An ontology-based system for intelligent matching of travellers’ needs for Group Package Tours

Dimitris N. Kanellopoulos

Department of Tourism Management,
Technological Educational Institute of Patras, Greece,
Meg. Alexandrou 1, GR 26334, Patras, Greece
E-mail: d_kan2006@yahoo.gr

Abstract: In recent years, the package tour industry has fostered a dependency on technology for its operational and strategic management. However, the existing tools for facilitating searches of availability of package tours do not provide intelligent matching between offered package tours and the personal requirements of the travellers. The objective of this research work was to design an intelligent web portal to serve as service provider in tasks related with package tours. This portal aims at helping people living in Europe to find package tours that match their personal travel preferences. For this purpose, the knowledge of the package tour domain has been represented by means of ontology. Additionally, the ontological component allows for defining an ontology-guided search engine, which provides more intelligent matches between package tours offers and travelling preferences.

Keywords: ontologies; Group Package Tours; GPT; travel agencies; web portal; semantic web.


Biographical notes: Dimitris N. Kanellopoulos holds a PhD Degree in Multimedia Communications from the Department of Electrical and Computer Engineering of the University of Patras, Greece. Currently, he is with the Department of Tourism Management at the Technological Educational Institute of Patras. He is also a member of the Educational Software Development Laboratory (ESDLab) in the Department of Mathematics at the University of Patras. His research interests include: multimedia communications, intelligent information systems, knowledge representation, web engineering and web-based education. He has authored many papers in international journals and conferences at these areas. He serves as a member of the editorial boards in several academic journals.

1 Introduction

Nowadays, semantic models are very significant in web applications, while many semantic web related applications and languages are being created. The semantic web is an extension of the current web in which information is given well defined meaning, better enabling computers and people to work in cooperation (Berners-Lee et al., 2001).
An ontology-based system for intelligent matching of travellers’ needs

It argues for a set of technologies and techniques that integrates artificial intelligence into the core of the web. In the semantic web, we can create complex applications such as intelligent browsers, intelligent software agents, global databases with data from the web etc. The semantic web provides enhanced information access based on the exploitation of machine-processable metadata.

Central to the vision of the semantic web are ontologies (Fensel and Musen, 2001). The term ‘ontology’ is derived from the Greek words ‘onto’, which means being, and ‘logia’, which means written or spoken discourse. Gruber (1993) defines an ontology as an explicit specification of a conceptualisation. Ontologies provide a shared understanding of a domain of interest to support communication among human and computer agents, typically being represented in a machine-processable representation language (Maedche and Staab, 2001). Ontologies present two main advantages, namely, shareability and reusability. These properties make them very attractive and powerful for representing domain knowledge. Consequently, the knowledge they contain can be used in different applications, for different purposes and by different people.

Currently, the continued rapid growth in package tours-related information volume makes it increasingly difficult to find, organise, access, and maintain the information required by consumers. At the same time, several travel web services are provided to serve many consumers who frequently navigate the web. Many of those services are developed by means of web portals, which have become easy and useful tools to capture and manage huge volumes of travel-related information. However, current web portals, which provide package tour information, are not able to answer queries concerning the contents of documents, not to mention presenting or combining facts which are found in different web documents. In particular, these portals do not make use of the advantages provided by semantics-related techniques.

In this work, a public web portal developed by using semantic web technologies (i.e., ontologies) is presented. A semantics-based design has been applied to the development of this web portal. Specifically, an ontology has been built to guide and support the development of that portal. This facilitates the implementation of a semantic matching engine in the application between offered Group Package Tours (GPT) and tourists’ desires. For this purpose an ontology-guided intelligent search has been used, enabling access to package tour offers that match user preferences. In particular, users interact with concepts and relations embodied in ontologies in a dialogue process that can be interpreted as a query. The ontology-guided intelligent search engine uses the ontology-enabled search tactics (e.g., super, sub, relate, contrary, record, select) described by García and Sicilia (2003). The referred web portal will facilitate the access to the package tour industry, especially for users from Europe. Actually, this web portal is a meeting point for tourists and travel agencies offering package tours in Europe. We focus on the development of a website, which can serve as a service provider in package tours-related issues. The goal of this portal is to enable travellers living in Europe to search for GPT. For this purpose, some requirements were defined in line with the package travel regulations 1992 (Grant, 1996).

The structure of the remainder of this paper is as follows. Section 2 analyses critical service features in GPT provided by travel agencies. Section 3 presents some issues on ontological engineering and the use of ontologies in web portals. Section 4 describes the ontological model used in our work. Section 5 presents the implemented domain ontology, and Section 6 discusses the design of the web portal. Finally, Section 7 draws a number of conclusions and offers directions for further work.
The package tour industry and the travel recommender systems

2.1 The package tour industry

In general, there are two types of package tours:

- a basic package tour which includes only transportation and accommodation
- an all-inclusive package tour which is defined as a trip planned and paid for in advance at a fixed price.

An all-inclusive package tour covers commercial transportation and accommodation, meals and sightseeing, and, sometimes, the services of an escort or guide. The ontology-based system we propose uses all-inclusive GPT.

