

Tourism Information System-Integration and Information Retrieval of Tourism Information Systems using Semantic web services

K.R.Ananthapadmanaban
Research Scholar,
SCSVMV University,
Enathur,
Kanchipuram 631 561
Tamilnadu, India.

H.Srimathi
Phd, Professor
Department of M.C.A
SRM University
Kattankalathur 603 203
Tamilnadu, India.

S.K.Srivatsa,
Phd, Senior Professor,
St.Joseph's College of Engg.,
Jeppiaar Nagar,
Chennai 600 064
Tamilnadu, India.

ABSTRACT

Travel and tourism are commonly known as information intensive domains where in online information places a crucial role for the whole lifecycle of the journey. The growth in the use of Internet has transferred the Web users to use the Internet to gather travel related information. Also, the recent development, the semantic Web [1] aims to extend the current Web standards and technology so that the semantics of the Web content are machine processable. Semantically, annotated Websites can not only be understood by the human readers, but also by machines. The backbone of semantic Web is formed by ontology. Ontologies can assist organizational, browsing, parametric search and in more general, more intelligent access to an online information and service. This research is intended to create a new platform on tourism ontology for Tamilnadu, India. It also aims to integrate tourism information from various Websites and effectively retrieve the right information of user needs using semantic Web services.

Keywords

Ontologies, Owl, Protégé tool, Semantic Web.

1. INTRODUCTION

Electronic tourism is one of the activities that have enjoyed important success in the Internet. Tourism is a data which has rich domain. Data is stored in many hundreds of the data sources. The quality of available information is large, and the problem is to find useful information without wasting time. Ontology based models are a promising approach since ontologies enable presenting data in a machine readable, thus offering easy exchange of data between different applications. Ontology has become one of the main components in knowledge management [2], [3], e-learning [4], medical models [5], knowledge in diagnostic systems [6] and the semantic web. Ontology aware systems provide the possibility to perform semantic search. The user can search the destinations using several criteria related to travelling such as accommodation, accommodation rating, expense, activity, interest, sport, shopping, etc. The main focus of this research is to develop a knowledge-based by constructing tourism ontology for Tamilnadu.

2. RELATED WORKS

Presentation of tourism destination on the web makes a huge amount of data. As quality of available information is large it is difficult for the tourist to find his desired destinations. Ontology has the potential to improve the process of searching appropriate destinations according to the customer preference. Here we are going to analyse various papers on tourism ontology.

Eleni Tomai, Stavros Michael and Poulicos Prastacos explored how the use of ontologies in a web-based environment can be used for tourism applications. Their methodology consist of building two separate ontologies, one for the user profile and another one concerning tourism information and data in order to assist visitors of an area to plan their visit. The user profile ontology is elicited by user responses by means of a form (interface) where user can fill the area of interest and tourism ontology developed by the service provider. In this paper, reference point for all users is taken to be the centre of city Heraklion. Concepts such as time needed to visit interested destinations, time needed to visit interested destinations from the point of reference, average time to see the place and comeback, accessibility of the spot, entrance fees if any and opening hours of the spot if applicable, are other crucial parameters [7].

Robert Barta, Christina Feilmayr, Christoph Grun and Hannes Werthner explored user preferences which according to them, it is fundamental to personalise information about tourist objects i.e. user preferences can be aggregated to set of tourist types. In this paper, concepts like opening hours of point of interest, opening days of point of interest, time pattern to model recurring events are included [8].

Waralak V.Siricharoen discusses some ontological trends that support the growing domain of online tourism. Concepts like date and time of travel, contact data, tickets and locations are explored. His focus is on description of accommodation, infrastructure i.e. many prefer to stay close to an accommodation that is close as possible to an infrastructure [9].

Weiwanag Et al. explored how the use of ontologies can assist tourist plan their trip in Web-based environment. It consist of two Ontologies on for user profile and the other concerning about tourism information and data in order to assist visitors of an area to plan their visit, Bayesian network is used to estimate the travelers preferred activities[10].

Danica Damljanovic and Vladan Devedzic illustrated how the semantic web technologies combined with traditional e-tourism application concepts such as user interest, activities, age group are used and concepts like vacation packages, types of vacation, traveler's types, accommodation, food service, transportation service and room type in hotel are the valid parameters [11].

