The Survey on E-tourism and Ontologies

Dr. Waralak V Siricharoen
Computer Science Department, School of Science,
University of the Thai Chamber of Commerce, 126/1 Dindeang,
Bangkok, Thailand
E-mail: waralak_von@utcc.ac.th, lak_waralak@yahoo.com

Abstract

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. E-tourism/e-travel Software adapted from original e-commerce, ready for creating instantly online reservation/booking. The Semantic Web relies heavily on the formal ontologies that structure underlying data for the purpose of comprehensive and transportable machine understanding. Therefore, the success of the Semantic Web depends strongly on the proliferation of ontologies. Ontologies can assist organization, browsing, parametric search, and in general, more intelligent access to online information and services.

This paper is particularly interested in the new possibilities afforded by Semantic Web technology in the area of knowledge management applied to the travel industry. This paper also discusses some ontological trends that support the growing domain of online tourism. The preview of e-tourism is introduced in general. The paper also present the example concepts of existing e-tourism using ontologies display in graphical model in ontologies editor tool called Protégé and show the example of e-tourism ontologies description in OWL and RDFS syntax. The last part of the paper is a briefly summary on the e-tourism ontologies projects.

Key Words: E-tourism, Ontologies, E-commerce, Semantic Web, Travel, Tourism

1. Introduction

Tourism is a networked industry. It contain a set of interrelated businesses, involving travel companies, accommodation facilities, catering enterprises, tour operators, travel agents, providers of recreation and leisure facilities [24]. Tourism has become the world's largest trade and its development shows a stable year-to-year increase. Kim addresses in the paper [11] that competitive benefit is no longer ordinary, but increasingly driven by science, information technology and innovation. The Internet is already the major source of tourist destination information for travelers. The travel services are typically offered via the Internet and are accessible from a variety of locations, from private personal computers at home or at work to electronic kiosks and other devices in public places. 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 [6].

This is show that the tourism business mission is to revolutionaries the traditional tourism industry to next generation e-tourism powered by Semantic Web technology. It will be realized by an advanced e-tourism Semantic Web portal which will connect the customers and virtual travel agents from anywhere at anytime with any needs and requests [5].

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.

Tourism must be treated as an information intensive industry because for tourism as a service industry, information is one of the most important quality parameters to support actions [26]. 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 site as see example of demo version of Thailand's e-tourism Web Site in Figure 1. in next page, to create tourism products and services.

Figure 1. The demo version of Thailand's e-tourism Web Site

E-tourism is a perfect application area for Semantic Web technologies, since information dissemination and exchange are the key backbones of the travel industry [7]. 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 (the example of interoperability problem will be given in Figure 2.) through the use of a shared vocabulary and meanings for terms with respect to other terms [8]. 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.

The challenge to develop software package [4] for online commerce is to find a solution to cope and integrate the non-standard way of defining e-tourism products and services. There are no standards or common criteria to express transportation vehicles, leisure activities, and weather conditions when planning for a vacation package, several ways can be found among all the existing Web Sites. To deal with the lack of standard and enable data integration, it can rely on the use of ontologies and semantic annotation. Simply say that, semantic annotation is information about what entities (or, more generally, semantic features) appear in a text and where they do. Formally, semantic annotations represented a specific sort of metadata, which provides references to entities in the form of URIs (Uniform Resource Identifiers) or other types of unique identifiers [25] as see in Figure 2.

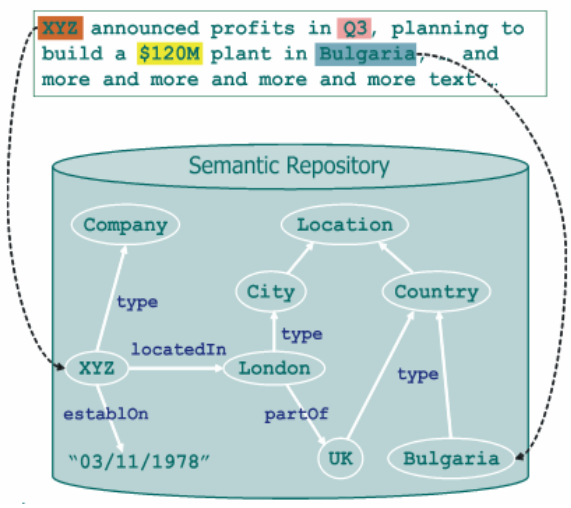

Figure 2. The example of semantic annotation [25]