With the popularity of GPT, potential profits accrue for travel agencies. According to the travel industry council (TIC, 2006), there are approximately 4200 travel agencies in Europe. The market is highly competitive. Nevertheless, all the package tours presented on the market are very similar. Itineraries offered by travel agencies possess high degrees of similarity with reference to entertainment and food. Besides, travel agencies are not only competing with other travel agencies in the market, their products are competing in the same product line or for the same destination.

Davies and Downward (2006) presented the results of a questionnaire research investigation to generate practitioners’ insights into the competitive behaviour of the UK package tour industry. Considering the advertisement domain, group package tourists react to the different combinations of advertising components. Wang et al. (2007b) identify the optimal numbers and combinations of advertising components on travel web pages and their contributions to advertising effectiveness, while Wang et al. (2002) examined how the effectiveness of advertising would be changed if tour leaders were used as endorsers for group package tour brochures.

The propensity of travellers to choose a GPT is somehow related to: travel risks; financial considerations; and the attractiveness of the package tour (Sheldon and Mak, 1987; Roehl and Fesenmaier, 1992; Tsaur et al., 1997). In the GPT decision-making process, typically, tourists have to take various factors into consideration while planning a GPT, such as: kind of GPT, destination country, departure date, optional tour, cost and length of tour, activities (sport and non-sport activities), travel risk, travel agency, tour guide etc. Wang et al. (2000) conducted an exploratory research to identify the critical service features of the GPT. Wong and Knong (2004) identified important selection factors for all-inclusive package tours as perceived by outbound tourists. By using factor analysis, they derived eight dimensions for package tours selection, namely, tour arrangement and service quality, attractions, hotels and airlines, TV promotions and customer care, routing, personal interests, word-of-mouth, and time. Wang et al. (2007a) developed and empirically validated an instrument (scale) which measures the GPT service. They employed multistage steps for investigation, and they used qualitative and quantitative approaches so as to develop a new Customer Comment Card (CCC) for GPT. Undoubtedly, travel-related behaviours play an important role in the package tours selection. Wang et al. (2004) examined how family role relationships, inclusive of parents and children, vary over decision-making stages for the GPT.
2.2 Travel recommender systems

Recommender systems in the travel industry reduce the burden of information overload and domain complexity for users. Franke (2003) considered recommendation and customisation techniques for tourist information systems. Hybrid recommendation strategies combine user preferences (content-based or social filtering) with knowledge and/or utility-based preferences (Ricci and Werthner, 2001). Loban (1997) presented a methodological approach to the implementation of a counselling system for travelling. His approach includes mathematical formulations for problems related to package tour selection. In general, an effective online recommender system must be based upon an understanding of consumer preferences and successfully mapping potential products onto the consumer’s preferences (Adomavicius and Tuzhilin, 2005). Gretzel et al. (2004) and Pan and Fesenmaier (2006) argue that this can be achieved through an understanding of how consumers describe, in their own language, a product, a place, and their experience when consuming the product or the place. However, making recommendations for travel and hospitality products in an online system is considered a more complex task than general consumer goods (Ricci, 2002; Fesenmaier et al., 2006). First, travel involves bundling of a large number of heterogeneous products and services and, thus, applying the content-based approach requires extensive domain knowledge to be built for the particular application. Second, because the consumption of travel products is individual-based and context-specific, it cannot be assumed that two identical trips will result in the same experience, even if two travellers go to the same destination and visit the same attractions. The two most successful recommender system technologies are Triplehop’s TripMatcher (used by www.ski-europe.com among others) and VacationCoach’s expert advice platform Me-Print (used by Travelocity.com). The DieToREcs recommender (Fesenmaier et al., 2002) is a case-based travel planning system that supports multiple decisions styles by letting the user enter the system through three main doors: iterative single-item selection, complete travel selection, and inspiration-driven selection. Hopken et al. (2006) proposed the etPlanner framework for a comprehensive support for travellers during all phases of their trip. The etPlanner provides high quality recommendations and efficient adaptation to the variety of mobile devices on the market.

Recently, semantic web technologies have been offering new possibilities in the area of knowledge management applied to the travel industry (Kanellopoulos et al., 2006). For instance, Kanellopoulos et al. (2004) designed a novel P2P management system of semantically enriched web travel plans in order for these to become manageable, effective and adaptive to the users’ needs. Over the last years, some researchers (Ricci, 2002; Berka and Plößnig, 2004; Kanellopoulos and Kotsiantis, 2007; Ponnada and Sharda, 2007; Kanellopoulos and Panagopoulos, 2008) have proposed intelligent Destination Management (DMS) and recommender systems. For example, the layered adaptive semantic-based DMS project (LA_DMS) provides semantic-based tourism destination information by combining the P2P paradigm with semantic web technologies (Kanellopoulos and Panagopoulos, 2008). In the LA_DMS project, a metadata model is proposed that encodes semantic tourism destination information in a RDF-based P2P network architecture to be adaptive to travellers’ preferences. The LA_DMS model combines ontological structures with information about tourism destinations and peers. Cardoso (2006) designed the architecture of a semantic dynamic packaging system based entirely on semantic web technologies. It is worth mentioning that dynamic packaging
enables consumers (or travel agent) to build a customised itinerary by assembling multiple components of their choices. Dynamic packaging produces one reservation, completes the transaction in real time and entails only one payment from the consumer. The idea of dynamic packaging is compatible with the idea of an all-inclusive package tour, where a package tour is paid for at a single price.

