Using Ontologies for E-tourism

WARALAK V. SIRICHAROEN
Computer Science Department
University of the Thai Chamber of Commerce
126/1 Vipavadee-Rangsit Rd., Bangkok
THAILAND
http://www.waralak-lak.blogspot.com

Abstract: The usual business nowadays has been recommended to change to digital business using the Internet and e-commerce. E-tourism/e-travel Software adapted from original e-commerce, ready for creating instantly online reservation/booking. It is the total solution of instant e-commerce for travel 24 hour. E-tourism is a perfect candidate for Semantic Web because it is information-based and depends on the World Wide Web, both as a means of marketing and transaction channel. Ontologies can assist organization, browsing, parametric search, and in general, more intelligent access to online information and services. Ontology-based information retrieval makes it possible to handle the known challenges in connection with Web-based information systems in a more efficient way. This paper discusses some ontological trends that support the growing domain of online tourism. The first part introduces the e-tourism implementation and preview of e-tourism in Thailand. The second part describes fundamental background of ontologies; definition, modeling, language and etc. The third part discusses the example concepts of existing e-tourism using ontologies and it is a briefly summary on the important e-tourism ontologies. Finally, the last part is the conclusion.

Key-Words: - Ontologies, E-Commerce, E-tourism, Semantic Web, Travel Ontology, Trip Ontology

1 Introduction

"The study of e-commerce in the tourism industry has emerged as a 'frontier area' for information technology [1]. Tourism Information Systems are a new form of business systems that provide and support e-tourism and e-travel organizations, such as airlines, hoteliers, car rental companies, leisure suppliers, and travel agencies. One class of these systems relies on travel related information sources, such as Web sites, to create tourism products and services. Tourism is an information-intensive sector and therefore has significantly benefited from ICTs. ICTs enable tourism businesses to make tourism products and services directly available to a large number of consumers, and to interact with them as well as with other tourism producers and distributors. The increasing number of consumers who use the Internet to plan leisure or business trips represents a major incentive for developing countries to organize and develop their tourism supply and its promotion over the Internet [2]. Information distribution and interactions are the key backbones of travel industry, which is currently mainly based on the printed brochures, posters, advertisements via television or limited web access. While Semantic Web will bring the revolution to this area by not only exponentially extending the dissemination and exchange channels with unlimited access, unlimited time and unlimited locations, but also assisting users with smart information searching, integrating, recommending and various intelligent services. Therefore, e-tourism is the one of the decent application areas for Semantic Web technologies and it is also a good test-bed to prove the efficiency and utility of Semantic Web technologies [3].

An ontology can be constructed for e-tourism. Tourism is a data rich domain. Data is stored in many hundreds of data sources and many of these sources need to be used in concert during the development of tourism information systems. The e-tourism ontology provides a way of viewing the world of tourism. It organizes tourism related information and concepts. The e-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). OWL was proposed by the W3C for publishing and sharing data, and automating data understanding by computers using ontologies on the Web. OWL is being planned and designed to provide a language that can be used for applications that need to understand the meaning of information instead of just parsing data for display purposes.

2 Semantic Web and Ontologies

Objective of the Semantic Web is to make the information on the Web understandable and useful

to computer applications in addition to humans. The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. The Semantic Web promises a solution to this problem: Semantically annotated websites can not only be understood by the human reader but also by machines. Enriching websites with machine-readable semantics will enable more intelligent and efficient searching and further processing of data without requiring the human user to interfere.

Ontology is a formal explicit description of concepts in a domain of discourse (classes/concepts). Slots/properties/roles Properties of each concept describing various features and attributes of the concept. And facets/role restrictions means restrictions on slots. An ontology together with a set of individual instances of classes constitutes a knowledge base. In reality, there is a fine line where the ontology ends and the knowledge base begins. Classes are the focus of most ontologies. Classes describe concepts in the domain [3]. A class can have subclasses that represent concepts that are more specific than the superclass. Slots describe properties of classes and instances.