As it has been recognized, the Semantic Web can considerably improve e-tourism [3]. The information from several travel, leisure, and transportation online sites, it point out the lack of standards in the tourism domain. Some of the differences founded among several sites are the following example. The street of tourism related address in travel domain is expressed in many different concepts [15] but they probably have same meaning as see in Figure 3. [24] that *Street:* can be defined as *StreetNumber* in tourism organization 1's Web Site or *Strt* in tourism organization 2's Web Site, that represented in different form what we called interoperability problem.

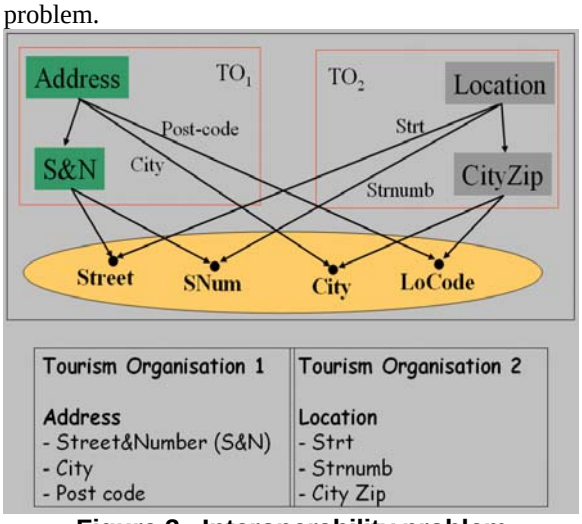

Figure 3. Interoperability problem

To finding a solution to improve on this lack of standards in the tourism field by automatically understanding the different ways of expressing

tourism products and services, extracting its relevant information and structuring. The sophisticated technologies, such as semantics and ontologies, are good candidates to enable the development of dynamic information systems [3].

2. Ontologies for E-tourism

The World Wide Web (WWW) as known it today is a huge collection of information or sometimes called WWW as the information superhighway. The number of websites on the WWW is growing daily. However, this expands of information is not as good as to an increase of functionality: information extraction has become a difficult task. Current technologies do not provide means to reuse existing information efficiently. Common search engines can perform keyword-based searches. Finding a certain part of information is very difficult. However, the number of results usually is enormous and not manageable by the human reader. For the human user it is simply impossible to go through all the websites that are delivered as results to a query.

Ontology is actually well known in philosophy research area more than 40 years. Ontology is mentioned by Tom Gruber which used to refer to "an explicit specification of a conceptualization [of a domain]. In other words, ontology refers to a formalization of the knowledge in the domain. From a technology point of view, it can be seen as a repository of concepts; much like a database represents a repository of data [15]. Ontologies enhance 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. An ontology structure holds definitions of concepts or we called classes (*Hotel, Rooms, Tourist, Accommodation, and Address*) as see in Figure 4., binary relationship between concepts and attributes. Relationships may be symmetric, transitive and have an inverse. A minimum and maximum cardinality constraint for relations and attributes may be specifies. Concepts and relationships can be arranged in two distinct generalization hierarchies [14]. Classes are the focus of most ontologies. Classes describe concepts in the domain [10]. A class can have subclasses that represent concepts that are more specific than the superclass. Slots describe properties of classes and instances. Concepts, relationship types and attribute abstract from concrete objects or value and thus describe the schema (the ontology) on the other hand concrete objects populate the concepts, concrete values instantiate the attributes of these objects and concrete relationship instantiate relationships [15].

