

Web ontology language applied to the tourism sector

Lenguajes de ontologías web aplicados al sector turístico

Ing. Alexis Messino Soza¹, Ing. Luis C. Tovar Garrido¹

1- Ingenieros de sistemas. Grupo de investigación IngeBioCaribe.
amessino@unisimonbolivar.edu.co

Recibido: 20/04/2010, Aceptado: 15/06/2010

RESUMEN

El presente artículo describe los resultados del proyecto de investigación titulado “Plataforma TIC para propiciar la organización y colaboración de los diferentes actores del sector turístico en el Caribe Colombiano” y desarrolla algunos conceptos sobre aplicaciones de Ontologías Web al sector turismo. Una ontología define los términos a utilizar para describir y representar un área de conocimiento. Las ontologías son utilizadas por personas, bases de datos y aplicaciones que necesitan compartir un dominio de información común e incluyen definiciones de conceptos básicos del dominio, y las relaciones entre ellos. El artículo inicia con el análisis del concepto de ontologías, los componentes que lo conforman y algunos lenguajes de ontologías web frecuentemente utilizados y explica cómo se pueden aplicar las ontologías al sector turístico. El artículo finaliza explicando cómo se construyó un buscador de atractivos turísticos basado en ontologías.

Palabras Clave: Ontología, OWL, Turismo, Buscador, Conocimiento.

ABSTRACT

This article describes the results of the research project entitled “ICT platform to promote the organization and collaboration of different actors in the tourism sector in the Colombian Caribbean” and develops some concepts about Web Ontology applications for tourism. An ontology defines the terms to use to describe and represent an area of knowledge. Ontologies are used by people, databases and applications that need to share a common information domain and include definitions of domain concepts and relations between them. The article begins with an analysis of the concept of ontologies, its components and some web ontology languages commonly used and explains how ontologies can be applied to the tourism sector. The article concludes by explaining how was built a seeker engine for sites based on ontologies.

Key words: Ontology, OWL, Tourism, Searcher, Knowledge.

1. Introduction

Speaking of ontologies is impossible to think of the semantic Web, which is based on the idea of adding semantic metadata and ontology to the world wide web. That information, -which describes the content, meaning and the relationship between data- must be provided in a formal way to human agents or automatons responsible for its processing.

Nowadays, Internet is a technology for daily use by all actors in society, comparable coverage to other traditional media such as radio, television or telephone, which are exceeded in many aspects. Internet is now an extraordinarily flexible and economic way of communication, trade and business, and entertainment, access to information and service offerings, among others. Internet's growth is the cause of the technologies on which it is based have undergone rapid evolution.

Some of these technologies, which enable the semantic Web, including languages for representing ontologies, parsers, query languages, development environments, management modules (storage, access, update) of ontologies, display modules, ontologies conversions, and other tools and libraries.

The paper is organized as follows: initially it defines some of the basic concepts. It then provides an overview of the use of OWL[®] (Web Ontology Language) in the tourism sector. Subsequently, we explore the implementation of a search engine based on ontologies and finally, conclusions are presented.

2. Basics

2.1. Ontology

In the context of information, a definition of "ontology" widely adopted is the following: An ontology is an explicit and formal specification of a shared conceptualization.[1].

The main function of an ontology is to represent concepts formally attributable to the application domain, ie, the portion of reality under consideration, the existing knowledge. Specify the conceptualization means assigning unambiguous meaning to the definition of knowledge in a specific domain. The shared word reflects the fact that the ontology captures consensual knowledge, or accepted by a group.

An ontology isn't a knowledge of the objective world, but represents the view shared by a group [2],[3]. It aims to

eliminate or at least reduce, the confusion of concepts and terminology, to have a shared understanding [3], [4].

Usually an ontology is composed by:

- Concepts that describe a specific application domain, such as tourism, the subject of this document.
- Definitions of concepts, through a formal language or natural language.
- The relations between the concepts expressed through a formal language. Relationships can be of several types: taxonomy (a), meronymy (part of), associations, etc.
- Generated concepts that express the main categories that is the world, which is the highest possible level of an ontology of the world. It happens in general, domain independent.

An ontology can occur in various forms and have different levels of formality, but always includes a vocabulary of terms with the description of its meaning [5].

