ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

A Novel Semantic Approach for Integrating Heterogeneous Data for Tourism Information.

GS Monisha*, Senduru Srinivasulu, and J Sapphire Paul.

Department of Information Technology, Sathyabama University, Jeppiaar Nagar, Chennai, India.

ABSTRACT

The role played by Tourism industry is crucial in economic progress in South East Asian countries and therefore rich domestic tourist information helping visitors is highly in demand. In these nations, there are specific deciding effect factors impacting visitor journey like constraints in laws, festivals and culture. Such kind of exclusive information have not been made available or applied on to the common search engines till now. In management of knowledge and Semantic web, Ontology plays more and more important role. Tourism data ontology has become an essential field in research activities in empire of data retrieval. The main aim of this paper is exploration of inherent task of FCA (Formal Concept Analysis) within web-oriented ontology building maintain scheme in the tourism empire. Tour planner is provided with an inbuilt intellect that allows generating travel plans. It does this through matching traveler needs with vendor offerings maintained in combination with travel ontology. Both the ontologies then get mapped by making use of the Formal concept analysis and Bayesian analysis for assessment of tourists' desires against data published through tourism data suppliers. Consider planning a particular tour of the city of New York and Hyderabad as an example for describing the ontology approach proposed. For ranking tourism appeal proposed by the FCA and ontology-based strategies and analytic hierarchy procedure (AHP) is employed.

Keywords: Tourism Ontology, Semantic network, FCA, Tour Planner.

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Ontology can be defined as conceptualization of a realm into a format that is machine-readable and human understandable and consists of axioms [1], relationships, attributes and entities. Ontology plays a progressively more important role part in management of knowledge. It is being help as standard information portrayal for Semantic network. With assistance of ontology, it is possible for uses to communicate with one another by employing a common awareness of a realm. Ontology has earned the spot as natural and important means to represent real-world intelligence for developing information scheme (e.g., retrieval of information). Yet, most of the ontology constructions are not executed either automatically or systematically [2]. Furthermore, construction of ontology is one tough and tedious task. Previous research about tourism data requirements at Guangxi Standard Foundation in China has indicated that tourism data demands much, including contents on history, meals, culture, geography, transportation, accommodation, tourism itinerary, scenic places and so on. This, for building tourism data look for platform has to depend on ontology regarding this field. Moreover, tourism data available over the Web is lively and renewing, hence it is of paramount importance to find a way to properly ensure the accuracy and updating of tourism data. It becomes crucially significant to monitor the oncology construction in the field of tourism. Different methods of ontology construction are found to be presented during the recent times [3].

An ontology constructed by Khan et al. making use of an altered self-organization map (SOM) clustering design using bottom up style [4]. Yoshinaga et al. created ontology automatically with relations [5] and keywords. Zhou et al. suggested one customizable cooperative scheme for constructing the realm ontology. Maedche and Staab introduced an ontology studying scheme that consisted of extraction, refinement, pruning evaluation and ontology import [6][7]. An integrated strategy of web ontology engineering and learning that can build and approach the realm ontology for integrating data intelligently within certain virtual participant community was introduced by MissiKoff [8]. Navigli made use of the SemCor and WordNet for interpreting complex phrases semantically, thus avoiding the semantic ambiguation removal problem [9]. Hotho introduced different clustering mechanisms for viewing text documents by using assistance of the ontology [10]. Andreasen developed a system and procedure for content-oriented query of the texts founded on the ontology [11]. Elliman et al. suggested a technique to construct ontology for representing a group of given web pages in a particular site by making use of SOM for creating the hierarchy [12]. Metais and Lammari put forward a group of algorithms for constructing and maintaining Ontologies [13].

In this study, we have suggested a method of constructing ontology called TOCM (Tourism Ontology Construction Method) which can produce ontology that is founded on the theory of Formal Concept Analysis. As per Tom Gruber [14], Ontologies form a portion of W3C grades stack for Semantic Web, wherein they may be used for specifying normal conceptual terms in which activities such as exchanging data among the systems, providing services for replying to queries, publishing reusable intelligence bases and offering services for facilitating interoperability over manifold, heterogeneous schemes and databases. So, this study aims at developing intelligence built upon Web application through creation of ontology for implementing cultural tourism on Dusit district selected as a case study. In fact, Ontology is one type of knowledge depiction illustrating conceptualization of any domain. Ontology indicates a vocabulary that consists of key terms, their linguistic interlinking and a few inference rules [15] [23].

