

# The Use of Terminological Theory Approach in the Development of Ontology Based Bilingual Terminology System on College Campus of Taiwan

Hsing-I Wang

Department of Information Management

The Overseas Chinese University, Taichung, Taiwan, 407

hsing@ocu.edu.tw

**Abstract**—For the majority college students in Taiwan, learning and using terminologies of a specific domain between Chinese and English interchangeably are quite a challenge. Most of the students seek for assistances from library resources or search for answers on web. Unfortunately, the students would not be able to identify the correctness of their findings, or the worse, the students cannot choose the right translation for their queries. The lack of a well designed bilingual terminology system would hinder the efficiency and the advance of learning of college students.

Terminologies are the keys to agglutinate the knowledge. To understand the terminologies properly would wider the door to the core of the knowledge field of that domain. However, most of the terminologies references or dictionaries were simply linguistics or lexicographic oriented. This research adopts the terminological theory and proposes a framework of an ontology-based bilingual terminologies system on campus.

**Index Terms**—Terminological theory, terminology system, ontology

## I. INTRODUCTION

For the majority college students in Taiwan, learning and using terminologies of a specific domain or learning terminologies between Chinese and English interchangeably are quite a challenge. Most of the students seek for assistances from library resources or search for answers on web. Unfortunately, the students would not be able to identify the correctness of their findings, or the worse, the students cannot choose the right translation for their queries. The lack of a well and suitable designed bilingual terminology system would hinder the efficiency as well as the advance of learning of college students.

Terminologies are the keys to agglutinate the knowledge of a specific domain. To understand as well as to use the terminologies properly would wider the door to the core of the knowledge field of that domain. To

standardize terminologies of a field would also eliminate ambiguities and set the status of science for the terms. Scholars have urged the importance of the establishment of terminology as a discipline for all practical purposes [1]. However, most of the terminologies references were edited while terminological theory was still in the collective stage, most of them were simply linguistics or lexicographic oriented. Moreover, the definition of “terminology” is, most of the time, mixed with terms and words [2, 3]. Therefore, the terminologies of a specific domain are simply explained in few words or terms which would not illustrate the implications comprehensively.

The purpose of this research is to develop an ontology-based terminology system. The structure of the ontology is organized in accordance with the terminological theory. The paper is organized as follow: section 2 first explains the importance of defining and modeling terminologies. The terminological theories as well as the concept of ontology are reviewed in section 3 and section 4. Section 5 explains how the ontology is structured in this research and in section 6, this paper presents the architecture, the database structure and the search flow of the system. Section 7 gives the conclusion of this paper and further work is discussed.

## II. THE IMPORTANCE OF DEFINING AND MODELING TERMINOLOGIES

To demonstrate the importance of defining and modeling terminologies, fig. 1 and fig. 2 give the examples of the query of “ontology” on two popular web systems.

Apparently, the result of web dictionary (fig. 1) is too simplified. Although the knowledge system (fig.2) provides multiple explanations and allows users to look into other related categories, for students such as those who are majored in information systems, the results would not give sufficient and proper references. In fact,

