implementation stage which is the process to transform the knowledge model into preferred knowledge representation language, Ontology Web Language (OWL). OWL can be best developed by using the ontology development editor called Protégé version 4.2. All classes, subclasses, relationship and instances which have been identified in previous section have been transformed as ontology visualization by using the Protégé ontology development editor. The step by step practical guideline in developing ontology by using Protégé 4.2 was referred to Horridge et. al (2007).

IV RESULT

This section describes the results of this research. The preparation methods of herbs can be represented by six classes: *Decoction, Drying, Extraction, Direct Usage, Peel and Incised*. Each class is comprised of instances that have similar preparation purpose. The identified classes and instances are validated by the expert for accuracy. From the findings, the unorganized information on Malaysia indigenous preparation methods of herbs knowledge is organized in structured form shown in table 2.

Table 2. Sample of Structured Information for Preparation Methods

Bil	Local Name	Medicinal Uses	Preparation Method	Preparation Category
TREE				
1	Limau Purut	Dandruff	Squeeze	Extraction
2	Mengkudu	High Blood Pressure	Eaten Fresh	Direct Usage
		Diarrhoea	Smearing, Seared	Direct Usage, Dried
		Increase White Blood Cells	Boiled	Decoction
		Cancer	Eaten Fresh	Direct Usage
		Diabetes	Eaten Fresh	Direct Usage
		Skin Defect	Smearing	Direct Usage
		Cough	Blended	Extraction
		Hypertension	Eaten Fresh	Direct Usage
3	Tongkat Ali	Reproductive	Boiled	Decoction
		Tonic	Boiled	Decoction
		Antiviral	Boiled	Decoction
4	Tea Tree	Wound	Crushed	Extraction

The classes of plant part uses in preparing the medicinal herbs were then structured. It was found that each herb has its own capability to cure ailments by using some of the herbs' part or the whole part, which being categorized in six categories namely, *Leaf, Root, Stem, Bark, Fruit and Whole Plant*. The research also found that each plant part may require specific and special preparation method to make medicinal herbs. The overall representation of Malaysian indigenous herbs knowledge is shown in table 3. To conceptualize the relationship between the herbs part uses and its preparation methods, a knowledge

model is developed. The conceptualize model in the form of knowledge model is created by applying meaningful relationship between the classes or concepts. The relationship is based on existing relationship defined by Soergel et al. (2004), Schmitz-Esser (1999) and Food and Agriculture Organization (FAO) (2011). The selected relationship is shown in the table 4.

Table 3. Sample of Structured Information on Herbs Part Uses and Its Preparation Method

Bil	Local Name	Medicinal Uses	Plant Part	Preparation Method	Preparation Category
Tree					
1	Limau Purut	Dandruff	Fruit	Squeeze	Extraction
2	Mengkudu	High Blood Pressure	Leaf	Eaten Fresh	Direct Usage
		Diarrhoea	Leaf	Smearing, Seared	Direct Usage, Dried
		Increase White Blood Cells	Root	Boiled	Decoction
		Cancer	Fruit	Eaten Fresh	Direct Usage
		Diabetes	Fruit	Eaten Fresh	Direct Usage
		Skin Disease	Fruit	Smearing	Direct Usage
		Cough	Fruit	Blended	Extraction
		Hypertension	Fruit	Eaten Fresh	Direct Usage
3	Tongkat Ali	Reproductive	Root	Boiled	Decoction
		Tonic	Root	Boiled	Decoction
		Antiviral	Root	Boiled	Decoction
4	Tea Tree	Wound	Leaf	Crushed	Extraction

Table 4. Concepts Relationship in Malaysia Indigenous Herbs Knowledge Model

Relationship	Inverse Relationship	Source
hasSubclass	isSubclassOf	Protégé 4.2
part of	hasPart	FAO (2011)
instrumentFor	performedByInstrument	Schmitz-Esser (1999)
processFor	UsesProcess	FAO (2011)
subprocessOf	includeSubprocess	Soergel et al. (2004)
treatedWith	treatmentFor	Schmitz-Esser (1999)