Ou, S., Pekar, V., Orasan, C., Spurk, C., Negri, M explored on spatial temporal entities which have both temporal and spatial context. It comprised activities, which might be performed at a certain destination and events which refer to an occurrences with specific content at a specific location with a certain time period [12]. Reservations, booking or any kind of tickets are temporal entities as they are valid for certain period of time.

Dellerba, Foder, Hopken and Werthner stated about integrated semantic web technology and they tried to merge tourist electronic markets using ontology as mediator [13].

Siorpaes and Bachlechner 2006; Pranter, 2004 built a prototype system from scratch and stored their data. They created knowledge based on the ontology with the limited amount of concepts and relations [14].

Cardoso presented a system that creates vacation packages dynamically using previously annotated data in respect to the ontology [15].

Hepp et al. explain that there are not enough data in the domain of tourism available on the web. His experiment revealed that existing data on the web are incomplete [16].

On observations, challenges remain for users in reusing available ontological information. Lack of standards is very vital issue, and parameters vary from one ontology to another and also mostly on location based services. Due to heterogeneity of tourism sector, the process of developing tourism ontology on various parameters would be also tedious. In e-tourism different ontologies have been developed for different areas and they might not meet the needs to regional destination for any specific area. In this paper, we are going to construct tourism ontology for Indian scenario especially with parameters for Tamilnadu tourism.

3. E-TOURISM ONTOLOGY AND CREATION BASICS

The ontology has been developed using protégé for e-tourism. Protégé [17] is an extensible, platform independent environment for creating and editing ontologies and knowledge base and the reasoner used is Fact++ [18]. It is a tool which allows users to construct domain ontologies,

having various storage formats such as RDF[19],OWL[20] and XML. The tool displays and edits ontology in graphical mode. It also helps in building OWL-DL ontology and using description logic reasoner to check the consistency of the ontology and automatically compute the ontology hierarchy.

A class hierarchy is defined by stating that a class is a subclass of another. The class Thing is the class that represents the set containing all individuals because of this; all classes are subclass of Thing. The class Trekking, Cycling, Swimming are subclasses of Activity. It is also possible to assert that two classes are disjoint from each other. For example class Budget Hotel is disjoint with the class Star Hotel. This ensures that an individual which has been asserted to be a member of one of the classes in the group cannot be a member of any other classes in the group, i.e. it would not make sense for an individual to be Budget Hotel and a Star Hotel. There are two main types of properties in OWL. The first one, called the object property, can be used to relate two individuals e.g. Accommodation [hasactivity Swimming]. The second is referred as data type property which related an individual to XML schema data type value or an RDF literal. Example: Visitors [canspend 1500].

In OWL properties, may have a domain and range specified. Properties link individuals from the domain to individuals from range. For example in our tourism ontology if the object property is [isinterested] then the domain will be tourist and range will be Wildlife Sanctuary. Data type properties allow instances to have data values in specific data types. The OWL uses the data types borrowed from XML Schema such as "xsd:string", "xsd:int" and "xsd:boolean". Properties can have several characteristics in representing the relation between classes. Properties can be symmetry, if the pair (X, Y) is an instance of the symmetric property P then the pair(Y, X) is also an instance of P. For example, the cousin relationship or friend relationship can be asserted to be a symmetric property. Properties may be stated to be transitive. If the pair of (X, Y) is an instance of transitive property P and the pair(Y, Z) is an instance of P. For example the ancestor relationship can be asserted to be a transitive property. Properties can state to have unique value using functional property whose minimum cardinality is zero and maximum cardinality is one. For example [hasAge] can be stated to be a functional property. Properties also have restrictions so that an instance can satisfy a certain condition; some of restrictions are all values from, some values from, hasvalue mincardinality, maxcardinality and cardinality. Cardinality Restrictions-Meaning: The property must have at least/almost/exactly X values. Example: A Family destination is a destination that has at least one Accommodation and at least 2 Activities.

allValues From Restrictions-Meaning: All values of the property must be of a certain type Example: Hiking is a sport that is only possible in National Parks.

Some Values from Restrictions – Meaning: At least one value of the property must be of the property must be of a certain type example: A National Park is an urban area that has at

least one Campground and offers at least one Hiking opportunity.

hasValueRestrictions-Meaning: At least one of the values of the property is a certain value. Example: A part of Coimbatore is a destination where one of the values of the ispartof property Coimbatore.