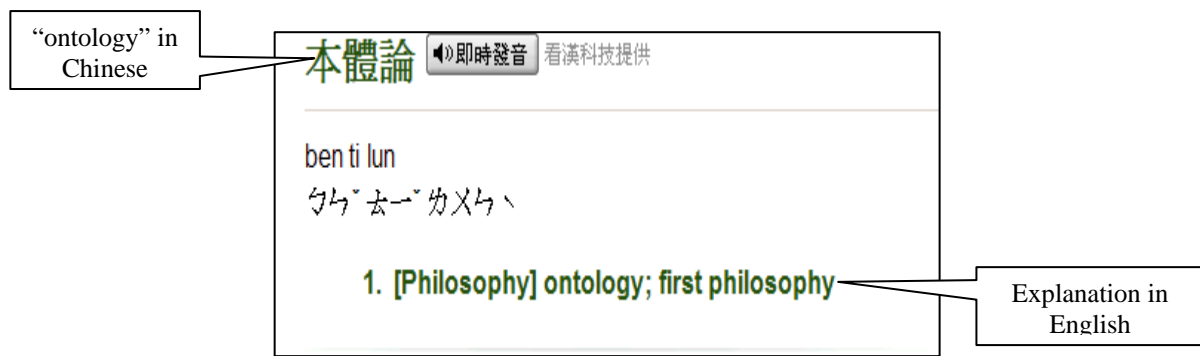


Figure 1. Result of query "ontology" using a web dictionary



Figure 2. Result of query "ontology" using a web knowledge system

an experiment had shown that only 57% of the results of terminology queries on current web resources were considered accurate [4]. In addition, the terms used for the experiment were proved not able to be obtained from a general-purpose dictionary [4].

A terminology system should be able to describe and explain the terminologies in the full awareness in the sense of linguistic, thematic and situational context [1]. Even better, the system would describe how new special knowledge is produced [1]. In the above "ontology" example, ontology is a theory about the nature of existence in philosophy; in artificial intelligence (AI), ontology defines the relations among entities in terms of the prosperities and classes. In semantic web, ontology captures and integrates heterogeneous domain knowledge [5], and in turn, the ontology provides the capabilities of enabling better search and discovery. A terminology, hence, may imply different meanings due to various

applications, disciplines and/or contexts. A unitary design of terminology system would obviously not meet the aforementioned needs.

### III. TERMINOLOGICAL THEORY

Wüster (1898-1977) is the pioneer of terminological theory. He devoted his life to promote theory of terminology to eliminate ambiguity for better communication, to help users benefit from standardization, and to establish terminology as a discipline [1]. Although the terminological theory is still in its infancy, its importance has never been neglected.

In the past, terminology has constantly been mistaken as terms. Bessé et al. [6] defined terms as "a lexical unit consisting of one or more than one word which represent a concept inside a domain" and terminology as "the vocabulary of a domain or a subject field, i.e. a set of terms in a domain or a subject field". Clearly,

terminology is by no means equivalent to terms or words. In fact, in average, about 80 percent of the terminologies in scientific domains are made up of complex terms [7]. Nevertheless, terminologies are connected to each other through shared terms [7].

“Ref. [8]” proposed that to integrate terminology with linguistics is quite important especially the bound between onomasiology and conceptology. In the literatures, most researchers seem to share the same opinions. As pointed out in [9] study, he noted that almost all terminology studies started with concepts and took an onomasiological approach. The findings suggest that besides domain, the concept delivered within terminology should be taken into consideration while building the terminology database.

Terminological unit is another issue that has been discussed in terminological theory. From the above definitions, we notice that terminology is related to knowledge, language and communication. Therefore, a terminological unit actually covers a cognitive component, a linguistic component and a socio-communicative component [1]. However, this definition does not make terminology distinguished from language unit, [1] then specified terminological unit in more detail with respect to the three perspectives:

- They are context oriented; occupy a precise place in a conceptual structure and the meaning is determined by their place in this structure.
- They can have lexical and syntactic structure and belong to one of the broad semantic categories which are characterized by entities, events, properties or relations.
- They are acquired through a learning process and hence are handled by specialists in their field.

Finally, terminology is considered a set of needs, a set of practices to resolve these needs, and a unified field of knowledge [1].

IV. THE ONTOLOGY

Traditionally, “ontology” is about philosophical study of the nature of being or existence [10]. It was then borrowed to AI to illustrate the relationships of the set of concepts and terms. “Ref. [11]” defined ontology as an explicit specification of a conceptualization. It allows us to describe resources formally.

Ontology can be classified into terminological ontology, information ontology and knowledge ontology. In recent years, ontology has been widely adopted in the areas such education [12], e-learning [13], knowledge management [14, 15] or even managerial fields [16].

