Constructing Digital Archive of Architectural Material with ontology

Norio TOGIYA Akira BABA

University of Tokyo, Japan

http://www.chi.iii.u-tojyo.ac.jp

Abstract

We developed an ontology concerning the knowledge information described in the architectural documents, and created archives based on these conceptual systems. By conceptualizing knowledge information within the architecture documents using the ontology, it becomes easier to clarify the relationship among documents. In the end, we have found that digital archives using ontology technique is extremely useful for user. In this paper, we describe about the how to construct archive and use in digital archive.

Keywords: ontology, digital archive, architectural material

1. Introduction

We are in the process of establishing digital archives for digitizing and utilizing documents related to residential architecture built in Japan in the 1920s.The term digital archive specifically means a “content system that stores and uses digitized cultural resources and heritage” Architectural documents digitally archived include many technical terms, and a great many names of related persons are described in such documents. To this end, an archive could contain documents in which the content may be difficult to understand unless the relationship among such terms, individual names, and organizations is defined. To solve the problem, we decided to build ontology related to the knowledge and information in the architectural documents and subsequently established an archive based on the conceptual system of the ontology. Ontology-based techniques have been widely introduced in industry, medical services, and legislation, but few examples have been applied to historical documents to date. Under such a background, a digital archive was established and subsequent considerations were given in this study while focusing on the following important

points of discussion:

1. In order to unify technical and historic terms, dose the ontology function as intended in the archive?
2. What problems exist in terms of ontology in an archive that handles historical documents, and what issues should be identified in order to solve any problems?
3. For an archive of architectural documents, what problems exist and what issues need to be addressed?
4. Dose the use of ontology also work effectively in an archive that handles other historical documents?

We will establish and implement a digital archive, while focusing on the above-stated points as major issues to be discussed.

2. Construction ontology

2.1. Content of material

We will primarily carry out digital archiving of the following construction documents. (1) Construction documents: A. construction diagrams, B. estimates, C. bills, D. receipts, E. other construction documents; (2) Photos and images (3) the survey drawing that we made in 2004. These documents, especially construction documents, contain many technical terms and a variety of organizations and persons' names concerning the construction. However, if we do not understand the relationship among these documents, it is difficult to understand importance of the documents and we cannot appreciate their relationship with other documents. Consequently, it is necessary to systematize how the terms can be divided into different concept categories. For this purpose, the project team isolated all the terms that were used in the documents, analyzed knowledge information within the documents, and developed ontology of the construction documents. Ontology means to systematize a variety of knowledge information based on class structure. At present, it is often used to structuralize each concept and term systematically in a variety of fields such as law, industry, and

medicine. However, there are few cases where ontology techniques have been applied to historical documents. The project team decided to establish their ontology by focusing on the construction documents of the Tsuboi house.

2.2 constructing ontology

For constructing ontology related to architectural documents of the former residence of Tsuboi, the project team employed and implemented the following steps:

1: Select words used in the documents.

2: Analyze the words and the background meaning and concepts.

3: Order the underlying conceptual system of the architectural documents into a hierarchy based on the considerations made in Step 2 above to establish ontology.

First, we made a start at Step 1 above. Step 1 is to select the terms used in the architectural documents first and grasp what terms are used. After finishing Step 1, we examined the concepts used in the architectural documents. The examination revealed that the basic ontology classes can be roughly classified into information on places, information on structure of the residence, information on persons and organizations, information on behaviors, information on things, and information on units and values.

The scope of information on places and open spaces includes information on positions within the residence, the addresses of the construction companies in Tokyo, etc. which are described in the architectural documents. The information on structure of the residence means the respective rooms, such as the kitchen, the bedroom, the reception room, etc. The information on persons and organizations includes concepts on names and titles of the construction companies, carpenters, craftsmen, etc. involved in the construction. The information on behaviors includes concepts related to such behaviors as clerical work related

to construction, including the variety of work in involved in construction projects, estimation, placing orders, etc. concerning. The information on things includes the concept of architectural materials, tools required for construction, etc. Finally, the scope of the information on units and values includes information and concepts related to the numbers and prices of materials, tools, etc. used for construction.

The ontology was established by using these six major pillars. The ontology thus established are shown as a tree-structured overall diagram, wherein, according to each concept, the definition, the dominant/subdominant conception, terms belonging to the concept, etc. were stated. The tree was named the Conceptual Tree as a function of the archive.

2.3 Features of Ontology in this Archive

For constructing ontology concerning architectural documents of the Former Residence of Tsuboi, one of the differences from other fields is that the object handles historical materials. Therefore, in some examples, archaic terms may be used for similar concepts, or the same terms may be used for different concepts. To solve the problem, we decided to create a tree-structured diagram according to concepts first and then associate the terms and images that belong to the respective concepts. With such an arrangement, terms whose use differs according to the times, geographical conditions, and differences in assigned groups could be unified into one concept. As stated above, regarding historical documents that contain many terms whose use differs according to the time, geography, and assigned group, unification of concepts would be important.