4. DESIGN AND IMPLEMENTATION

The main challenge in implementation of effective retrieval of tourism information system is

- Availability of semantic Web content.
- Ontology-based information retrieval

For effective retrieval, the existing Websites of tourism should be made semantic. The architecture is applied to accommodation service, amusement parks and cinema theatres services in the state of Tamilnadu, India tourism domain. The semantic interface architecture is given in Figure-1.

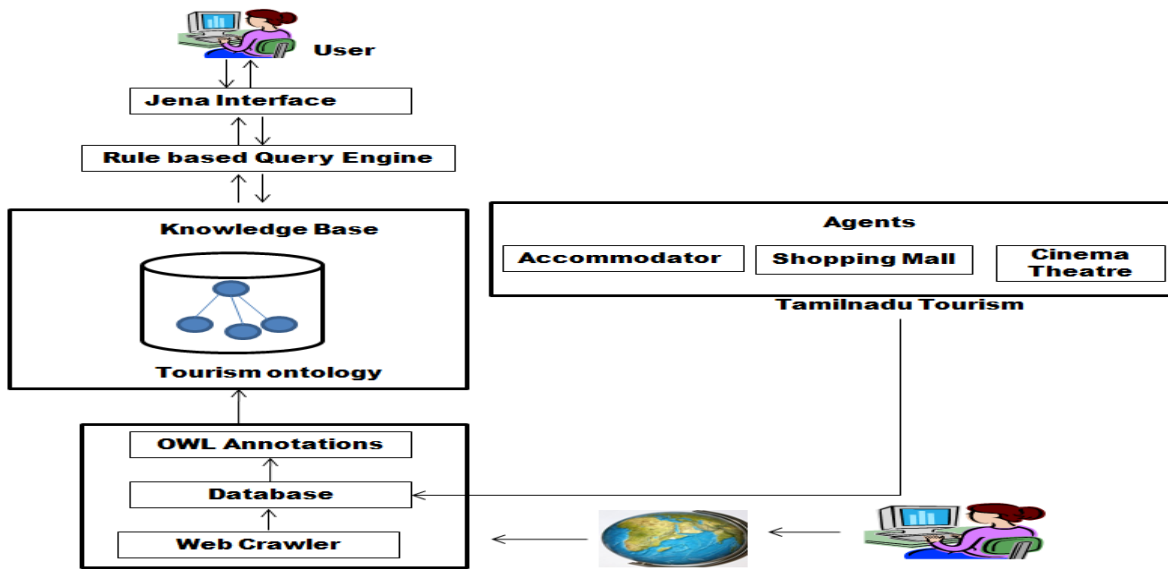


Figure 1: Semantic Interface Architecture

4.1 Getting user Preferences:

This section presents in detail the characteristics of user preferences. The user profile is created in order to extract the user personal information, needs and interests under the context of personalisation. The user can fill in the form their

interests, accommodation needs and other details. This is stored in a database with the help of jena interface and it is used to compare with the tourism ontology for querying the desired results. The screen shot of interface of user profile is given in Figure 2.

Figure 2: User Profile

4.2 WEBCRAWLER

The World Wide Web is an inter linked collection of billions of documents formatted using HTML. Ironically the very size of this collection has become an obstacle for information retrieval. Web crawlers are the heart of search engines. Web crawlers continuously keep on crawling the web and find any new web pages that have been added to the web, pages that have been removed from the Web. A focused crawler is an agent that targets a particular topic and visits and gather only relevant web pages where it will find that keyword. In our work the crawler searches the Internet and finds potentially interesting sites with details about tourism information related to Tamilnadu. This can be done with the help of IN and OUT search web crawler. The results are stored with the URLS in the relational database and URLS redirected to annotation manager. The screenshot figure 3 shows the IN and OUT search.