RELATED WORK

Exhibiting tourism destination over the web produces a large chunk of data. Since quality of the available data is huge, it becomes tough for tourists to identify their preferred destinations. Ontology is supposedly having the potential for improving the task of searching suitable destinations in accordance with the preference of customers. We will analyze different studies about tourism ontology here.

Poulicos Prastacos, Eleni Tomai and Stavros Michael probed about how usage of ontologies on a web-oriented environment may be utilized in tourism applications. The method suggested included constructing two different ontologies, one with relation to user profile whereas the other was concerning tourism data and information for assisting potential visitors of a region to plan about their visit. User outline ontology is extracted through responses of users by using one form (interface) wherein user fills the region of his interest and the tourism ontology created by service supplier. In this study, the center of Heraklion city has been taken as the reference point pertaining to all users.

The other vital parameters [16] are concepts like average time for seeing a place and returning, time taken for visiting preferred destinations, time required for visiting desired locations from the reference point, approachability of the place, entrance charges, if any, and if applicable, opening timing of desired spot.

Christina Feilmayr, ChrisophGrun, Hannes and Robert investigated preferences of users that is according to their view is fundamental in personalizing information of tourist motives. In other words, user preferences may be amassed to a group of tourist classifications. In this study, concepts such as hours of interest points, opening period of interest points, and pattern of time for modeling the recurring events have been covered [17] [22].

Waralak debates certain ontological tendencies which back the growing realm of the online tourism. Ideas such as contact data, time and date of travel and spots are examined. He focuses on describing infrastructure and accommodation many people desire to stay at places which are as close to infrastructure as possible [18].

Weiwanag et al. examined how ontologies usage can help tourists to plan about their trips in the Web-oriented environment. It includes two ontologies, one regarding user profile and the other, regarding tourism data and information for assisting visitors of a spot in planning their visit. Bayesian network has been employed for estimation of preferred activities pertaining to travelers [19].

Vladan Devedzic described how semantic web strategies in combination with conventional e-tourism implementation concepts like activities, age group, user interest are utilized and theories such as vacation types, holiday packages, type of travelers, food service, sheltering room type and transportation service with the hotels are certain acceptable parameters [20].

SYSTEM ARCHITECTURE

The system architecture represents the system flow and mechanism of the overall process that involves in tourism information activity.

Tourists

Tourist is the one who is willing to plan a tour to spend their leisure time. Before planning a tour through a website, the tourist has to register their details to login. Once the tourist login into the system, user can plan for their trip with respect to tourist information.

Service Providers

Service provider is the one which provides information about the certain place or attractions which the tourists requests. The main role of the service provider is to provides respective services to the tourists.

Local Ontologies

Local ontologies are the one which stores the information about the tourists and service providers. The data's stored in local ontologies are usually in XML format as RDF file.

Services for Tourists

It provides certain services for the tourists based on tourist preferences. Generally service information are redirected from the service providers once the tourist requested for certain services.

Final Planned Tour

The tour plan is generated based upon the preferred activities that are registered by the tourists at the time of registration. Later the respective accommodation and dinning services are chosen based on their stay at a particular tourist spot.