3 The role of ontologies

According to Hendler, an ontology is

“a set of knowledge terms, including the vocabulary, the semantic interconnections, and some simple rules of inference and logic for some particular topic.” Hendler (2001, p.30)

Existing ontologies are classified into four major categories:

1. meta-ontologies
2. upper ontologies
3. domain ontologies
4. specialised ontologies.

Ontology languages such as RDF, RDFS, DAML+OIL, (DAML, 2001), OWL (Web ontology language: http://www.w3.org) are actually meta-ontologies themselves; and their instances are semantic web ontologies. Upper ontologies provide a high level model of the world using the ontology constructs provided by meta-ontologies. It is well known that OpenCyc can be used as the basis of a wide variety of intelligent applications. OpenCyc (2006) is the open source version of the Cyc technology, the world’s largest and most complete general knowledge base and commonsense reasoning engine. The entire Cyc ontology, containing hundreds of thousands of terms, along with millions of assertions relating the terms to each other, forms an upper ontology whose domain is all of human consensus reality. Specialised ontologies concentrate on a set of basic and commonly used concepts. Domain ontology is the main classification.

This work focuses on domain ontologies, which refer to the detailed structuring of a context of analysis with respect to the sub-domains, which it is composed of. Domain ontologies define domain specific conceptualisations and impose descriptions on the domain knowledge structure and content. In this project, ontology is treated as a specification of a domain knowledge conceptualisation (van Heijst et al., 1997).

An ontology comprises the classes of entities, relations between entities and the axioms, which apply to the entities of that domain (Mizoguchi, 2004). It is made up of the following parts:

- **Classes and instances.** For example, an ontology modelling the package tour domain structure may contain classes such as ‘destination country’ or ‘package tour environment’. Usually, instances are used to model elements and belong to classes. For example, the instance “5 days for your children at Disneyland” belongs to the class ‘Child-centred package tours’. Classes are usually organised in a hierarchy of subclasses. For example, the concept ‘man’ can be defined as a sub-class of an existing concept ‘person’ in WordNet vocabulary.
An ontology-based system for intelligent matching of travellers’ needs

(http://xmlns.com/2001/08/wordnet/). If class A is a subclass of class B, instances of class A also belong to class B.

- **Properties.** They establish relationships between the concepts of an ontology. For example, the property ‘BelongTo’ associates an object with its owner.

- **Rules.** They model logical sentences that are always true. Rules provide us with high expressiveness and they make for more complex reasoning with the ontology, which can give rise to scalability issues.

Ontologies having no rules are called **lightweight ontologies**. The simplest type of lightweight ontology is called **taxonomies** and they are made up of a hierarchy of classes representing the relevant concepts in the domain. Studer et al. (1998) considered taxonomies as full ontologies, while Brewster et al. (2004) analysed current and future problems, and challenges for ontologies.

In Artificial Intelligence (AI), theories fall into two broad categories: mechanism theories and content theories. Particularly, two uses of ontologies can be identified: vocabulary and content. Ontologies are essentially content theories about the sorts of objects, properties of objects, and relations between objects. Intelligent mechanisms such as rule systems or fuzzy logic cannot do much without a good content theory of the domain in which they must work. As Chandrasekaran et al. state:

“Once a good content theory is available many different mechanisms might be used equally well to implement effective systems, all using essentially the same content.” (Chandrasekaran et al., 1999, p.23)

The terms provided by the domain ontology can be used to assert specific propositions about a domain or a situation in a domain. In the GPT domain, we can represent a fact about a package tour: The “5 days for your children at Disneyland” GPT has destination country ‘France’, where “5 days for your children at Disneyland” is an instance of the concept ‘GPT.Name’ and ‘France’ is an instance of the concept ‘destination country’. Once we have the basis for representing propositions, we can also represent knowledge involving propositional attitudes such as “During this tour at Disneyland, we shall have educational/historical activities”. This results, as defined earlier, in: 5_Days.Disneyland. Non sport activities = Educational/Historical.

An effective ontological domain analysis clarifies the terminology, which enables the ontology to work for coherent and cohesive reasoning purposes. Terms are associated with concepts and relations in the ontology. It is noteworthy that shared ontologies can form the basis for domain-specific knowledge representation languages. Ontologies are considered essential in two of the current most relevant research areas: knowledge management and the semantic web. Thus, in Kanellopoulos et al. (2006), the importance of ontologies is highlighted for performing efficient knowledge management in the travel domain. The authors declare that describing the travel domain in ontological terms facilitates an intelligent access to the travel related information.

Ontologies can improve the accuracy of web searches, because the search engine can look for specific terms semantically rather than syntactically. The semantic annotation makes it much easier to develop programmes that can tackle complicated questions whose answers do not reside on a single web page.
Certainly, the development of web portals guided by ontologies facilitates the development of the semantic web. There are mainly two different approaches for using ontologies to guide the development of semantic web portals:

- The Knowledge Annotation initiative of the Knowledge Acquisition community (KA2: http://ka2portal.aifb.uni-karlsruhe.de) which is an example of a semantic portal guided by ontologies (Staab et al., 2000). In this project, a community web portal was developed. KA2 is based on ontology as a semantic backbone for accessing information on the portal, for contributing information, as well as for developing and maintaining the portal.