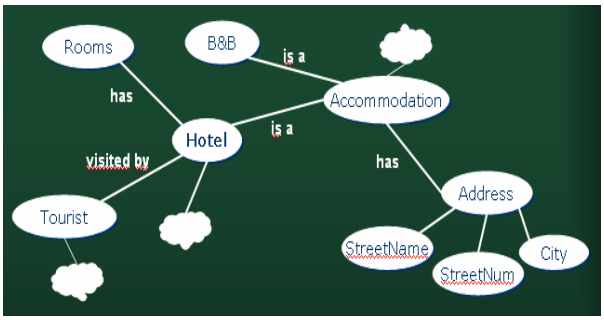

Figure 4. The e-tourism ontologies

Ontology can be constructed for e-tourism because tourism is a data rich domain. Data is stored in many hundreds of data sources and many of these sources need to be used during the development of tourism information systems. Ontologies for the Semantic Web are characterized as RDF ontologies, and are being built using OWL (Web Ontology Language) and other languages based on RDF such as DAML (DARPA Agent Markup Language), OIL (Ontology Inference Layer) and etc.

Current attention to the Semantic Web and the language standardization it offers has resulted in the single most outstanding 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 [12]. Shared ontologies allow for different systems to come to a common understanding of the semantics of 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.

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 most ontologies that can be used in the tourism context are written in DAML and OWL. 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 [9].

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 [9]. 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.

One of a good example of using ontologies for e-tourism is KIM project [25]. The semantic annotation in KIM consists in discovering named entities like persons, companies, locations etc. in the documents to annotate. Then a hyperlink is created to the relevant entity in the knowledge base and to the closer concept in the ontology. The annotations are created in a separated document. However, IBM research has developed a semantic annotation platform [2] that annotates the web at a large scale and provides embedded annotations. It is just arguing that the semantic web development will come with the availability of a huge amount of annotated web pages. The majority of the current annotation tools like OntoMat [19], SHOE [18], MnM [20], SMORE [21], Melita [22] are doing the same. Only the Annotea [23] project separates annotations from the annotated document [17].

Jakkilinki [16] also provide an overview of the development methodology and applications for tourism ontologies. Ontologies are created using ontology development tools, such as Protégé [13]. A Java-based ontology editor with OWL Plug-in: that means that it allows ontology implementation as an applet on the Web. This permits multiple users to share the ontology. The W3C has recently finalized the OWL as the standard format in which ontologies are represented online. With OWL it is possible to implement a semantic description of the travel domain by specifying its concepts and the relationships between the concepts.

Another example of e-tourism ontology see in Figure 5 developed by DERI[1], 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 large 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. 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.

The travel intelligent agents and more to come each day could then make suggestions on consumers; make arrangements in consideration of consumer preferences. For these agents, the Semantic Web infrastructure would be based on core travel ontologies that would be published on fixed URI's as OWL files. Ontologies would allow these providers to publish metadata about their travel services and contact information [13].

In Figure 5. show the basic information of this ontology, that there are 12 classes, 9 object properties, 15 data properties, and 16 individuals (instance).

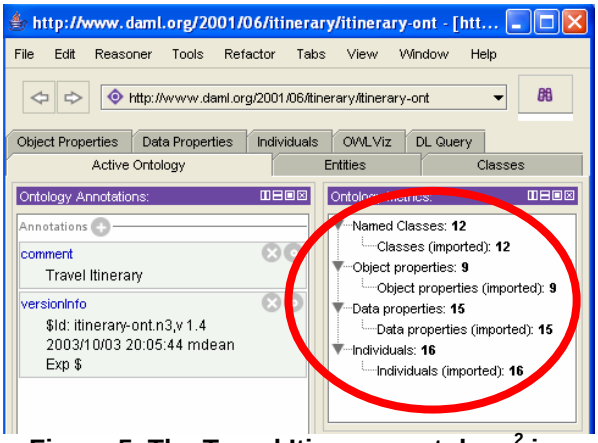

Figure 5. The Travel Itinerary ontology[2] in Protégé

Also when user select the classes window, it shows created classes in graphic form by Protégé as see in Figure 6., which presented classes are as following: *Aircraft*, *Airline*, *AirportCode*, *Class*(implied as class of the plane's service {*First Business*, *and Coach* see also in Figure 9.}; not classes of ontologies), *Flight*, *HotelReservation*, *Itinerary*, *Meal*, *RecordLocatorNumber*, *RentalCar*, *and nonNegativeInteger*.

