The 10th International Conference on e-Business (iNCEB2015)
November 23rd - 24th 2015

A framework for Taxonomic Ontology of Prehistoric Pottery in Thailand

Thantham Yeesarn[1], Bunthit Watanapa[2], Kriengkrai porkeaw[3]
School of Information Technology, King Mongkut's University of technology Thonburi
126 Pracha Ulthit Rd, Bang Mod, Thung Khru, Bangkok, 10140.
54442029@st.sit.kmutt.ac.th[1],bunthit@sit.kmutt.ac.th[2], porkeaw@sit.kmutt.ac.th[3]

***Abstract*- The aim of this paper is to propose a conceptual framework for developing an archeological knowledge system of prehistoric potteries found in 35 important sites in Thailand. The structural design for the knowledge system will be presented by a taxonomic ontology of 23 artistic and cultural classes using OWL, XML, and web technology. The system, once complete, shall enable effective archeological knowledge management with web-based semantic searching ability. Performance evaluation schemes are also proposed.**

Keyword: Ontology, Prehistoric potteries taxonomy, querying system.

I. PREFACE

The study of prehistoric potteries is treated as a very important subject in prehistoric archeology [1], especially in Thailand. Information encoded in the pottery bodies and the contextual surroundings of the sites can help to implicitly indicate lifestyle, prosperity and the settlement of people in the prehistoric age. Other area of study include social formation such as beliefs and rituals, technology development, and art evolution of concerned ethnics can be also explored.

Currently, prehistoric archeology is no longer anomaly and on going studies keep generate voluminous and comprehensive information. Fortunately, the complexity of this knowledge management can be supported by information technology. Ontology is a method invented for such a purpose of representing a network of concepts of a certain domain, organizing entities and establishing their relationship for querying or logic operation.

The aim of this study is to propose a framework for developing a taxonomy of prehistoric pottery found in Thailand using static ontology [2] this will defining the existing prehistoric pottery excavated from the selected 35 sites throughout Thailand. Accessing the knowledge entities is based on certain semantic relations [3]. A well-developed ontology representing taxonomy of prehistoric potteries shall establish a basis for supporting the study of archeology in Thailand. It will be constantly updated to reflect new discoveries evidences, for the ontology is extendable and reusable by nature.

The paper is presented as follows. The section II presents a review of ontology theory. Section III. illustrates the framework for developing taxonomic ontology. Then the evaluation scheme for the designed ontology is proposed in Section IV. Finally, the conclusion and work direction are provided in Section V and VI.

II. ONTOLOGY THEORY

Ontology is a scientific method designed for managing knowledge. Gruber [4] defines Ontology as structural conceptual knowledge which explains knowledge by class, relation, axiom and instant.

Generally, there are 3 types of ontology according to formal classification based on tools and languages supporting ontology development [5]

A. Vocabulary–taxonomy ontology. Vocabulary-taxonomy is tree structure as shown in Fig.1. This ontology type could be beneficial for explaining static knowledge which is not over complicated. The vocabulary–taxonomy ontology represents knowledge in a top-down fashion where the domain or top-level class, e.g. domain of animals, is at the top (root node) and details of different levels can be assigned to the lower levels in the hierarchical order until reaching the lowest level of classes (leaf nodes), e.g. Fish as a vertebrate animal.

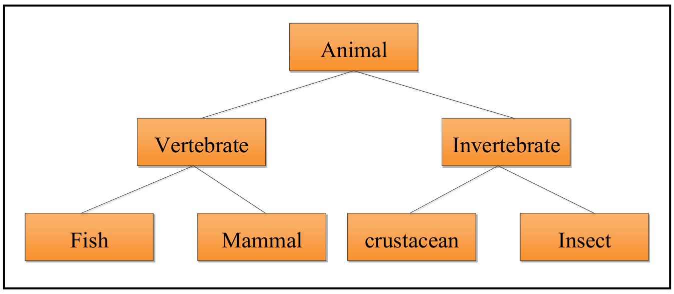

Fig. **1**. Vocabulary–taxonomy Animal ontology.

