*The 8th International Conference on e-Business (iNCEB2009)*
*October 28th-30th, 2009*

# APPLYING ONTOLOGY AND METADATA FOR TOURISM INFORMATION

*Patitta Suksomboon Garcia, Atchara Leeraphong, Nanthiya Ariyapichai*

Information and Technology Management Department
Faculty of Commerce and Management
Prince of Songkla University
Muang, Trang 92000 Thailand
{patitta.s, atchara.l, nanthiya.a}@psu.ac.th

## ABSTRACT

The WWW is one of the greatest repositories of information which is available to anyone at anytime with the Internet connection. This advantage has attracted plenty of users, including tourists who look for travel information. However, the more the amount of information grows, the right and proper information is needed. This paper presents tourism information representation architecture to represent tourism information that can be used for tourists to search for tourism packages which is relevant to their needs, and also for tourism providers who propose their tourism packages through the Internet. Ontology is used to represent the tourist instance while metadata describes tourism packages. We try to combine the advantages of using ontology and metadata to represent tourism information; as well as tourists and tourism providers' feedbacks; this will lead to improve the efficiency and precision of tourism information retrieval system.

***Index Terms*** — Ontologies, Metadata, Tourism information

## 1. INTRODUCTION

Recently, due to the continued increasing number of travelers, the vastitude of information such as the tourist attractions and much kind of activities in tourism packages are posted on the Internet to attract the tourists. Nevertheless, it is not comfortable for a tourist to search the exact information that really wants from this huge information from such airlines, hoteliers, car rental companies and travel agencies, which are available on the Internet. Because this information is constructed with unstructured data on web; consequently, if there is tourism ontology to help both travel agencies and tourists for organizing that information, it is certainly swift the speed of tourist's getting the traveling information.

Definition of ontology is a conceptualization of a domain into a human understandable, machine-readable format consists of entireties, attributes, relationships, and axioms. It is also with the characteristic of the reusability, which makes it very attractive and powerful for representing domain knowledge [2].

The related ontology applications involve in many research fields. Especially, tourism ontology-based systems provided intelligent matching in order to facilitate semantic matching between variable tourism sites and tourist keyword, a specific vocabulary of the tourism domain such as user type, time range and traveling place is needed.

Since Semantic Web contributes exponential benefits to this industry by not only supplying lower constraints (access, time, and location) with flexible mean for exchanging diverse information, but also accelerating users searching process, with multiple intelligent services: incorporative and advised information. Hence, it is challenging to assess Semantic Web performance and capability with the emerging applications in e-Tourism area. In addition to humans, Semantic Web is used to make the web information more understandable and useful to computer applications.

The tourism ontology provides a way to achieve integration and interoperability through the use of a shared vocabulary and meanings for terms with respect to other terms. The e-Tourism ontology was developed using OWL (Web Ontology Language) [11]. OWL was proposed by the W3C for publishing and sharing data, and automating data understanding by computers using ontologies on the web. OWL was planned and designed to provide a language that can be used for applications that need to understand the meaning of information instead of parsing data for display purposes.

While tourism participant groups created a number of tourism ontologies, they can hardly fulfill this goal because of their mainly focus on limited domain concepts. There are some efforts to find an alternative approach for converting the semantic space of tourism through the

integration of modularized ontologies, such as user, W3C Time or W3C Geo , that center around a core domain ontology for the tourism sector [11].

This paper is organized as follows: Section 2 briefly summarizes the related work emphasize on tourism ontology. Section 3 describes about metadata which apply in various domains. Section 4 roughly sketches the system architecture. Section 5 concludes the paper. Finally, Section 6 presents our further work.

## 2. RELATED WORK

Different working groups have developed disparate ontologies for supporting their segregate tourism information utilization through the webs. In addition, some ontologies are hosted by industry group, while others are created within specific collaborative projects or academia. In general, these ontologies share some common imperfections, including the limitation of vocabulary in tourism concepts, the disintegration of existing domain-independent ontologies, and the variety of concepts among them. This results in the interoperability problems [11]. As a consequence, several efforts have been made to establish international standards in order to streamline the information exchange process among heterogeneous data sources [7].

