

## Design for Ontology Knowledge Base Based on Structural Members

Qian Li-Ping<sup>1</sup>, Yang Xiao-Ping<sup>2</sup> and Song Yu<sup>3</sup>

<sup>1</sup>*School of Information, Renmin University of China, Beijing 100872  
College of Electrical and Information Engineering, Beijing University of Civil  
Engineering and Architecture, Beijing 100044*

<sup>2</sup>*School of Information, Renmin University of China, Beijing 100872*

<sup>3</sup>*College of Electrical and Information Engineering, Beijing University of Civil  
Engineering and Architecture, Beijing 100044*

<sup>1</sup>*qianliping@bucea.edu.cn, <sup>2</sup>yang@ruc.edu.cn, <sup>3</sup>songyu@bucea.edu.cn*

### Abstract

*By studying the existing ontology design method and integrating the feature of domain knowledge which has significant component characteristics, this paper propose a method to build ontology knowledge base basing on structural member. The process includes the following six steps: requirement analysis, framework designing, coding, ontology evaluation, ontology evolution, document filing. Each step of the task is further decomposed into several detailed operation. An ontology evaluation method is also proposed. Through applying the domain ontology into the practices and evaluating the result before and after, the rationality of relationship is reversely inferred out. Ontology language is used to describe the concepts and the relation among them. It is also an ideal selection to describe the ancient buildings knowledge. The result of instant study confirms the approach is feasible and effective.*

**Keywords:** *Domain ontology, structural member, ontology evaluation, ancient building*

### 1. Introduction

Knowledge model and knowledge base is a popular research topic in knowledge engineering field. Many institutions and scholars, both at home and abroad, have done much study about it [1-2]. It will be easier to organize, describe and induct the knowledge if we can build a well-formed knowledge model. The most important thing is how to represent the knowledge effectively. Knowledge representation is to describe and store the knowledge with the method that computer can accept and manage. It is a data structure and expresses the relationship between the data. It is also the premise and foundation of knowledge organizing and knowledge utilizing. Method selected for knowledge representing is critical to the efficiency of reasoning and acquisition of new knowledge. Commonly used knowledge representation methods include predicate logic, production rule, semantic net, framework, and so on. Predicate logic is mainly used for automatic theorem proving. Production rule focuses on stating interactions among procedure knowledge. Framework is a kind of hierarchy structure to store all the relevant knowledge of an object or event. Semantic net has strong expression ability and is flexible, which express the concepts and relationships between concepts as knowledge network.

Knowledge representation has been improved by the introduction of knowledge ontology. Now the ontology theory has been widely used in the fields of artificial intelligence, knowledge engineering, biomedicine, economics and so on [2-4].

This paper studies construction method of domain ontology knowledge base based on structural member. Then the detailed operation and feasibility of the proposed method are illustrated through designing and realizing of the domain knowledge base of ancient buildings based on structural member.

## 2. Related Work

### 2.1. Ontology

There are so many achievements about ontology both at home and abroad. In the field of health care, Institute of science and technology of New Jersey has developed Object-Oriented Healthcare Vocabulary Repository (OOHVR). It has about 5000 concepts in the semantic net which are all stored in the object-oriented database. SNOMED and Unified Medical Language System (UMLS), the other two large vocabularies, are all used in the medical field. In the area of product and service, United Nations Development Program classified the terminology of the product and service. CYC is a general ontology. It devoted itself to integrate all kinds of ontology and commonsense knowledge together, and on this basis to realize knowledge reasoning. Princeton University, Berkeley developed the well-known WordNet [14-17].

Chinese computer scientists have also done a lot of researches on ontology. Lu Ruqing, academician of Chinese academy of sciences, leaded to build a large commonsense knowledge base PANGU, which had been utilized to resolve natural language understanding in the machine translation [5, 18].

Jin Zhi, a researcher from Chinese academy of mathematics, has researched method to access to the software requirements through combining the enterprise ontology and domain ontology and using the model reuse technology to create a system model [6]. Cao Ronggeng, a researcher from Chinese academy of sciences, has researched a large intelligent NKI (National Knowledge Infrastructure). It is a large knowledge system which includes multidisciplinary such as geography, medical science, chemistry, *etc.* [7].

