Prospect of Establishing a Database of Nonwritten Cultural Materials Based on Ontological Principles

NOTO Masato and KINOSHITA Hirotsugu

1. Introduction

Writing has played an immense role in the annals of science and civilization. Since time immemorial, humans have gained knowledge from a plethora of information and passed it on from one generation to another. The advent of writing allowed humans to document and disseminate knowledge, making letters the lifeblood of civilization. Meanwhile, the birth of the computer opened up a new way to use knowledge and allowed people to analyze abstract concepts without arbitrary interpretations. Now, the worldwide expansion of the Web is generating new ways to organize knowledge.

Computer capacity and information networks have been expanding so fast that it is becoming the norm to distribute and acquire knowledge via the Internet. As a result, engineering studies of knowledge have moved into the spotlight. As human intelligence has underpinned scientific progress, technology to accumulate knowledge, as well as its implications that cannot be represented by symbols alone, is instrumental especially in a knowledge system for cultural science. Today, researchers pursue ontological studies to systematize knowledge encompassed in various fields [1].

Needless to say, it is not easy to accumulate and share nonwritten cultural resources, such as material on folk culture, through conventional ontological approaches. Therefore, by defining the resources of folk cultural research that cannot be described by existing means as “nonwritten cultural materials” based on ontological principles, and by expanding concepts into suitable forms for ontological databases on nonwritten cultural materials, we may be able to establish the ideal database for such materials. Moreover, nonwritten cultural materials exist in myriad forms in the field of cultural anthropology, as researchers have recorded them in their own ways. Consequently, we face two major obstacles to sharing and distributing nonwritten materials:

- The relation between one material and another cannot be characterized clearly.
- Due to poor interoperability, these materials are difficult to exchange among researchers.

The key to clearing these hurdles is to explicate implicit elements by linking diverse concepts.

Based on ontological principles, this study reveals conceptual properties essential for an ontological definition of nonwritten cultural materials. As a step toward achieving the sharing and distribution of nonwritten materials, we aim to handle these properties — which have been treated in an ad hoc and unstructured manner — in a systematic, categorical framework.

2. Nonwritten Cultural Materials

2.1 Definition and Structure of Nonwritten Materials

"Nonwritten Cultural Materials" in this report is intended to cover the wealth of folk culture that has been handed down without ever having been recorded in writing. The nonwritten history of human culture includes areas as diverse as ideologies, knowledge, wisdom, and the activities of men, and is quite incomparable to written history. In the history of folk culture, nonwritten cultural materials are resources on human life that have never been recorded in writing, and which have been handed down from generation to generation. Research methods in folk studies include anything from hearings and fieldwork to studies of written records, architecture, daily commodities, and folk tradition. In some cases, the fields of research closely overlap with history, cultural anthropology, sociology, and religious studies, so that one research extends over two or more disciplines.

Moreover, many researchers personally keep and control their own research materials personally, so that in order to allow a smooth distribution of research materials, standardizing the methods to organize and itemize them is of critical importance. It is, however, difficult to "standardize" studies in folk culture, as research is mostly speculative, and methods are idiosyncratic and atypical. Standardizing research materials should thus be avoided, as there is a risk of homogenizing folk studies, which could cause an overall deterioration in research quality. On the other hand, it is true that with the Internet now widely used, information and research materials on folk studies should be made public, and in doing so, information technology must be introduced to offer systematic support for researchers. Solving such conflicting needs concurrently is a major issue.

Based on this background, a research environment where a database on nonwritten cultural materials is constructed and put out on the Internet is necessary, in order to broaden the base of research by offering objective measures for analysis.

2.2 Metadata

Metadata is the "data of data", and the Dublin Core [2] is the international standard of rules in defining metadata. The Dublin Core is applied on the assumption that the information resources of the target can be fixed on to existing media, such as library catalogues. We, however, believe that the nonwritten cultural materials mentioned in this report cannot be fully defined by the Dublin Core due to their idiosyncratic nature. The data stored in nonwritten cultural materials include important information, comprised of metadata as well as the data itself. Nonwritten cultural materials include numerous bits of data which are incompatible with the Dublin Core; the "true information" of the research target can only be determined by interpreting the relationship between metadata and the "true data", making them an important source of research on folk culture studies.

