Research on ontology mapping of tourism information resources based on description logic

Jianhou Gan[1*], Bin Wen[2], Jinxu Li[2]

[1]*Key Laboratory of Education Informalization for Nationalities, Ministry of Education Yunnan Kunming, China*

[2]*School of Information Science and Technology, Yunnan Normal University, Yunnan Kunming, China*

Received 12 June, www.tsi.lv

Abstract

In this paper, the Ontology mapping of tourism information resources is discussed. Ontology mapping tries to find the corresponding relationships between two entities, and to achieve interoperability and information sharing in heterogeneous Ontology. We researched the Ontology mapping of tourism information resources and introduced the Description Logics to solve the Ontology mapping. The key relationships and determining relationships among Ontologies in description logic based Ontology mapping is first described. Then a description logic based mapping model is proposed, which can solve the problem of Ontology semantic heterogeneity.

Keywords: tourism information resources, resource integration, ontology isomerism, ontology mapping, description logic

1 Introduction

Information technology contributes greatly to the development of tourism. With the rapid development of global travel e-commerce, as a basic platform for tourism, tourism websites also play an important role in the development of travel e-commerce. However, most of the existing travel information sites have some shortcomings, such as do not update information timely, classifying information illogically, describing resources subjectively, and displaying content incompletely or with limited exhibition ways. Therefore, clients may misunderstand the information or even be misled by the information provided by the websites. The basic reason of these problems is the disorganized information. The current tourism information providers cannot provide interactive and participatory contents for the users; they are also lack of contents, which are systematic and integrative. Problems of the existing tourist information resources are scattered, for example, content sharing is usually poor, low development and utilization level, and inferior quality of services still exist. These problems seriously restrict further development of regional tourism development. In this article, based on Description Logic, we proposed an Ontology mapping model of tourism information resources to overcome the existing problems in tourism information applications [1-3].

An ontology is a "formal, explicit specification of a shared conceptualisation"[4]. An ontology provides a shared vocabulary, which can be used to model a domain, that is, the type of objects and/or concepts that exist, and their properties and relations. Ontologies are the structural frameworks for organizing information and are used in artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise bookmarking, and information architecture as a form of knowledge representation about the world or some part of it.

In this paper, we consider the tourism information resource ontology building way and built them, as well as put forward the ontology mapping way based on description logics, which better resolved the problem of tourism information resource ontology semantic heterogeneity.

2 Building tourism information resource Ontology

2.1 TARGET OF CONSTRUCTING TOURISM INFORMATION RESOURCE ONTOLOGY

Tourism involves six elements, such as catering, accommodation, sightseeing, transportation, shopping and entertainment. The target of constructing tourism information resource Ontology is to make tourism information integrative and intelligent. This will make online searching become easier for visitors, which can simplify and speed up the process of tourism trade, and also can provide more information about tourism services. The intelligent process can provide wider and more accurate travelling information and personalized services for users.

2.2 PROCESS OF CONSTRUCTING TOURISM INFORMATION RESOURCE ONTOLOGY

To develop Ontology of tourism information resources, several processes are included:

1) Determining fields and categories of tourism information resource Ontology;

* ***Corresponding author*** e-mail: kmganjh@yahoo.com.cn

2) Listing important terms, concepts in tourism information resource Ontology;
3) Establishing framework of tourism information resource Ontology;
4) Defining classes and classes hierarchy system;
5) Defining attributes slots and value types in classes;
6) Encoding domain Ontology and making it be formal.

2.3 DETERMINING SCOPE AND TERMS IN ONTOLOGY

Before building an Ontology, main contours of the target Ontology is described, which makes the purpose, scope, representation and application of the Ontology clear. The intermediate result is used to describe detailed instruction of Ontology. In this work, related terms and concepts based on determining range of Ontology are listed.

In tourism information resource Ontology, important terms and concepts include: people, organization agencies, scenic spots, travelling routes, transportation, accommodation, traveling, location, specialties, amusements, national customs, travel agencies, scenic spot management agencies, transport enterprises, accommodation enterprises, tourism bureau, insurance companies, specialty enterprises, entertainment industries, hydrological landscape, physiographic landscape, cultural landscape, history heritages, national intangible cultural heritages, full lines, pick lines and so on.