- On the other hand, Ontoportal (http://ontoportal.org.uk) uses ontological hypermedia principles to enrich the linking between concepts within a research community, allowing researchers to not only position a concept within the context of the entire community in which they work but, more importantly, pose intricate research queries.

It is noteworthy that Lausen et al. (2005) give a state of the art survey for semantic web portals and evaluate portals like Esperonto, OntoWeb, Empolis K42, and Modeca ITM.

4 The ontological model

Our work was inspired by the work of Garcia-Sanchez et al. (2006) who designed and developed an ontology-based portal for employment. We represent ontologies by means of Multiple Hierarchical Restricted Domains (MHRD) in a similar manner to that employed by Eschenbach and Heydrich (1995). Especially, we term MHRD to a set of concepts holding the following:

- Concepts are defined through a set of attributes. The value(s) of each attribute must belong to a set of possible values for this attribute.

- There can be taxonomic relations among concepts, so that attribute (multiple) inheritance is permitted. There can also be mereological relationships among the concepts. The different (sub)types of mereological relations are distinguished in our model.

The ontology representation schema adopted here includes ‘structural’ axioms, that is, axioms drawn from the proper structure of the ontology. With the term ‘axiom’ we mean all the rules and/or constraints among the elements considered in the model: concepts, attributes, relationships and attribute values. Defining ontologies without non-structural axioms does not mean that these sort of axioms cannot be defined as a (part of the) specification of a conceptualisation. Our semantic model contains attributes whose values are restricted to the ones included in a set of possible or allowed values. In this application, most of the attributes are restricted. So, the system can avoid errors due to wrong values introduced by a user and can check for their correctness. The proposed intelligent system consists of three primary components: the user interface, the knowledge base, and the inference engine. In deductive inference, the knowledge base of our system consists of a set of axioms. Each of the axioms is analysed to determine the value of a statement containing a variable. For example, consider the set of axioms
presented in the following pseudo-code, while the desired activities are included in a package tour.

```
desire(GPT.activity, nightlife) OR desire(GPT.activity, dining)
NOT (desire(GPT.environment, peace & quiet))
GPT.environment(nightlife, sun & sand)
GPT.environment(dining, peace & quiet)
```

Applying deductive inference to determine that nightlife is the preferred activity for this tourist is straightforward in this simplified case.

## 5 Development of the ontology

### 5.1 Domain requirements

The purpose of the work presented here was the development of a website, which can serve as a service provider in package tours-related issues. In particular, the goal of this portal is to enable travellers living in Europe to search for package tours. For this purpose, some requirements were defined in line with the Package Travel Regulations 1992. The EC Package Travel Regulations 1992 (SI 1992 No. 3288) became law in December 1992. They impose a comprehensive consumer protection regime on ‘organisers’ (tour operators) and ‘retailers’ (travel agents). The main requirements are that organisers provide financial security to consumers in the event of their insolvency; that they provide extensive information to consumers before they enter into package holiday contracts; and that organisers accept strict, or near strict, liability for breach of contract.

The proposed portal should offer to users search guidance for GPT. This portal would serve as a meeting point among travel agencies. In other words, this portal has been conceived to help both travellers to find a desired package tour, and travel agencies to find specialised travellers. Therefore, this is a website that allows travelling search and package tours offers. In this portal, the information contained is dynamically updated.

The main information items we allow to be introduced in the portal are:

- **Travel agencies**: The main information provided by agencies is announcement for package tours availability.
- **News**: Travel agency-related news.
- **Requests for package tours**: It is allowed to record travellers’ preferences for GPT (i.e., traveller’s profile).

We have used common descriptors for offers and demands (the profile introduced by a traveller has similar attributes to the one introduced by a travel agency in an announcement of a package tour availability). This allows matching the traveller against announcement of package tour availability in a better way. The objective of the system is not to make a tracing of the matching the travellers have against the travel agencies’ announcements. On the contrary, the system simply makes travellers know about the special offers of every registered travel agency. Likewise, it lets travel agencies know about the possible responders to a group package tour offer. A brief scheme of the developed ontology is shown in Figure 1.
5.2 Ontology development

The main concepts in this ontology are ‘Traveller’ (the person who is looking for a group package tour), ‘Travel Agency’, ‘Offer’ (the distinct offers made by a travel agency) and ‘Profile’ (the requisites for a package tour offer or the merits certified by a traveller). According to Loban (1997) selection criteria for package tours may be:

- safety of the tour
- the tour operator
- the tour mode (e.g., air, cruise, rail, bus)
- the tour cost
- the tour time
- the tour accommodation
- the tour environment
- the destination country
- the tour activities
- the tour events.

The full version of the ontology incorporates these attributes and it appears in the Appendix A. The ontology is composed of several taxonomies, designed in order to categorise some concepts. These taxonomies are also presented in the Appendix A. All the main concepts, attributes, axioms and relationships relevant to these concepts are detailed in the following sections.