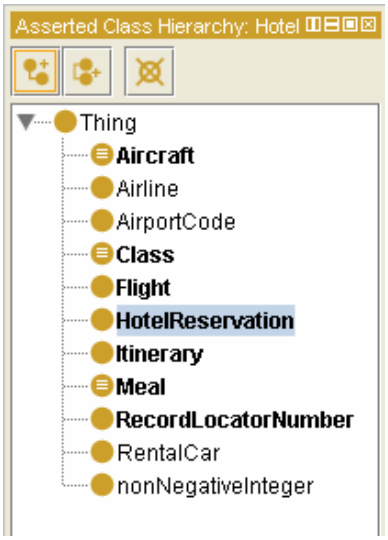

Figure 6. Class Hierarchy in Travel Itinerary ontology[3] in Protégé

[1] http://e-tourism.deri.at/

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

In the Figure 7., describe Object Properties which are *aircraft, airline, class, destination, flight, meal, origin, rentalCar, and rln.* If we select class *Flight*, and it show Data Properties which are *address, arrive, checkin, checkout, confirmation, depart, duration, hotel, hotelName, miles, passenger, rate, rloc, seat, and smoking* in Figure 8.

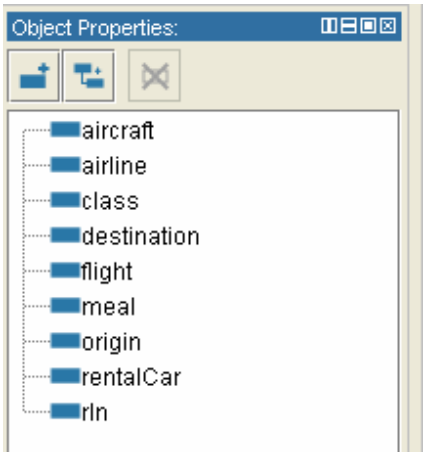

Figure 7. Object Properties of Travel Itinerary ontology[4] in Protégé

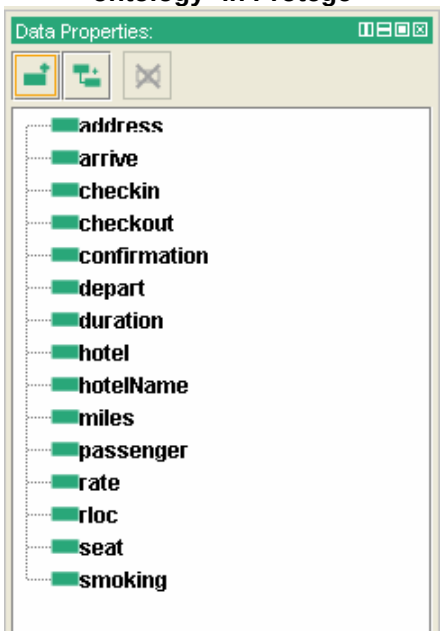

Figure 8. Data Properties of Travel Itinerary ontology[5] in Protégé

For Individuals is the specific instance of the classes, in this ontology show the individuals as see in Figure 9. that *A300* is the instance (Types) of class *Aircraft*, *Breakfast* is type of class *Meal*. And also *First* and *Business* are type of class *Class*.