2.2 Components of the Ontology

Ontologies have the following components used to represent some domain knowledge [6]:

- **Concepts:** These are basic ideas that try to formalize. The concepts can be classes of objects, methods, plans, strategies, thinking processes, etc.
- **Relations:** represent the interaction and liaison between the concepts of the domain. They tend to form taxonomy of the domain.
- **Functions:** they are a specific type of relation where one element is identified by calculating a function that considers several elements of the ontology.
- **Instances:** used to represent objects of a certain concept.
- **axioms** are theorems that state relations to be satisfied by elements of ontology. For example: "If A and B are class C, then A is not subclass of B", "For every A satisfying the condition C1, A is B", etc.

These latter components -the axioms- permit, along with the legacy of concepts, knowledge inference that is not explicitly stated in the taxonomy of concepts.

Figure 1. OWL ontology language DL
Figura 1. Ontología en el lenguaje OWL DL

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:lugaresGeograficos="http://www.owl-ontologies.com/Ontology1218023717.owl#"
  xmlns:transportes="http://www.owl-ontologies.com/Ontology1218023205.owl#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:proveedores="http://www.owl-ontologies.com/Ontology1218019535.owl#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:precios="http://www.owl-ontologies.com/Ontology1218024004.owl#"
  xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns="http://www.owl-ontologies.com/Ontology1218125991.owl#"
  xml:base="http://www.owl-ontologies.com/Ontology1218125991.owl">
  <owl:Ontology rdf:about="">
  <owl:imports rdf:resource="http://www.owl-ontologies.com/Ontology1218023717.owl"/>
  </owl:Ontology>
  <owl:Class rdf:ID="Republicanos">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
  >representa todos los sitios arquitectonicos que son republicanos.</rdfs:comment>
  <rdfs:subClassOf>
  <owl:Class rdf:ID="SitiosArquitectonicos"/>
  </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="CentrosTecnicosCientificos">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
  >representa todos los sitios que son tecnicos-cientificos</rdfs:comment>
  <rdfs:subClassOf>
  <owl:Restriction>
  <owl:onProperty>
  <owl:DatatypeProperty rdf:ID="horario"/>
  </owl:onProperty>
  <owl:minCardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
  >1</owl:minCardinality>
  </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
  <owl:Class rdf:ID="sitiosInteres"/>
  </rdfs:subClassOf>
  </owl:Class>
```

2.3. Web Ontology Language

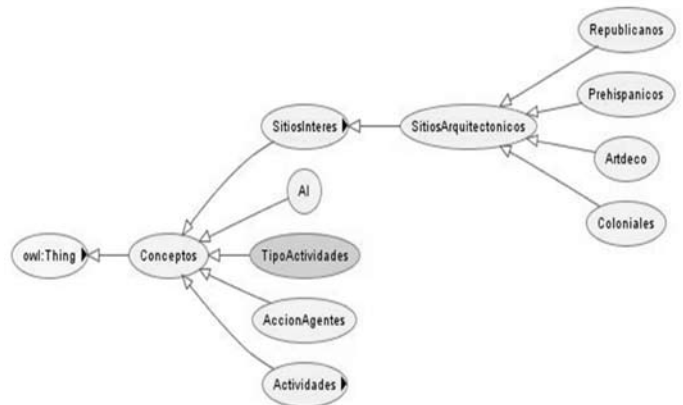
Ontologies can be developed in different languages like RDF ®, RDF SCHEMA ®, XML ®, XML SCHEMA ® and OWL ®. The latter is the most widely used because of their greater capacity for expressing meaning and semantics, OWL ® goes beyond the other languages mentioned in its ability to represent a computer interpretable content on the Web [2].

Currently, OWL ® has three versions ® OWL Lite, OWL DL and OWL Full ® ®; these versions incorporate different functionalities. OWL Lite ® is constructed so that any sentence can be resolved in finite time, the most complete version of OWL DL ® may contain infinite “loops” and is based on descriptive logic, while OWL Lite ® is based on a less expressive logic [2]. An example of OWL ® generated code can be seen in Figure 1.

3. Applied Ontology Tourism Sector