5.2.1 Concepts

5.2.1.1 Traveller

This concept represents the person who is looking for a package tour. A traveller can also find in the system other information (s)he may be interested in, such as interesting related news concerning the GPT domain. The attributes defined for traveller are:
- **E-mail and password**: They are used to identify, uniquely, each traveller. So, one precondition for being a traveller is to possess an e-mail account.

- **Contact phones**: The fastest way a travel agency can contact a traveller is by phone. The traveller may introduce two different phone numbers: home and mobile.

- **Country, province and city**: The travellers’ geographical data may be very important for the travel agency to know.

- **Date of birth**: To know the age of the traveller.

- **Gender**: male or female.

- **Times of joining GPT**: This attribute states how many times the traveller has joined GPT. The available attribute values are: 0, 1–3, 4–6, 7–9, >9.

- **Travelling partners**: This attribute can take the values: travel alone, spouse, friends, family and relatives, family and relatives and spouse and friends.

- **Travel with child(ren)**: This attribute can take two values ‘yes’ or ‘no’.

- **No of children travel with**: 1, 2, 3, >3.

- **No of travelling partners**: 1, 2–4, >4.

- **GPT.No of Days**: This is the number of days of the desired package tour. The values in this attribute can be: 2–3, 4–6, 6–7, >7.

- **GPT.QoS_Operator**: This is the service quality the traveller prefers from the GPT operator. The service quality has three values: low, medium, high.

- **GPT.Kind**: There are available the following values: ‘child-centred’, ‘friends-centred’, ‘family and relatives family centred’ and ‘others’. Figure A3 depicts the GPT kinds.

- **GPT.Safety**: The safety of the tour may be low, medium or high.

- **GPT.Mode**: The mode of a GPT that the traveller prefers. The mode can be: air, cruise, rail or bus. Figure A4 depicts the GPT modes.

- **GPT.DepartureDate**: This is the departure date that the traveller prefers for his or her package tour.

- **GPT.DeparturePlace**: This is the place that the traveller prefers to departure.

- **GPT.Accommodation**: This is the accommodation (e.g., hotel, motel, condominium, other) offered during a package tour.

- **GPT.Environment**: It is the environment (e.g., sun and sand, scenic beauty, weather/climate, peace and quite) of the package tour that the traveller prefers. Figure A3 depicts GPT environment.

- **GPT.DestinationCountry**: It is the European country of the GPT that the traveller prefers.
• **GPT.Non-sportActivities**: These are the non-sports activities (e.g., educational/historical, cultural sites, sightseeing, naturism, spas/hot springs, shopping, dining, nightlife) that the GPT should include.

• **GPT.SportActivities**: These are the sports activities (e.g., surfing, white-water rafting, water skiing, snorkelling/diving, other water sports, golf, tennis, biking, hiking, equestrian, hunting, angling, snow skiing) that the traveller desires.

• **GPT.Events**: The traveller may desire sporting events, cultural events or natural events.

• **Max_price**: It is the maximum price the traveller is willing to pay for a package tour. The search engine which can match GPT offers against suitable travellers uses this information. As we shall see we must have: \( \text{Max_price} \geq \text{GPT.cost} \) in order to book a package tour.

The travellers have to specify their preferences in order to create their own profile. The traveller’s profile will be matched against the profiles specified by the different travel agencies’ offers. The profile is also helpful to know the news that might be of interest for the traveller. The profile concept will be described later in this section.

### 5.2.1.2 Travel agency

A travel agency is the entity that makes offers about available GPTs described in Appendix A. When a travel agent wants to use the system to search for travellers, one must sign up and be validated by the system administrator. Generally, a travel agency is a business that sells travel related products and services, particularly package tours, to end-user customers on behalf of third party travel suppliers, such as airlines, hotels, tour companies, and cruise lines. In addition to dealing with ordinary tourists, most travel agents have a special department devoted to travel arrangements for business travellers. Indeed, some agencies specialise in that work. Some agencies also serve as general service agents for foreign travel companies in different countries. The attributes concerned with the ‘travel agency’ concept are:

• **E-mail and password**: They are used to identify each travel agent.

• **Name and TAC (entity code)**: The name of the travel agency and a code by means of which the system can automatically (or the system administrator) or manually check for the existence of the travel agency.

• **GPT.Kind**: This attribute indicates the GPT kind. The purpose of this attribute is to show travellers the available GPTs offers.

• **Phone number**: If any person wants to contact the travel agency, a quick way is by phone. This number must be the contact phone number.

• **Province, city, address and postal code**: These attributes specify the place where the travel agency is located.

• **Contact information**: This is a set of attributes describing the contact details (name, position, phone number, e-mail and fax) of the travel agency for travellers.
5.2.1.3 Offer
This concept represents all the elements that the travel agency can make public on the web site. Amongst these elements, the most important is the announcement of GPT’s availability. The discussion now turns to a description of the ‘announcement of GPT’s availability concept.