B. A Frame-Base ontology system is similar to the class diagram of the object-oriented concept. It is more complicated than Vocabulary–taxonomy ontology. This kind of ontology has attributes to define or describe property of the class and facets to define the condition of the attribute. Frame-based ontology has 2 types of relations between each class: (i) Taxonomy relation, such as Part of, and (ii) Non-Taxonomy relation, such as Is-a, Has-a etc. The first relation is useful for emphasizing or giving details of a class, e.g. earthenware is part of pottery as shown in Fig.2.

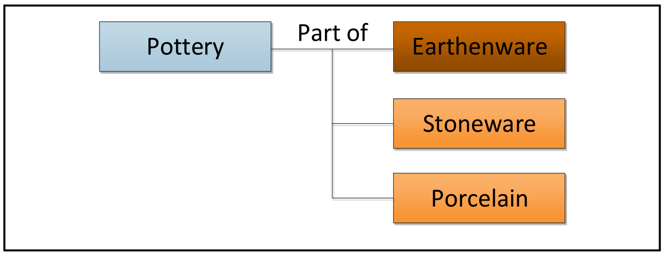

Fig. 2. Example of taxonomy relation.

Non-taxonomy relation allows linking between different taxonomies. It can help to connect or link two or more knowledge subsystems. For example, "This earthenware was used for contain food and liquid" can be represented by a link of two different classes and lead to new more meaningful knowledge as shown in Fig.3.

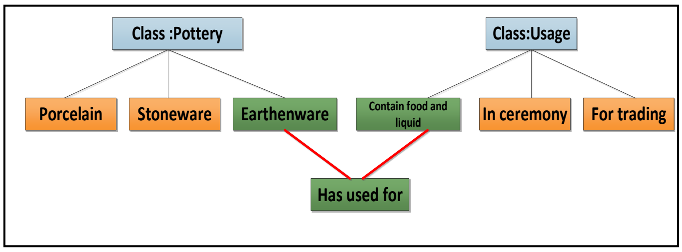

Fig. 3.An example of a non-taxonomy relation that links 2 classes.

C. Descriptive logic ontology is the ontology of compiling targeted questions and answers to build a system of concepts and roles with logical relations presented by symbolic expressions. For example, the concept of the sentence "**the pottery is the earthenware type**" will be represented by the expression of

[Pottery] → (is) → [Earthenware] which is possibly transformed into interchange format (CGIF) of **(exists((?x pottery)(?y earthenware)(is ?x ?y)).** Such a form of CGIF is manageable by a tool such as the KIF (Knowledge information format) that can operate predicate logic or calculus [6].

Ontology Benefit

When being compared with a data base, there are several benefits when applying Ontology to develop a knowledge system. Two outstanding benefits [7] are discussed below.

A) Portability. A complete and reliable ontology developed on web service technology can be operated or shared across platforms.

B) More insightful. The ontology can explicitly encode roles and relation whereas the database contains only data. Given this ability, searching by ontology provides more insight and accuracy than a database.

C) Reusable. It greatly benefits future development. Ontology development is structured and inheritable which leads to time saving and integrity for new or derived ontology development. [8], [9].

Presently the ontology is used to manage various domains of knowledge such as engineering, medical and agricultural. It can also be applied to create artificial intelligence such as the ontology to manage data for rice products [10], or the ontology for individual nutrition intake [11].

Next, the detail of the development of the ontology process are presented

III. THE PROCESS OF ONTOLOGY DEVELOPMENT

For the development of this ontology, we plan to follow the process defined by Natalya F.Noy and Deborah L.McGuinness [12] which can be classified into 6 steps as shown in Fig.4.