One problem in systematizing nonwritten cultural materials is that, in the field of folk studies, information from resources is not currently being shared or disseminated because research materials that have been accumulated and made into current databases are owned and controlled individually. In order to systematize the information stored in nonwritten cultural materials, the targets and fields of research in folk culture studies

must first be digitized and distributed on the Web. By constructing an ontology that explicitly defines and expresses the concept of the target world, the sharing and dissemination of nonwritten cultural materials will be made possible. With nonwritten cultural materials, various concepts can coexist within one target object, because the data characteristics of nonwritten cultural materials are naturally open to various interpretations, based on the social, historical, and cultural backgrounds of the person in charge of digitizing each concept. It is therefore important to establish an ontological architecture in order to organize the conceptual system. By constructing ontology, the sharing and dissemination of common research targets will be made possible among researchers.

3. Ontology

3.1 Basic Theory

Lately, ontological research is frequently seen in the field of knowledge processing. Ontology means describing the process of conceptualization; by constructing an ontology of the target knowledge system, we may expect a great leap forward in the sharing and dissemination of knowledge obtained from the target concepts. “Ontology” is originally a term from philosophy, meaning “a systematic approach to the study of existence”, but in information engineering, it is defined as “the explicit definition of concepts”[3]. Of the several definitions of “ontology”, its essence as used in information engineering is as follows: “That which not only exists as a target, but is also the actual result of the study of an existing target, and that which explicitly shows the concepts and relationships brought to light, and provides a clear semantic definition.” When modeling after the real world with computers, we are actually acknowledging the existence of “the existence of concepts” in the real world, and extracting the concept itself, or “conceptualizing”. When doing so, the concept is characterized and the meaning of the concept is unconsciously assimilated by identifying the difference from, and relationship to, other concepts. Normally, the explicit describing of an unconscious and implicit process of “conceptualization” such as this is called ontology. In this study, we have noted the role of ontology in the sharing and dissemination of nonwritten cultural materials, with the understanding that it is the condition where “previously implicit knowledge and basic concepts of ontology is segmented and systematized.”

3.2 Top-Level Ontology

Ontology usually consists of the definitions of concepts, taxonomy, and the relationship between concepts. Concepts are most often categorized in a hierarchy. Top-level ontology is the descriptive component that commonly constitutes ontology of the target world, and has a high level of generality and abstraction [4, 5]. As the upper level of the conceptual hierarchy, basic concepts such as substance, event, time, space, concept, relationship, and attributes are categorized and given meaning. Top-level ontology is constructed from basic relational principles such as the identity of an object, and the definition of “parts” as opposed to the “whole”.

In literature [6], top-level ontology is categorized into the following three groups, and provides 12 basic categories made from the combinations of their subcategories.

（1）Category 1: "Physical", "Abstract"

（2）Category 2: "Independent", "Relative", "Mediating"

（3）Category 3: "Continuant", "Occurrent"

3.3 Concepts in the Establishment of Ontology

Ontology is defined as describing the process of conceptualization. An implicit process of conceptualization is necessary to describe a target, and established ontology shows the relationship between the conceptualized targets. Concepts in the establishment of ontology are as follows:

- Generation of ontology from an arbitrary point of view

  Commonality and consensus are qualities necessary in ontology. Ontology needs to be created based on consensus, but on the other hand, this very idea serves to limit the scope of its individuality and specifity. Therefore, with targets where information and objectives are diversified, it is useful to create several ontological structures generated from an arbitrary point of view.
- Integration and conversion among ontological structures

  When conceptualizing a target with several objectives in mind, different ontological structures are established from a single target. Upon creation of such ontological structures, integration and conversion shall take place between them as necessary, according to each objective, and researchers shall provide further information. Therefore, the integration and conversion among ontological structures needs to be fully discussed as an important issue in identifying, clarifying, and defining of its target.