ISSN: 0975-8585

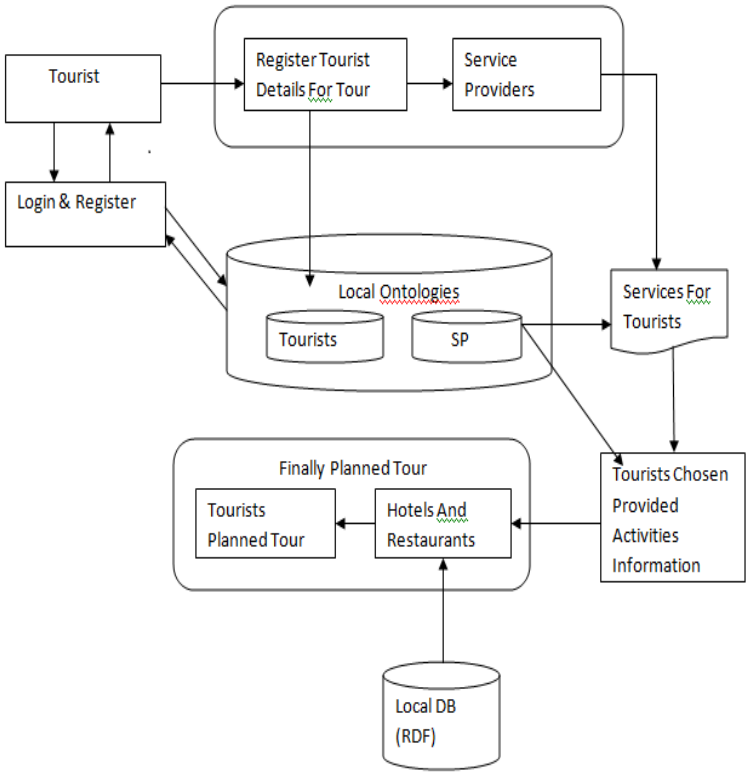

Figure 1: Architecture

SYSTEM FLOW MECHANISM

STEP 1: A tourist wishes to plan a tour online, firstly the tourist has to login the system using respective username and secret code which is generated after registration. Suppose if the tourist is a new user, the user has to register their personal information for further processing. Those login and register details are stored in local ontologies which is inbuilt in application.

STEP 2: After login into the system with respect to username and secret code, the tourist can able to plan their trip based on given tour schedule, tourist place, budget and transportation in order to plan their trip easily.

***STEP 3*:** After planning the tour, the tourist has to select their appropriate service domain either domestic or international services. Suppose if the tourist chooses wrong domain that is not same as the tourist selected, the system will indicate that they have chosen wrong domain.

STEP 4: We have implemented four service providers, two for domestic and two for international services. Suppose if the tourist chooses an international domain, the system analysis the 2 service provider using Formal concept analysis. Finally the best service provider will be provided to the user based on Bayesian analysis.

STEP 5: After choosing the best service provider, the user can proceed their tour plan based on their preferred activities.

STEP 6: Once the trip is planned for tourist attraction, the user can select their favorite hotel for accommodate, which provides the details about location, staying date and hotel rent details. Those data's are stored in local database as rdf file format.

STEP 7: Finally the user has planned their tour trip successfully.

ISSN: 0975-8585

PROPOSED WORK

Ontologies and FCA for Tourists

Ontology happens to be a precise stipulation about the ideas in a realm and also the various relations among them, and it offers a formal lexicon regarding exchange of information. Integration of information via various sources requires sharing through a clear knowledge about the concerned domain. Awareness portrayal formalisms offer structures to organize this intelligence, but furnish no schemes to share it. Ontologies contribute common lexicon for support sharing as well as knowledge reuse. The extracted tourist ontology consists of the primary concepts mentioned below at the initial stage: Desired tour budget, aspired time of tour, activities desired by tourist, and desired mode of transportation. Each idea would have several-level sub-ideas and properties related to both sub-ideas and ideas. These concepts indicate the elements which are ranked consistently as highest related to recurrence of their utility in the study cited. The inference agrees with normal findings of research that choice of tourist appeals frequently rely on only a few common factors like budget, desired activities and time.

Ontology is existence nature philosophy to integrate the heterogeneous data. The researchers give several ontology descriptions. Ontology explicit specification represents relationships and concepts in abstract model definition and explicit names. The ontology allocated and general domain understanding that communicated between communities. Relationships and concepts are common components in ontology. For relationships and concepts web documents plays vital source. Online tourist attractions information is the concept of source to derive relationship with tourist ontology. In this paper we considering New York and Hyderabad tourism attraction place. In ontology we investigated 'Location' like geographic information, 'Attraction' like name, 'Activity' (visiting attraction location) and 'Admission' Fees' like cost. Each attributes has sub attributes for example 'Open Hours', 'Open Times', 'Minimum Time to Stay' and 'Closed Dates'. Tourist attraction may present more activities and unique location. Every activity has various admission fees, based on traveler's occupation and age group. The domestic ontologies progression includes more online text quantity. This learning focuses on tourism attractions in New York and Hyderabad city. Through tourist attractions the information about tourist will include in website. Through location and open hour information in websites to be execute.