The knowledge model of Malaysian indigenous herbs is shown in figure 2. It shows the relationship between the herbs part uses to make herbal medicine along with the preparation method. It also adopted the herbs classification by Mamat et al. (2008) and Medicinal values categories by Nor Nadiah (2010). Figure 3 shows the class hierarchy. Figure 4 shows the Protégé relationship view identified which connect the classes and instances based on knowledge model developed above. Figure 5 shows the overall view of herbs ontology based on identified classes and instances. The ontology shows all herbs domain classes or concepts found in this research along with the subclasses and instances. Furthermore, figure 6 shows an example of relationship between *'Kunyit'* and *'Cekur'* that share the same preparation method of *'Boiled'* by using '*Whole'* plant parts to cure diseases (eye irritation and women after give birth). This relationship view supported the Knowledge Model defined in figure 2, whereby there are relationship between classes and instances in the knowledge model which looks like a rectangle shape. Since the ontology has concepts and relationship, this research is able to design a prototype system that provide search box that allows user to find information by keywords-search. The inference engine infers the input based on the if-else rules to derive desired result from the database. From the query result, the system will suggest the related information based on the same concept of the existing result search selected. For example, when user search for the keyword "Mengkudu", the result will shows that Mengkudu falls under herbs category "Tree" and is able to cure "High Blood Pressure". Therefore, the system will suggest all herbs that fall under category Tree and all herbs that can cure High Blood Pressure.

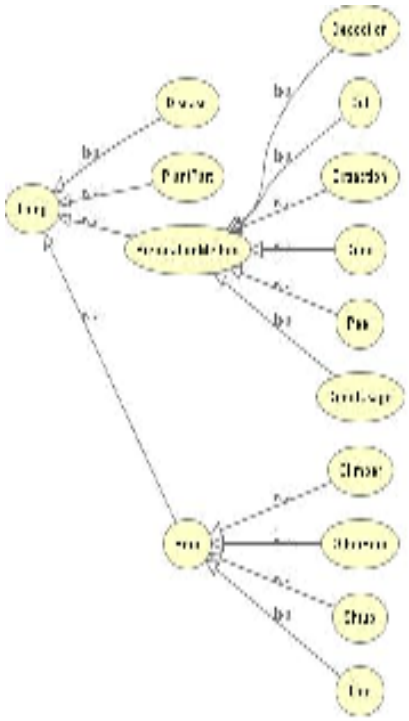

Figure 3. Class Hierarchy

The prototype of Malaysia Indigenous Herbs System (MIHES) is developed as the evaluation process of knowledge model and ontology design. The system consists of three main components which are user interface for user interaction, inference engine which will derive the conclusion from the relational databases (encoded from OWL Ontology) based on the matches index that occurs

