

A SURVEY ON EXISTING TOURISM ONTOLOGIES

Archana Mathur¹, Akshatha K², Apoorva Shastry³, Anitha J⁴

^{1,2,3}Department of Computer Science and Engineering, PESIT BSC, Bangalore-560100

archanamathur@pes.edu

⁴Dayanand Sagar Academy of Technology and Management

anitha.jayapalan@gmail.com

Abstract

Many studies have been conducted, to indicate extensive use of Tourism websites by vacation planners, to analyze travel related information and data. Semantic web technologies have been used extensively to build and enhance Analytical Engines for tourism industry over the past decade. Various Ontologies, as a powerful knowledge base, have been used for representing the various tourism concepts. These Ontologies can then be further integrated to handle the much needed requirements of resolving the heterogeneity problems arising in e-tourism. On the Flip-side, this integration of Ontologies may complicate the whole task of simplifying the search for the relevant e-tourism data. This survey paper intends to explore all dimensions of publicly available tourism Ontologies, and put-forth a new area of research, to identify and mitigate the shortcomings of existing Ontologies and there integration.

Keywords: Tourism, Semantic Web, Ontology

I INTRODUCTION

Tourism is one of the most diverse and dynamic industries in the world. It is a sector that generates a lot of business and revenue in a nation and contributes in economic growth of a country. There exist a huge amount of tourist destination information available on WWW. An immense number of travelers use Internet to plan their trips or vacation posing a big challenge in efficiently accessing and managing the tourism related information. The documents on web are predominantly in an unstructured format. To serve various user queries, the data needs to be collected from heterogeneous sources, integrated and presented to a user as response.

Semantic Web, a concept given by Tim Berners-Lee, is an effort to enable current web so that even machines can access the information present on web, integrate it and help users to find required information. It's a collaborative effort led by World Wide Web Consortium(W3C). Over a period of time, the W3C had published different standards for semantic web formats, facilitating the exchange of semantically rich information. RDF, OWL, SPARQL, RDF query language are few amongst these. Resource Description Framework (RDF) is a data representation format based on triples *subject-predicate-object* forming a graph of data. An ontology defines concepts of a domain and also specifies their relationship. The Web Ontology Language(OWL) is used to describe and author ontologies. SPARQL is like SQL to query the RDF data.

Ontology is a method to formally define concepts and their relationships designed with an intention of knowledge sharing and reuse. Various ontologies are built to support and integrate the tourism related heterogeneous information available on web, which can further provide useful

suggestions or may help travelers to plan their itinerary. Ontology related to tourism are created modified and integrated with other technologies to make the tourism search efficient and easy. This integration of technologies can complicate the whole task of simplifying the search for E-tourism data. Hence an extensive study is a must to make aware of which all ontologies are built and how are these used in t.

This paper presents a preliminary survey of existing ontologies which are publicly available and makes a survey of how it is efficiently utilized in various researches of tourism domain.

II RELATED WORK

The area of tourism is highly dynamic and diverse. The major section of the existing Tourism Information System uses the available technologies to cater the travelers need At the same time, the right set of data that a traveler requires and its detailed information becomes unavailable owing to the vastness of the information on the World Wide Web. On the other hand, semantically annotated web resources can be understood by humans as well as by machines which can subsequently enrich websites with semantics and enable more intelligent and efficient searching coupled with further processing of data without human intervention.

There continues to be extensive research work in the field of Ontology development for tourism. Waralak V. Siricharoen [1] in his paper laid stress on usage of ontologies in the process of building a semantic-enabled E-tourism software and packages. While discussing some ontological trends the author summarizes various E-tourism ontologies like Harmonise, Hi-Touch, SATINE, ImagineIT and explores every aspect of its usage. The paper reveals the importance

of using ontologies in building semantic-enabled E-tourism packages and explains the usage of RDF to describe a resource in web. RDF is an xml like metadata language.

K.R.Ananthapadmanaban, H.Srimathi, S.K.Srivatsa [2] aims at creating an ontology exclusively for TamilNadu Tourism. They had built a core tourism ontology formalizing the concepts of accommodation, site seeing, transportations, leisure activities etc. The work is exclusively carried out suiting a requirement for a specific region which may not serve well for the other areas in need.

Wei Wang[3] had developed a mechanism to recommend tourist destinations based on personalized recommendations. Two ontologies are created, one on user profiles and another on tourism information. Applying Bayesian Network, the author combines the traveler's preferences and the behavior of other travelers of same taste, to generate suggestions/recommendations for a tourist. A tourism ontology was built using protégé and Ruby on Rails application was used to build the complete system. Netica tool is used to calculate preferred activities for a traveler.

Fatima Faiza Ahmed, Fatima Saima Ahmed [4] observes the lack of standards in Tourism Information System and proposes a mechanism where a user, willing to travel can submit a query in his natural language. A wrapper is developed to extract the online hotel information. This information is used to populate the MyHotel Ontology created by the author. Their work is predominantly to search a desired hotel that meets all the requirement of a traveler.