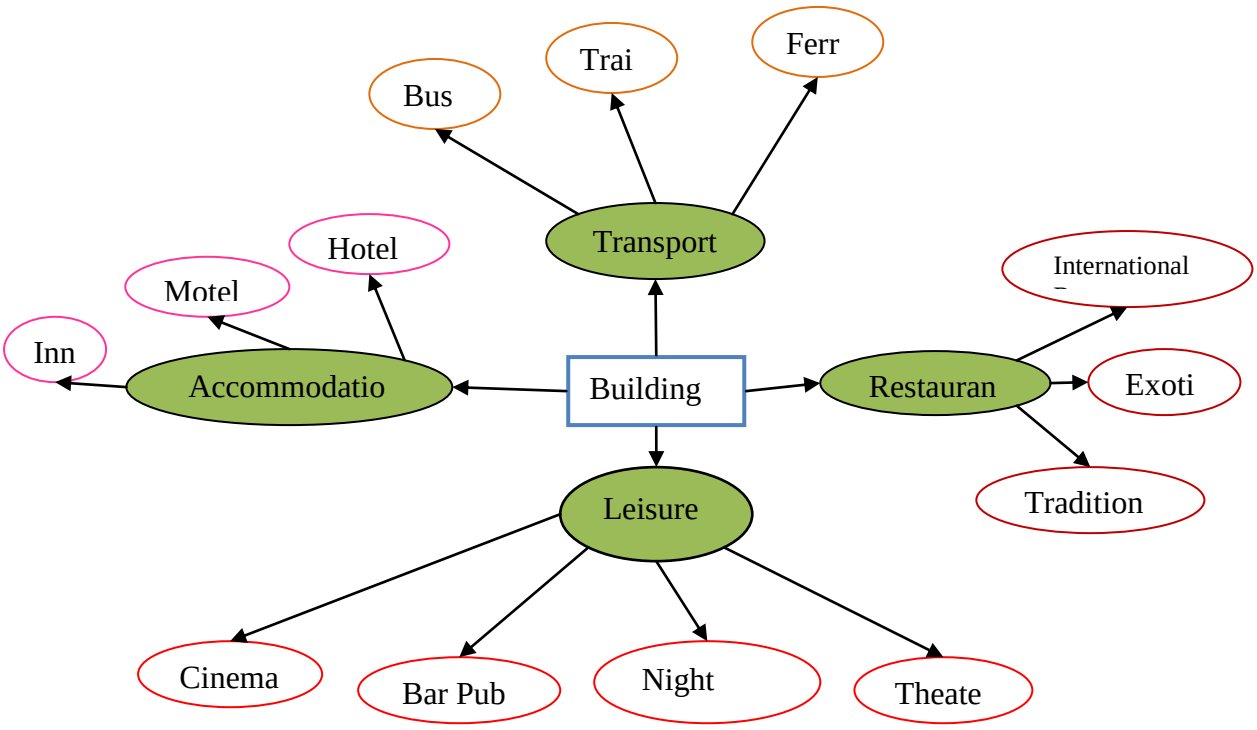

Figure 2: Ontology Process

The evolution of domestic ontologies includes a huge quantity of the online text. For analyzing the said text at some reasonable level, this study focuses on data regarding attractions for tourists in Hyderabad and the city of New York. Those destinations have been chosen as they happen to be among certain very large metropolitan regions in the globe and they also offer varied attractions. Websites about tourism get selected first via tourist attractions. To begin with, such websites should include information of a great number of tourist attractions. Next, such information present in these sites has to be exhibited through obvious titles like open hours and location. Sites with information presented in the multiple levels regarding titles are the most preferred ones. Ideas and their sub-ideas have been described under: 1) The concept of open hours includes three primary sub-ideas: open time (open hours and day), closed dates (holidays and day) and the least time needed for touring some attraction (minutes and hours). 2) The idea admission fees is illustrated in terms of dollars.