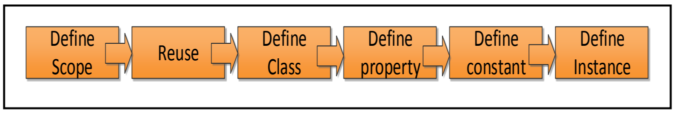

Fig. **4**.Process of the ontology development.

a) Scope: The information of prehistoric earthenware pottery in this study will be collected from 35 important sites around Thailand which includes 4 archeological sites in the north, 12 sites in the north east, 10 sites in central Thailand, 5 sites in the west, another 2 sites in the east, and the last 2 sites in the south. All information is authentic and has been empirically analyzed by excavators [13]. We have defined a set of questions for this ontology based on the pre-historic archaeological research interests. By this framework, all archaeological information should be analyzed to define classes, properties, slot and instance e.g. types, shapes, and techniques. In fact, the obtained information can infer even human culture and tradition, food, and settlement etc. [14].

To gain an insightful knowledge system of prehistoric pottery, a preliminary survey was conducted on the subjects who were archeologists to gather questions of interest for querying this ontology. The obtained scope of questions is shown in Fig. 5.

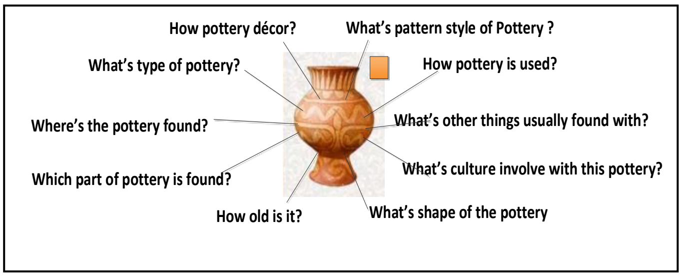

Fig. 5. The Question scope of querying.

b) Reuse: This step is to reuse or extend the existing ontology for ensuring standardized terminologies and classes. This is beneficial for the archeological domain due to the nature of archeological information which constantly change and is updated according to the emergence of new information or evidence.

c) Defining Class: In this ontology, there are 23 main classes defined in accordance with archeological information and interests. Fig.6. shows all 23 classes with potential cross-referencing.

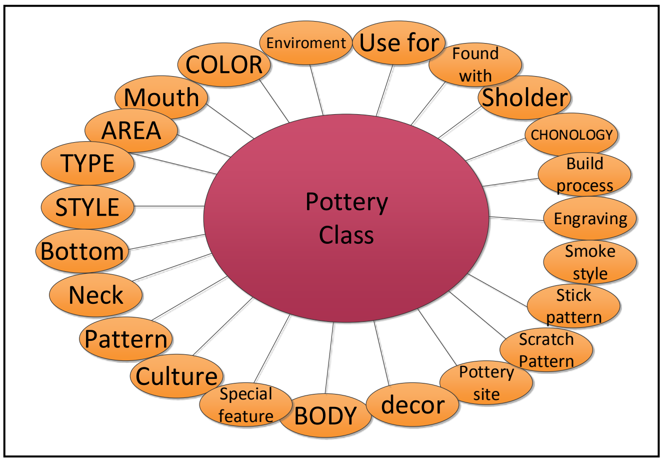

Fig. 6. Shows 23 main classes of this ontology.

In the study of archeology, there are two types of meaning for specifying insight into a class of interest. The first is direct meaning, and the second one is transitive meaning.

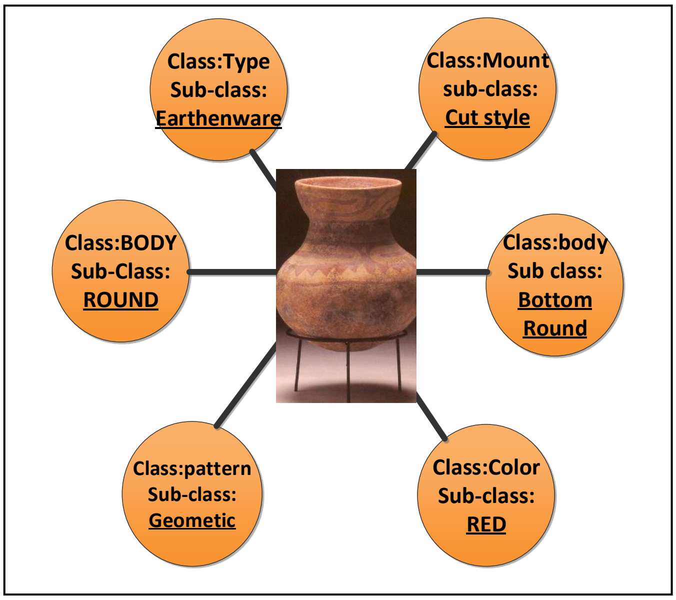

Fig. 7. Earthenware pottery and the main classes indicating direct meanings.