5.2.1.4 Announcement for GPT availability
This concept represents a specific type of offer that may be made by a travel agency. A GPT availability announcement is made public when the travel agency is looking for GPT travellers. Each GPT availability announcement has a profile, similar to that for travellers. Besides the profile, a GPT availability announcement has the following attributes:

- **Name**: This attribute is inherited from the concept Offer. It could be seen as a short description of the offer.
- **GPT.Kind**: It indicates the GPT kind in which the offer is included.
- **Departure place**: These attributes establish the place where the offer is valid.
- **GPT.Cost**: This is the cost of the package tour.
- **Number of individuals**: This is how many individuals could complete the GPT offer.
- **Expiration date**: When the announcement becomes invalid.
- **GPT_DepartureDate**: It indicates the departure date of the GPT.

5.2.1.5 Profile
This concept is central to the application developed because it allows matching the travellers who are looking for GPTs against the offers by the travel agencies. A profile is like a set of travelling preferences for a traveller and like a requirement list for an announcement of the GPT’s availability. Then, in order to determine whether a GPT offer is suitable for a traveller or not, the system will only have to check for their respective profiles. The ‘profile’ concept is composed of several types of GPT characteristics such as: GPT_NoofDays, GPT.QoS/Operator, GPT.Kind, GPT.Safety, GPT.Mode, GPT.Accommodation, GPT.DepartureDate, GPT.DeparturePlace, GPT.Environment, GPT.DestinationCountry, GPT.Non-sportActivities, GPT.SportActivities, GPT.Events, GPT.Cost.

5.2.2 Relations
Two types of inter-conceptual relations have been used in the system, namely, ‘Part-of’ and ‘IS-A’. A ‘Part-of’ relation means that one concept is a part of another one. The ‘IS-A’ relation is the taxonomic one, used for creating categorical structures. In this way, for example, a GPT offer is modelled as a part of a travel agency because it is a part of the data introduced by a travel agency in the system. The same can be said for profile – announcement of GPT availability (the profile is a part of the information needed in order to complete an announcement of GPTs availability) and for
profile – traveller (the profile is seen as the set of preferences of the traveller and uses part of the information needed). On the other hand, when we say that an announcement of GPTs’ availability is an offer we mean that the announcement of a GPT is a specific type of offer. Hence, the ‘the announcement of GPT availability’ concept incorporates all the attributes of the concept ‘offer’.

5.2.3 Axioms

Several types of axioms have been implemented in the system. On the one hand, there are domain-independent axioms, which are not specific for the GPT domain but derived from general conditions or situations (e.g., the format of an email address). On the other hand, there are domain-dependent ones, which are derived from a proper structure of the ontology and from proper conditions and situations given in the GPT domain. Several types of axioms have been defined from the possible values of the attributes linked to each concept. These axioms can affect the value of a single attribute or the values of various attributes owning to the same or different concepts. The referred types of axioms are, respectively, called: format constraints, restricted set of allowed values, law constraints, and dependent attributes.

5.2.3.1 Format constraints

There are intrinsic restrictions over the values of some attributes related to their format. For example, an e-mail address must retain a specific format: ‘____@____.__’. So, the value assigned to this attribute has to preserve this format. Other examples of this type of axiom are passwords (these must contain at least six characters in our case), phone numbers (the length must be between 9 and 12 and all the characters must be numbers) and dates (a date must have the format dd/mm/yyyy). More examples of this type of axiom are the values of the attribute ‘postal code’ or the values of ‘TAC’ (an identification code of a travel agency; it is composed of a letter indicating the entity type and eight numbers).

5.2.3.2 Restricted set of allowed values

The model contains attributes whose values are restricted to the ones included in a set of possible or allowed values. In this application, most of the attributes are restricted. So, the system can avoid errors due to wrong values introduced by a user and can check for their correctness. There are a lot of examples to illustrate this point. For example, the value for the attribute ‘destination country’ is restricted to the list of all the official European countries. In particular, if the user selects Greece as the country, then the value of the attribute ‘Province’ will be restricted to the list of all the provinces belonging to Greece.

In the same way, if the user selects Peloponnesus as the province, then the value of the attribute ‘City’ will be restricted to the list of all the cities and towns within Peloponnesus. On the other hand, there are several attributes restricted to a short set of allowed values such as gender (‘male’ or ‘female’), GPT.accommodation (‘hotel’, ‘motel’, ‘condominium’ and ‘other’). There are some attributes whose set of allowed
values is very large. In this case, it is convenient to build a categorisation to facilitate the user to select the concrete value. For example, in the case of the attribute ‘GPT.activity’, a multi-level taxonomy has been implemented to easily determine the activity that the traveller prefers. A portion of this taxonomy is shown in Figure A2.

5.2.3.3 Law constraints

This type of axiom is identified in the rules established by official package travel regulations. An example can be found in the traveller’s date of birth. The law establishes that a person, in order to be able to travel alone, must be, at least, 17 years old under specific conditions. So, this must be taken into account for not allowing signing up any person younger than 17. Another example is the ‘TAC’. This identification code for a travel agency has a control digit that can be calculated using the other seven numbers. This could be used to check whether a given TAC is valid. There are operational constraints established by the Travel Industry Council (TIC).