Ontology is the technology used in Semantic Web to describe formally a domain of discourse. In general, ontology consists of a finite list of terms and relationships between these terms. The terms denote importance concepts (classes of objects) of the domain. Ontology provides a common vocabulary to support the sharing and reuse of knowledge [13].

Ontologies are commonly used to construct knowledge bases [6] and have been proposed as a tool for marking up data on the Semantic Web [12]. An ontology specifies a conceptualization of a domain in terms of concepts, attributes, and relations [3]. The concepts are typically organized into a taxonomy tree where each node represents a concept and each concept is a specialization of its parent.

Next, we will briefly summarize key tourism ontologies which have been actively used by different groups. We shall see the root causes of problem in this area.

### 2.1 The OnTour and Mondeca Tourism Ontologies

The OnTour [9] ontology was developed under the DERI's project which concentrates on accommodation and infrastructure domains. This ontology allows users to inquire their designated packages through a tourism portal in that it returns users the relevant accommodation and infrastructure. In addition to these concepts, it supplies more information about leisure activities and geographical locations.

Unlike the OnTour's, Mondeca ontology integrates leisure activities definition in detail. Several dimensions are included such as tourism object profiling, tourism and cultural objects, tourism packages, and tourism multimedia contents [9].

However, both ontologies are similar in that they include tourism domains from the WTO (World Tourism Organization) thesaurus, an international ontology standard. Nevertheless, the OnTour project was developed with a wide range of terms; sometimes it is too broad whereas other times it is deep [14].

### 2.2 The QALL-ME ontology and TISCOVER

The QALL-ME ontology, which was funded by EU, permits users to query tourism objects in multiple natural languages. On the one hand, users can input their questions into the system by various means; and it later returns a group of related answers. With 122 classes and 107 properties, it answers the questions in such destinations, events, and transportation [11].

Like the QALL-ME's, TISCOVER, the largest Austrian web-based tourism platform, enables users to inquire about accommodation with natural languages [14].

### 2.3 The Harmonize Ontology

The Harmonize ontology was established under the Harmonize project using RDF. With an ontology-based mediation concept, distinguished tourism organizations can exchange data without changing any local data structures and information systems. Additionally, local-diverse ontologies will be mapped into central single ontology. Semantic annotation is used to insert in web pages' tags, and this would account for a B2B integration and e-Tourism collaboration. As a consequence, an interoperability issue for e-Tourism communication among isolated ontologies in this industry can be resolved.

### 2.4 cDOTT - The Mosaic Ontology Model

Since there is no existing tourism ontology that could deliver enough information for trip planning, as well as changes during the trip. The Mosaic Ontology was recently constructed to connect different modular ontologies together by linking them to central core domain ontology, cDOTT. The cDOTT concept is to provide common vocabulary of tourism sector and can be extended its usage by other tourism ontologies [11]. Modularized domain-independent ontologies (time, weather, geo-spatial, user, and currency) and fine-grained ontologoies (gastronomy, hotel, transportation, and event) are located surrounding the inner model. Including of these ontologies is reasonably chosen based on purpose of usage.

## 3. METADATA

Metadata is data about data. It creates a new representation where it contains meta-information that usually does not appear in the original resource, that is, metadata about the original information (data) [4]. Metadata have been used widely in many domains, for example, in e-Learning. Learning materials associated with metadata facilitates the interoperability between learning object repositories. The e-learning community has seen fruitful initiatives in the standardization of learning object metadata by IEEE called LOM (Learning Object Metadata) [8] and SCORM [1] which proposed the emerging educational specification for learning content.

In our work, we try to describe tourism package with metadata. Each package can be described its attributes or specifications such as, *package name, destination, departure time, departure place, trip duration or tourism provider's name.* These metadata is useful for tourist to specify their search or make advance searching to retrieve their proper tourism packages.