Ontology can be expressed in the form of taxonomic tree; however, the kernel is much complicate. The representational primitives of ontology often include the information about the meaning and constraints on the logically consistent application [17]. Therefore, building a learning or knowledge system with ontology will enhance sharing and reuse knowledge; the advance and

the improvement of the knowledge of a certain domain will be verified as well.

In the review of terminological theory, literatures indicated that terminologies are connected to each other through shared terms. Therefore, an ontology-based terminological database would be more appropriate in the development of a terminology system.

V. THE DEVELOPMENT OF THE STRUCTURE OF THE ONTOLOGY

Based on the contentions stated in section 3, this research concludes the following key notions that should be included in modeling the ontology of a web-based terminology system:

- Category: The broader classification of the fields.
- Domain: The special knowledge needs.
- Context: The activities related to the domain.
- Concept: The thoughts that a terminology tries to deliver.
- Description: Description is the center of the ontology. The description of a terminology must contain a cognitive component, a linguistic component and a socio-communicative component. Meaning that the descriptions are context oriented and are handled by specialists in their fields.

Fig. 3 shows the specialized tree structure of the terminology system. The building of the ontology in this research is based on terminological theory. The process depends heavily on the contribution of domain experts as well as native English speakers or English teachers.

The categories are determined according to the schools that a college has. The domains of the terminologysystem for college campus can then be divided in consistent with (but not limited to) the departments of each school. Contexts and concepts are basically determined by experts or teachers. Notice that a terminology may appear in more than one route.

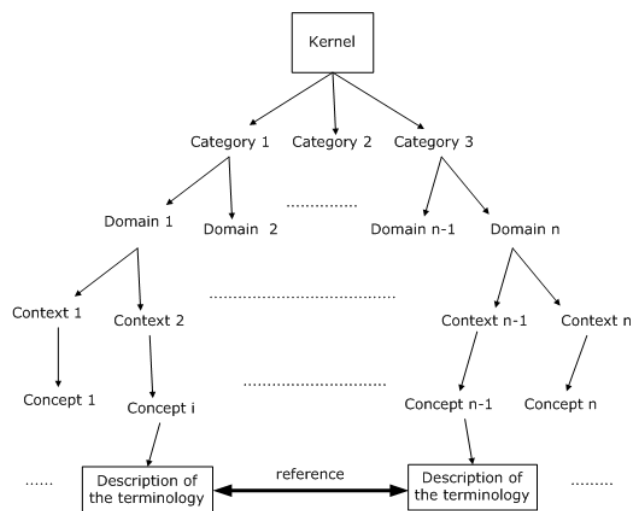


Figure 3. The specialized tree structure of this research

Fig. 4 shows the ontology tree of the terminology “ontology”. The example will be adopted throughout this paper.

## VI. THE BILINGUAL TERMINOLOGY SYSTEM

### *The framework and data structure*

The web-based bilingual terminology system proposed in this research consists of three layers including the presentation layer, the evaluation layer and the ontology layer. The framework is shown in fig. 5. The presentation layer represents the interface between the users and the system. The second layer is the diagnosis layer. The diagnosis layer functions as both the pre-process as well as the assessment station. The experts and translators are responsible for the job by interacting with the system. Masters or experts are the designate teachers from each different or field. The main responsibility of the masters is to determine the logical route in which the terminology should reside on the architecture of an ontology tree. The translators are responsible for translating the already well-defined terminologies into English. Basically, the process of the translation must follow the mandated standard of the module of the ontology as well.

The third layer, which is the ontology layer, keeps the whole structure of the terminological ontology. After the terminology is orientated in the ontology module, the information is stored in the database.