2.4 DEFINING CLASS AND CLASS HIERARCHY SYSTEM

Class is used to describe abstract entities objects, representing a kind of instances objects, which have common characters. Class has the property of inheritance, formed by hierarchical organization structure. Top class represents the most abstract entity concept, and subclass inherits abstract properties from their parents, which represent entities concepts, which are more specific or narrower than their parent. Adopted top-down approach to define class hierarchy, part of class hierarchy is described in Figure 1.

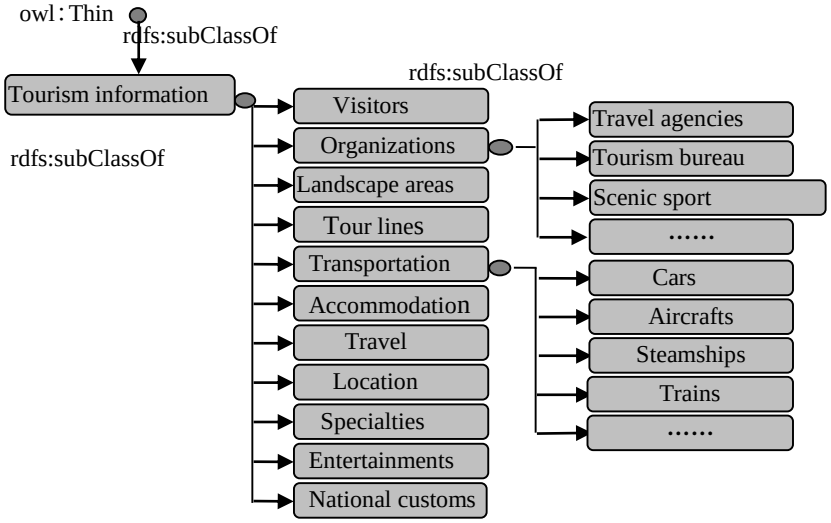

FIGURE 1 Part figure of class hierarchy

2.5 DEFINITION ATTRIBUTES OF CLASS

It is not enough to perform knowledge of field if only having name of each class. Therefore, attributes (properties) for each class are needed. Each class may have a large number of properties. The principle is to define properties of class in corresponding field based on demands. For example, the attributes of scenic spot and tourists can be expressed as:

Tourists (name, gender, ID, age, travel category, hobbies, phone, E-mail).

Scenic spot (name, attractions degree, management agency, category, address, capacity, phone)

2.6 ONTOLOGY ISOMERISM

Nowadays, Ontology has been wide used, and plays a vital role in many applications, such as database integration, P2P system, electronic commerce, Web service and social networking. As web is the decentralization for cross-cutting areas even in the same field. Therefore, there are always multiple individuals and there are differences produced among different Ontology, which is called heterogeneity [5]. Since multi-Ontology is needed to assist the work, Ontology isomerism will occur through Ontology mapping. Problems of sharing and reusing knowledge, between different Ontology will be solved.

3 Description Logic based on Ontology mapping

3.1 THE RAISE OF ONTOLOGY MAPPING PROBLEMS

Ontology mapping tries to find the corresponding relationships between two entities, and to achieve interoperability and information sharing in heterogeneous Ontology. Ontology mapping is one of the key technologies in Ontology, aims to achieve information

integration based on Ontology, where the most important process is to find semantic associations.

The task of Ontology mapping can be simply described as follows: suppose there are two Ontology, A and B, for every concept in A, we try to find a counterpart concept in B, which has the same or similar properties with the concept in A [6]. Ontology mapping is not a unified expression of Ontology and data, but to achieve conversion between instances according to semantic relation on conceptual level.

Ontology mapping process can be divided into three parts, which are finding mapping, expressing mapping and executing mapping. "finding" is to find related, similar concepts, properties and relationships from two different Ontology through the methods, which can be manual, semi-automatic or automated. "expressing" uses language to express previous mapping relationships; "executing" completes the conversion from source Ontology instances to target Ontology instances, based on mapping relationships.

The types of Ontology mapping are: mapping between Concept-Concept, Property-Property, Property-Concept and so on [7]. Mapping relationship between Concept-Concept, refers to two concepts in different Ontology expressing the same types of information, they are consistent; Mapping relationship between Attribute-Attribute refers to two properties in different concepts indicating the same information; Mapping relationship between Attribute-Concept, refers to the properties of a concept in Ontology express the same information with a concept in another Ontology [8].