## 4. SYSTEM OVERVIEW

This section presents the components of the system. These consist of tourism information from agencies and tourism experts which are describes by metadata and represented in ontology-based. The information retrieval system processes users' travel requirement and returns related tourism information to tourists. Figure 1 depicts our proposed system architecture.

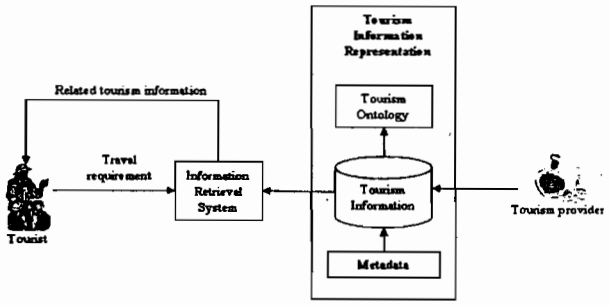


Figure 1: System overview

The tourist puts their requirement through user interface for eliciting the travel's characteristics. The interface poses queries to elicit information concerning the travel information.

The information retrieval system searches for tourism places or tourism instances that are relevant user's queries. Information retrieval system tries to identify tourism information that can cover all parts of the travel's characteristics from user input.

The tourism information representation module describes tourism information by metadata and has linked to tourism ontology.

In [5] defined the class or concept of the ontology is travel and a class hierarchy is an IS-A hierarchy expressed in OWL in our system we use OnTour [3] which defined class hierarchy for the e-Tourism ontology.

The tourism information representation architecture is shown in figure 2. There are two levels of tourism information; 1.) Tourism instances and 2.) Tourism packages.

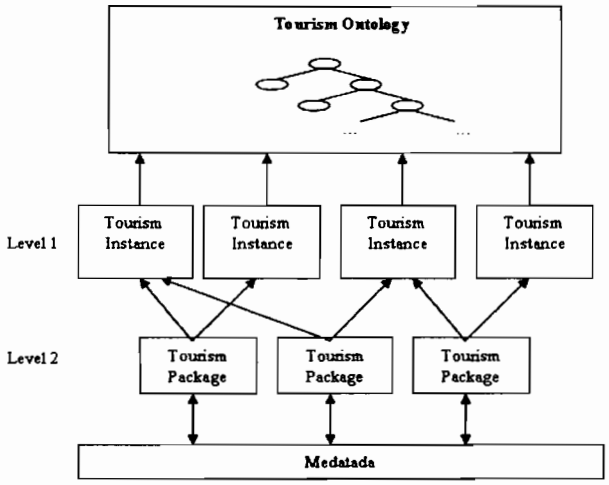


Figure 2: Tourism information representation architecture

Level 1; Tourism instance represents the tourist places. It can be the name of the place or the name of accommodation. Each tourism instance links to the node in tourism ontology.

Level 2; Tourism package is described by metadata. It can be composed of many tourism instances. Tourism packages are different depend on the tourism provider. Some tourism packages propose only plane ticket and accommodation while some propose organized trip include local transportation, meals or entrance tickets.

## 5. CONCLUSIONS

Considerate types of tourism information are unstructured provided in the Internet, and this posts significant barrier to both tourists and tourism providers. For tourists, it would be difficult to find relevant information in an effective mean, in terms of being time and cost obstacles. For tourism providers, many participants, such as airlines, hoteliers, car rental companies, and travel agencies, could not supply satisfactory tourism packages based on tourist's interests and preferences.

Even many existing ontologies have been actively used in this industry, there is no single alternative to solve

these incurred problems in both tourists and tourism providers' perspectives.

This paper proposes tourism information representation architecture which is relied on tourism ontology, the OnTour version 8, and described by tourism metadata. First, to solve the mentioned problems, disseminated tourism providers' information shall be standardized and integrated in a centralized database. Second, to understand this information, metadata is efficiently used to explain tourism packages such as package name, destination, departure time, departure place, trip duration or tourism provider's name. In addition to this explanation, an interrelation among each tourism instance would be described by tourism ontology.