Mapping between Local Ontologies

Tourists' integrated ontology regarding tourism data have been mapped for evaluating tourists' desires against data that is supplied through tourism information suppliers. Mapping process among two ontologies is being performed through two levels [21]. While the first one maps the ideas among them, the properties regarding a set of given mapped ideas are matched in second level. Even though the ontology-oriented method illustrated in this study pays attention on tourism appeals, it is also extended to other services which are associated normally in planning of tour, like dining services and accommodation. Akin to proposing tourism appeals, a couple of ontologies, first one to do with service providers and the second, for service users are required and mapped then, among them, for selecting suitable restaurants or hotels. To estimate tourist predilection against tourism data provider published information ontology is used for tourist and for tourism data provider integrated ontology is used. The attributes like arrival time, tour date, transportation mode, total tour hours, activities, travel distance and costs. From tourism literature these concepts extracted. To choose attractions consider both tourism information providers and tourists perspectives. To map two ontologies consider tourism information providers and tourists perspective. Mapping is used to collect the information about tourism like distance, cost and accommodation details.

Resource Description Framework

RDF offers a way to add semantics to any document without the need to make any assumptions of the documents structure. It sums Meta data to the Web record. The framework of resource description tries to address the semantic limitations of XML. It offers an easy model which may be used in representing any type of data. Such data model includes nodes joined through labeled arcs, in which the web resources are represented by the nodes and properties of those resources are represented by the arcs. In RDF database the information about accommodation details will store in the XML format. Whenever the user want accommodation and cost of accommodation details the information will retrieve from the RDF database.

ALGORITHM

Bayesian Algorithm

Bayesian algorithm is probabilistic graphical representation to signify causal relationship between variables. It contains set of directed arcs and nodes. The arcs signify direct causal influences among linked nodes and nodes signify the variables. Arc situate from parent node (F) to child node (C). Child node depends on parent node and temporarily independent others. Condition probability P(F|C), presents how parent node F influence probability distribution over child node C.

1. Formulate Tourism Ontology Construction Method as set of probability distributions condition on various values for method limit $\theta \in \Theta$.
2. Arrange information about θ over Θ of probability distribution.
3. Collect information and include into relations of distributions known in step 1.
4. Use Bayes' theorem to compute new information of θ.
5. Assess a model.

ISSN: 0975-8585

Analytic Hierarchy Process

Begin
Input the initial solution location *S1*
Calculate the objective rank *l0'* for input location using equation maximize L(x) $\frac{1}{1+l(x)}$
Set *iter =1*
best_frank = f0
best_sol = S2
While *iter<= iter_max*
Create initial set *'S'* of randomly generated location (*Si,S, i=1,2,3,…,n*)
Calculate objective rank*'li'* for solution string *Ti*
Compute *ϑ = (li-l)*
if*ϑ*> small random no. (1.0000e-006)
best_sol= Si
best_lrank= li
Else
Perform swap move (exchanging ranks of two random
Suppliers of *Si*)
Check for intention location rank *'li'*
Go to step 10
Accept collection
Else removeleast rank location
Stop if most iteration has been accomplished

RESULT AND DISUSSION

FCA Approach

Figure 3: FCA Approach

Figure 3 shows FCA approach. Formal Concept Analysis provides the tourism information like tourism location, open hours, entry fee and phone number details.