For the direct meaning, as shown in Fig.7, each main class in this ontology can be described more specifically using direct meaning sub-class, e.g. class of BODY is with the sub-class of ROUND to define a round body shape. See Fig.7.

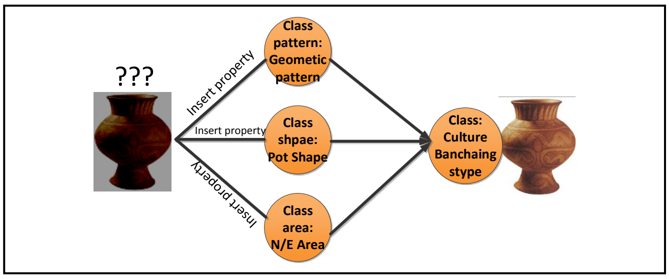

Fig. 8. Example of transitive meaning sub-class.

Fig.8. displays the class transitive meaning. Class transitive meaning uses the transitive property for making the relation between classes [15]. As an example, an earthenware pottery vessel having 3 properties (pot shape, geometric line and found in the north east of Thailand) could be concluded that this pottery is from Ban Chaing.

Relying on the direct meaning and transitive meaning, therefore means that the ontology allows users to input partial data for querying and the entire related information can be retrieved.

d) Defining property: This is to define property (slot) of a specific class, e.g. Earthenware pottery as shown in Table I. These properties can help to define an instance out of a specific class as shown in Fig.9.

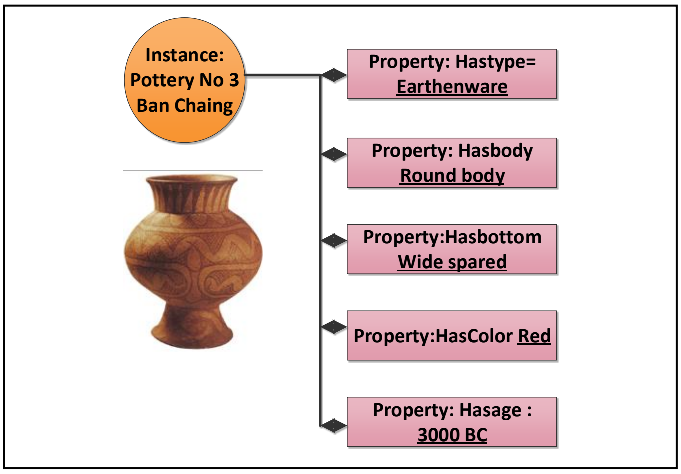

Fig. 9. A sample instance with properties.

Table I Example of Property for each type of earthenware

	Property	Description
1	(IsFoundon)	Found area
2	(HasTechniqueof)	Hasbuildfrom technique
3	(IsShapeOf)	Pottery shape
4	(HasColor)	Pottery color

Moreover, well-defined properties can help in searching across domains (or classes).