### *The data structure*

In the ontology layer, the database contains thirteen tables. There are two identical tables for each level in the ontology tree except that the contents are recorded once in English and again in Chinese. In addition, two tables are designed to keep the full path of each terminology in the ontology tree. The last table is used to store the terminology that is new to the database and will be processed later. The thirteen tables are Category\_English, Category\_Chinese, Domain\_English, Domain\_Chinese, Context\_English, Context\_Chinese, Concept\_English, Concept\_Chinese, Terminology\_English, Terminology\_Chinese, Terminology\_path\_English, Terminology\_path\_Chinese, and Unresolved\_terminology respectively. The schemas are given in table I. Attributes that are underlined is the key of that table. The same rule applies to the Terminology\_path\_Chinese table as well. E-R models are shown in fig. 6. The E-R models demonstrate the relationships between each level of the ontology tree. In addition, the relationship between tables “Terminology\_English” and “Terminology\_path\_English” (or between “Terminology\_Chinese” and “Terminology\_path\_Chinese”) indicates that a terminology may appear in more than one route. Therefore, the primary key in Terminology\_path\_English table contains all attributes in the table to assure the uniqueness of each record and furthermore, the full path of the terminology in the ontology tree is recorded. Once a query is sent to the system, all paths that contain this

TABLE I.

Schema of each table in database

Table name	Category_English
attributes	{ <u>category_id</u> , category_name, category_description }
Table name	Domain_English
attributes	{ <u>domain_id</u> , category_id, domain_name, domain_description }
Table name	Context_English
attributes	{ <u>context_id</u> , domain_id, context_description }
Table name	Concept_English
attributes	{ <u>concept_id</u> , context_id, concept_description }
Table name	Terminology_English
attributes	{ <u>terminology_id</u> , terminology, terminology_description }
Table name	Terminology_path_English
attributes	{ <u>terminology</u> , <u>concept_id</u> , <u>context_id</u> , <u>domain_id</u> , <u>category_id</u> }
Table name	Category_Chinese
attributes	{ <u>category_id</u> , category_name }
Table name	Domain_Chinese
attributes	{ <u>domain_id</u> , category_id, domain_name }
Table name	Context_Chinese
attributes	{ <u>context_id</u> , domain_id, context_description }
Table name	Concept_Chinese
attributes	{ <u>concept_id</u> , context_id, concept_description }
Table name	Terminology_Chinese
attributes	{ <u>terminology_id</u> , terminology, terminology_description }
Table name	Terminology_path_Chinese
attributes	{ <u>terminology</u> , <u>concept_id</u> , <u>context_id</u> , <u>domain_id</u> , <u>category_id</u> }
Table name	Unresolved_terminology
attributes	{ <u>id</u> , terminology, concept_id, context_id, domain_id, category_id }

specific terminology will be retrieved from Terminology\_path\_English (or Terminology\_path\_Chinese) and the descriptions of concept, context, domain and category can also be retrieved by referencing to the related tables. Based on the example shown in fig. 4, table II lists all the related records stored in Terminology\_path\_English table.

### *Search flow*

In the presentation layer, the ontology tree is presented similar to the directory structure of Windows that most people are familiar with. There are two ways to locate the terminology. The first and probably the easier one is that the user interacts with the system and explores each directory in depth until he/she finds the node (or the