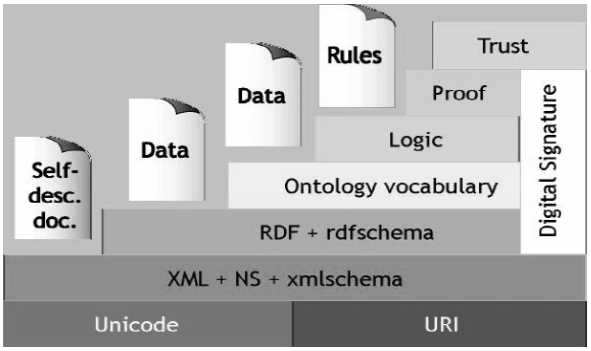

Fig.1 The Semantic Web architecture [5]

From the Fig.1 above showing, at the bottom layer XML (eXtensible Markup Language [6]) a language that lets one write structured web documents with a user-defined vocabulary. XML is particularly suitable for sending documents across the Web, thus supporting syntactic interoperability. XML allows users to add arbitrary structure to documents without saying what these structures mean. RDF (Resources Description Framework: [7]) allows meaning to be specified between objects on the Web and was intentionally designed as a metadata modeling language. RDF is a basic data model for writing simple statements about web objects (travel-related web resources). The RDF data model does not rely on XML, but RDF has an XML-based syntax. RDF Schema provides modeling primitives for organizing web objects into hierarchies. RDF Schema is based on RDF and can be viewed as a primitive language for writing travel-related information ontologies. But there is a need for more powerful ontology languages that expand RDF Schema and allow the representation of more complex relationships between web objects. Ontology languages, such as OWL (Web Ontology Language [8]) by WWW Consortium (W3C) are built on the top of RDF and RDF Schema. The logic layer is used to enhance the ontology language further, and to allow writing application-specific declarative knowledge. The proof layer involves the actual deductive process, as well as the representation of proofs in web languages and proof validation. Finally, trust will emerge through the use of digital signatures, and other kind of knowledge, based on recommendations by software agents we trust, or rating and certification agencies and consumer bodies.

Two important technologies for developing the Semantic Web are XML and RDF. Also now, the OWL (Web Ontology Language) is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics. As following is the example of OWL used for describing travel system ontologies according to class and property hierarchy from DAML (DARPA Agent Markup Language) ontology library [9]. These examples uploaded by BBN Trip Report Ontology [10] and Travel Itinerary Report Ontology [11] shown as following.

Example1: Trip Report Ontology :

Class Hierarchy

- *Expense (amount, date)*
- *Trip (airfare*, car*, end, hotel*, parking*, start, total)*

Property Hierarchy

- *airfare, amount, car, date, end, hotel, parking, start, total*

Example 2:Travel Itinerary Report Ontology :

Class Hierarchy

- *Aircraft ()*
  - *instance A300*
  - *instance A340 …*
- *Business ()*
- *Class ()*
- *Coach ()*
- *First ()*
- *Flight (aircraft, airline, arrive, class, depart, destination, duration, flight, meal*, miles, origin, seat*)*

- *HotelReservation (address, checkin, checkout, confirmation, hotelName, rate, smoking?)*
- *Itinerary (flight*, hotel*, passenger, rentalCar*, rln*)*

Property Hierarchy

- *address, aircraft, airline, arrive, checkin, checkout, class, confirmation, depart, destination, duration, flight, hotel, hotelName, meal, miles, origin, passenger, rate, rentalCar, rln, rloc, seat, smoking*

3 Ontologies for e-tourism

Usually, there is several tourism ontologies were considered for reuse, before considering built the new ontology. In e-tourism different ontologies have been developed for different areas. However, sometimes in different countries or regions around the world, the existing ontologies might not meets the needs to describe regional distinctions for any specific areas. An international standard is the thesaurus on Tourism & Leisure Activities of the World Tourism Organization (WTO). It is a very extensive collection of terms related to the area of tourism. The example of e-tourism ontology developed by DERI [12], the project started with a list of terms that should be included in the ontology. On the one hand it was helpful to have a voluminous collection of terms, and on the other hand it was misleading because the broad range of terms sometimes led to too detailed concepts, which had to be taken out in a later stage of the development [12]. However, after identified relevant parts of the WTO thesaurus according to the categories that had chosen before. Then ontology had expanded by adding relations and properties. Ontologies enhances the semantics by providing richer relationships between the terms of a vocabulary. The three major uses of ontologies are:

- *To assist in communication between humans.*
- *To achieve interoperability and communication among software systems.*
- *To improve the design and the quality of software systems.*

Together with DERI, “OnTour” [13] project came up with four main categories. The ontology’s main focus is the description of accommodations and infrastructure. The intention of a user to query a tourism portal is to find a package of relevant accommodations and infrastructure. Many tourists prefer to stay at accommodations that are as close as possible to an infrastructure where they can partake in activities. Therefore, we defined a category for activities and a category for infrastructure respectively. Additionally, a category for events was introduced because tourists might also be interested in visiting events during their stay. However, the ontology clearly focuses on accommodation and infrastructure. Additionally, more general concepts had to be specified, such as concepts to describe date and time, contact data, tickets and location. The concept location comprises not only of postal addresses but also GPS coordinates which are very useful when it comes to the computation of distances between entities. This is also a feature which is provided by the OnTour [13]. DERI is constructing an ontology for the tourism industry. The DERI e-tourism ontology [12] contains many properties to describe the defined concepts in more detail and gives information that might be relevant for searching for a package of accommodation and infrastructure. Finally, relations among classes and subclasses were specified.