3. Designing Functions by Using Ontology

3.1 Linkage to the concept tree

After constructing the ontology, we linked concept tree and corresponding material.

The architecture documents that we access from graphic interface or text searches display republication data along with graphics of digitalized historical documents. The republication data is the full text of all the contents of the architecture documents and it is described in XML format. Each term is linked with each class on the ontology tree. Consequently, when a user clicks on the link for each term, an ontology map will be displayed in another window and the class tree diagram will be displayed to which each term belongs as an instance. On these tree diagrams, the other terms that belong to the same class and the names of the documents in which these terms are appeared will be displayed. Consequently, we can find terms with close meanings and the other documents including these terms on these tree diagrams.

3.2 Creation of Diagram showing Relationship among Documents

After establishing the association between the conceptual tree and the documents, we created a diagram showing the relationship among documents based on the major classifications of the conceptual tree **(Fig. 1)** . First, concerning the information on places and open spaces, documents that related to respective places are shown on the residence's floor plan, which changed according to the age of housing. Further, in the information on persons and organizations, the names of persons concerned are defined in the conceptual tree in terms of the respective human relationships among persons who placed orders, persons who received orders, primary receivers, secondary receivers, and so on. To this end, these relationships are graphically shown in chronological order, and documents based on those

persons concerned can be indicated. In addition, for the information on behaviors, the behaviors related to specific issues by respective persons and organizations could be indicated in chronological order.

3.3 Concept search

On the other hand, besides a visible diagram, we make sure to search each document even if we carry out text-based search. In particular, even in the search by text, we devise a search method taking advantage of the ontology tree. In an ontology search, the concept tree of five basic concepts appears in the component box **(Fig. 2)**. Even if we do not know the actual terms described in the document , selecting the concept category among them in accordance with our necessary degree of depth, we can search on the concept level. Because construction documents contain many technical terms and synonyms, we make sure that we can carry out a search without knowing the exact terms by searching appropriate documents on the concept level. We can construct AND search among categories. By combining categories, we can make a more accurate search. For example, in an abstract search such as searching documents that deal with “fixtures” in a “kitchen” that “carpenters” engaged in the construction in “1932,” we select “kitchen” in the component box of place and space, “1932” in the “information about units and figures” component box, “carpenter” in the “information about people and organizations” component box, and “fixture” in the “information about things” component box. Then, we can carry out a search for appropriate multiple documents.

4. Conclusion

As stated above, the establishment of an architectural archive utilizing ontology was completed. When considering the points of discussion as referred to earlier from what has

been stated in the above, the following can be obtained:

First, regarding the effectiveness of ontology in establishing the present digital archive, the following points were confirmed:

(1) Unification of terms with different meanings was possible;

(2) Relationships among documents were visualized on the basis of the concept of ontology;

(3) Accessing the respective documents was possible simply by choosing a concept without knowledge of the terms used.

From the above, utilization of ontology contributes to elimination of various barriers concerning terms and to an improved environment in which a user can access the desired documents more easily. For this purpose, for the present archive that handles various technical and historical documents, establishment of ontology is very helpful.

Next, concerning problems in constructing an ontology that handles historical terms, since terms are used in various ways according to the time and geography/place, it is necessary to set up a concept that covers more than the time and geographical conditions, and unification of terms should be made in a manner that associates the concepts. For historical documents, usage of a concept differs according to differences in the times, but even in the contemporary times, many cases exist where usage differs in different places. Consequently, it is important for an ontology handling historical terms to eliminate the dissociation between concepts and terms for the above-stated two conditions. In addition, regarding the establishment of ontology concerning historical materials for architecture, architectural documents describe not only the materials and construction methods concerning the architecture, but also various places, relationships among persons and organizations, and the exchange of money and objects.

Therefore, in an archive of architectural documents, it is necessary to expand our scope and establish a concept in a more comprehensive range based on the classification of places, structure of residence, persons and organizations, behaviors, things, etc. The six classifications used in the present archive play the role of basic pillars that unify the different types of information and are helpful in constructing ontology for other architectural archives.

Finally, in an archive that handles historical documents and professional areas, as well as documents from different fields, the establishment of ontology would be very effective in unifying various terms. In particular, in a digital archive that is digitized and released to the public on the Internet, etc., it is assumed that not only specialist users, but also general users from various fields would benefit. To this end, the establishment of ontology is very important and necessary to allow users without professional knowledge to access the desired documents. Based on the considerations stated above, introduction of ontology is a very helpful tool in a digital archive. However, to demonstrate the effectiveness of such ontology, ontology should further be applied to various other resources in order to generate more case studies.