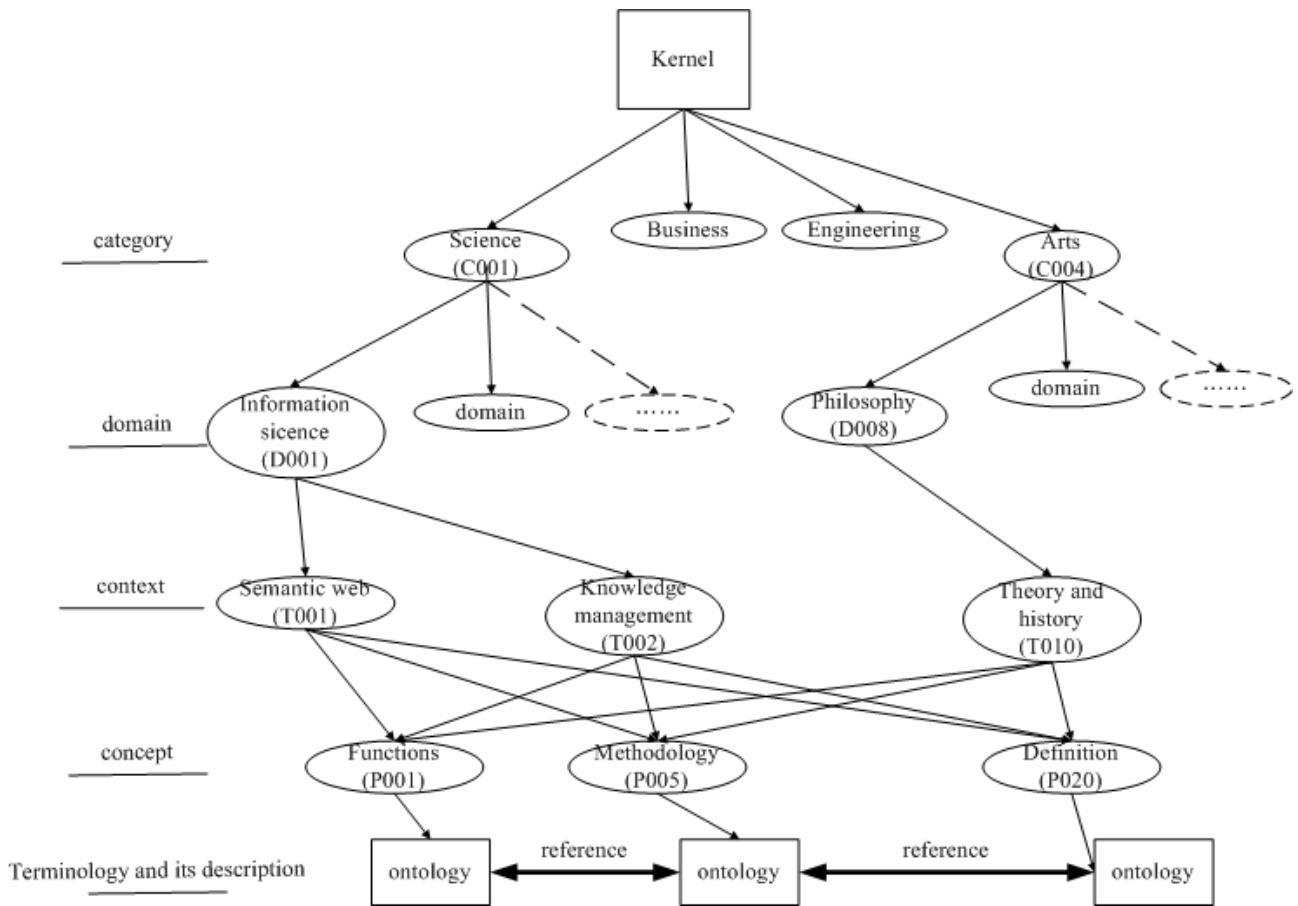


Figure 4. The ontology tree of terminology "ontology" example

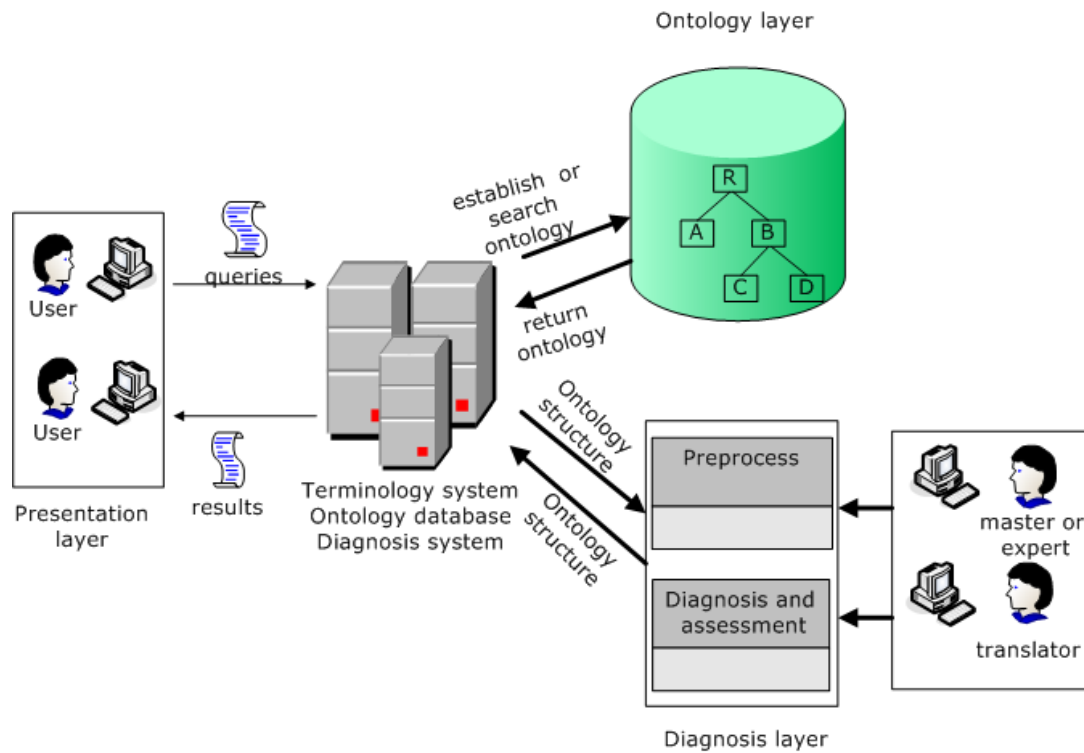