HowNet is a knowledge base. It is developed by the Chinese Academy of Sciences Enderle Group. It is based on the English-Chinese bilingual represented concept and the characteristics of concept. HowNet can better reveal the relations between the concept and the characteristics [21].

Ontology description also plays an important role during the process of building ontology. Ontology languages are usually used to describe ontology. The common ontology languages include RDF, OIL, DAML, OWL, SHOE, XOL, *etc.* Web Ontology Language, or OWL, is a W3C recommendation. The W3C Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. OWL is a computational logic-based language such that knowledge expressed in OWL can be exploited by computer programs [22]. It can express knowledge base with a clear hierarchy of concepts and the relations among concepts. OWL has three sub-language, OWL Lite, OWL DL and OWL Full. OWL DL already has implemented reasoning ability.

Domain ontology has stronger domain background. It is more urgent to have domain ontology to support vertical search and big data analytics.

### 2.2. Method to Build the Ontology

So far, there are many methods to build ontology [8-9]. Some widely used methods include IDEF-5 method, skeleton methodology, TOVE enterprise modeling method, Cyclic Acquisition Process and Seven steps, to name a few. Seven steps are proposed by School of Medicine, Stanford University. The construction process of ontology is divided into seven steps: (1) Investigate and survey the applied range of the ontology; (2) Examine whether the existing ontology can be reused; (3) List important and basic terms

and concepts; (4) Build the framework of the ontology; (5) Define every class in the ontology, class hierarchy or layer architecture; (6) Define the attribute and type of the value of the attribute; (7) Create instance. Skeleton methodology is reached from the modeling process of Enterprise Ontology, the constructing process is: determine the applied range of the ontology, analyze ontology, represent ontology, construct ontology and evaluate ontology.

Through the comparison of several kinds of traditional ontology construction method, we can conclude that domain ontology construction is still in the exploratory stage. There are still some shortcomings more or less in these methods, such as inadequate demand analysis, poor extensibility, limited applicable area and so on. Since there are so many structural members in ancient buildings, and relations among them are complex, previous methods are not sufficient and applicable.

### **3. Technique to Build Ontology Knowledge Base-Based Structural Members**

#### **3.1. Basic Theory**

Ontology can be formally described as:

$$L = \{ E, R \}$$

$K$  denotes ontology.  $E$  denotes set of knowledge element. And  $R$  denotes a collection of relations between the knowledge elements.

It is a very controversial question about the concept of knowledge element [10,13]. The general concept is knowledge unit that has perfect knowledge express. Different areas have different knowledge structure, and so different element. Structural member in this paper refers to structural composition of an object. It is a fundamental unit, but it is not the unit in the last place. Sometimes it can be called Knowledge-ware or knowledge element. Here is the express of the structural member:

$$E = (N, P, V)$$

In the above expression,  $E$  denotes structural member,  $N$  is the name of the structural member,  $P$  is the attribute feature set of the structural member and  $V$  is the value of the attribute feature set. Ontology knowledge base based on structural member is constructed through the relations among the structural members.

#### **3.2. Method to Build Ontology Knowledge Base**

This paper draws on the experience of and integrates some traditional ontology construction method, and then proposes a new method to build domain ontology knowledge base. It is developed over skeleton method and seven-step method, follows the demand of soft project standard, and is based on object-oriented programming [19]. The method can be used to build domain ontology knowledge with many structural members.

The building process of ontology based on structural members is divided into six phases, as shown in Figure 1.



For example, we use the ontology to expand query for search engine, and then compare the search results returned by search engine before and after using ontology. The basic idea of similarity measure of query expansion based on domain ontology is that for each query word, through the use of semantic similarity computation method, automatically find a scalable owl files for extensible concepts of the relationship and add them in the queries collection [10]. It is depicted in Figure 2.

Equivalent relation: Extended concept is synonymous of query concept, that is the value of the semantic relative term is 1 or approximately 1.

Parent-son relation: In the hierarchy, extended concept and query concept is parent node and son node.

Sibling node: Extended concept is sibling node of the query concept. They have the same parent node.

Subtree node: Extended concept is sub-tree node of the query concept.