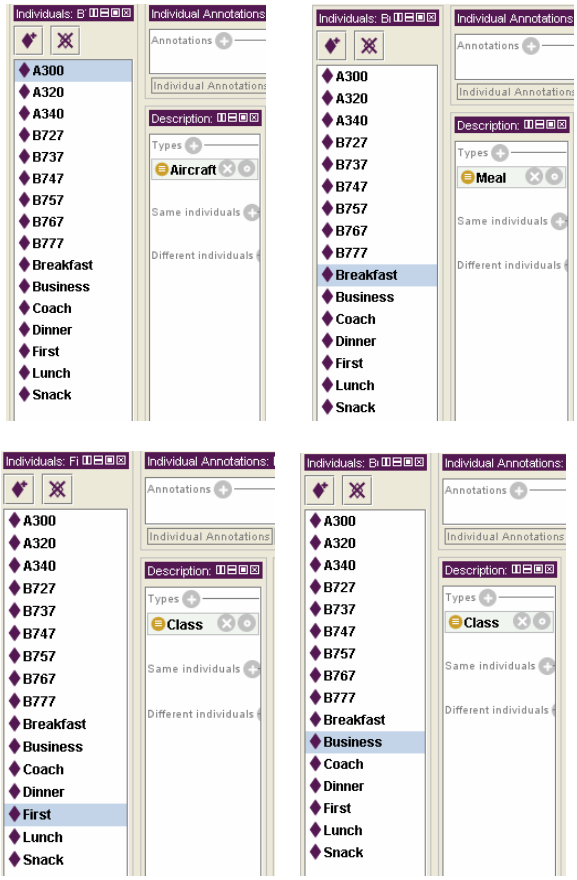

Figure 9. Individuals of Travel Itinerary ontology[6] in Protégé

The following is the example of OWL description of Travel Itinerary which shows in Protégé previously. The beginning of ontologies is declared the related URIs as following:

```
<rdf:RDF
    xmlns="http://www.daml.org/2001/06/itinerary/itinerary-ont#"
    xmlns:airline="http://www.daml.org/2001/06/itinerary/icao#"
    xmlns:airport="http://www.daml.ri.cmu.edu/ont/AirportCodes.daml#"
    xmlns:log="http://www.w3.org/2000/10/swap/log#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
```

[3] http://www.daml.org/2001/06/itinerary/itinerary-ont
[4] http://www.daml.org/2001/06/itinerary/itinerary-ont
[5] http://www.daml.org/2001/06/itinerary/itinerary-ont
[6] http://www.daml.org/2001/06/itinerary/itinerary-ont

The example of classes of Travel Itinerary Ontology in Figure 10., can be referred as "**owl:Class**". For example, class *Aircraft*. The other class is *Flight*, which has *origin* as a property can be referred as "**owl:onProperty**" which collect the defined universal airport code value from *AirportCode* referred as "**owl:allValuesFrom"**.

```
<owl:Ontology rdf:about="">
<rdfs:comment>Travel Itinerary
</rdfs:comment>
   <owl:versionInfo>$Id: itinerary-
ont.n3,v 1.4 2003/10/03 20:05:44 mdean
Exp $</owl:versionInfo>
    </owl:Ontology>

<owl:Class
rdf:about="http://www.daml.org/2001/06/it
inerary/itinerary-ont#Aircraft">
        <owl:oneOf
rdf:parseType="Resource">
            <rdf:first
rdf:resource="http://www.daml.org/2001/06
/itinerary/itinerary-ont#A300"/>
            <rdf:rest
rdf:parseType="Resource">

…
<owl:Class
rdf:about="http://www.daml.org/2001/06/it
inerary/itinerary-ont#Flight">
  <rdfs:subClassOf
rdf:parseType="Resource">
            <rdf:type
rdf:resource="http://www.w3.org/2002/07/o
wl#Restriction"/>
            <owl:allValuesFrom
rdf:resource="http://www.daml.ri.cmu.edu/
ont/AirportCodes.daml#AirportCode"/>
            <owl:onProperty
rdf:resource="http://www.daml.org/2001/06
/itinerary/itinerary-ont#origin"/>
        </rdfs:subClassOf>

<owl:allValuesFrom
rdf:resource="http://www.w3.org/2001/XMLS
chema#dateTime"/>
            <owl:onProperty
rdf:resource="http://www.daml.org/2001/06
/itinerary/itinerary-ont#depart"/>
…
```

Figure 10. The OWL description of Travel Itinerary[7] Ontology

In addition, there is the Mondeca´s tourism ontology, which includes tourism concepts from the WTO[8] thesaurus. At this writing this ontology has 1000 concepts that describe accommodations and transportation and a few other secondary elements related to geography, health and immigration[9]. There are over ten elements on the list of domain specific ontologies that can be useful for the tourist sector, including geographic ontologies, means of transportation ontologies, gastronomy ontologies, etc [1]. General – or sometimes called upper – ontologies also exist and aim to gather definitions and concepts that together make up what is known as unspecialized common knowledge. One of the best known of these is WordNet – more appropriately referred to as a lexical reference system[10] which was extended from solely English into other languages such as through the EuroWordNet[11].

Looking at the ontologies for hotels and tourist destinations (the organizations examined for the feasibility analysis of the approach proposed here) the following facts emerged. Since the concept of hotel is part of common knowledge, the notion is present in WordNet.