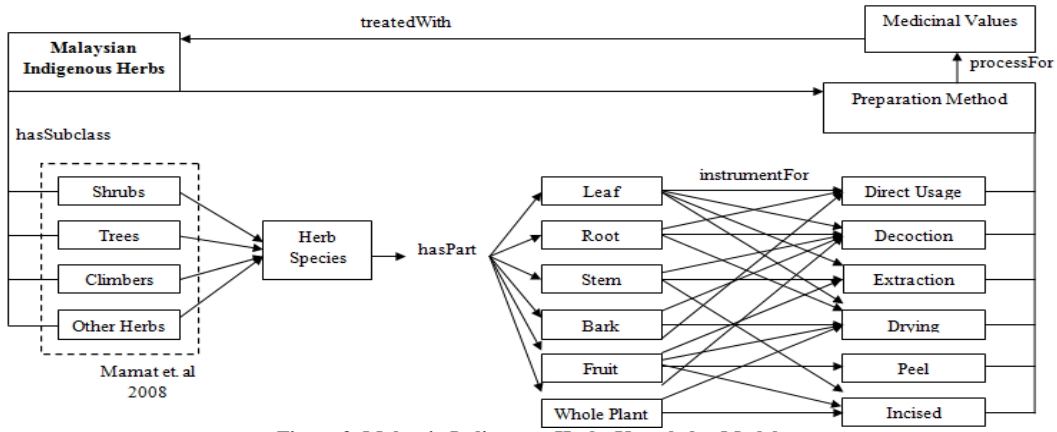

Figure 2. Malaysia Indigenous Herbs Knowledge Model

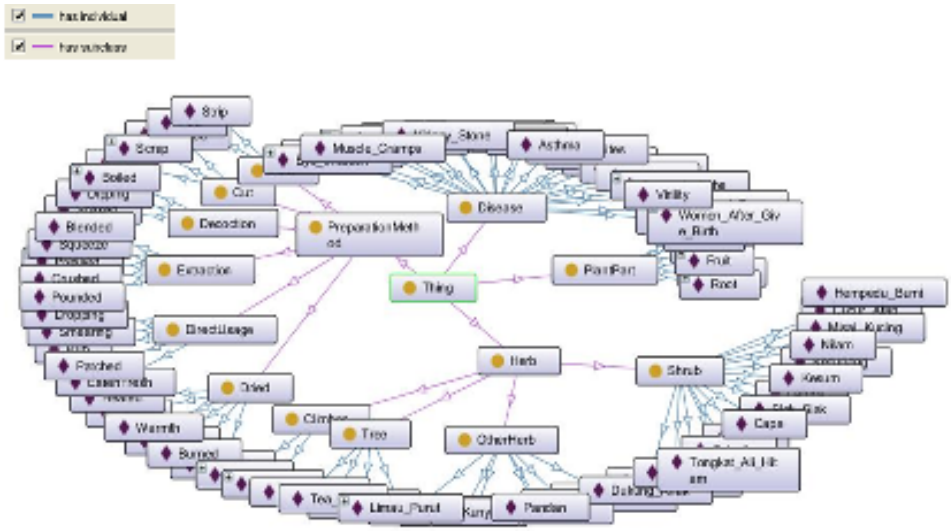

Figure 5. Implementation of Knowledge Model as Ontology

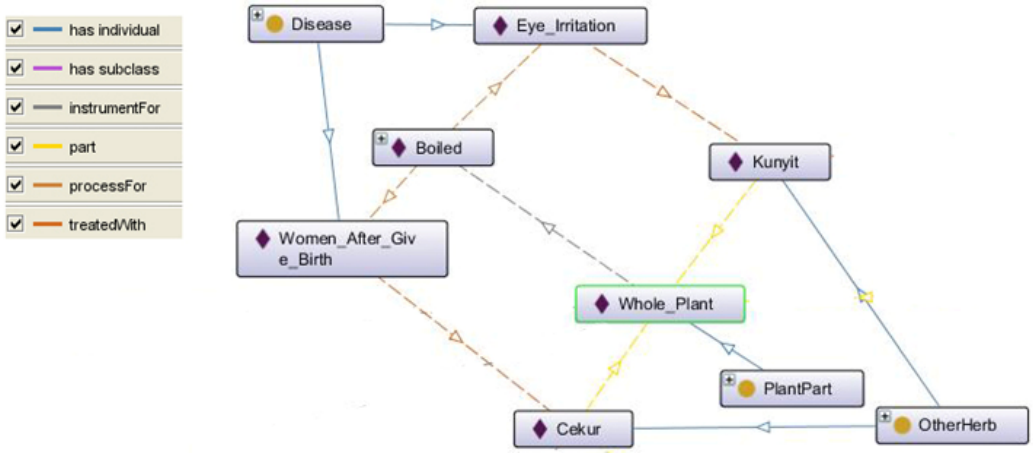

Figure 6. Sample of Herbs Relationship (To Support Figure 2)

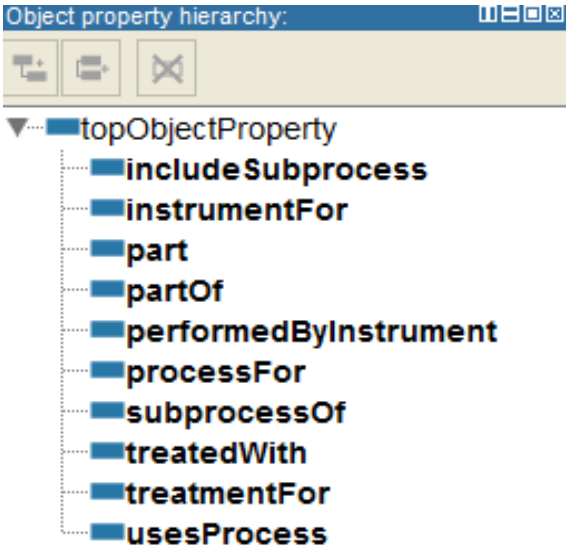

Figure 4: Relationship views

in an AJAX file and relational database. The ontology knowledge representation developed is encoded in knowledge based system which allows the inference engine to examine the rules.