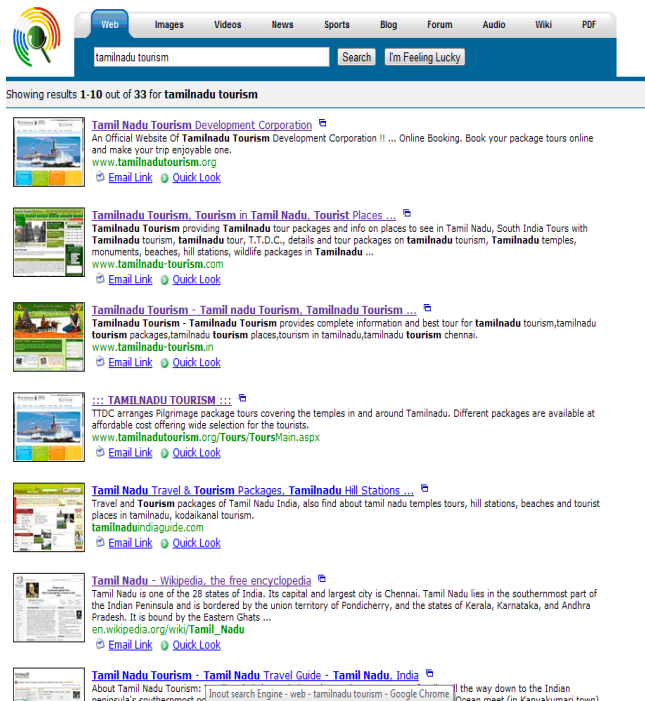


Figure 3: In and Out Search

4.3 ANNOTATION

Annotation manager performs annotation process based on the predefined tourism ontology which will refer to the new concepts/ instance and could be further used to enrich knowledge base. In our work we are using OntoMat for annotation. OntoMat-Annotizer is an interactive web page annotation tool which supports the user with the task of creating and maintaining ontology based markups by creating of instances, attributes and relationships. It includes an ontology browser for the exploration of the ontology and instances and a HTML browser that will display the annotated parts of the text. It is a java based tool and provides plug-in interface for extension. By using OntoMat annotizer user can enrich their knowledgebase. Instead of manually annotating the page with a text editor, OntoMat allows the annotator to highlight relevant parts of the webpage and create new

instances via drag and drop interactions. The Figure 4 shows Annotizer Antro Mat.

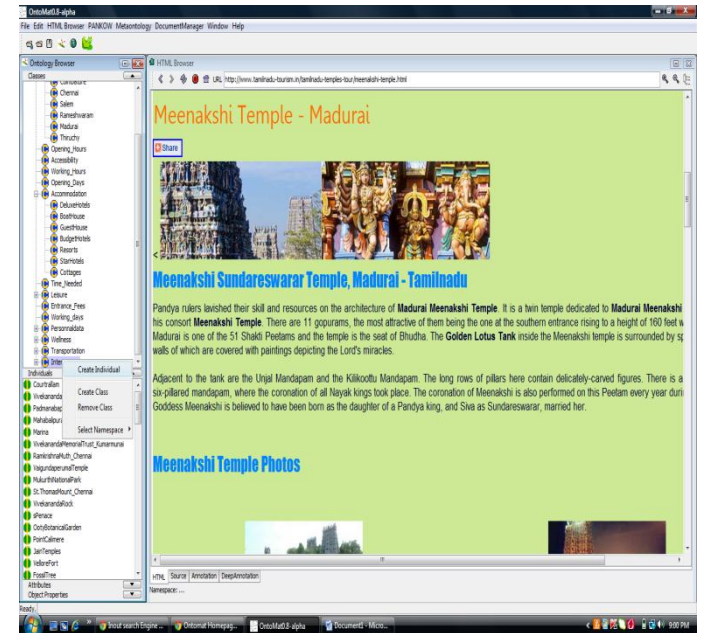


Figure 4: Onto Mat

4.4 Tourism Ontology

The tourism ontology for Tamilnadu is built to conceptualize tourism spots, accommodation and activities during travel. The ontology contains eight important classes, each class has subclasses.

The figure 5 below shows the core concepts of tourism ontology, Table 1 shows the description of main classes in tourism ontology and Table 2 shows the properties used in tourism ontology.