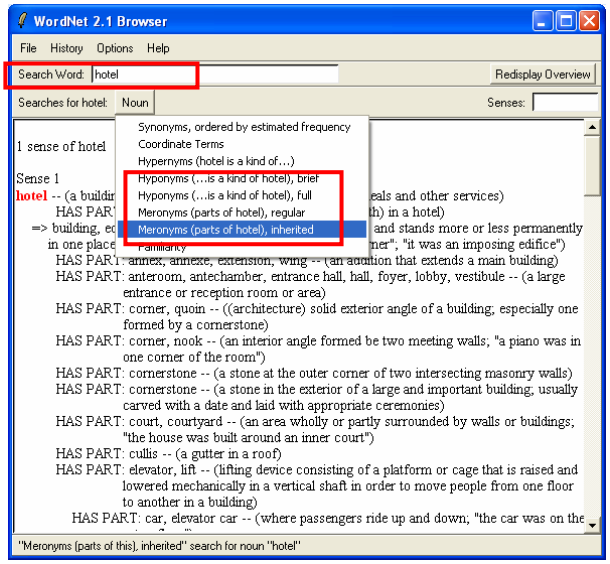

Figure 11. The WordNet searched result of "hotel" concept with "part of" relationships with other concepts

More specifically, for each concept – in this case *hotel* – WordNet gives information see in Figure 11. on the generalization, specialization and also on "part of" relationships (*(parts of hotel), inherited*). A description of the linked concepts is available at Answer.com[12]. The definition of tourist destination and the classification of diverse types of destination show how the necessary concepts belong to diverse domains. WordNet gives no treatment to the concept of tourist

[7] http://www.daml.org/cgi-bin/hyperdaml?http://www.daml.org/2001/06/itinerary/itinerary-ont

[8] www.world-tourism.org

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

[10] http://www wordnet.princeton.edu/

[11] http://www.globalwordnet.org/gwa/wordnet_table.htm)

[12] www.answer.com

destination. It is thus necessary to use different ontologies to describe the geographic area, the different attractions, sports, transport, etc. On the other hand, general ontologies (upper ontologies) contain a lot of information that is not useful because it is related to concepts that differ greatly from those used in the tourism context. In addition, general ontologies for tourism cover only some of the necessary areas (for example, Mondeca[13] gives good treatment to accommodations and transportation) but not for some indispensable aspects of a destination (examples being cultural, natural and artistic attractions or events).

3. Conclusions

Ontologies will play an important role as they promise a shared and common understanding of traveling concepts that reaches across people and application systems. 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. The travel industry is facing rapid changes with the advent of the Semantic Web technologies. There is now the need for developing an infrastructure to manage the online travel information and deliver to consumers what they want. New superior consumer services can be deployed such as travel market overview and price comparison.

It is commonplace that the delivery of tourism e-services is determined by information technology, organizational designs and the preparation of information mainly delivered online. With regard to the area of tourism, key companies such as the tourism industry, the information and communication technology industry and the public authorities mainly structure the service offer and the service quality.

One of the most important objectives of the works mentioned in this paper is to explain the brief introduction of e-tourism ontologies. As the idea to better the lack of standards in the tourism field by automatically understanding the different ways of expressing tourism products and services, extracting its relevant information and structuring. The refined technologies, such as semantics and ontologies, are good candidates to enable the development of e-tourism information systems.

However, the standard still not available at the time, there are many ontologies projects working in the tourism area. They are emphasizing the different point of view and different language of tourism. There are many e-tourism projects going on that are not mentioned in this paper. However, they try to cooperate and make e-tourism complete and practical. Soon we will have mature and available e-tourism ontologies enough for world wide reusability.

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

5. Reference

[1] www.crt.dk

[2] www.nua.com/surveys

[3] Awad E. and H. Ghaziri, Knowledge Management, New Jersey, Pearson Education. 2004.

[4] K. Siorpaes, , OnTour System Design, 2004, [Online] Retrieved: http://e-tourism.deri.at/ont/docu2004/OnTour%20-%20System%20Design.pdf.

[5] J. Cardoso, E-tourism: Creating Dynamic Packages using Semantic Web Processes, 2005.

[6] L. Carton, 2006. [Online] Retrieved: http://www.etourismnewsletter.com/ecommerce.htm.