4. Establishing the Ontology of Nonwritten Cultural Materials

4.1 Conceptual System of Folk Studies in Categorizing Folk Implements

In the field of folk studies, folk implements differ according to the region and culture of their origin, and various concepts coexist for the target concerned. In regard to categorization methods, numerous systems based on different objectives and intentions may coexist for the same target. Table 1 shows a part of the categorization system. "Categorization by application" is otherwise known as the consumer classification system, and is based on the efficacy and purpose for consumers. On the other hand, "categorization by function" is a classification system often used as the standard in folk study statistics. The difference of categorization systems concerning a single target is a result of different processes of conceptualization, depending on which attribute the objective or target focuses on.

Target	Application	Function
Bowl	Japanese tableware	Container for solid food
Plate	Western tableware	Container for liquid food
Chopsticks	Japanese tableware	Carries food to mouth
Spoon	Western tableware	Carries food to mouth

Table 1: Categorization by "Application" and "Function"

4.2 Ontological Structure in the Categorization of Collected Folk Implements

In the study of folk culture, research targets are a collection of concepts. In defining the target, conceptualization needs to be carried out implicitly, to indicate the relationship between the target concepts, which are then systemized into ontological structures. In other words, the overall concept in ontology is the concept of the entire target world, whereas the relational concept of ontology is the concept of relationships between different target worlds.

It is important to systematize the collection method of folk tools, which is one exemplary area of research in folk culture, in order to organize and understand the underlying concept. In folk studies, the “bowl”, a simple object, is made up of two concepts: it is “a container of food”, and at the same time, it “carries food to the mouth”. Therefore, the concept of a “spoon” is also comparable here. In other words, within the category of a bowl, there are two semantic concepts. Such discrepancies in the recognition of concepts hold an important meaning within the category of folk implements in folk studies.

As mentioned before, a single target may have multiple concepts. The systematization of ontology is established by the defining of the concept, and the target knowledge is shared and disseminated by the creation of ontology. Figure 1 shows the categorization of folk implements.