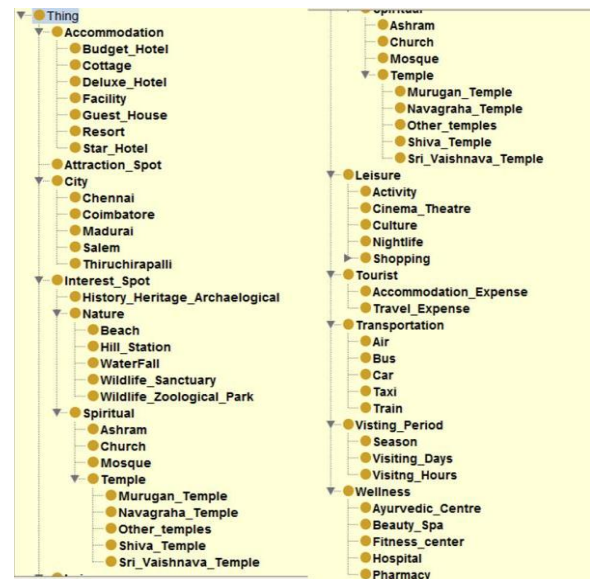


Figure 5: Core concepts of Tourism Ontology

Table 1. The description of main classes in tourism ontology

S.No.	Class	Description
1	Tourist	Tourist is the visitor. All the constraints are based on the tourist preferences. All the classes and subclasses are related to the Tourist.
2	Interest_Spot	The places of interest which the tourist visits, having three subclasses – History_ Heritage_ Archaeological, Nature and Spiritual. Nature consists of five subclasses - Beaches, HillStations, WildlifeSanctuary, WildlifeZoologicalPark and Waterfalls. Spiritual consists of four subclasses - Temples, Church, Mosque and Ashrams.
3	Accommodation	The places providing accommodation service having six subclasses – Budget Hotels, Deluxe Hotels, Resorts, GuestHouse, Cottage and StarHotels.
4	Wellness	This class is about the hospitals and health. It consists of five subclasses namely Hospitals, Pharmacy, Ayurvedic_Centre, Beauty_Spa and Fitness Room.
5	Leisure	This informs about Leisure activities or services, having five subclasses - Activity, Nightlife, Cinema Theatre, Culture and Shopping.
6	Transportation	It provides means of transport to the tourist. It has five subclasses and they are Air, Bus, Rail, Car and Taxi.
7	Attraction_Spot	It consists of subclasses indicating the attraction spots in the interest spots.
8	City	It has the subclasses of major cities that has more interest spots in and near it in Tamilnadu

Table 2. The important properties which links the classes and their usefulness.

S.No.	Property Name	Description
1	hasattractionspot	This property links Attraction spot and Interest spot. It helps the ontology to identify attraction spot in and near the tourist interest spot.

2	hasaccommodation	This property links Accommodation and Interest spot. It helps the ontology to identify accommodation in and near the tourist interest spot.
3	hasactivity	This property links the Accommodation, Interest spot with the Activity. It helps the ontology to identify activities that can be utilized by the tourist in the accommodation and in the interest spot.
4	Hastransportation	This property helps to identify the transportation available during the tour.
5	isnearerto_or_at	This property links the city nearest to tourist spot
6	Hasfacility	It identifies the facility available in the Accommodation place for the Tourist
7	visit_season	It identifies the ideal time of visit to the tourist place. E.g. Ideal season of visit to the tourist place
8	Shoppingat_or_near	It links the shopping places nearer to the interest spots.
9	hospital_at_near	It links the hospitals in and around the tourist spot
10	isopenon	It indicates the days opened of that particular tourist spot.
11	Cinema_theatreat_n ear	It links the Cinema theatres in and around the nearest city of the interest spot
12	Visit_hour	It helps to know the time of visiting hours of the tourist spot.
13	Camera_fee	It is a data type property which helps to know the fee for camera to take inside a particular spot
14	transportin	It helps to connect the places through transportation

From the point of application, the classes cannot exist independently and they are related to one another by means of properties. An interest spot may have attraction spots (i.e. hasattractionspot some interestspot). The tourist should have some accommodation (i.e. hasaccommodation some Tourist).

The figure 6 shows the relationship of the City Chennai with the necessary amenities like transportation, accommodation, hospitals, shopping places and attraction spots.

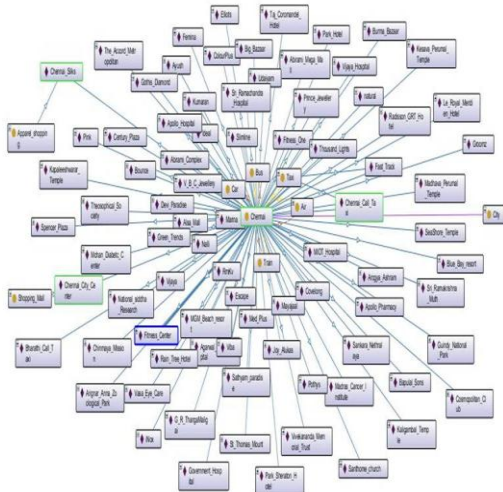


Figure 6: Classes relationship using ontograf.