[7] N. F. Noy, and D. L. McGuinness, Ontology Development 101: A Guide to Creating Your First Ontology, DERI, E-tourism Working Group. 2005, DERI International, 2001, [Online] Retrieved: protege.stanford.edu/publications/ontology_development/ontology101.pdf

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

[9] W.V. Siricharoen, Using ontologies for E-tourism, in proceeding of The 4th WSEAS / IASME International Conference on ENGINEERING EDUCATION (EE'07), Agios Nikolaos, Crete Island, Greece, July 24-26. 2007.

[10] M. Klein, D. Fensel, F. V. Harmelen, and I. Horrocks, The Relation between Ontologies and Schema-languages: Translating OIL-specifications in XML-Schema, 2000, [Online] Retrieved: http://www.cs.vu.nl/~mcaklein/papers/oil-xmls.pdf

[11] C. Kim, E-TOURISM: AN INNOVATIVE APPROACH FOR THE SMALL AND MEDIUM-SIZED TOURISM ENTERPRISES (SMTES) IN KOREA, College of Hotel & Tourism Management, Kyunghee University, Korea, 2000, [Online] Retrieved: www.oecd.org/dataoecd/56/13/34268048.pdf.

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

[13] Protégé, what is protégé-owl?, Stanford Medical Informatics, 2007, [Online] Retrieved: http://protege.stanford.edu/overview/protege-owl.html
[14] R.Volz, D.Oberle, and R. Studer, Views for light-weight web ontologies, in Proceeding of SAC 2003, Melbourne, Florida, USA.
[15] W. Vongdoiwang, and D. N. Batanov, Similarities and Differences between Ontologies and Object Model, in proceeding of CCCT 2005, Austin, Texas, USA.
[16] R. Jakkilinki, N. Sharda, and I. Ahmad, Ontology-based Intelligent Tourism Information Systems: An overview of development methodology and applications. In proceeding of TES2005: Tourism Enterprise Strategies: Thriving – and Surviving – in an Online Era, Melbourne, Australia, 2005.
[17] F. Scharffe, D13 v0.1 KIM evaluation and possible applications to e-tourism E-Tourism Working Draft 15 February 2005, [Online] Retrieved: http://e-tourism.deri.at/2005/d13/v0.1/D13v0.1.pdf
[18] J. Heflin, J. Hendler, S. Luke, and Q. Zhendong, SHOE: A Knowledge Representation Language for Internet Applications, Technical Report CS-TR-4078 (UMIACS TR-99-71), 1999.
[19] S. Handschuh and S. Staab, Authoring and Annotation of Web Pages in CREAM, in proceeding of WWW 2002.
[20] M.V. Vera, E. Motta, J. Domingue, M. Lanzoni, A. Stutt and F. Ciravegna, MnM: Ontology Driven Semi-Automatic and Automatic Support for Semantic Markup, in Proceeding of the 13th International Conference on Knowledge Engineering and Management (EKAW 2002), ed Gomez-Perez, A., Springer Verlag, 2002.
[21] A. Kalyanpur, B. Parsia, J. Hendler, and J. Golbeck, SMORE - semantic markup, ontology, and RDF editor, [Online] Retrieved: http://www.mindswap.ord/papers/.
[22] F. Ciravegna, A. Dingli, D. Petrelli and Y.Wilks, User-System Cooperation in Document Annotation based on Information Extraction, in Proceeding of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW 02), 2002.
[23] A. Barstow, J. Kahan, M. Koivunen, R. Swick, Annotea: A Generic Annotation Environment using RDF/XML, W3C Document, 2001.
[24] M. Dell'Erba, Harmonise. An ontology-based approach to semantic interoperability in the tourism domain, eCommerce and Tourism Research Laboratory, ITC-irst, [Online] Retrieved: http://ectrl.itc.it
[25] Semantic Technology Lab: OntoTextLab, The KIM Platform: Semantic Annotation, 2007, [Online] Retrieved: http://www.ontotext.com/kim/semanticannotation.htm
[26] M. Gratzer, and W. Winiwarter, A Framework for competitive advantage in etourism, 2003, [Online] Retrieved: http://homepage.univieac.at/werner.winiwarter/enter2003.pdf