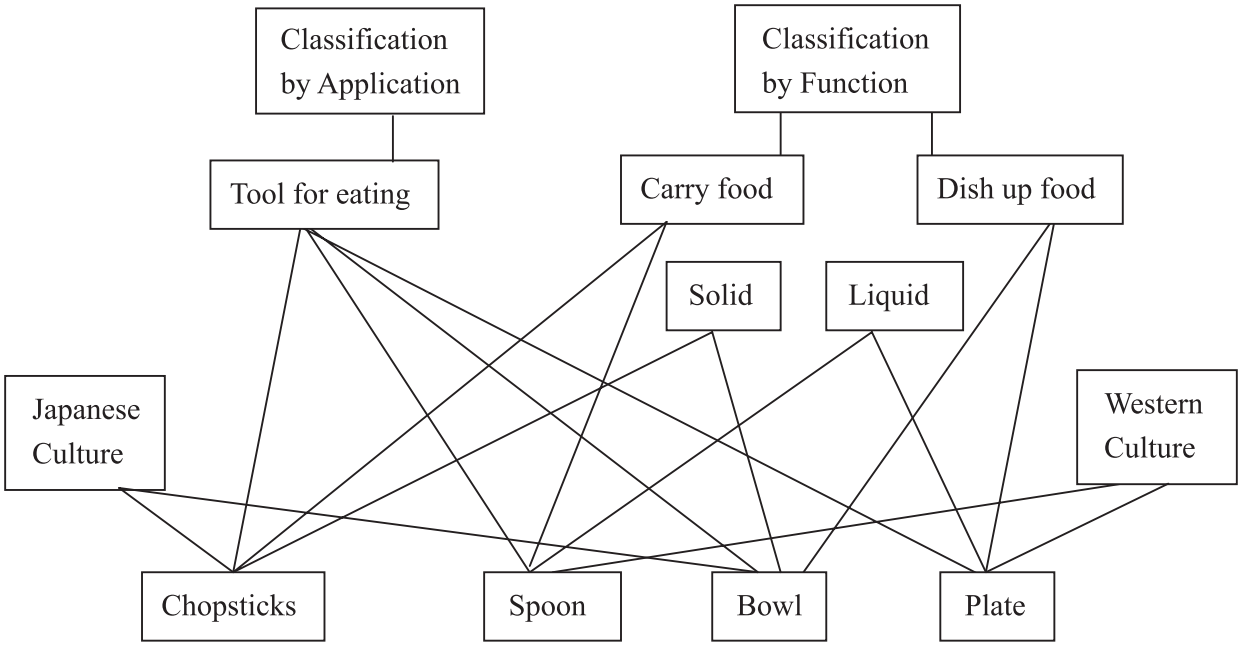

Figure 1: Categorization of Folk Implements

The ontological structure was created by expanding the traditional top-level ontology of the categorization of folk implements. Figure 2 shows the ontological architecture of the collection of folk implements. The “application” and “function” seen in the following figure correspond to the same words as shown in Figure 1.

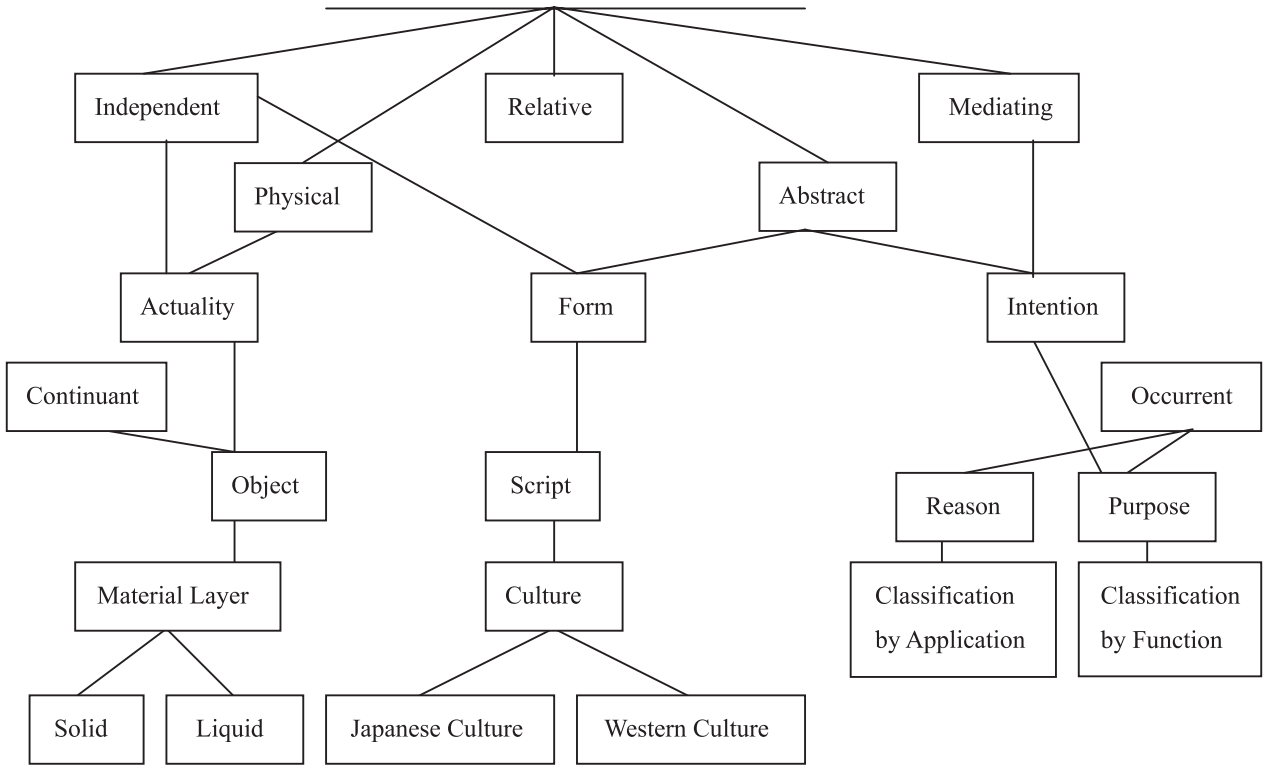

Figure 2: Ontological Structure in the Collection of Folk Implements

By establishing an ontological structure using the top-level ontology based on basic ontological principles, the following points have been made clear in the categorization method of folk implements.