An instance or an individual of a class can be related to another class. An instance, a tourist spot Parameshwara Vinagaram is a worshipping temple and also can view as Archaeological site. Similarly some temples are of archaeological importance and also attract tourist as a worshipping temple. E.g. Varaha Nithya Kalyana Perumal temple built by Pallava dynasty exhibits the archaeological importance and also has deities worshipped by people and is a spiritual spot. Brihadeeswarar temple built by King Raja Raja Cholan I can be viewed as Archaeological site as well as worshipping temple. Such diversity in tourist spots can be interpreted by multiple tourism resources. This is especially helpful when two or more people of different choices decide a destination together. The above spots are given as instances of subclass Temple of superclass Spiritual and are related to another class History_Heritage_Archaeological. Thus, multidiversity in tourism is handled for the delicate retrieval of tourism information.

In our Work the agents interact to perform tasks such as:

- 1) Crawling the Internet for Tourism Websites in Tamilnadu at regular intervals to search for any new Accommodation, Amusement Park, Cinema theatres, shopping Mall or any other details uploaded.
- 2) If any thing found new by the system admin those URL are redirected to annotation manager-(Onto Mat) and also stored in the database for future reference.
- 3) Annotation manager in turn analyse the content and information is added to predefined ontology , thus constant updates made to the predefined ontology.
- 4) The GUI is accessed remotely by an end user searching for the information in the same way as conventional search engine.

5. RESULTS AND DISCUSSION

The main objective of this paper is to generate enough knowledge by creating ontology for tourism so that users can

- 5) User request are passed to the web agents which in turn, formulate the query plan.
- 6) Inference is performed on ontology schema information and instance data by the activation of a reasoner, which is the component of middleware.
- 7) SPARQL queries are formulated and processed by the agents in conjunction with Jena and results displayed to the end user via GUI.

The most unique aspect of our system design is its generic reasoning and SPARQL querying capabilities. The system was specified to allow any OWL DL(Description Logic) ontology to be loaded into a Jena supported backend, classified with a reasoner and SPARQL queries run over the inferred version of ontology. This design represents a significant advancement in presently available technology because it allows information to suit different user needs and at the same time providing access to a SPARQL query engine that processes inferred knowledge.

4.5 REASONING OVER ANNOTATIONS

A reasoner, also known as an inference engine, is an application that can read an ontology model with its schema and instance data and derive additional knowledge about the domain. For example, if a person wants to search for cinema theatre within the premises of shopping mall in Chennai, by applying a reasoner to the tourism ontology, it infers not only the cinema theatre, but also the facilities available in the shopping mall since the theatre is within the premises of the shopping mall.

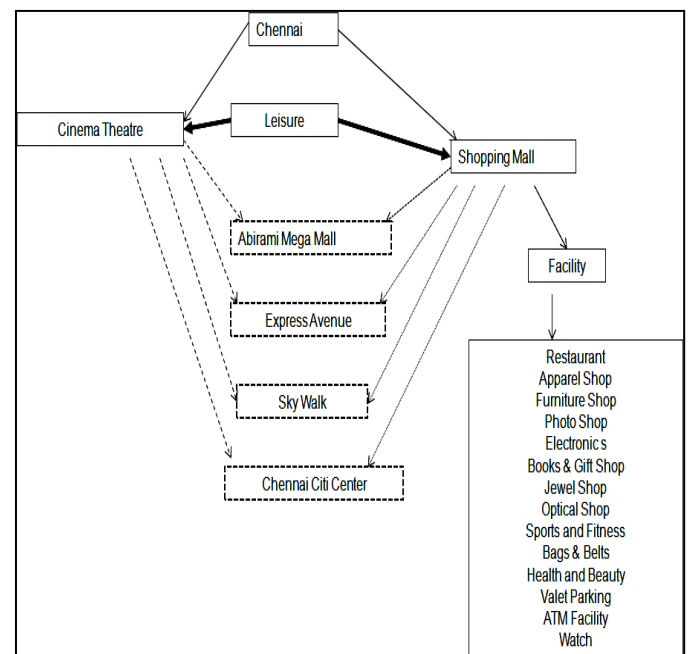


Figure 7: Inference model of tourism ontology

plan for their interest spot by giving their requirements. This can be tested using DLQuery and the same is displayed in GUI using Jena interface.

Query 1: Traveller searching for a resort which is near to Mahabalipuram beach with the facilities of Currency Exchange, Conference Hall and Travel Desk and with activity swimming.