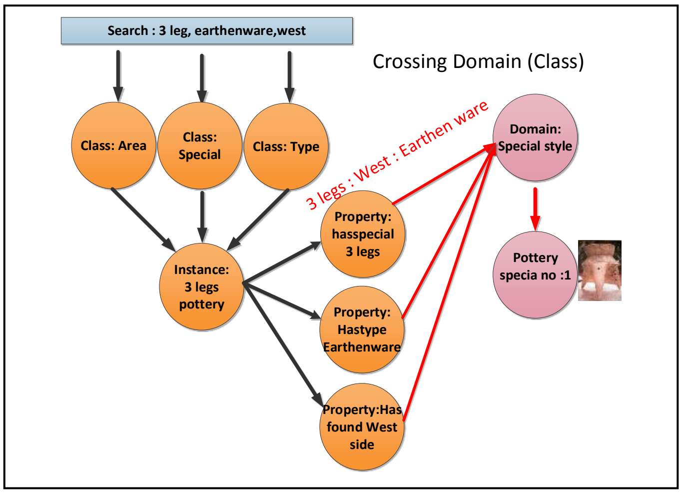

Fig. 10. Example of Domain-Cross Searching

According to Fig.10. a search with the key words of three legs, earthenware and west will indicate properties of an instance that inherited from a class of special pottery no. 1 triple leg pot. The traceability of relationship enables effective cross searching ability which is powerful when the network of knowledge is complex and contains many domains or classes.

e).Defining Constant. This is to define constant value in ontology. The constant value is information which is definite such as pottery types found in Thailand are only 3 types include Earthenware, Stoneware and Porcelain [16].

f). Defining instance. This is to define an element of the object such as specific piece of earthenware. Properties of each pottery instance must be declared by the expert, for it is the underlying knowledge of the ontology.

IV. Architecture of Ontology

The architectural elements of this ontology are in accordance with web application standard as illustrated in Fig.11. There supposes to have 3 integral elements in the system: Operation server, the web service and user interface, and Ontology storage.

a) Operation server: this is the main server where the management of ontology and querying logic operates of. The system is supposed to comprises (1) uploaded ontology which could be developed using as open-source platform such as the Protégé by Stanford University, USA [17], and (2) programs to manage the logical flow and control the operation, incl. pre- and post- processes. Java could be a good candidate.

b). Web server and interface: This provides the communication channel for the users to obtain services from the operation server. Available web and semantic web technology, e.g. web browser, Web Ontology Language (OWL) and Net framework which is the tool for developing semantic web using C# language could be utilized. The designed user interface must ensure friendliness in facilitating users to connect and search information in the developed ontology [18]

C). Ontology Storage: It collects, manages data, and restores information in XML format which is the suggested platform by W3C for semantic layer. XML format has several benefits eg. XML is self-described data using definable tags. XML is also platform independent, and XML supports multi lingual documents including Thai language. [19]

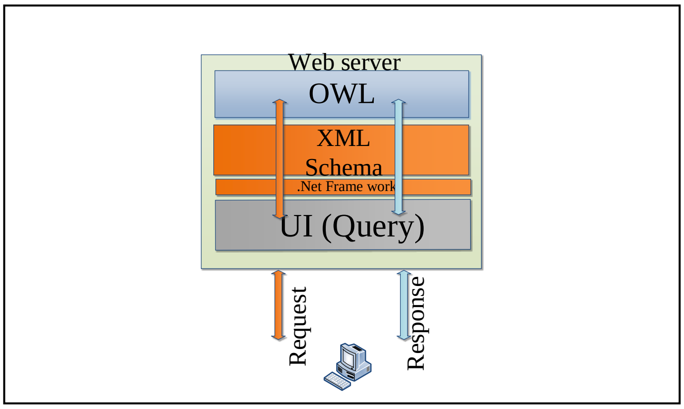

Fig. 11. Architectural structure of the proposed ontology.

V. System Testing.

The main target users of this study are those archeologists working in the field and others interested in prehistoric earthenware pottery of Thailand. The search system should be tested against the planned scope of questions of interest. Measurements in terms of the (i) semantic variance and (ii) users satisfaction are thus proposed here. The querying process could be established as shown in Fig. 12 where the users input key words and the system will facilitate the XML parsing and work through the OWL process to acquire answers from ontology through class (concept), role (property) and finally the relevant instances.