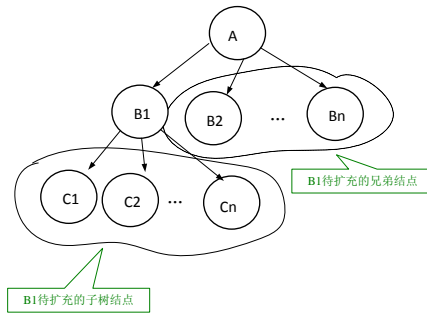


Figure 2. Expanse Query

Querying results returned for each query string are measured by degree of similarity between query and the texts of retrieval result. The texts of the retrieved result are expressed as VSM (Vector Space Model). After word segmentation and the removal of stop words, improved TF/IDF method [19] is used to compute the weight of each feature. If TF is term frequency, IDF is inverse document frequency, weight between feature t and document d is computed as:

$$w_{t,d} = \frac{tf_{t,d} \times \log(n/n_t + 0.01)}{\sqrt{\sum_{t_i \in d} [(tf_{t_i,d} \times \log(n/n_{t_i} + 0.01))]^2}} \quad (1)$$

$tf_{td}$  represents the number of occurrences of feature t in document d, n represents the number of total texts in search results,  $n_t$  represents the number of text that includes feature t.

Degree of similarity between query and the texts of retrieval result is calculated using cosine theorem. Suppose the  $i$ th a query string  $d_i = \langle w_{i1}, w_{i2}, \dots, w_{im} \rangle$ , the  $j$ th text  $d_j = \langle w_{j1}, w_{j2}, \dots, w_{jm} \rangle$ , then similarity  $S(d_i, d_j)$  is:

$$S(d_i, d_j) = \frac{\sum_{k=1}^m (w_{ik} \times w_{jk})}{\sqrt{(\sum_{k=1}^m w_{ik}^2)(\sum_{k=1}^m w_{jk}^2)}} \quad (2)$$

Retrieval result is measured through precision MP, recall MR and M-Measure (MF):

$$MP = (TP / (TP + FP) + TN / (TN + FN)) / 2 \quad (3)$$

$$MR = (TP / (TP + FN) + TN / (TN + FP)) / 2 \quad (4)$$

$$MF = (MP * MR * 2) / (MP + MR) \quad (5)$$

We use macro-precision (MP), macro-recall (MR) and macro-f1 (MF) with (3) - (5) to compute the classification performance metrics. Let TP, FP, TN and FN be the true positive, false positive, true negative and false negative respectively.

The fifth stage is ontology evolution. After the initial version of the ontology is built, we can evaluate and debug the ontology by using it in problem-solving. We may need to revise the initial ontology. This process of iterative design will likely continue through the entire lifecycle of the ontology. We also need to add new conceptions into the ontology. On the other hand, we need to add instances into the ontology. The method of ontology evolution includes automatic evolution, semi-automatic evolution and manual evolution. Manual evolution is to add and edit the ontology manually. This method is more time-consuming, but it can bring high rate of accuracy. Automatic evolution often gets help from machine learning, concept extraction and knowledge discovery technology. Semi-automatic combines manually adding and editing with techniques of automatic machine recognition.

The sixth stage is document filling. Related processes information including requirement analysis, important concept and relations among concepts should be logged from the beginning to end during ontology construction and construction document should be filled.

#### **4. Model Design of Ancient Architecture Ontology Knowledge Base**

The knowledge background of Chinese ancient buildings belongs to Subject of Architecture. Now, Domestic construction for building ontology research focuses on the protection of ancient buildings technology [11], mainly from the perspective protection of ancient buildings to get concept of field - related property definitions. These works focus on building construction period, the geographical location, description of the current protection status and a description of the level of protection. Thus, it cannot fully express the building scene and its cultural connotation.

Building components are the elements that constitute the building, just like product components. The main components in the buildings are: root, wall, pillar, base *etc.* China's ancient architecture is made up of various components skillfully whose roof style and material can be divided into many components. From the original society, in the long history, ancient buildings come down in one continuous line with its peculiar structure way, and the composition of the building components in the same historical period style is basically the same. Ontology has a strong power of expression in semantic and knowledge acquisition, thus can express ancient architecture.