Figure 5. The framework of the ontology-based terminology system

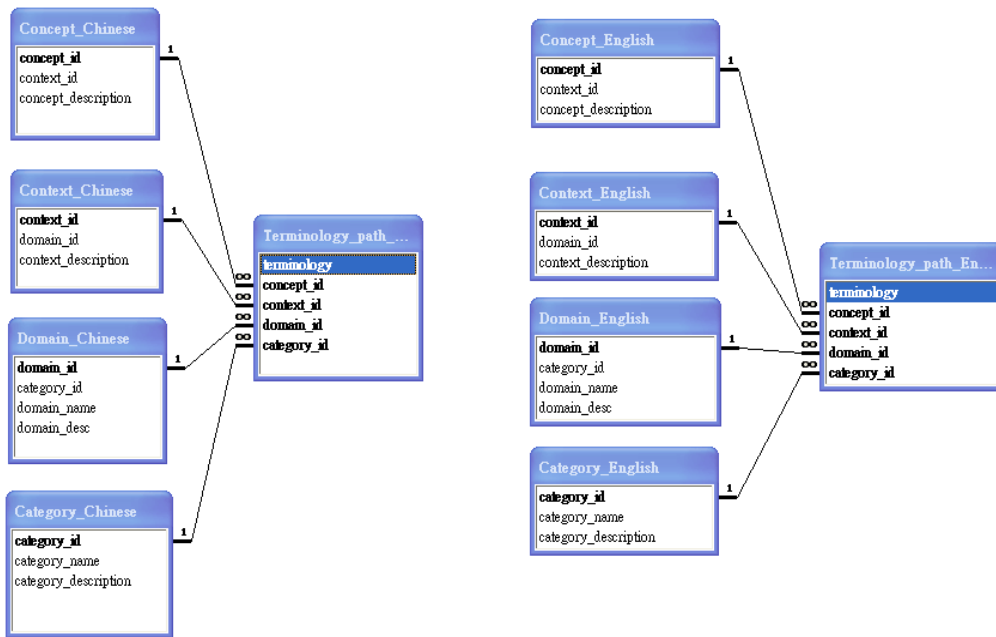
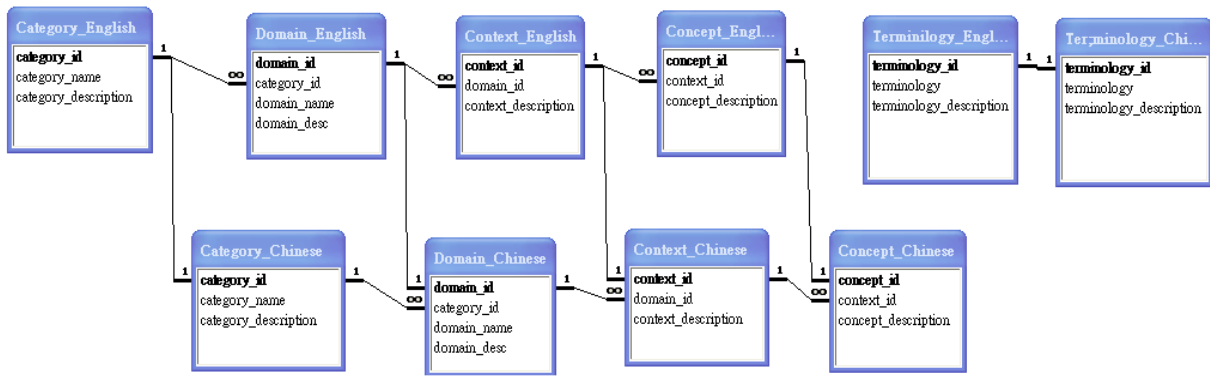