There is research on how we describe the concepts related to the field of tourism in an ontology. It is clear that this area is broad and the boundaries are not well defined. The purpose of a tourism ontology is to facilitate the description of the attractions that are offered as a web service, therefore, the concepts that have been chosen to be considered are primarily hunters sights, hotels, tour operators and activities that can be done in a certain place [7], as shown in the diagram in Figure 2.

Figure 2. Turistic Concepts of an Ontology
Figura 2. Conceptos de una Ontología Turística



In regard to the discovery of contents to identify concepts in the field of tourism ontology, we studied the entire “life cycle” of a trip considering all the needs of a tourist. For this reason, we analyzed the vehicle booking services and facilities where tourists can stay, also services such as weather forecasting, route calculation, information about countries and cities, among others.

Some difficulties are related to certain concepts of tourism that are not well defined, for example, with regard to accommodation, in some countries there is no national law, but regional standards.

To resolve these difficulties were consulted the most authoritative sources identified and extracted concepts that were common [8].

During the identification and analysis of related concepts in the tourism domain, it concluded that it is not possible to isolate these other concepts that have to do with the so-called “common sense” and are independent of particular domain concerned. Examples are: those related to time (day and date), the location of the space (place, address, etc.), Among others.

That familiarity can be designed and connected in a different way. If we analyze the different meanings of the terms used in English to describe them, it is found that a term that seems intuitively one-way detail in the description of different concepts is different from one another. It then becomes necessary to make decisions, trying to identify between the different interpretations, which are most relevant to the field of tourism [8].

The meaning of concepts is mainly expressed through the definitions in natural language. For this purpose you can use OWL DL ®. Through this language expresses the relationships between different concepts. The use of an upper ontology, especially when there is extensive use of multiple inheritance, has the advantage of providing a manual navigation of the ontology and therefore simplify the recovery of concepts that can serve as a description. The construction of an ontology is a task even more delicate than the choice of domain concepts [7], [8].

4. Building a Search Engine Based on Ontology Attractions

The traditional way to find search engine results is through robust text comparisons made between the sites html pages already stored and indexed in its database (if the website is not very important, it only saves one to two pages); less complex procedures search engines use the timely information sites (titles, keywords) as a criterion of result or the existence of a domain, ie, how many times it is accessed (implemented by Google). This way of handling the search without any ranking, but only with a reference plain text, creates inaccurate results in many cases [9].

The ontological search based on conceptual schemes generate consistent results with the user's criteria on one or more domains of information. The domain is the set of concepts defined by the ontology. Here are the steps followed for the implementation of an information domain created under the OWL ontology language on Colombia's Caribbean tourist attractions in the web search portal "Somos Caribe Colombiano... Magica realidad" ("We're Colombian caribbean... Magical reality") (www.somoscaribe.com) and the results obtained from this experience.

4.1. Understanding the Concept Domain Information