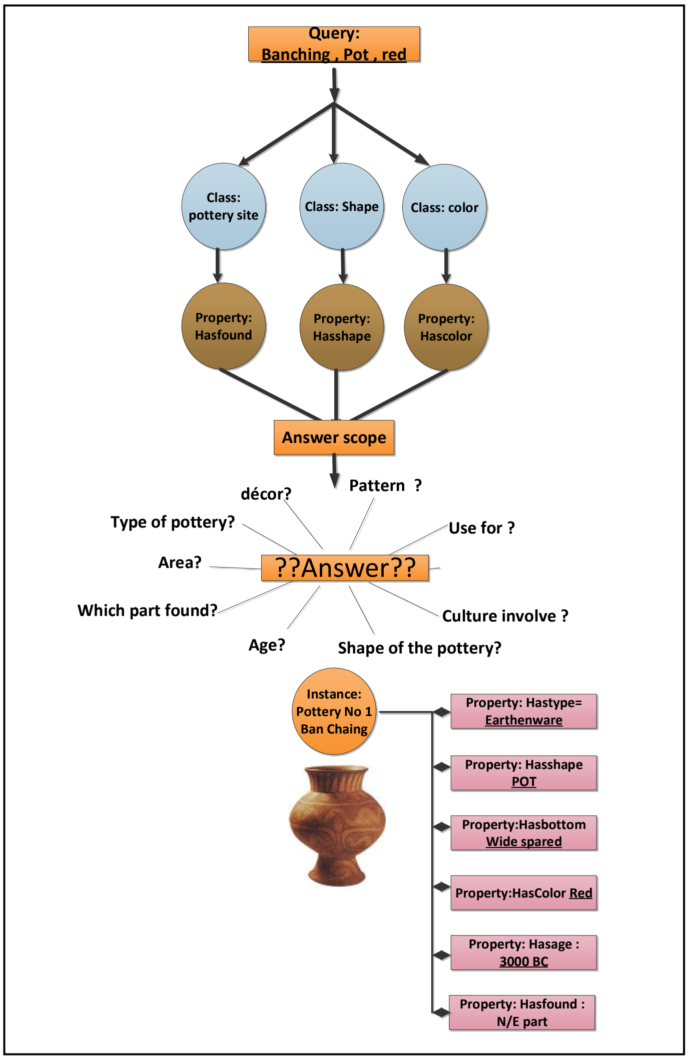

Fig. 12. Example searching process

Semantic variance measures effectiveness of the architectural design of the ontology [20]. The measure is based on the statistical concept of class dispersion with regard to the center of the depth and breadth of the ontology structure as shown in Equation 1

$$\text{Semantic\_Variance}(O) = \frac{\sum_{c_i \in C} d(c_i, Root(O))^2}{|C|} \qquad (1)$$

where O represents an ontology containing a set of concepts (C) with cardinality of $|C|$. d is the distance from the root node, $Root(O)$, to a concept, $c_i \in C$.

User satisfaction can be measured by surveying opinions of experts in the fields of Archeology and Information Technology, and also of those who are interested in this knowledge system. The measure could be examined by descriptive statistics [21] of arithmetic mean (Equation 2) and standard deviation (Equation 3) of subjects' opinions collected via questionnaire and quantified in Likert's 5-scale system (see Table II) [22].

$$\bar{x} = \frac{\sum_{i=1}^{n} xi}{n} \qquad (2)$$

Where x_i is the obtained scores for question i and n is the number of evaluators.

$$SD = \frac{\sqrt{(x-\bar{X})^2}}{n-1} \qquad (3)$$

The t-test scheme could be applied on the acquired opinions of different groups of respondents to gain insights into the hypothetical distinguishableness in their ideas

.

Table II: Scheme of scoring based on Likert scale

Level of score	Evaluation
5	High Distinction
4	Distinction
3	Credit
2	Fair
1	Most dissatisfaction

VI. Conclusion and Potential Development

This paper proposes how to systematically develop the ontology for prehistoric earthenware pottery found in Thailand and the semantic search in the knowledge system. The platform of open-source, web-based and standardized OWL are recommended and the framework for development is illustrated for effective interoperability and portability. The evaluation of semantic variance and user satisfaction are suggested as key evaluation schemes for they can help to predict the accuracy and usability of the ontology design once completed. The ontology system can serve as underlying system for the development of knowledge management (KM) or semantic web of the domain of archeological study in Thailand as the archeological information keeps constantly updated or changed reflecting new evidence discovered.