Travel and tourism comprise the leading application field in business-to-consumer (B2C) e-commerce, representing approximately half of the total worldwide B2C turnover. Even in the 1960s, travel applications (i.e., computerized airline reservation systems) were at the forefront of information technology. The product is a confidence good, consumer decisions rely on information available beforehand, and the industry is highly networked, based on worldwide cooperation between stakeholders of different types. The latter factor and the related problem of interoperability represent a major challenge for IT solutions.

Another e-tourism project “Harmonise” [14], a European project based on a Semantic Web approach and utilizing a Web services infrastructure, deals with business-to-business (B2B) integration on the "information" layer by means of an ontology-based mediation. It allows tourism organizations with different data standards to exchange information seamlessly without having to change their proprietary data schemas. The "weak" coupling takes into consideration the specific industry context, with its majority of small or medium-sized enterprises (SMEs) and with many different, also legacy, solutions. Real-world business tests show that this approach meets industry expectations and can facilitate the necessary network effect in order to create an interoperable e-tourism marketplace.

The e-tourism ontology provides a way of viewing the world of tourism. It organizes tourism related information and concepts. The ontology will allow achieving interoperability through the use of a shared vocabulary and meanings for terms with respect to other terms [15]. In an early stage of Harmonise project, a partial ontology for the e-tourism was created using Protégé [16] and the OWL language. This was a very time-consuming task since it was necessary to find out information

about real tourism activities and infrastructures on the Web and feed them into the knowledge base.

Fig.2 E-tourism ontology dimensions

Parallel to OnTour work, another approach is called Hamonise [3] differs, it is objective-oriented - the ontology is able to answers four types of questions Fig.2 that can be asked when developing a dynamic package. These questions involve the predicates What, Where, When, and How. Examples of typical questions are:

- What can a tourist do while staying at the Phuket Province.
- Where are located the interesting places to see and visit.
- When can the tourist visit a particular place?
- How can the tourist get to its destination to see or do an activity?

In Harmonise project, it relies on semantic annotation [14] to resolve the differences among the data present in distinct e-tourism Web sites. Semantic annotation is the process of inserting tags in documents, whose purpose is to assign semantics to the text between the opening and closing tags. Unstructured Web pages are annotated using the e-tourism ontology. The semantically annotated pages are then stored in a knowledge database. Final dynamic package processes are created using conditional planning, ranking, and selection. The main objective of the planning is to schedule an appropriate timeframe during which the tourist will realize a particular activity referenced by a dynamic package.

The ontology using in e-tourism project, is usually written in OWL. This decision to use OWL instead of RDF was made because the ontology contains inverseOf*** properties. More important, OWL was more suitable regarding extensibility of the ontology: as OWL supports some attributes RDF does not support, such as cardinality constraints and data types, the decision to use OWL was made. The following is the example of owl ontology concerning the travel ontology. Ontologies for the Semantic Web are characterized as RDF ontologies, and are being built using OWL and other languages based on RDF. Current attention to the Semantic Web and the language standardization it offers has resulted in the single most prominent change in ontology editors since the original survey in 2002. This growth in direct support for RDF and various species of OWL has created some controversy [17]. Shared ontologies allow for different systems to come to a common understanding of the semantics of learning concepts. The present the required ontology model including the formal expression of ontology, mapping to XML representation and the corresponding system architecture for binding web services. This can see that ontology model work well together almost the same with object model because it is the fundamental of system architecture [28].