According to the method proposed above, with the building component as knowledge unit, we design the main frame of domain ontology knowledge base of the component-based ancient building. The ontology has been divided into several sub-ontologies, and developed by different groups. Finally, these sub-ontologies are integrated into a targeted ontology.

The following gives a detailed description of the construction process of the component-based ontology knowledge base.

##### **4.1. Requirement Analysis**

Detailed requirement analysis is the key to the success of ontology building. The ontology that we build will be used to improve performance of information retrieval and other information processing systems. It will also be used to offer queries and management of the Multi - cultural relics.

It is required that we can retrieve the information of the single member in the ancient building ontology and the information of the members of the designed ancient building. It is also required that the ontology possesses certain reasoning capability.

## 4.2. Framework Design

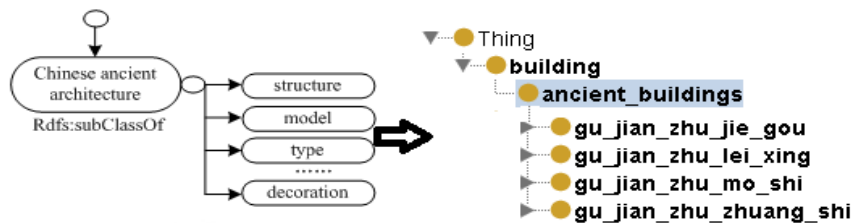
By studying books and literature about ancient buildings such as Dictionary of Ancient Chinese architecture literature and searching large amounts of information provided on Internet by domain experts, we have acquired rich knowledge of the specified domain knowledge. Domain ontology terms of ancient building include structure, type, structural member, feature, material, decorate, cultural background and historic significance and so on. Structural Members include roof, wall, pedestal, pillar, handrail, pavement, tile, beam, color painting, dougong, sparrow brace, ceiling, door, window, and other concepts. By determining the importance of domain ontology, a core set of concepts are set up [12]. For these terms, we follow the De facto standard expression. Taking into account the convenience of information processing, when the term corresponding to the concept of information processing, the name of concept will use English word or Chinese pinyin. Table 1 lists part of them.

**Table 1. Representation of Concept**

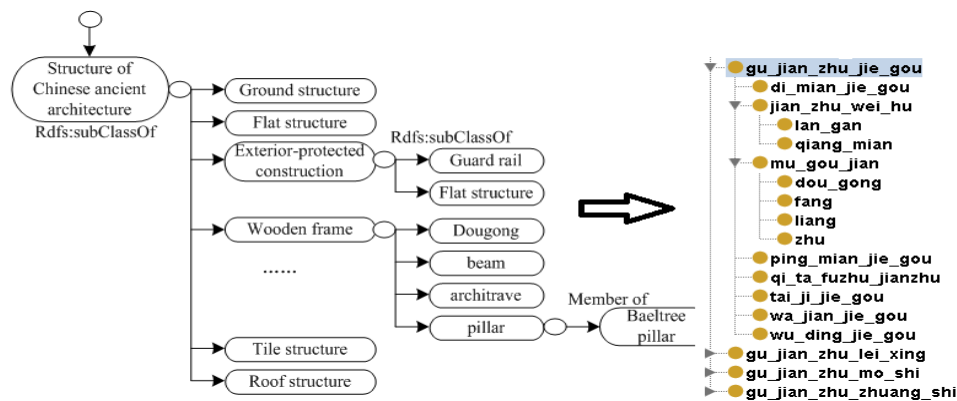
Term	Expression	Representation
Palace	Palace construction, one of ancient architecture	gong_shi_jian_zhu
Mausoleum	Tomb construction, one of ancient architecture	ling_mu_jian_zhu
Temple	Religious construction, one of ancient architecture	miao_yu_jian_zhu
Bridge	Road & bridge construction, one of ancient architecture	qiao_liang_jian_zhu
Temple pagoda	Religious construction, one of ancient architecture	si_ta_jian_zhu

Domain ontology is a subclass of the top ontology (“thing”). It is required to consider the relationship between the domain ontology and the top ontology. The building ontology is a subclass of thing and ancient ontology is a subclass of the building.