The next stage of this study is to materialize the ontology system using the proposed framework and platform. All involved archeological data as required by the scope of questions are already acquired from archeologists and the literature. We also plan to publicize the ontology on web and extend the functions for being a learning support system for archeological domain.

References

[1] Professor Pasut Intrawut, "Index The pottery Thadhamravadee," Bangkok, Sipakorn University press, 1995.

[2] Aarek Obitko, "Modularization of Ontology," [online] Available:http://www.obitko.com/tutorials/ontologies-semantic-web/modularization-of-ontologies.html, July 7, 2015

[3] Atanas Kiryakov, "OntoText Lab," Sirma AI Ltd.,Hr. Botev 38A, 1000 Sofia, Bulgaria, 2010.
[4] T. R. Gruber,"Towards Principles for the design of ontologies used for knowledge sharing," In N. Guarino and R. Poli, editors, Formal Ontology in Conceptual Analysis and Knowledge Representation Kluwer Academic Publishers, Deventer, The Netherlands, 1993.
[5] Assistant Professor Somchai Prakanchalern "Ontology: alternatives of knowleadgebase development in content base form" [online] Available: http://202.44.34.144/nccitedoc/admin/nccit_files/NCCIT-20110504195051.pdf, July 7, 2015, P.5-7.
[6] KIF (Knowledge Interchange Format) [Online] Available: http://www.ksl.stanford.edu/knowledge-sharing/kif/#ontologies, July 7, 2015.
[7] Mike Uschold and Michael Gruninger, "Ontologies: Principle, Methods and applications" The university of Edinburgh.1996.
[8] Modularization of ontology: Reusing of knowledge [online] Available:July 18, 2015.
[9] Sirirut Pakitikornchai,"Ontology design of herb" Thesis Master degree computer science, school of science, Bangkok, King Mongkul's University of Technology North Bangkok, 2007.
[10] Rosarin Aknit, "Ontology develop for knowledge management rice product"Agriculture technology department, Bangkok: Faculty of agriculture industry, Kasatsart University, 2007.
[11] Napat suksom, Malut Buranarut, Tappachai subniti, Pornlidee Natisonakul, "Ontology for introduces nutrition for person" Bangkok: Information technology department, King Mongkul's university of technology North Bangkok, 2010.
[12] Natalya F. Noy and Deborah L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology"Stanford University, Stanford, CA, 94305.
[13] PreechaKanjanakom, "Introduction to archaeology" Bangkok:Sipakhon press.1995.
[14] Pichasini kitwattathawon, Thara Aungsakul, Jitimon Aunsakul "Knowleage quering for online hotel journal from fuzzy logic" Bangkok: King Mongkul's university of technology North Bangkok, journal Vol 2 May-August, 2013.
[15] Transitive function [Online] Available: http://www.w3.org/TR/owl-ref/, July 18, 2015
[16] Supapol natapintu, 2007,"Ancient Thailand: Culture development of pre-historic"Bangkok:Matichon Press,2007.
[17] Protégé Stanford university [online] available in: http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library#OWL_ontologies, July 18, 2015.
[18] MSDN Web server Development [online] available: https://msdn.microsoft.com/en-us/library/aa347613(v=vs.90).aspx, July 18, 2015.
[19] Benefit of XML [online]Available:http://www.mulberrytech.com/papers/HowAndWhyXML/slide009.html, July 18, 2015.
[20] David Sánchez, Montserrat Batet,Sergio Martínez,Josep Domingo-Ferrer ,"Engineering Applications of Artificial Intelligence" Universitat Rovira i Virgili,2014.
[21] Natthapol ruangpayungsak "The development of knowleadgebase system about revenue department general inforamtioncase study by using ontology technology" Bangkok: King Mongkul's university of technology North Bangkok,2007.
[22] Likertscale[online] Available: http://core.ecu.edu/psyc/wuenschk/StatHelp/Likert.htm.July 19, 2015