Summary Ontologies in the Travel Industry:

For the Semantic Web to succeed it will require methodologies for extracting and defining the knowledge that is to be represented. The success of the Semantic Web in the travel industry will depend heavily upon the creation of suitable ontologies. An ontology comprises the classes of entities, relations between entities and the axioms, which apply to the entities of that domain. The development of ontologies is fundamental to allow machine-supported travel-related data interpretation and integration. Furthermore, ontologies offer a promising infrastructure to cope with heterogeneous representations of travel web resources. Data heterogeneity is solved, if semantic reconciliation with respect to the travel ontology is provided between the different information systems. Travel ontologies contain knowledge about the domain of travel for developing intelligent information systems. The most important ontologies are the following ones [5]:

- The *OnTour** project [13] deployed the e-tourism ontology using OWL. This ontology was based on an international standard: the *Thesaurus on Tourism & Leisure Activities* of the World Tourism Organization [2]. It describes the domain of tourism and it focuses on accommodation and activities [35].
- In the travel industry there have been efforts in defining semantics, such as the Harmonise and Satine projects. The *Harmonize* project [14] allows participating tourism organizations to keep their proprietary data format and use ontology mediation while exchanging information. *The Satine project* developed a secure semantics-based interoperability framework for exploiting web service platforms in conjunction with P2P networks in tourism industry.

- *Mondeca's* tourism ontology [19] defines tourism concepts based on the WTO thesaurus. These concepts include terms for tourism object profiling, tourism and cultural objects, tourism packages and tourism multimedia content.
- The development of another comprehensive and precise reference ontology named *COTRIN* (Comprehensive Ontology for the Travel Industry) is described by Cardoso [20]. Major airlines, hoteliers, car rental companies, leisure suppliers, travel agencies and others will use COTRIN to bring together autonomous and heterogeneous tourism Web services, Web processes, applications, data, and components residing in distributed environments. The objective of COTRIN is the implementation of the semantic XML-based OTA (Open Travel Alliance) specifications.
- In the *LA_DMS project*, a comprehensive ontology for tourism destinations was deployed in order Destination Management Systems (DMS) to become adaptive to user's needs concerning information about tourism destinations [14].
- Some travel ontologies can be found within the *DAML language* portal shown as the example of OWL ontology implemented on previous page. The itinerary ontology represents travel itineraries and involves definitions of basic terms (e.g. aircraft, class, flight). Ontologies are created using ontology development tools, such as Protégé 2000 [16]. A Java-based ontology editor with OWL *Plugin*: that means that it allows ontology implementation as an applet on the Web. This permits multiple users to share the ontology. The *EU-IST project SWAP* [21] demonstrated that the power of P2P computing and the Semantic Web could actually be combined to share and find "knowledge" easily with low administration efforts, while participants can maintain individual views of the world. Maedche and Staab [22] analyzed the advantages of Web Semantics and P2P computing for service interoperation and discovery in the travel industry.
- The largest Austrian web-based tourism platform *Tiscover* in cooperation with developing a query interface exploiting the intuitiveness of natural language. The prototypical system allows for searching for accommodations throughout Austria via queries posed in natural language, such as "I am looking for a double room in a hotel in Innsbruck having a sauna and a swimming pool. It should furthermore provide a baby sitting service". After automatically determining the language of the query (currently English and German), the relevant terms are extracted from the query and semantically tagged [23].
- Finally, Semantic Web methodologies and tools for intra-European sustainable tourism were developed in the *Hi-Touch projec*t [24]. These tools are used to store and structure knowledge on customers' expectations and tourism products. In the near future, providers of travel-related services will advertise their services on the Semantic Web, so that intelligent software agents can find them dynamically.

These intelligent agents and more to come each day could then make suggestions on consumers; make arrangements in consideration of consumer preferences. Ontologies would allow these providers to publish metadata about their travel services and contact information.

4 Conclusions

Revolutionizing and publishing information are the most important part of tourism business. Now mostly based on brochures, advertisements via medias or limited web access. Therefore, the tourism domain is the decent area for Semantic Web technologies by supporting customers and travel agencies with information searching, integrating, recommending and many intelligent services.

Ontologies and ontology-based information retrieval have the potential to significantly improve the process of searching information on the World Wide Web. Concept search and browsing can ease the burden of searching the web using keyword-based techniques. This is especially important in information-based business, such as e-tourism. Ontologies will play an important role as they promise a shared and common understanding of traveling concepts that reaches across people and application systems. Semantic Web technology pave the way to enhanced Knowledge Management (KM) solutions that are based on semantically related knowledge pieces of varying granularity. Although early semantic web-based KM approaches and solutions have shown the benefits of ontologies and related methods, there still exist a large number of open research issues that have to be addressed in order to make Semantic Web technologies fully effective when applied to KM solutions. Delivering the Semantic Web to the travel industry depends upon: (1) the syntactical and semantic mark-up of travel content; (2) the development of better knowledge analysis and modeling tools; (3) widespread adoption of interoperable knowledge representation languages, and (4) the construction of suitable ontologies.