Figure 6. E-R models of the tables in this research

TABLE II.

Records of “ontology” that are recorded in Terminology\_path\_English (based on the examples shown in fig. 4)

terminology	concept_id	context_id	domain_id	category_id
ontology	P001	T001	D001	C001
ontology	P005	T001	D001	C001
ontology	P020	T001	D001	C001
ontology	P001	T002	D001	C001
ontology	P005	T002	D001	C001
ontology	P020	T002	D001	C001
ontology	P001	T010	D008	C004
ontology	P005	T010	D008	C004
ontology	P020	T010	D008	C004

terminology). However, this will take a lot of time if the user is not certain which category or domain or context or concept the terminology belongs to. The second way, which is the most likely case, the user enters the terminology to query. The ontology tree will be searched once a query is proposed and the system lists all the possible routes for user to choose.

For example, if the user searches for the terminology "ontology". Based on the ontology tree shown in fig. 4, we will find the following records in Terminology\_path\_English as shown in table II. Records in table II indicate that "ontology" can be found thru:

- "Science" → "information science" → "semantic web" → "functions".
- "Science" → "information science" → "semantic web" → "methodology".
- "Science" → "information science" → "semantic web" → "definition".
- "Science" → "information science" → "knowledge management" → "functions".
- "Science" → "information science" → "knowledge management" → "methodology".
- "Science" → "information science" → "knowledge management" → "definition".
- "Arts" → "philosophy" → "theory and history" → "functions".
- "Arts" → "philosophy" → "theory and history" → "methodology".
- "Arts" → "philosophy" → "theory and history" → "functions".

The bold arrows in fig. 4 indicate that in this design, terminology can also share the knowledge across fields.

The results will be listed on the screen. Since the users may choose any one from the list, they will learn more about this terminology than they expected in the beginning.

If no related records are located, the system will return a message and ask users to choose the possible category, domain, context and concept so that the request can be stored. Later, the evaluation as well as the process will be carried out by masters or experts in the diagnosis layer.

## VII. CONCLUSION

This research contributes a different thinking in the development of a web-based terminology system. The system different from traditional terminology systems in two ways: firstly, the answers to a query of a terminology are the search results from an ontology-based terminology database. Secondly, the structure of the ontology-based terminology database is established according to the rules that are derived from the terminological theory.

The approach this research adopts assures the following: first of all, the explanations of the terminologies are presented not simply in the forms of words or terms, but are described to fit the context and

the domain. Therefore, the results will make more sense to the queries.

Secondly, the ontology is built in accordance to the terminological theory and the rules are derived from the theory. Each terminology is assured to occupy a precise place in the conceptual structure. In addition, the ontology structure, both the Chinese and English version, is fixed and recognized with the help of the expert of the domain.