General framework for Ontology of ancient architecture is divided into three different levels according to level of abstraction to the study of ancient architecture [12]: feature level, structural level and instance level. Feature level describes the structure, model, type, indoor and outdoor decorative of Chinese ancient architectures and structural relations among all kinds of buildings. For the convenience of description, these features are described separately. Structural level describes every structural member and relations among them. Instance level describes feature of structural member of instance and relations associated with the structure level. Figure 3 and Figure 5 is a diagram of the three levels.

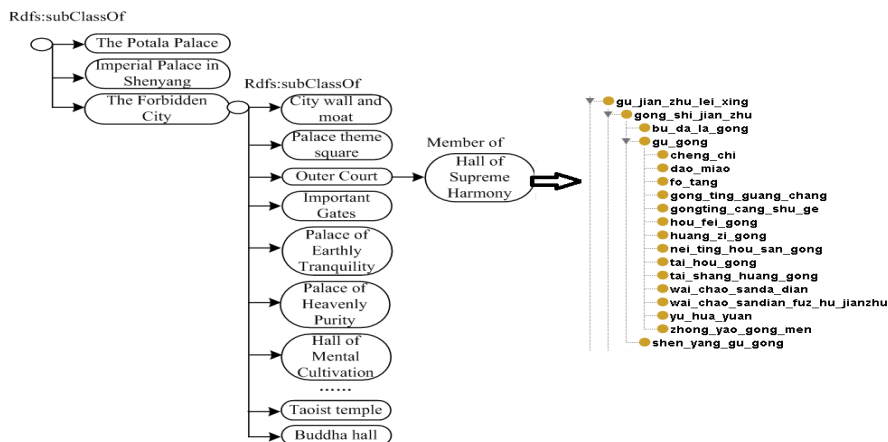


**Figure 3. Feature Level Diagram**



**Figure 4. Structural Level Diagram**

Building types of Chinese ancient building are: palace, dwelling, pagodas, bridge, tomb, city *etc.* For example, in the building process of pagodas, starting from individual component, we should refer to the structural of the other domain ontology. When building classifications, we should consider the universality in the pagodas domain, enhance the scalability and sharing. We divide the pagodas into four parts: tower\_location, tower\_adornment, tower\_model and tower\_structure. The location of the tower is according to the province of classification. The adornments of the tower include stele, fresco, Buddha, furnish and decorate, horizontal inscribed board and the outdoor display. The building structure of the tower is classified and described from all parts of the tower. Building model of the tower is the core of the tower.



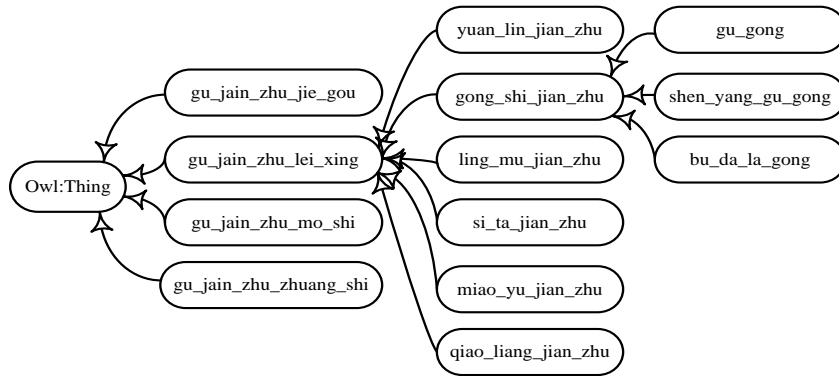
**Figure 5. Instance Level Diagram**

### 4.3. Coding

The most widely used ontology building tools is Protégé. It is developed by Stanford University. Its description language of ontology is OWL. In accordance with the architectural discipline classification, this paper implements ontology description of ancient architecture using Protégé. It takes the ancient architecture as root node, and extends down to the top-level class. Every top-level has its own subclass. Figure 6 is a diagram showing the class level.

An example of the OWL description is shown as follows. In Chinese information processing, information content is Chinese. The method we use is to express the ID with Chinese pinyin or English, but with the comment in Chinese.