Figure 7. Relational Database Schema (ERD) from OWL language

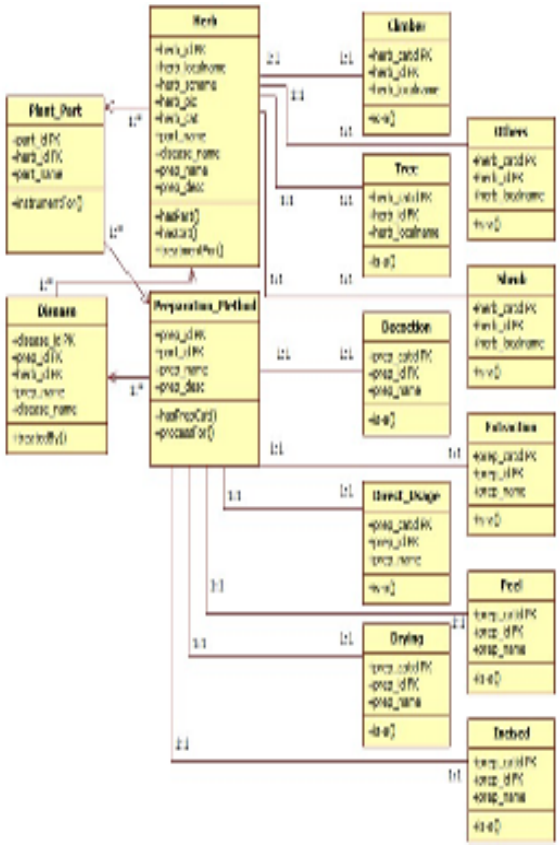

Figure 7 shows the relational database schema from the OWL ontology developed in this research.

The conversion from ontology KR to relational databases is based on OWL2DB algorithm by Vysniauskas & Nemuraite (2006). OWL2DB algorithm allows us to convert class and subclass structure into relational tables in the form of parent and child table by using Primary Key (PK) and Foreign Key (FK). The relationship is converted from OWL object property or relationship to table relationship in the database by using the PK and FK.

V CONCLUSION

This research discusses on how the available indigenous herbs knowledge data found from ethnobotany research and expert interview can be represented by using ontology technique. The ontology is able to identify and represent the commonly terms uses in herbs domain as classes or concepts and the relationship between all concepts. The classes were arranged in hierarchical design and instance of each class is defined accordingly. A structured form of data can be developed based on this classes and instances. The research continues to design the knowledge model by applying several meaningful relationships. The knowledge model which shows the relationship between all classes and its relationship is implemented in Web Ontology language (OWL). The implementation process shows the overall view of indigenous herbs knowledge representation as ontology. Finally, the knowledge model and ontology design is evaluated by developing a knowledge-based system. The system requires conversion of ontology design to relational database using OWL2DB algorithm. The findings allow knowledge engineer, developer and user to gain better understanding on the concept and classes contain in the domain by looking at the ontology design. This research could be extended by focusing into other herbs knowledge classes using the same ontology developed from this research. This enables the reuse of domain knowledge in the form of ontology.

REFERENCES

Ahmad, F. (2008). Robotic Technology Malaysia. Retrieved November 2011,fromssc.undp.org/uploads/media/Robotic_Technology_Malaysia.pdf

Ahmad, F., & Ismail, G. (2003). Medicinal Plants Used by Kadazandusun Communities Around Crocker Range. ASEAN Review of Biodiversity and Environmental Conservation (ARBEC).

Brachman, R. J., & Levesque, H. J. (2003). Knowledge Representation and Reasoning.

Food and Agriculture Organization (FAO) (2011). AGROVOC Thesaurus. Retrieved November 2011, from http://www.fao.org/agrovoc

Grenier, L. (1998). Working With Indigenous Knowledge; A guide For Researcher. Ottawa, Canada: IDRC.

Gruninger, M. and M.S. Fox. 1995. Methodology for the design and evaluation of ontologies, pp 51-60. *In* Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal.

Horridge, M., Jupp, S., Moulton, G., Rector, A., Stevens, R., Wroe, C., (2007). *A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools (1st Ed.)* The University of Manchester.

Kulip, J. *2003. An ethnobotanical survey of medicinal and other useful plants of Muruts in Sabah, Malaysia*. Telopea *10(1): 81–98*.

Lin, K. W. (2005). Ethonobotanical Study of Medicinal Plants Used by the Jah Hut Peoples in Malaysia *Indian Journal of Medical Sciences*, Vol 59. 156-161.

MARDI (2010). Retrieved Februari, 2010 from http://www.mardi.gov.my/web/guest/skop-penyelidikan15

Mamat, A., Erlin, & Rahman, A. A. (2008). Organising Herbs Knowledge: Is an Ontology or Taxonomy the answer? *Proceedings of the International Symposium on Information Technology (ITSIM'08).*

M.R. Quillian. Semantic memory. In M. Minsky, editor, Semantic Information Processing, pages 227–270. MIT Press, Cambridge, MA, 1968.