Granularity of Ontology mapping have 1:1, 1:*n* and *n*:*m*. Mapping case of 1:1 is often called simple mapping, other cases are called complex mapping.

3.2 FRAMEWORK FOR ONTOLOGY MAPPING MODEL

When we input two isomerism Ontologies, establishing mapping relationship from the source Ontology to the target Ontology, this is the task of Ontology mapping system. Mapping process is an iterative process, as illustrated in Figure 2.

1) Feature extraction: Parsing document, then, extracting vocabularies of Ontology (including concepts, properties, relations and so on).

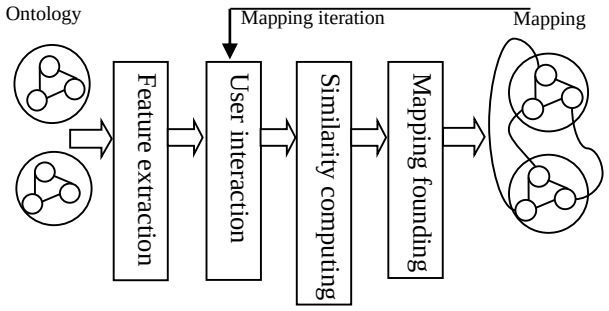

FIGURE 2 Framework for Ontology mapping model

2) User's interaction process (optional process): Ontology mapping system supports an optional user interaction process. Through interaction, users can pre-specify one or more mapping relations, before automatically mapping can also correct error mappings, which are found in mapping system, after automatic discovery of mapping, or create missing mapping relations. User's interaction actions will transmit mapping which impacting other associated elements, with the result, exerting an influence to mapping of the entire Ontology to achieve the purpose of improving the accuracy of mapping.

3) Similarity computing: Consider different methods to calculation according to differences between conceptual similarity and attribute similarity in conceptual similarity calculation. It starts from four aspects: name, attributes, structure and instances. However, attribute similarity calculation only comes from three aspects of matches, which are name, domain and range, to determine the similarity, in calculation process, the value of similarity is limited in [0,1].

4) Mapping discovery: Mapping discovery is based on similarity value after iteration. It aims to select the best mapping relations between elements of Ontology according to some selection strategies, constraints of Ontology and contextual.

5) Mapping: Algorithm outputs mapping table. Each item in the table corresponds to a mapping relation, each item contains four elements set: the first one is element set in source Ontology O_1 $\{e_{i_1}\}$, the second one is element set in target Ontology O_2 $\{e_{i_2}\}$, the third one is corresponding relationship between elements, and the fourth one is similarity values of relationship.

6) Mapping iteration: using the obtained similarity values to conduct iterative operation, to receive comprehensive predictive values of candidate mapping.

3.3 THE CONCEPT RELATIONS OF ONTOLOGY MAPPING

Concept relations in Ontology mapping include: equivalent, generalization, specialization, deviation, and intersection.

1) Equivalent: Assuming two concepts C and D, the interpretation in a certain domain and the corresponding interpretation set of C is equivalent with the corresponding interpretation set of D, so C and D are equal.

2) Generalization (Specialization): Assuming two concepts C and D, the interpretation in a certain domain and the corresponding interpretation set of C is a subset of corresponding interpretation set of D, called D is a generalization of C or C is a specialization of D.

3) Deviation: Assuming two concepts C and D, the interpretation in a certain domain and the corresponding interpretation set of C and the corresponding interpretation set of D has empty intersection , this is called C and D are in deviation.

4) Intersection: Assuming two concepts, the interpretation in a certain domain and the intersection of corresponding interpretation set of C and corresponding interpretation set of D is not empty, called C and D are intersection.

3.4 ONTOLOGY MAPPING BASED ON DESCRIPTION LOGIC

Description Logic is built on concepts and relations, it explains concepts as a collection of objects, and expresses relations as a binary relationship between objects. A Description Logic system contains four basic components: Construction set, which is used to express concepts and relationships. Tbox contains assertions, Abox instance assertions, Reasoning mechanism on Tbox and Abox. Choices and different assumptions on the above elements, decided capacity and reasoning ability of a Description Logic system[9].

A knowledge base K=<T,A> consists of two parts: Tbox T and Abox A. Tbox is a finite set which included assertions, has form: C⊑D, C and D are concepts, usually use C≡D as abbreviation of C⊑D and D⊑C, Abox is a finite set of instance assertions, form is C(a),C is a concept, a is the name of an individual, or the form is P(a,b).P is a primitive relation, a and b as names of two individuals [10,11].