**Figure 6. Hierarchical Structure of Ontology**

```

<owl:Class rdf:ID="nei_yan_zhuang_shi">
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    内檐装饰</rdfs:comment>
<rdfs:subClassOf>
    <owl:Class rdf:ID="gu_jian_zhu_zhuang_shi"/>
</rdfs:subClassOf>
</owl:Class>
    
```

#### 4.4. Ontology Evaluation

It is a complicated project to evaluate the constructed ontology. This paper applies domain ontology of ancient architecture to information retrieval system of ancient architecture. We use it to expand query, and evaluate whether the concepts of ontology are correct and whether the relations are clear and correct through results of information retrieval before and after using the ontology.

We crawl and clear up 120 relevant pages with ancient architecture and 80 irrelevant but easy to confuse pages from the website by search engines. Using 10 query strings and querying respectively to the traditional retrieval algorithm and query expansion algorithm based on domain ontology, we compute the results through statistical analysis. By averaging eventually, we get the result of the experiment, as shown in Table 2.

**Table 2. Analysis of Result**

	Ontology-based	Conventional query
MR(Recall ratio)	65%	50%
MP(Precision ratio)	58%	39%

We can see from the Table 2, query expansion algorithm based on domain ontology can increase the accuracy and recall of queries. A simple example is that when querying “故宫” (The Imperial Palace), we can also get the information of “紫禁城” (alias of The Imperial Palace). Using the rule of reference in OWL, we can get more complex information to expand the query. For example, when we query “the building that has horsehead walls”, we not only get the result of “Hui style architecture”, but also get the information of “Jiangxi sent building”.

The experiment also shows that the concepts of ontology are correct and the relations are clear and correct.

#### 4.5. Ontology Evolution

Domain ontology evolution is an important part of the domain ontology construction. There are so many domain concepts in the ancient architecture area and some of them are

quite different from modern lexicon. The same building or structural member may have different names during different periods and in different places. For example, the current "National Palace Museum" is known as "the Forbidden City" during Ming and Qing dynasty, and ornament "Wen" is called "Big Wen" in the north or "Lingwei" in the south. There haven't had a very complete books or dictionary yet which collects entirely the ancient architecture domain knowledge, so we need to improve constantly on the basis of the core ontology. Since different types of ancient buildings are not likely to fill in the short term, we need to constantly add later.

Emerging concepts and relationships also need to be added into the ontology under supervision of the experts. For instance adding, we introduce the use of automatic information extraction based on Baidu (a Chinese search engine service provider) encyclopedia and other network resources.

#### 4.6. Document Generation

The process of document generation runs through the whole process of ontology construction. The generated documents during the ancient ontology building process include: ancient ontology building requirements specifications, ancient ontology building designing reports, implementation and evaluation reports, development reports of ancient ontology construction. In the design reports of ontology, there are concept tables, overall framework of ontology model, definitions of object properties, definitions of data, definition of relations among concepts, illustration of model, hierarchy among and so on.

### 5. Conclusions

It is a cumbersome and meaningful activity to construct an ontology base. It not only needs certain theory, but also needs to abide by some standards and norms. Suitable methods are convenient for constructing ontology in different field because of the different characteristics.

This paper studies construction method of domain ontology knowledge base based on structural members and applies it to the information process to test and verify the relationship among the concepts. The method proposed is also applicable to and has a certain reference value for ontology construction in other domain.

### Acknowledgements

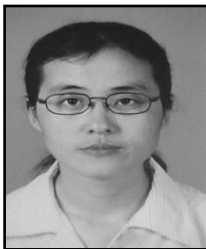
This work is supported by National Nature Science Foundation under Grant 71271209.