ISSN: 0975-8585

Bayesian Analysis

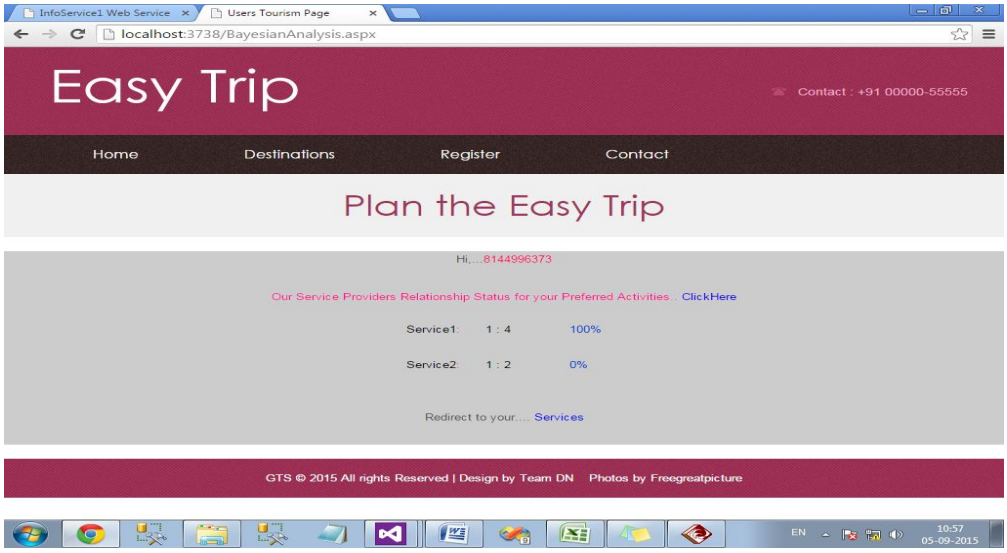

Figure 4: Bayesian Analysis

Figure 4 shows Bayesian analysis. Two service providers are there. We can suggest user any one of best service. Each service provider contains information about tourism.

Tourism Analysis

Table 1: Tourism Analysis

Services	User Interest (%)
Accommodation	60
Travelling	72
View	80
Security	78
Cost	55

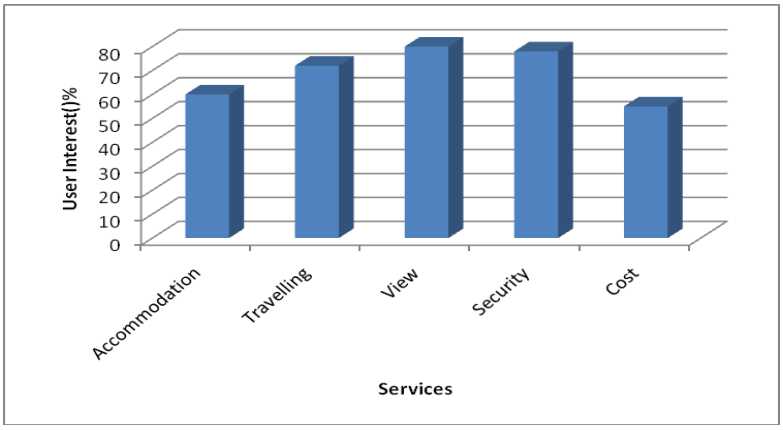

Figure 5: Tourism Analysis

Figure 5 shows tourism analysis. Based on the service the users interest will be varying.

ISSN: 0975-8585

Table 2: Cost Factor

Cost Factor	User Interest
10k	4000
20k	3500
30k	2500
40k	1000

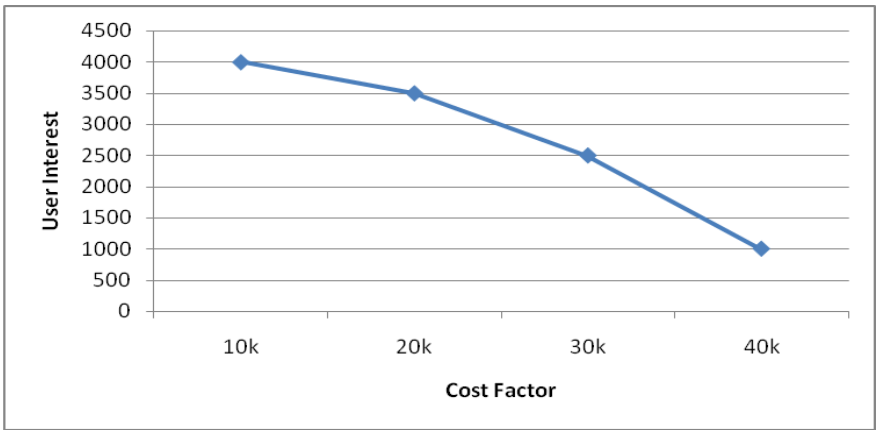

Figure 6: Cost Factor

Figure 6 shows Cost Factor. Based on cost factor user's shows the interest in tourism.