5.2.3.4 Dependent attributes

By dependent attributes we mean a rule like this:

\[
\text{VAL}(C, X) = X1 \rightarrow \text{VAL}(\hat{C}, A) \in \{A_1, A_2, \ldots, A_k\} \subset \{A_1, A_2, \ldots, A_n\}
\]

where \(C\) and \(\hat{C}\) are concepts (maybe \(C = \hat{C}\)), \(X\) and \(A\) are attributes of the concepts \(C\) and \(\hat{C}\), respectively. The function \(\text{VAL}(\text{concept}, \text{attribute})\) returns the value of the attribute ‘attribute’ of the concept ‘concept’. For example, the following rule indicates that if the value of the attribute ‘destination_country’ = ‘Greece’ concerning the concept ‘GPT’, then for the same concept the value of the attribute ‘environment’ must take values from the subset \{Sun & Sand, Scenic Beauty\} which belongs to the whole set.

The following rule indicates that in a ‘Sun and Sand’ GPT environment, the GPT sport activities are mainly related with water sports.

6 Design of the GPT web portal

The goal of the proposed web portal is to provide travellers with easy access to the package tour market, mainly to those living in Europe. Thus, the portal allows travellers to contact European travel agencies. The target audience for this service are the travellers with problems in finding a GPT because of the lack of knowledge about GPTs. For these reasons, it was found necessary to set up a new service that could approach travel
agencies to travellers requesting for GPTs. The domain ontology was developed with the use of the Protégé\textsuperscript{2} tool, which serves as a rapid environment in which ontology designers can instantly create individuals of their ontology and experiment with semantic descriptions. The system can generate user interfaces that can be further customised for knowledge acquisition in the GPT domain. The Protégé OWL plugin (Knublauch et al., 2004) provides an open testing framework in which code can be executed at any time. Another feature of Protégé is code generation: i.e., the system takes an OWL ontology and creates the corresponding Java classes from it. The ontology is published on fixed Universal Resource Indicators (URI’s) as OWL files. It allows travel agencies to publish metadata about their offers. Travel agencies instantiate the classes from the ontology and publish the resulting individuals as OWL files. The inference engine reasons on crawled and newly created instances and on the domain ontology. It is used to query whether, and which, instances already exist in the semantic web and it serves the ontology browser, because it allows querying for existing classes, instances and properties. The inference engine is implemented with the use of the Racer\textsuperscript{1} tool, while the rules are expressed via the Semantic Web Rule Language\textsuperscript{5} (SWRL).

6.1 The software architecture

Figure 2 illustrates the software architecture of the application that finds the GPT for the user. Our application consists of two separated but linked layers. The semantic web layer makes ontologies and interfaces available to the public, whereas the internal layer consists of the control and reasoning mechanisms. The functionality of this application is made available to software agents through a Web Service Interface. Input to this service is a collection of data objects about an announcement of GPT availability. The output is a list of suggested GPTs. The Web Service Description Language (WSDL) provides the technical specification of the service operations, that is, the types of input and output messages and the URL where the service can be invoked. The information on what the service is doing, the possible properties the service might have and also the meaning carried in its messages are not available through the WSDL descriptions of the services. It is noteworthy that WSDL does not provide semantic information for the announcement of GPT availability. The input and output data structures are formally represented in terms of OWL ontologies, so that external software agents can correctly use the service. Much of the application logic itself is implemented in Java, a conventional object-oriented language. The objects that are exchanged between the application and other service should be represented as Java objects. For this purpose, a typical implementation would employ an OWL parsing library such as Jena\textsuperscript{6}. Jena provides a dynamic object model in which OWL classes, properties and individuals are stored using generic Java classes like OntClass and Individual. The knowledge encoded in the external ontologies is used by reasoning engines (e.g., a rule execution engine).
6.2 Users

From a usability perspective, three types of users can be distinguished:

- **Travellers**: Users who are looking for a GPT. Travellers have specific information needs (as GPT offers) and have committed to give some private information about themselves in order to find exactly the GPTs they prefer. They are the main actors of the system.

- **Travel agents**: They are the financial actors of the system and who really pay for using the tool to look for the most suitable consumers for their GPTs. They usually enter the web portal with a well-defined GPT profile to discover what the travellers are interested in. This is the reason why the searching process is based on the personal preferences of the travellers for GPT.

- **Administrator**: She or he is responsible for publishing the adequate information in the portal, verifying the reliability of the information provided by travel agencies and travellers, and rejecting those travellers or travel agents who make incorrect use of the tool. The administrator is also in charge of managing enhancements in the tool and maintaining it according to the last trends in travel agencies portals.

6.3 Security

The proposed system has a security sub-system based on three profiles: external level, private level and administrator level:

- **External level**: Everybody can use the system at this level without any password. Its main goal is to give access to potential users. This access allows for browsing the pages contained in the website by using simple searches.

- **Private level**: It allows access to the private information pertaining to the traveller. It also permits to make changes to one’s private profile. This kind of access is password-protected.