### References

- [1] D. Zhihong, T. Shiwei, Z. Ming, Y. Dongging and C. Jie, "Overview of Ontology", *Acta Scientiarum Naturalium, Universitatis Pekinensis*, vol. 38, no. 5, (2002), pp. 730-738.
- [2] D. Xiaoyong, M. Wenfeng and W. Wenjuan, "Construction and Evolution of Discipline Domain Ontology –A Case Study for Economics Domain Ontology", *New Technology of Library and Information Service*, vol. 3, (2007), pp. 7-12.
- [3] Zhou X., Wu Z. and Yin A., "Ontology development for unified traditional Chinese medical language system", *Journal of Artificial Intelligence in Medicine*, vol. 32, no. 1, (2004), pp. 15-27.
- [4] H. Keqing, H. Fei and L. Bing, "Research on Service-Oriented Ontology & Meta-Modeling Theory and Methodology", *Chinese Journal of Computers*, vol. 28, no. 4, (2005), pp. 524-533.
- [5] C. L. Jinzhi, "Constructing Virtual Domain Ontologies Based on Domain Knowledge Reuse", *Journal of Software*, vol. 14, no. 3, (2003), pp. 350-355.
- [6] J. N. Zhi, "Ontology-Based Requirements Elicitation", *Chinese Journey of Computers*, vol. 5, no. 23, (2000), pp. 486-490.
- [7] C. G. Cao, Q. Z. Feng and Y. Gao, "Process in the Development of Nation Knowledge Infrastructure", *Computer Science and Technology*, vol. 5, (1993), pp. 523-534.
- [8] M. Gawich, A. Badr, A. Hegazy, H. Ismail, "A Methodology for Ontology Building", *International Journal of Computer Applications*, vol. 156, no. 2, (2012).

- [9] A. De Nicola, M. Missikoff and R. Navigli, "A software engineering approach to ontology building", *Information Systems*, vol. 134, no. 2, (2009), pp. 258-275.
- [10] T. Xuan, D. Xiaoyong and L. Haihua, "Computing Term-Concept Association in Semantic-Based Query Expansion", *Journal of Software*, vol. 19, no. 8, (2008), pp. 2043-2053.
- [11] L. Qifeng and C. Bingguo, "On ODM-Based Design Approach for Ontology-Based Knowledge Base of Ancient Buildings Protection Technology", *Computer Applications and Software*, vol. 27, no. 4, (2010), pp. 654-659.
- [12] S. Yu and Q. Liping, "Ontology Model of Ancient Chinese Architecture Members and Culture Design and Application", *Journal of Beijing University of Civil Engineering and Architecture*, vol. 28, no. 4, (2012), pp. 48-51.
- [13] W. Youkui and J. Yuying, "Knowledge Discover Based on Knowledge Element", Xi'an Electronic Science & Technology University Press, Shaanxi, (2011).
- [14] G. Antoniou and F. V. Harmelen, "Web Ontology Language: OWL, Edited S. Staab and R. Studer, Handbook on Ontologies in Information Systems", Springer-Verlag, (2003), pp. 76-92.
- [15] M. Schliefnig, "The Semantic Web Trust Layer: Legal and Social Effects", *Frisch WS 2008 G6/F Information Society & E-Governance: Digital Identity & Privacy Developments*; Vienna, Austria, January 20, (2009).
- [16] Fluit C., Sabou M. and Harmelen, "Ontology based information visualization", Vladimir Geroimenko, Chaomei Chen, Editors. *Proceedings of Visualizing the Semantic Web*, Springer Verlag, (2002), pp. 546-554.
- [17] J. Bezivin, "A M3-Neutral infrastructure for bridging model engineering and ontology", *Proceedings of the First International Conference on Interoperability of Enterprise Software and Applications*, Geneva, Switzerland, February 23-25, (2005).
- [18] R. Q. Lu and S. M. Zhang, "PANGU-an Agent-oriented Knowledge Base", *Proceeding of Conference on Intelligent Information Processing*, Beijing, China, August 21-25 (2000).
- [19] W. Lidong and Q. Liping, "Exploiting, Exploiting RFC2828 as a Domain Vocabulary for Identifying IT Security Literature", *Proceedings of Fifth International Conference on Information Assurance and Security*, Xi'An, China, August 18-20, (2009).
- [20] Hownet knowledge. [http://www.keenage.com/html/c\\_index.html](http://www.keenage.com/html/c_index.html).
- [21] Smith M. L., Welty C., Deborah McGuinness L. "OWL Web Ontology Language Guide", <http://www.w3.org/TR/2004/REC-owl-guide-20040210/>.

## Authors



**Qian Liping**, born in 1971, Ph.D. candidate. Her current research interests include knowledge Engineering, intelligent information processing.

**Yang Xiao-Ping**, born in 1956, professor and Ph.D. supervisor. His current research interests include information system engineering, knowledge engineering.

**Song Yu**, born in 1983. Her current research interests include information retrieval, information processing.