Thirdly, the system provides two ways to locate a terminology. The users may explore the ontology tree and choose the proper explanations. Or, the users may submit a query about a terminology to the system, and the system returns all possible explanations to the users. Since the results will contain the full path including the category, the domain, the context and the concept of each answer, the users may choose a proper answer with more definite guidance. Moreover, the users may learn more other applications about a terminology in the system.

Finally, this system is built on campus for college students. The contents are collected either as the request of the students or by teachers in the related fields. The contents are believed more suitable for students in terms of self-learning, degree of profundity, and degree of usefulness.

This paper assumes that users use a simple terminology to search in the database. Neither combined conditions nor nature language queries are considered in this paper. Further studies may take the issues into considerations. The efficiency of the systems as well as the satisfaction of the users toward the results of the system has not been investigated. Future work will focus on the above subjects.

## REFERENCES

- [1] M.T. Cabre, "Theories of Terminology", Terminology, vol. 9, pp. 163-199, 2003.
- [2] J.C. Sager, "In Search of a Foundation: Towards a Theory of Term", Terminology, vol. 5, pp. 41-57, 1998/1999.
- [3] K. Kageura, "Theories of Terminology: A Quest for a Framework for the Study of Term Formation", Terminology, vol. 5, pp. 21-40, 1998/1999.
- [4] W. H. Lu, H. J Lee, and L. F. Chien, "Term Translation Extraction Using Web Mining Techniques", dissertation, Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, 2003, unpublished.
- [5] B. Sridharan, A. Tretiakov, Kinshuk, "Application of Ontology to Knowledge Management in Web-Based Learning", IEEE International Conference, pp. 663-665, 2004.
- [6] B. de Bessé, B. Nkwenti-Azeh, and J. C. Sager, "Glossary of Terms Used in Terminology", Linguistics, vol. 35, pp. 861-877, 1997.
- [7] K. Kageura, "An Analysis of the Motivatedness Structure of Japanese Terminologies", Mathematical Linguistics, vol. 26, pp. 241-263, 2008.
- [8] J. Myking, "Against Precriptivism? The 'Sociocritical' Challenge to Terminology", IITF Journal, vol. 12, pp. 49 - 64, 2001.

- [9] K. Kageura, "Toward the Theoretical Study of Terms", *Terminology*, vol. 2, pp. 239-258, 1995.
- [10] Wikipedia, <http://en.wikipedia.org/wiki/Ontology>.
- [11] N. Guarino, "Formal Ontology, Conceptual Analysis and Knowledge Representation", *International Journal of Human-Computer Studies*, vol. 43, pp.625-640, 1995.
- [12] M.-H. Abel, A. Benayache, D. Lenne, C. Moulin, C. Barry, and B. Chaput, "Ontology-based Organizational Memory for e-learning", *Journal of Educational Technology & Society*, vol. 7, pp. 98-111, 2004.
- [13] D. Sampson, C. Karagiannidis and F. Cardinali, "An Architecture for Web-based e-Learning Promoting Re-usable Adaptive Educational e-Content", *Journal of Educational Technology & Society*, vol. 5, pp. 27-37, 2002.
- [14] A. Maedche, B. Motik, L. Stojanovic, R. Studer, and R. Volz, "Ontologies for Enterprise Knowledge Management", *IEEE Intelligent Systems*, vol. 18, pp. 26-33, 2003.
- [15] J. Davies, D. Fensel and F. V. Harmelen, *Towards the Semantic Web: Ontology-driven Knowledge Management*, New York: Wiley, 2003.
- [16] V. C. Storey, "Real World Knowledge for Databases", *Journal of Database Administration*, Winter, pp. 1-19, 1992.
- [17] T. Gruber, "Ontology", *Encyclopedia of Database Systems*, Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, 2008.

**Hsing-I Wang** is an Associate Professor in the Department of Information Management, Overseas Chinese University, Taichung, Taiwan. She holds a PhD degree in management information systems from the National ChengChi University, Taipei, Taiwan. Her previous research areas include e-government, object oriented system design, applications of artificial neural network, knowledge management, and business managerial issues.