The main strategies to gain an accomplishment are accelerating and increasing electronic information,

building tourists' and entrepreneurs' trust and confidence in e-commerce, developing e-commerce capacity, marketing via electronic devices, and promoting the development of e-marketplace for tourism. The area of tourism especially e-tourism seems to a perfect application area for Semantic Web technologies. As information dissemination and exchange are the key backbones of travel industry, the Semantic Web can considerably improve e-tourism. Thai entrepreneurs have to adapt themselves well to suit the e-tourism environment such as serving niche markets, becoming subsidiaries of large enterprises to handle domestic business or niche products, and developing web sites to connect directly with their customers. Support from the government will accelerate the e-commerce development. The government should provide space for SMEs to set up their web sites at low cost, develop a portal for the tourism industry, provide sufficient facilities, training, and information, and resolve the macro problems such as law, tax, and security.

Reference:

[1] Turban, E., Lee, J., King, D., & Chung, H.M. (2000). *Electronic Commerce: A Managerial Perspective*. p 451.

[2] http://www.etourismnewsletter.com/ecommerce.htm. [Online]. Retrieved October 7, 2006,

[3] Cardoso, J., (2005) E-tourism: Creating Dynamic Packages using Semantic Web Processes [Online]. Retrieved October 3, 2006. www.w3.org/2005/04/FSWS/Submissions/16/paper.html

[4] http://www.w3.org/XML/

[5] Siorpaes, K., (2004). *OnTour System Design*. [Online]. Retrieved October 6, 2006. http://e-tourism.deri.at/ont/docu2004/OnTour%20-%20System%20Design.pdf.http://www.cs.vu.nl/~mcaklein/papers/oil-xmls.pdf

[6] http://www.w3.org/RDF/

[7] http://www.w3.org/OWL/

[8] www.daml.org/ontologies/

[9] http://www.daml.org/2001/06/expenses/trip-ont

[10] http://www.daml.org/2001/06/itinerary/itinerary-ont

[11] DERI, E-tourism Working Group. 2005, DERI International

[12] Fodor, O., Werthner, H., *Harmonise: A Step Toward an Interoperable E-tourism Marketplace*. International Journal of Electronic Commerce Volume 9, Number 2 / Winter 2004-5 page11 – 39. Hannes Werthner. Harmonise - A Step towards an Interoperable eTourism Market Place. http://www.itit2003.itc.it/download/werthner.pdf, 2003.

[13] Volz, R., Oberle, D., & Studer, R., (1999). *Views for light-weight web ontologies.* Proceeding of SAC 2003, Melbourne, Florida, USA.

[14] Uschold, M., Gruninger, M. (1996). *Ontologies: Principles, methods and applications*. Knowledge Engineering Review, 11(2).

[15] Protégé, Protégé. 2005, Stanford Medical Informatics.

[16] McGuinness, D. L. and Harmelen, F. V. (2004). OWL Web Ontology Language Overview. [Online]. Retrieved July 29, 2006. http://www.w3.org/TR/owl-features/.

[17] Wouters, B., Deridder, D., & Paesschen, E. V., (2000). *The use of Ontologies as a backbone for use case management.* This research was partially supported by Wang Global and the Brusseles Capital Region (CCOOS Project), Belgium, to appear.

[18] http://swap.semanticweb.org/

[19] Kiryakov, A., et al., (2005). *Semantic Annotation, Indexing, and Retriev*al. Journal of Web Semantics (Elsevier), 2(1).

[20] Cardoso, J., Sheth, A., (2006). *Semantic Web Services, Processes and Applications*. *Springer*. [Online]. Retrieved August 29, 2006. dme.uma.pt/jcardoso/Books/Semantic-Web-Services-Processes-and-Applications/Slides/Ch01.ppt

[21] Maedche, A., Staab, S. (2003). *Services on the Move: Towards P2P-Enabled Semantic*. Web Services, In Proc. of ENTER 2003, pp. 124-133, Helsinki: Springer.

[22] Dittenbach, M., (2006). *Natural Language Search (e-tourism)*. Retrieved August 30, 2006, from www.ifs.tuwien.ac.at/~mbach/research/powersearch.html

[23] Hi-Touch Working Group (2003). Semantic Web methodologies and tools for intra-European sustainable tourism [Online]. Retrieved April 6, 2004, from http://www.mondeca.com/articleJITT-hitouch-legrand.pdf/

[24] http://etourism.deri.at/ont/

[25] http://www.mondeca.com