Table 3: Tourist Visits

TOURISTS		
Country of origin of visitor	**Number of visitor**	**Average length of stay**
Italy	51732	41
China	308452	47
UK	734244	34
USA	456084	24
Canada	109843	42
New Zealand	1075794	14

CONCLUSION

In this paper, we have explained about a structure of ontology building approach in tourism data realm which interconnected the linguistic intelligence data with context intelligence data, by making use of FCA. Under the framework of the proposed ontology building approach, knowledge engineers will be able to achieve a chunk of both context-oriented intelligence and linguistic data which have been described as helpful in supporting their ontology building tasks in connection with tourism data realm. In future, these proposed methods and algorithms can be implemented to integrate two or more different cancer specialized hospitals information around the major cities according to the user's individual disease like cancer, who are in search of cancer hospitals and their hospitals related information like location, visiting hours, contact number can be collected from different databases are uniformly represents the cancer hospital information based on user's individual disease using ontologies and semantic approach.

REFERENCES

[1] Lu RQ, World Scientific Publishers, 1992.
[2] Lin CY, Soo VW. Conference on Artificial Intelligence and Applications 2001; pp: 241-246.
[3] Wang M, Kao Y, Kuo Y, Lee C. Data & Knowledge Engineering 2007; 60(3): 547-566.
[4] Luo F, Khan L. IEEE International Conference on Tools with Artificial Intelligence 2002; pp. 122-127.

[5] Zhong N, Yoshinaga K, Terano T. International Conference on Knowledge-Based Intelligent Information Engineering Systems 1999; pp. 62-65.
[6] Zhang D, Booker QE, Zhou L. Hawaii International Conference on System Sciences 2002; pp. 957-965.
[7] Staab S, Maedche A. IEEE Intelligent Systems, 2001; 16 (2): 72-79.
[8] Velardi P, Navigli R, Missikoff M. IEEE Computer society 2002; 35 (11): 60-63.
[9] Gangemi A, Velardi P, Navigli R. IEEE Intelligent Systems 2003; 18 (1): 22-31.
[10] Staab S, Madche A, Hotho A. IJCAI-2001 Workshop Text Learning: Beyond Supervision 2001; pp. 48-54.
[11] Jensen PA, Andreasen T, Paggio P, Nilsson JF, Thomsen HE, Pedersen BS. 2004; 48: 199-219.
[12] Pulido G, Rafael J, Elliman D. European Conference on Object Oriented Programming 2001;
[13] Metais E, Lammari N. Data and Knowledge Engineering 2004; 48: 155-176.
[14] Ontology Definition, available: http://tomgruber.org/writing/ontology-definition-2007.html
[15] Huhns MN & Singh MP. Processes, Agents, Wiley, 2005;
[16] Spanakim, PoulicousPrastacosKavouras M, ElenTomai. International Workshop on Semantic-based Geographical Information Systems 2005; pp. 3-4.
[17] Christina Feilmayr, Robert Barta, Hannes Werthner ChristophGrun. Workshop on Context, Information and Ontologies 2009;
[18] Waralak Siricharoen V, Engineerng Education, Greece, 2007, pp 24-26.
[19] Guosun Zeng, Wei Wanag, Yu Huang, DongqiZhang, Xi aojunwang, YufengQiu, IEEE, Computer Society, 2008.
[20] VladanDevedzic, Danica Damljanovic. Chapter X, Springer, pp. 243-263.
[21] Yuxia Huang, Ling Bian, IEEE Transactions on Emerging Topics in Computing, 2015; 3 (2): 172-184.
[22] Senduru Srinivasulu, Sakthivel P, Contemporary Engineering Sciences 2013; 8 (5): 197-207.
[23] Senduru Srinivasulu, Sakthivel P, Ramya V. International Review on Computers and Software 2014; 10 (4): 415-423.