Suppose K is a knowledge base. T is Tbox, C and D are concepts:

1) If exit an explanation ^{I} make $C^{I}\neq\varnothing$,then C is satisfied. I shows also known as model C.

2) With each explanation $\cdot^{I}$ if have $C^{I}\subseteq D^{I}$, then C contained in D, which denotes as C⊑D;

3) With each explanation $\cdot^{I}$ if have $C^{I}=D^{I}$.then C and D is equal, denoted as C≡D. that is C⊑D and D⊑C;

4) With each explanation $\cdot^{I}$ if have $C^{I}\cap D^{I}=\varnothing$,then, C and D are non-intersect [12].

Elements corresponding relationship in Ontology mapping can express by using Description Logic:

R_1: concepts C and D equal, if and only if C⊑D and D⊑C;

R_2:concepts C and D satisfy generalization, if and only if D⊑C;

R_3:concepts C and D satisfy specialization, if and only if C⊑D;

R_4: concepts C and D deviate, if and only if C⨅D⊑⊥;

R_5:concepts C and D deviate, if and only if C⨅D⋢⊥.

In the process of Ontology mapping, the above remarks can use following process to judge the relationship between concepts C and D:

Step1: Determine the inclusion relationship between C and D, there are four cases:

a) if C⊑D and D⊑C. according to <R_1>.C and D are equivalent;

b) if D⊑C. and C⊑D not established, according to <R_2>.C and D satisfy generalization;

c) if C⊑D. and D⊑C not established, according to <R_3>.C and D satisfy specialization;

d) if C⊑D and D⊑C not established. C and D are deviation or C and D satisfy intersection.

Step 2: The case of the previous step appears(4)can be considered;

i) if C⨅D⊑⊥ established, according to <R_4>.C and D are in deviation;

ii) if C⨅D⊑⊥ not established, C and D are intersection.

3.5 EXAMPLES OF ONTOLOGY MAPPING

In order to build Ontology in tourist information resource field, some concepts are refined.

The main top concept: people, organization agencies, scenic spots, travelling routes, modes of transportation, room and board, travel routes, location, specialties, amusements, national customs and so on.

Organization agencies have sub-concept: travel agencies, scenic spot management agencies, transport enterprises, accommodation enterprises, tourism bureau, insurance companies, specialty enterprises, entertainment industries and so on.

Scenic spots have sub-concept: hydrological landscape, climate biological king, physiographic landscape, cultural landscape, history heritage, other scenic spots and so on.

Travel routes have sub-concept: full lines and pick lines and so on.

Transportation methods have sub-concept: cars, trains, aircrafts, steamships, tour buses, ropeways and so on.

Description Logic expresses some information as defined below.

3.5.1 Describe concept of tourism knowledge

Tour agencies ⊑ organization agencies, scenic spot management agencies ⊑ organization agencies, transport enterprises ⊑ organization agencies, accommodation enterprises ⊑ organization agencies, tourism bureau ⊑ organization agencies, insurance companies ⊑ organization agencies, specialty enterprises ⊑ organization agencies, entertainment industries ⊑ organization agencies.

Hydrological landscape ⊑ scenic spots, climate biological king ⊑ scenic spots, physiographic landscape ⊑ scenic spots, cultural landscape ⊑ scenic spots, history heritage ⊑ scenic spots, other scenic spots ⊑ scenic spots.

Full lines ⊑ travel routes, pick lines ⊑ travel routes.

Cars ⊑ transportation methods, trains ⊑ transportation methods, aircrafts ⊑ transportation methods, steamships ⊑ transportation methods, tour buses ⊑ transportation methods, ropeways ⊑ transportation methods.

3.5.2 Describe relationship

Group (travel agencies, tours) by group (tours, travel agencies), management (scenic spot management agencies, scenic spots) be operating (scenic spots, scenic

spot management agencies), group = by group$^-$, management =be operating$^-$.

How to determine relationship between organization agency and management agency: scenic spot management agencies⊑organization agency is true. However, organization agency⊑scenic spot management agencies, is false. According to R_2, organization agency and scenic spot management agency have generalization relationship. Scenic spot management agency ⊑organization agency is true. But organization agency⊑scenic spot management agencies, is false. According to R_3, organization agency and scenic spot management agencies have specialization relationship.