$Q(A) \leftarrow Accommodation(A). Resort(A, R). isnearerto_or_at Mahabalipuram(A, M). isnearerto_or_at Beach(A,B). hasfacility Conference_Hall(A,H). hasfacility Travel_Desk (A,T). hasactivity Swimming(A,P)$

where A = Accommodation, R = Resort, M = Mahabalipuram, B = Beach, H = ConferenceHall, T = Travel desk and P = Swimming.

Conjunctive Query: $Q(A) \leftarrow Accommodation(A) \wedge Resort(A,R) \wedge isnearerto_or_at Mahabalipuram(A,M) \wedge isnearerto_or_at Beach(A,B) \wedge hasfacility Conference_Hall (A,H) \wedge hasfacility Travel_Desk (A,T) \wedge hasactivity Swimming(A,P)$

The corresponding SPARQL query is as follows:

PrefixQ: <http://

www.semanticchennaiatourism.com/tourism.owl#>

Select ? Business Name? URL

Where { ? Accommodation Q: Resort

? Accommodation Q : isnearerto_or_at Q : Mahabalipuram
? Accommodation Q : isnearerto_or_at Q : Beach
? Accommodation Q : hasfacility
Q : Conference_Hall
? Accommodation Q : hasfacility Q : Travel_Desk
? Accommodation Q : hasactivity Q : Swimming }

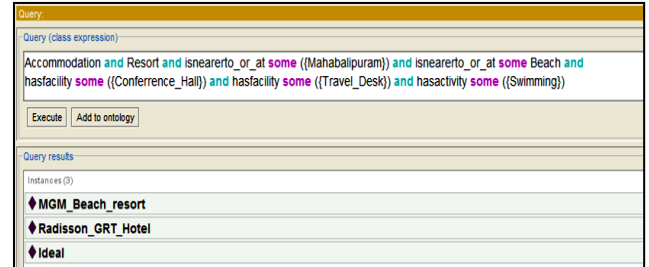


Figure 8: Shows the DL query 1 result



Figure 9: Output of Figure 8 using GUI Interface

Query 2: Some people may visit the tourist spots specifically on Sunday. If tourist X wants to visit a Zoo with a family, then the requirements of the tourist are

a) Is it open on Sunday?

b) Which Zoo is nearer to the city Chennai?

$Q(Z) \leftarrow Zoo(Z). isnearerto_or_at(Z,C). isopenon(Z, S)$

where Z = Wildlife Zoological Park, C = Chennai and S = Sunday

Conjunctive Query: $Q(Z) \leftarrow Zoo(Z) \wedge isnearerto_or_at(Z,C) \wedge isopenon(Z, S)$