Fig. 1: Diagram showing relationship among documents

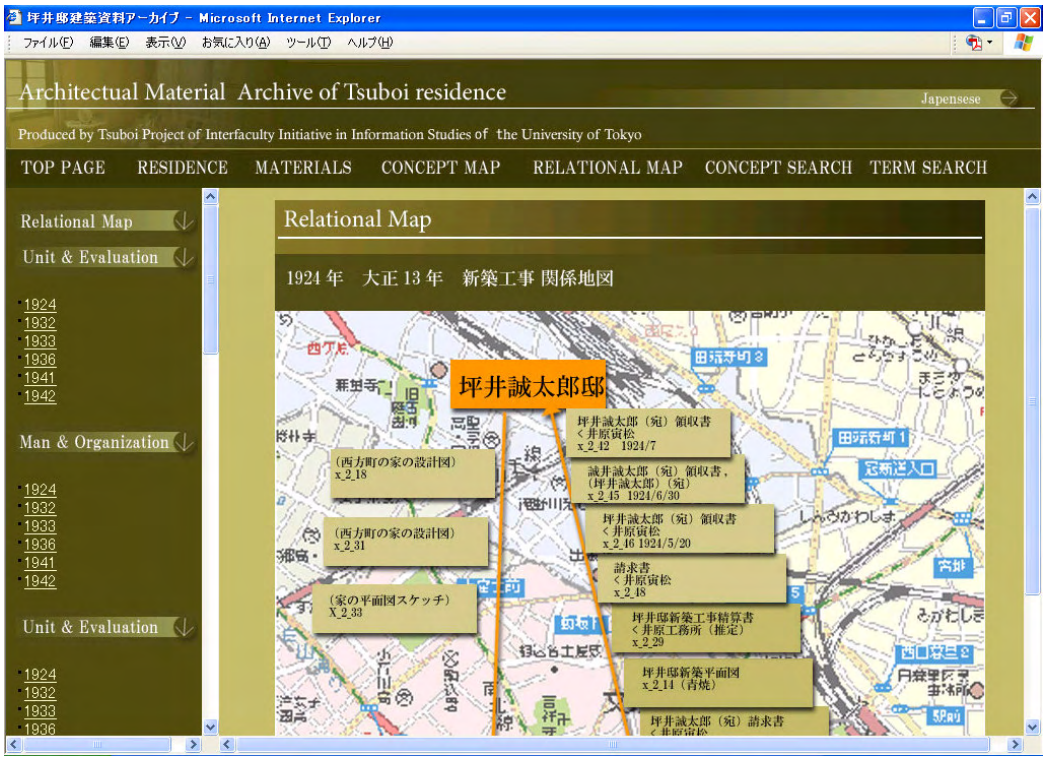

Fig. 2: Concept search window

References

Sowa,J .(2000). Knowledge representation;logical,Philosophical and computational foundation . Boston: PWS Publishing Comp.

Staab,Steffen .(2004). Handbook on Ontologies (International Handbooks on Information Systems) . New York: Springer.

Gomez-Perez, Asuncion, Corcho,Oscar, Fernandez-Lopez, Mariano.(2004). Ontological Engineering : with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web . New York: Springer.

Handschuh,Siegfried . Staab, Steffen (Editor).(2003).Annotation for the Semantic Web. Ios Pr Inc.

Mizoguchi,Riichiro. (1995). Ontology for Modeling the World from Problem Solving Perspectives, *Proc.of IJCAI-95 Workshop on Basic Ontological Issues In knowledge Sharing,* 1-12.

Chasndrasekaran, B.,J.R.Josephson ,R.menjamins (1999). Why are ontologies, and why do we need them?. *IEEE Intelligent System.20-26*

Author's Biography

Norio TOGIYA is a doctor course student of the Graduate School. of Interdisciplinary Information Studies , University of Tokyo. He was born in 1974 in Yokohama city in Japan.. In 2002 he entered the graduate school of University of Tokyo. He has studied about archives and museums informatics.
e-mail:togiya.norio@iii.u-tokyo.ac.jp
url:http://lifetheque.web.infoseek.co.jp/togiya/

Akira BABA is a professor of the graduate school of Interfaculty Initiative in Information Studies and Interdisciplinary Information Studies, the University of Tokyo. In 1958 He was born in Ibaraki prefecture in Japan. He graduated from Waseda university graduate school of letters. After graduation, he became an assistant professor of University of Tokyo, associate professor, and professor successively. His specialties are economic history, science history, historical information studies, digital archival science and game producing study.
e-mail: baba@iii.u-tokyo.ac.jp
url: http://www.chi.iii.u-tojyo.ac.jp