How to determine relationship between hydrological landscape and car: hydrological landscape ⊑car is not true. At the same time, car⊑hydrological landscape is also untrue. Hydrological landscape⨅car ⊑⊥ is true. According to R_4, hydrological landscape and car have deviation relationship.

How to determine relationship between hydrological landscape and history heritage: hydrological landscape ⊑history heritage is not true. And history heritage ⊑ hydrological landscape is also untrue. And hydrological landscape ⨅ history heritage ⊑⊥ is untrue. Then, hydrological landscape and history heritage have Intersection relationship.

4 Conclusions

Integrating and synchronizing tourism information resources make it become orderly and standardized, which achieves the purposes of services to the tourism industry and promotes economic development. This is the target of information retrieval, knowledge management and knowledge services. In the future, the development of travel information resources services platform will be proceeded, which is based on Description Logic.

Acknowledgements

The work is supported by the National Natural Science Fund Projects (61262071), the General Project of Social Development in Yunnan Province (2011FB044), the humanities and social science research projects of the Ministry of Education (12YJCZH053).

References

[1] Yuan Z 2008 The discussion about integrating resources of regional tourism information *Market Modernization* **21** 248-9 *(in Chinese)*
[2] Lu Y, Kaibo T 2007 Personalized retrieval of tourism information resources based on agent *Researches in Library Science* **3** 71-3 *(in Chinese)*
[3] Shen L 2008 The application of Ontology in tourist information resources sites *Library Information Service* **32**(9) 110-2 *(in Chinese)*
[4] Gruber T R 1993 A translation approach to portable ontology specifications *Knowledge Acquisition* **5**(2) 199-220 *(in Chinese)*
[5] Hu W, Qi Y 2007 Ontology matching: Set up heterogeneous Ontology *Computer World* **45** B12 *(in Chinese)*
[6] Ehrig M, Sure Y 2004 Ontology Mapping-An Integrated Approach *Proceedings of the 1st European Semantic Web Symposium* **3053** of Lecture Notes in Computer Science Christoph Bussler Greece Springer Verlag 76-91
[7] Maier A, Aguado J, Bernaras A, Laresgoiti I, Pedinaci C, Peña N, Smithers T 2003 Integration with Ontologies *Proceedings of WM2003* Luzern 1-15
[8] Zhang Y 2005 Research of Ontology mapping in data integration *Central South University* 12-6 *(in Chinese)*
[9] Dong M 2006 A Description Logic with default reasoning *Chinese Journal of Computers* **26**(6) 729-36 *(in Chinese)*
[10] Gan J, Xu T 2007 Semantic web and its application-based on Ontology, Description Logic, Semantic networks *Yunnan Science and Technology Press* 78-85 *(in Chinese)*
[11] Gan J, Wen B 2008 The research of conceptual modeling based on Description Logic *Journals of Yunnan Normal University* **28**(6) 13-6 *(in Chinese)*
[12] Gao J, Zhu M, Ming H 2004 Foundation of artificial intelligence *Higher Education Press* 231-5 *(in Chinese)*

Authors

Jianhou Gan, born on August 13, 1976, Yunnan, China

Current position, grades: Ph.D candidate of Metallurgical Physical Chemistry, professor and supervisor of master student in Yunnan Normal University.
University studies: master's degree in Basic Mathematics from the school of information Yunan Normal University in 2004.
Scientific interest: education informalization for nationalities, semantic web, database, intelligent information processing.
Publications: more than 40 papers.
Experience: on education informalization for nationalities, semantic Web, database, intelligent information processing.

Bin Wen, born on March 15, 1978, Yunnan, China

Current position, grades: Ph.D. of Application Technology of Computer, Lecturer in Yunnan Normal University.
University studies: Ph.D. degree in Application Technology of Computer from School of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing in 2013. Master's degree in Computer Software and Theory from the school of information, Yunan Normal University in 2005.
Scientific interest: intelligent information processing, emergency management.
Publications: more than 20 papers.
Experience: a strong experience on intelligent information processing and emergency management.

Jinxu Li, born on October 11, 1972, Yunnan, China

Current position, grades: Master of Computer Software and Theory, lecturer in Yunnan Normal University.
University studies: Computer Software and Theory in Yunnan Normal University 2005.
Scientific interest: information safety, intelligent information processing
Publications: more than 10 papers
Experience: a strong experience on information safety and intelligent information processing.