With our proposed architecture, tourists' demand would be satisfied by closed to relevant tourism packages elicited from our system based on their characteristics. At the same time, it would encourage tourism providers' revenue by increasing a number of sold tourism packages and acknowledging which is a good or bad one. In another word, learning the tourists' feedback would help tourism organizations to agilely develop adaptive tourism packages that are best fit to tourist customers; which, as a result, generate a sustainable income even the dynamic change in tourist demands and external factors of the industry.

## 6. FURTHER WORK

Future work concerns the integration of information retrieval system with ontology in order to improve tourism information system. We will conduct system evaluation from both tourists and tourism providers. For tourist aspect, inputs gathered from tourists would assure that our system provides the results cover their needs. For tourism providers, tourist requirement, such as tourism package categories, will be included in this process, in order to guarantee our system effectively functions as their revenue generation channel. After that we might use the result from the evaluation process to modify the OnTour ontology using OWL for system implementation. Consequently, to solve the problem which a tourist can not search rapidly the exact and precision information that really wants from the numerous travel information. Moreover, not only it can help travel providers offer their tourism packages over the Internet but also it facilitates tourists to-select the retrieved information that they need from the Internet.

## 7. REFERENCES

[1] A. Merceron, C. Oliveira, M. Scholl, and C. Ullrich, "Mining for Content Re-use and Exchange – Solutions and Problems", Poster in Proceedings of the 3rd International Semantic Web Conference, ISWC2004, Hiroshima, Japan, pp. 39-40, 2004.

[2] C.S. Lee, Y.C. Chang, and M.H. Wang, "Ontological recommendation multi-agent for Tainan City travel," *Expert Systems with Applications*, Elsevier Ltd, pp. 6740-6753, Volume 36, Issue 3, Part 2, April 2009.

[3] D. Fensel, "Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce", Springer-Verlag, 2001.

[4] H. Song, L. Zhong, H. Wang, R. Li, and X. Hongxia, "Constructing an Ontology for Web-based Education Resource Library," Poster In the Proceeding of International Workshop on Applications of Semantic Web Technologies for e-Learning (SW-EL@ K-CAP'05), Banff, Canada. October 2-5, 2005.

[5] H.W. Toby and S.T. Lee Raymons, "iJADE FreeWalker: An Ontology-Based Tourist Guiding System," KES 2006, Part II, LNAI 4252, Springer-Verlag Berlin Heidelberg, pp. 644-651, 2006.

[6] J. Broekstra, M. Klein, S. Decker, D. Fensel, F. Van Harmelen, and I. Horrocks, "Enabling Knowledge Representation on the Web by extending RDF schema", *Proceedings of the Tenth international World Wide Web Conference*, pp. 467- 478, 2001.

[7] K. Prantner, Y. Ding, M. Luger, Z. Yan, and C. Herzog, "Tourism Ontology and Semantic Management System: State-of-the-arts Analysis".

[8] IEEE Learning Technology Standards Committee, "LOM," Available: http://ieeeltsc.wordpress.com [accessed August 2, 2009].

[9] STI Innsbruck, "OnTour," Available: http://www.deri.at/research/projects/eTourism [accessed July 5, 2009].

[10] W3C, "OWL Web ontology Language," Available: http://www.w3.org/TR/owl-ref/ [accessed June 7, 2009].

[11] R. Barta, C. Feilmayr, B. Proll, C. Grun, and H. Werthner, "Covering the Semantic Space of Tourism – An Approach based on Modularized Ontologies," *Proceedings of the 1st Workshop on Context, Information and Ontologies*, Heraklion, Greece, June 1, 2009.

[12] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web", Scientific American, pp.35-43, May 2001.

[13] W. Chunhua, Z. Pengfei, and W. Cong, "A comprehensive Tourism Information System", *Proceeding of The 6th IEEE International Conference on Industrial Informatics*, pp. 720-724, Daejeon, Korea, 2006.

[14] W. Siricharoen, "Using Ontologies for e-Tourism," *Proceedings of the 4th WSEAS/IASME International Conference on Engineering Education*, Agios Nikolaos, Crete, Greece, July 24-26, 2007.