Nor Nadiah, Y. (2010). Knowledge Modelling for Malaysian Medicinal Herbs (Master Dissertation of Universiti Teknologi MARA, 2010).

Noy, N. F. & McGuinness, D. L.(2000). Ontology Development 101: A Guide to Creating Your First Ontology, Stanford University.

Obomsawin, R (2002). Indigenous knowledge and sustainable development. International *Journal of Information Management,* Volume 26, Issue 3, June 2006, Pages 224-233

Ong, H. C., & Nordiana, M. (1999). Malay Ethno-Medico Botany in Machang, Kelantan, Malaysia. Fitoterapia vol 70, 502-513.

Pinto, H.S. and J.P. Martins. 2004. Ontologies: How can they be built. Knowledge and Information System 6: 441-464.

Poonam Tanwar et. al. / (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 07, 2010, 2274-2281.

RC Chakraborty, 2010. Knowledge Representation: AI Course Lecture 15-22. Retrieved December, 2012, from www.myreaders.info/html/aritificial_intelligence.html

Rosenberg, J. M. (1986). Dictionary of Artificial Intelligence and Robotic. New York:Willey.

Samuel, A. J. S. J., Kalusalingam, A., Chellapan, D. K., Gopinath, R., Radhamani, S., Husain, A., Muruganandham, V., & Promwichit, P. (2010). Ethnomedical Survey of Plants Used by the Orang Asli in Kampung Bawong, Perak, West Malaysia. *Journal of Ethnobiology and Ethnomedicine.* Retrieved November 2011, from http://www.ethnobiomed.com/content/6/1/5

Sujaidi, D. (2009). Mengenal 200 Jenis Herba, Kuala Lumpur, Malaysia.

Sandhya B, Thomas S, Isabel W and Shenbagarathai R (2006). Ethnomedicinal Plants used by the Valaiyan community of Pairanmalai Hills (Reserved forest), Tamilnadu, India-A Pilot study. Afr. J. Trad. Complement. Altern. Med., 3(1): 101-114.

Soergel, D., B. Lauser, A. Liang, F. Fisseha, J. Keizer and S. Katz. 2004.Reengineering thesauri for new applications: the AGROVOC example. Journal of Digital Information 4(4). Article No. 257, 2004-03-17. Available Sources : http://jodi.tamu.edu/Articles/v04/i04/Soergel/ November 15, 2007.

Studer, Benjamins, Fensel. Knowledge Engineering: Principles and Methods. *Data and Knowledge Engineering*. 25 (1998) 161-197 Information Week (2007), Defining Text Analystics. http://www.informationweek.com/software/blog/archives/2007/02/defining_text_a.html

Thunkijjanukij, A., (2009). Ontology Development for Agriculture Research Knowledge Management: A Case Study for Thai Rice (Doctoral Dissertation, Kasetsart University, 2009).

Uschold, M. and M. Gruninger. 1996. Ontologies: Principles, Methods and Applications. AIAI-TR-191. The University of Edinburgh.

Vysniauskas, E. & Nemuraite, L. (2006). Transforming Ontology Representation from OWL to Relational Database. *124th* Information Technology and Control, Vol 35, 333-34.

Warren, D. M. 1991 Using Indigenous Knowledge in Agricultural Development; World Bank Discussion Paper No.127. Washington, D.C.: The World Bank.

World Bank (1998). Indigenous Knowledge For Development: A Framework For Action. African Region.

WHO (2002). Traditional medicine- growing needs and potential. WHO Policy perspectives Med., 2: 1-6.

Schmitz-Esser, W. (1999) "Thesaurus and Beyond: An Advanced Formula for Linguistic Engineering and Information Retrieval". *Knowledge Organization*, Vol. 26, No. 1, 10-22

Schmitz-Esser, W. (1999) "Gedankenraumreisen - neue Thesaurusstrukturen, multimedial präsentiert, machennregung, Spielen, Lernen, Finden möglich für jedermann". In *Proc. DGI-Jahrestagung*, Hamburg (Frankfurt/M.: DGI), pp. 347-353.

Yunayuan, W., Rujing, W., Xue, W. & Yimin, H. (2010). Ontology-Based Knowledge Representation for Agriculture Intelligent Information System.Management and Service Science (MASS) International Conference. P. 1-4.