Robert Barta, Christina Feilmayr, Christoph Grün [5] used the concept of modularizing the ontologies, wherein the domain independent ontologies are integrated with fine grained domain ontologies to facilitate interoperability between different tourism services. The core domain ontology given by the author is Mosaic Ontology.

S. Saraswathi, Jemibha P, Sugandhi M, Mathimozhi M, Lourdu Sophia A [6] created a Semi Automatic Ontological Tree for Pilgrimage places in Tamil Nadu. Their work satisfies the tourism related information for a state and can be extended to cover a still larger area crossing boundaries.

Eleni Tomai, Stavros Michael, Poulicos Prastacos [7] created two ontologies here. A User Profile Ontology is build to create profile of user's interest, budget he may spend, duration of his stay etc. Users are prompted to fill a sequence of questionnaire to know their requirements. A Tourism Ontology encompasses concepts like sightseeing, leisure activities, archaeological places and various means to connect these places for a visitor. A matching algorithm runs in the background to match the requirements of the user with the tourist services available.

Preeti R. Dodwad, L. M. R. J. Lobo [8] had built a context aware Recommender system for Android device supporting GPS. User profiles and interests are gathered and matched with the already-built concepts of Tourism Domain

Ontology. The system not just recommends the desired spots and accommodations, but also makes available the GPS locations and weather conditions of that location. The dynamic variations of user preferences may affect the recommender system to not give the required results.

Most of the above mentioned works may recommend a spot of interest, or an affordable accommodation to a traveler, a very few can make a complete, end to end travel plan for a traveler. This proves that the ontologies are not used to its potency and still a lot of work can be done to efficiently use the travel domain ontology to help in developing better Travel Recommender Systems.

III EXISTING ONTOLOGIES: A REVIEW

The ontologies that are still in use and that had motivational affect on various researcher are as follows:

1) Mondeca Ontology:

The most prominent Ontology, built on the concepts of tourism is given by World Tourism Organization(WTO) thesaurus. The concepts given are object profiling, tourism packages, multimedia content related to tourism and description of archeological objects along with other concepts. Intelligent Topic Manager is proprietary software developed by Mondeca to maintain its travel ontology.

2) Harmonise:

The ontology built as a part of the Harmonise project financed by European Commission under the 5th Framework RTD Programme, mainly focuses on creating a network for exchange of data in tourism industry. The project was developed to cater the need to tackle Data Heterogeneity problem. Programs to handle such problems are expensive in terms of time and money. So, an ontology mediated integration of tourism system was developed. Key concepts of this ontology are travel related events, accommodation, tourist attractions and food styles. The information contained in ontology provides awareness in terms of cooking styles, currency used, local language spoken, coordinate location and date and time information of events. The ontology formalizes the information about local transport services facility a tourist may get for local movement. It not only covers the basic concepts but also encompasses the detailed terminologies to make it complete and concrete.

3) QALL-ME Ontology:

The ontology funded by EU-funded project covers several aspects of tourism sector which are covered in other ontologies as well. It includes accommodation, tourism sites, events, transportation etc. The ontology is mapped with two foundational ontologies – WordNet and SUMO. WordNet, a lexical database for English, groups English words into sets of synonyms providing short definitions and usage examples for all the words. SUMO or Suggested Upper Merged Ontology is the largest formal ontology consisting of 25000 terms and 80000 axioms. This ontology describes concepts

that are same across all knowledge domains. These ontologies are mapped and aligned with the core ontology thus creating a powerful model and lays a foundation of a Knowledge Base.

4) OnTour Ontology:

The eTourism Semantic Web Portal, to demonstrate the use of Semantic Web created a prototype called OnTour which describes the concepts of accommodation (hotels, guestrooms etc.) activities, date and time of certain events, languages, currency, location coordinates to enable tourism information to different tourists. Considering OnTour as an intelligent search engine, it allows a tourist planner to specify its constraints and preferences in terms of budget, level of comfort and recommends venue and date of travel to best suit his requirements. The project though provides a well structured representation of concepts but lacks too detailed information about certain other concepts.

5) HiTouch Ontology:

The development of HiTouch ontology was a part of IST/CRAFT European program aiming to develop software tools to be used by travel agents to cater the needs and expectations of prospective tourist. The ontology developed mainly by Mondeca, consist mainly of three classes – a document referring to any tourism related document, an object is a tourism related offer and a publication is a document resulted from an answer of a query. An object class can be further categorized as Activities, Environment, Ethics Imagination, and Logistics.

6) TAGA:

TAGA works on the platform compliant to Foundation for Intelligent Physical Agent(FIPA). The Travel Agent Games in Agentcities (TAGA) defines two domain ontologies for its working, one covers the basic concepts related to travelling like travel routes, services needed, reservations etc, and the other is dedicated to auctions and its protocols. Agentcities is an initiative to develop a heterogeneous network of Agent Services for various groups like Travel and Tour services, Business services etc.