Understanding the domain of the concept of information is the most important part in the creation of a web ontology, in this case to have the knowledge base (ontology and instances of classes) already created, you must understand the conceptual framework to use. The ontologic editors allow to understand the concept domain easily and clearly. There are various editors (swoop ®, semantic turkey ®, Sinaptyc ®, among others.), the tourist portal www.somoscaribe.com used the editor PROTEGE ® created by Stanford University (<http://protege.stanford.edu/>), which provides great support and it is built on an open platform.

PROTEGE ® delivers different ways to visualize the concepts and relation. These include: hierarchical tree of classes and subclasses of the ontology (Figure 4), plugins as Jambalaya ® and OWLViz ® that graphically show our domain (Figure 5, Figure 3).

Figure 3. Ontology Created using PROTEGE 3.4
 Figura 3. Ontología creada en PROTEGE 3.4

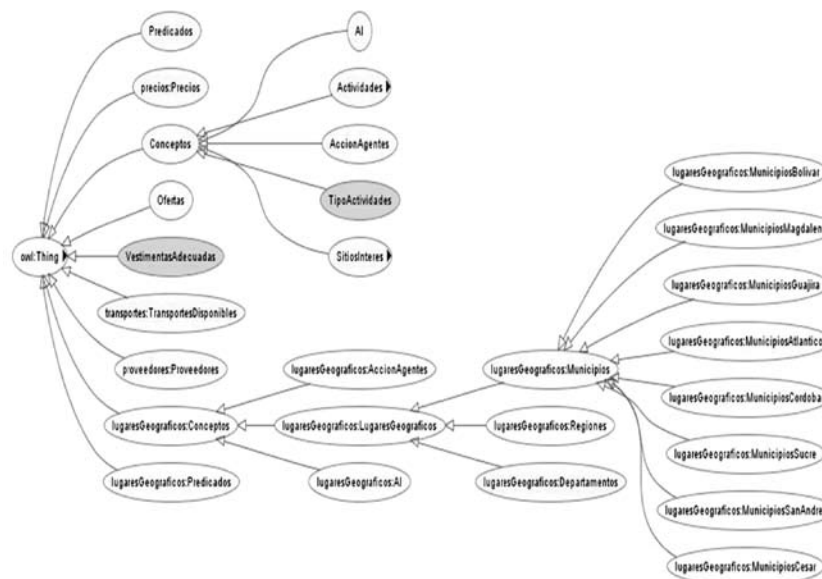
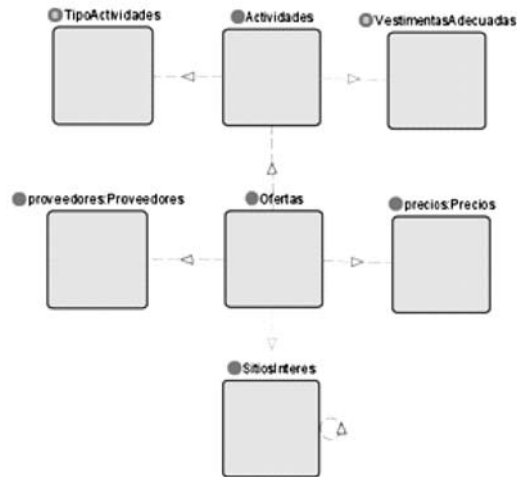


Figure 4. Hierarchical tree, ontological editor PROTÉGÉ®
 Figura 4. Árbol jerárquico, editor ontológico PROTÉGÉ®



Figure 5. Domain Graphics “atractivos turísticos” con Jambalaya® OWLViz®
 Figura 5. Gráficos del dominio “atractivos turísticos” con Jambalaya® y OWLViz®



Once well-understood the domain, we must limit the important concepts to achieve the desired objective.

4.2. Defining relevant concepts of the searches

Properly define what the essential search criteria are is important to create the ontological form. Placing trace data for properties that are irrelevant or vague can create inconsistent results. First of all, the purpose of the ontology must be clear. In the case of the tourist platform, its purpose is to show places of interest related to the name, municipality or department where they are located and its

attractive features. This allows to define relevant criteria for your search:

1. Name of tourist attraction.
2. Tourist attractions in a municipality (name of municipality).
3. Tourist attractions in a specific department (name of the department).
4. Tourist attraction synonyms (used in cases like Barranquilla and whose synonym is Puerta de Oro).
5. A brief deescription of the tourist attraction.

These are the 5-managed search criteria in the search for attractions and they are classified according to a range of importance. Within the ontology of tourist attractions are other properties of the instance (Expanded descriptions, date and time of care) that are not that important or they can create less accurate results.

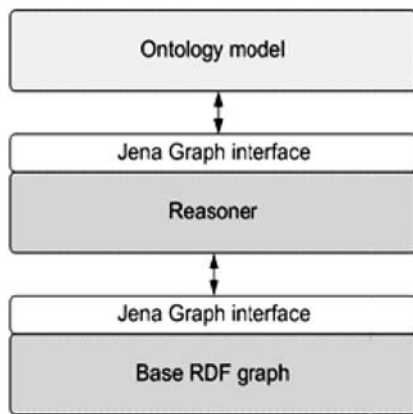
4.3. Using Tools for Ontological Management

The management of the ontology should be done through a tool to create the conceptual schema in memory, to reasoning, data query, modify and create new relation or concepts, among other processes. An appropriate Framework is JENA[®] from Java platform, which in its API (application programming interface) it contains different types of reasoners and various types of variables for the management of ontologies, such as ontmodel containing the ontology loaded into memory, or ontProperties for properties. It also enables the use of query languages in the semantic web (SPARQL[®]), as shown in Figure 6.

4.4. Query Generation through language SPARQL

To manage and query data in an ontology was used SPARQL[®] (SPARQL Protocol and RDF Query Language). The basic building block of a query SPARQL[®] is a triplet pattern. A triplet can have variables in place of springs in any of these three positions, subject, predicate and object. The variables are shown as symbols preceded by a special character: ? [10].

Figure 6. Ontology and reasoners in JENA
 Figura 6. Ontología y razonadores en JENA



An example of the ontology of turistic sites is:

```
PREFIX st: <http://purl.org/dc/elements/1.1/>
SELECT ?Names
WHERE (
?sites st: name ?Names
)
```

You can see the syntax used by SPARQL[®]. The PREFIX is used to store a URL that will reference an ontology, in this case is on the tourist attractions and have the prefix st. SELECT is the data to be returned by the search, in our case all the names of the attractions or tourist sites. WHERE are the conditions to be met for data to be in the results.

In general, you can define the query as: what objects (? Sites) that are within the ontology of sites have a name (dc: name) either (said to be either because the property is preceded by a variable ?Names). The result could be:

Nombres
"El Rodadero"
"Parque Tayrona"
"El Hoyo soplador"

The results from SPARQL[®] are just a literal type, ie only the data that have been inserted to instances of classes.

5. Conclusions

It could be observed that the tourism sector is an area in which ontologies can be applied anywhere. Its use is suitable for the development of automated services to users such as the specific case of tourism services.

There are currently several systems and prototypes that offer tourist services such as restaurants or museums. In these cases the services are usually treated independently and not integrated, as it would be if we use web ontologies.

The ontology-based search is an important step for generating more accurate results and higher levels of accuracy in the information, this applies to tourist domains. The use and investment in the development of ontology languages promises a convenient future [11].

Remarkably, we can say that there are well-developed ontologies in the internet created and improved by different people, which can be used for multiple purposes, the proposal is to organize a single universal domain and that contains other domains together in order to create a complete scheme of all the information on the web.

The main advantages of this new Internet revolution will be the development of applications with common data schemas and finding information with inference. To achieve these purposes is required to unify the semantic content using ontologies to formalize this knowledge in a consensus and reusable way. We need a web-based common language with sufficient expressive and reasoning power to represent the semantics of ontologies.

References

- [1] Thomas, G. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies* Volume 43, Issues 5-6, pp. 907-928.
- [2] Hendler, J. (2004). Frequently Asked Questions on W3C's Web Ontology Language (OWL).
- [3] Guizzardi, G. (2007). On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta) Models. Federal University of Espírito Santo (UFES), Vitória, Brazil. Laboratory for Applied Ontology (ISTC-CNR), Trento, Italy.
- [4] Meersman, R; Spyns, P, and Jarrar, M. (2002). Data modeling versus Ontology engineering. SPECIAL ISSUE: Special section on semantic web and data management. Vol 31, pp. 12-17.
- [5] Fernandez, M, and Gomez, A. (2003). The integration of OntoClean in WebODE. Faculty of Informatics. Universidad Politécnica de Madrid, pp. 2-8.
- [6] Honrubia, F. (2002). Introduction to Ontologies. Escuela Universitaria Politécnica de Albacete.
- [7] Batzold, M, Navarro J, Botti, V. et al. (2004). Development of tourist services to users. I Sistemas Informatics Department of Computing. Universitat Politecnica de Valencia, Spain.
- [8] Pandolfo, C. Torelli, F. (2004). Tourism domain ontology. Fund per la Ricerca di investimenti Base (Firbas). Programma per la Società abilitanti Stratego Tecnologie della conoscenza-ICT.
- [9] Cruz, I. Decker, S; Euzerat, J McGuinness, D (2002). The Emerging Semantic Web. City USA, pp. 154-155.
- [10] Allemnag, D. and Hendler, J (2008). Semantic web for the working ontologist. United States, pp. 85-87.
- [11] El mundo. is (2008). Twenty Brussels wants to lead the development of Web 3.0. Retrieved October 15, 2009 in <http://www.elmundo.es/navegante/2008/09/30/tecnologia/1222771475.html>
- [12] Lladó, E, Salamanca, I. (2004). Peer-To-Peer Knowledge Management in Tourist Destinations. IBIT Foundation. Palma de Mallorca (Balears, Spain).