- For one concept (bowl), multiple semantics consisting of two upper concepts (dish up food, carry food) may be assumed.
- The assumption of multiple semantics derived from the original concept (bowl) may include a totally different category (spoon).

5. Discussion

By refining the two systems of "application" and "function" in the categorization method of the collection of folk implements in folk culture, a top-level ontology was created based on basic ontological principles. Each definition was refined and classified based on discrepancies in conceptual characteristics. By introducing these concepts, the dissemination of conceptual definitions, which had previously gone through specialized processes, can now be treated in an explicit and versatile framework. The manifestation of the concepts, along with the organization and systematization of the target, is certain to contribute to the improvement of the sharing and dissemination of knowledge through ontology.

From the viewpoint of information engineering ontology, clarifying the structure of the conceptual system and the relationship between multiple systems are important when we envisage the common understanding of the target through the coexistence of information technology and men. Such concepts, which are the fundamental concepts of the target world, are normally presumed to be shared implicitly. Ontology is the

explicit definition of implicit concepts that underlie the target world, and works to improve the sharing and dissemination of knowledge. These implicit concepts show the knowledge and intentions concerning the target world. The target world of nonwritten cultural materials is unique to folk studies, which derives "true information" through reasoning of the multiple descriptions of relations between metadata and "true data". Multiple domain concepts of the target world do exist, but they should be set apart from the target area and the contents that need to be defined, since the necessary concepts would naturally be different if the nature and scope of the target are at variance. Also, even when regarding a single target, the contents may differ depending on the extent of description required for a particular concept. The following are items for discussion in defining the target when creating ontology suited for nonwritten cultural materials.

（1）Collection of specialized knowledge
（2）Analysis and systematization of specialized knowledge

The "collection of specialized knowledge" means to collect the necessary specialized knowledge necessary for the ontology about to be created, and to clarify the concepts that are to be defined in ontology. Numerous concepts coexist in the collection of expert knowledge, but the concepts defined in ontology differ according to the target range. Therefore, concepts need to be extracted from the myriad of collected specialized knowledge to suit the objectives. The "analysis and systematization of specialized knowledge" analyzes the contents of the collected information, and the extracted concepts are organized systematically. In order to make use of the ontology that has been defined through the process mentioned above, an ontological structure is indispensable, and will contribute greatly to the sharing and dissemination of the target world.

6. Conclusions

In this report, nonwritten cultural materials were defined, a brief summary of ontology was given, and examples of nonwritten cultural materials were used to show that ontology is effective in the sharing and dissemination of knowledge through establishing ontological structures. Even with the same target world, multiple ontological structures exist according to different objectives. Therefore, we need to accumulate and share existing ontological structures, and create new ones. The newly created ontological structures may be used to describe ontology-based knowledge. To describe knowledge based on ontology, it is necessary to create models based on ontology-described background information. In order to do so, it is important to share the contents and use of the information（patrimony of knowledge and wisdom）, instead of sharing formalized and standardized information. The sharing and dissemination of nonwritten cultural materials is the sharing and dissemination of knowledge itself, with very important stakes in handing down culture, and therefore has a recognized need for further research efforts. It would be a pleasure if this report serves to present useful information on folk culture research and to raise awareness on these issues.

Acknowledgement

We wish to thank Keiko Kinoshita, a student at Kanagawa University, Graduate School of Engineering , Course of Electrical, Electronics and Information Engineering, for her assistance in this research.

References

1 T. R. Gruber: A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition*, Vol. 5, No. 2, pp. 199-220 (1993).

2 http://www.dublincore.org/.

3 R. Mizoguchi: Ontological Engineering, *Ohmsha* (2005).

4 J. F. Sowa: Distinction, Combination, and Constraints, *Proc. of IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing* (1995).

5 J. Top and H. Akkermans: Computational and Physical Causality, *Proc. of 12th International Conference on Artificial Intelligence*, pp. 1158-1163 (1991).

6 J. F. Sowa: Knowledge Representation: Logic, Philosophical, and Computational Foundations, *Brooks/Cole* (2000).