7) GETESS:

The German Text Exploitation and Search System is a BMBF financed project aims at retrieving the tourism related information through tourism websites. This information can be queried by different users through natural language processing techniques. GETESS is an intelligent agent that gathers tourism related information from the web and uses Semantic Web methods and NLP to generate answers to human queries in a user friendly way.

Apart from the ontologies which were developed as a part of some funded projects, there are other customized ontologies being built to support varying requirement of a tourist. Robert Barta, Christina Feilmayr, Christoph Grün [5] in his work, created a modularized architecture of ontologies with an intension of handling unexpected events in a trip. The

domain independent ontologies like Geospatial, Time, Weather, accommodation are integrated with the core domain ontology cDOTT thus providing travel related information to tourist in a highly mobile environment.

Nalin Sharda, Roopa Jakkilinki, Mladen Georgievski and Mohan Ponnada [9] presented the Australian Sustainable Tourism Ontology (AuSTO) based on the specifications of the National Information Architecture for the Australian Tourism Industry (NIAATI). AuSTO consists of three classes – Foundational classes, the most generic classes – Intentional classes, built to provide functionality and control to tourism applications and – Operational classes, to create and use the generated operational knowledge by tourism application. The overall ontology architecture thus provides a concrete travel plan to a traveler takes care of his travel requirements and suggests a most optimal travel solution based on his needs.

Similarly Alisa Kongthon[10] in their work had built a question-answering system exclusively for Thailand Tourism Industry. The system is built by collecting the tourism related information from websites and by collecting natural language requests posted on public forums. An annotated corpus is created to serve the queries from traveler.

The concept of MultiLingual Ontologies was brought in existence by Chaves, Marcirio, Larissa Freitas, and Renata Vieira[11]. They had built a domain specific ontology called Hontology to support four language: English, Portuguese, Spanish and French. The ontology is aligned with the QALL-ME ontology and majorly serves the need to find the appropriate accommodation for any traveler.

IV. CONCLUSION

The survey presented in the paper is covering most of the Domain Ontologies built for tourism sector and explores its usage in the process of obtaining relevant information for tourists and traveler. Ontology along with the other Semantic Web tools ensures that a traveler receives the optimal plan suiting his requirements whether he is a vacation planner or a business traveler. Mostly authors in their work have used Protégé tool to create ontology. Ontology is queried using SPARQL, Simple Protocol and RDF Query Language which is a query language to query RDF data. Some of the domain ontologies like QALL-ME are used by few researchers with small modifications and alignment of concepts to cater their requirements. This survey covers almost every ontology that was built in recent past and examines the potential increase in information extraction by its usage.

REFERENCES

- [1] Waralak V. Siricharoen “Learning Semantic Web from E-Tourism”, KES-AMSTA 2008:516-525
- [2] K.R.Ananthapadmanaban, H.Srimathi, S.K.Srivatsa, “Tourism Information System-Integration and Information Retrieval of Tourism Information Systems

- using Semantic web services”, International Journal of Computer Applications (0975 – 8887) Volume 52–No.14, August 2012.
- [3] Wang, W., Zeng, G., Zhang, D., Huang, Y., Qiu, Y., & Wang, X. (2008, July). An intelligent ontology and Bayesian network based semantic mashup for tourism. In Services-Part I, 2008. IEEE Congress on (pp. 128-135). IEEE.
- [4] Fatima Faiza Ahmed, Fatima Saima Ahmed, “ Dynamic Tourism Information System Using the Semantic Web”, International Journal of Computer Science and Artificial Intelligence Sept. 2013, Vol. 3 Iss. 3, PP. 120-124
- [5] Robert Barta, Christina Feilmayr, Christoph Grün, “Covering the Semantic Space of Tourism – An Approach based on Modularized Ontologies”, CIAO 2009 Proceedings of the 1st Workshop on Context, Information and Ontologies.
- [6] S. Saraswathi, Jemibha P, Sugandhi M, Mathimozhi M, Lourdu Sophia A, A. Nagarathinam, “Semi Automatic Ontology Based Bilingual Information Retrieval System (Pilgrimage Tourism in South India)”, International Journal of Intelligent Systems and Applications (IJISA) ISSN: 2074-9058
- [7] Eleni Tomai, Stavros Michael, Poulicos Prastacos, “An Ontology-based Web-portal for Tourism”
- [8] Preeti R. Dodwad, L. M. R. J. Lobo, “Tourism Information Systems Integration And Utilization Within the Semantic Web” International Journal of Advanced Engineering and Nano Technology (IJAENT) ISSN: Z347-6389, Volume-1 Issue-10, September 2014
- [9] Nalin Sharda, Roopa Jakkilinki, Mladen Georgievski and Mohan Ponnada (2008), “Intelligent Visual Travel Recommender Systems Model For E-Tourism Websites” Australia, CRC
- [10] Kongthon, Alisa, et al. "A semantic based question answering system for Thailand tourism information." the KRAQ11 Workshop: Knowledge and Reasoning for Answering Questions. 2011.
- [11] Chaves, Marcirio, Larissa Freitas, and Renata Vieira. "Hontology: a multilingual ontology for the accommodation sector in the tourism industry", 2012.