The corresponding SPARQL Query is as follows

Prefix Q : <http: //

www.semanticchennaiatourism.com/tourism.owl#>

Select ? Business Name ? URL

Where { ? Nature Q:Wildlife_Zoological_Park

? Wildlife_Zoological_Park Q: isopenon Q : Sundays

? Wildlife_Zoological_Park Q: isnearerto_or_at
Q : Chennai }

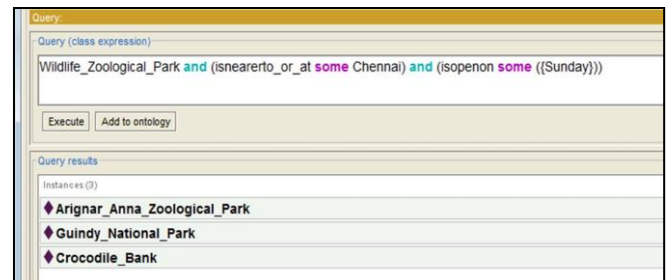


Figure 10: DL query showing the result of query 2



Figure 11: Output of figure 10 using GUI Interface

6. REFERENCES

- [1] N. Shadbolt, T. Berners-Lee, and W. Hall, "The semantic web revisited," *IEEE Intell. Syst.*, vol. 21, no. 3, pp. 96-101, Jan./Feb. 2006.
- [2] L. F. Lat "A knowledge engineering approach of knowledge management", "Inf.Sci.", vol. 179, no.19, pp.4072-4094, Oct.2007.
- [3] L.Zhou and D.Zhang, "An ontology-supported misinformation model: Toward a digital misinformation library." *IEEE Trans.Syst., Man, Cybern.A.Syst., Humans*, vol. 37, no. 5, pp. 804-813, Sep.2007.
- [4] J.T.Fernandez-Breis, D.C.Neives, and R.Valencia-Garcia, "Measuring individual learning performance in group work from a knowledge integration perspective," *Inf.Sci.*, vol.179, no.4, pp.339-354, Feb. 2009.
- [5] A.R.Tawil, M.Montebello, R.Bahsoon, W.A.Gray, N.j.Fiddian, "Interschema correspondence establishment in a cooperative owl-based multi-information server grid environment," *Inf. Sci.*, vol. 178, no. 4, pp. 1011-1031, Feb.2008.
- [6] M.Ulieru, M. Hadzie, and F.Chang "Soft computing agent for health in application to the research and control of unknown diseases, *Inf.Sc.*, vol, 176, no. 9, pp.1190-1214, May 2006.
- [7] Elen Tomai, Spanakim, Poulicous Prastacos Kavouras M. Ontology Assisted Decision Making-A Case Study in Trip Planning for Tourism, International Workshop on Semantic-based Geographical Information Systems, Cyprus, November 3-4,2005.
- [8] Robert Barta, Christina Feilmayr, Christoph Grun, Hannes Werthner. Covering the Semantic Space for Tourism-An Approach based on Modularized Ontologies, Proceeding of the Ist Workshop on Context, Information and Ontologies June 1, 2009, Heraklion Greece ACM 2009 ISBN: 978-1-60558-528-4.
- [9] Waralak V. Siricharoen, "Using Ontologies for E-Tourism", Proceedings of 4th WSEAS/IASME International Conference on Engineering Education, Agios Nikolaos, Crete Island, Greece, July 24-26,2007.
- [10] Wei Wanag, Guosun Zeng, DongqiZhang, Yu Huang, Yufeng Qiu, Xi aojun wang-"An Intelligent Ontology and Bayesian Network based Semantic Mashup for Tourism, IEEE, Computer Society 2008.
- [11] Danica Damjanovic, Vladan Devedzic : "Applying Semantic Web to E-tourism", Chapter X, Springer pp.243-263
- [12] Ou, S.,Pekar,V., Orasan,C., Spurk, C., Negri, M:Development and Alignment of a Domain-Specific Ontology for Question Answering, in European Language Resources Association (ELRA)(ed.): Proceeding of Sixth International Language resources and Evaluation (LREC'08), Marrakech, Morocco, May(2008).
- [13] Dellerba.M, Fodor, O.Hopken. W, Werthner.H (2005), Exploting semantic web technologies for harmonizing E-Markets nformation Technology & Tourism . 7(3-4). 201-219(9).
- [14] Siorpaes, K & Bachlechner , D. (2006). Ontour: Tourism Information Retrieval based on the Yars Demas and Posters of the 3rd European semantic web conference(ESWC 2006), Budhva, Montene.ro 11th -14th June 2006.
- [15] Cardoso. J.(2006). Developing dynamic packaging systems using semantic web technologies. Transaction on Information Science and Applications vol.3(4). 729-736
- [16] Hepp,M., Siorpaes,K., Bachlechner,D (2006). Towards the semantic web in E-Tourism:Can Annotation do the Trick? In proc. Of 14th European Conf. on information system (ECS 2006), june 12-14, 2006, Gothenburg, Sweden.
- [17] T.Tudorache, N.F. Noy, S.Tu, and M.A.Musen, "Supporting collaborative ontology development in protégé,"in Proc. Int. Semantic Web Conf., vol.5318, Lecture Notes in Computer Science, A.P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K.Thirunarayan.Eds., 2008, pp. 17-32
- [18] D.Tsarkov and I.Horrocks, "Fact++ description logic reasoner: System description," in Proc.JCAR.vol.4130, Lectures Notes in Artificial Intelligence,2006, pp.292-297.
- [19] F.Manola and E.Miller , "Rdf primer." in Proc. World Wide Web Consortium W3C Recommendations 10 Feb 2004, Tech.Rep. 2004. [Online]. Available: <http://www.W3.org/TR/rdf-primer/>
- [20] D. L. McGuinness and F. van Harmelen, "Owl web ontology language, overview." in Proc. World Wide Web Consortium W3C Recommendations 10 Feb 2004, Tech.Rep. 2004. [Online]. Available: <http://www.W3.org/TR/owl-features/>