- **Administrator level**: At this level it is possible to check for, change or modify the user information, and to detect the inadequate use of the tool so that users’ confidence can be ensured. Moreover, it is possible to publish GPTs.
7 Conclusions

In this work, a domain ontology has been implemented and a package tour web portal has been designed and partially implemented. The main differentiating component for this portal is the use of an ontology for representing Group Package Tours. This ontology has been used to guide the design of the application and to supply the system with semantic possibilities. An advantage of the application is its simplicity of use for the user and its capacity to help users to add information quickly. The use of ontologies helps to model the domain in a reusable, shareable and more efficient way. The restrictions included in the ontology facilitated the system’s capability to predetermine the allowed values for most attributes. This simplifies the user GPT reservation. The use of ontologies as modelling instruments implies some other important advantages. On the one hand, it allows the system to check, automatically in most cases, for the consistency of the information introduced by the user, thus preventing it from accepting information inconsistencies. On the other hand, the ontology model allows us to define an intelligent search engine, that is, an ontology-guided search engine. To be precise, the matching between the traveller profiles and GPTs offers is semantically performed, as are the searches for travellers and offers. Provided that the concepts ‘Traveller’ and ‘GPT availability announcement’ are related to the ontology by means of a common concept, namely ‘Profile’, the search process to match suitable GPT availability announcements with a concrete traveller can be performed more efficiently by returning solely such offers as the traveller can really be interested in. In the near future, the proposed web portal will operate at the pilot stage, and will be evaluated under certain criteria related to usability issues.

References


OpenCyc (2006) OpenCyc v. 1.0, Available at: http://www.opencyc.org/


**Notes**

1A GPT kind is ‘Child-centred’ package tour (e.g., 5 days for your children at Disneyland).


3http://www.sts.tu-harburg.de/~r.f.moeller/racer/

4http://www.daml.org/200/11/swrl

5http://jena.sourceforge.net
Appendix A

In this annexure, we present the GPT ontology. The first levels of the full ontology are shown in Figure A1. In this annexure, the concepts and relations of the ontology are briefly described. The core of the ontology is conformed by the concepts ‘Traveller’, ‘Profile’, ‘Announcement for GPT availability’, ‘Offer’ and ‘Travel Agency’ and their respective relations.

A.1 Concepts

<table>
<thead>
<tr>
<th>GPT</th>
<th>It corresponds with a group package tour</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPT.Name</td>
<td>It is the name of the GPT</td>
</tr>
<tr>
<td>GPT.Kind</td>
<td>For example, child-centred GPT</td>
</tr>
<tr>
<td>GPT.Safety</td>
<td></td>
</tr>
<tr>
<td>GPT.QoS/Operator</td>
<td></td>
</tr>
<tr>
<td>GPT.Safety</td>
<td></td>
</tr>
<tr>
<td>GPT.Operator</td>
<td></td>
</tr>
<tr>
<td>GPT.Mode</td>
<td></td>
</tr>
<tr>
<td>GPT.Cost</td>
<td></td>
</tr>
<tr>
<td>GPT.Destination</td>
<td></td>
</tr>
<tr>
<td>GPT.Environment</td>
<td></td>
</tr>
<tr>
<td>GPT.DestinationCountry</td>
<td></td>
</tr>
<tr>
<td>GPT.NonSportActivities</td>
<td>This is a set, which may include various non-sports activities</td>
</tr>
<tr>
<td>GPT.SportActivities</td>
<td>This is a set, which may include various sports activities</td>
</tr>
<tr>
<td>GPT.Events</td>
<td></td>
</tr>
<tr>
<td>Traveller</td>
<td>This concept models the person who is looking for a GPT</td>
</tr>
<tr>
<td>E-mail</td>
<td>It is the e-mail address the traveller can be contacted through</td>
</tr>
<tr>
<td>Password</td>
<td>It is a code established by the traveller</td>
</tr>
<tr>
<td>Phone number</td>
<td>It is the traveller home phone number</td>
</tr>
<tr>
<td>Cellular phone</td>
<td>It is the traveller mobile phone number</td>
</tr>
<tr>
<td>Country</td>
<td>It is the country where the traveller resides</td>
</tr>
<tr>
<td>Province</td>
<td>It establishes the place (province) the traveller resides</td>
</tr>
<tr>
<td>City</td>
<td>It establishes the place (city) the traveller resides</td>
</tr>
<tr>
<td>Birthday</td>
<td>It is traveller’s birth date</td>
</tr>
<tr>
<td>Gender</td>
<td>It indicates whether the traveller is male or female</td>
</tr>
<tr>
<td>Times of joining GPT</td>
<td>This attribute states how many times traveller has joined GPT.</td>
</tr>
<tr>
<td>Travelling partners</td>
<td>The available attributes values are: 0, 1–3, 4–6, 7–9, &gt;9</td>
</tr>
<tr>
<td></td>
<td>This attribute can take the values: (a) travel alone, (b) spouse, (c) friends, (d) family and relatives, (e) family and relatives and spouse and friends</td>
</tr>
<tr>
<td>Attribute</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Travel with children</td>
<td>This attribute can take two values: Yes or No</td>
</tr>
<tr>
<td>No of children travel with</td>
<td></td>
</tr>
<tr>
<td>No of travelling partners</td>
<td></td>
</tr>
<tr>
<td>No of days</td>
<td>This is the number of days of that the package comprises. The values can be: 2–3, 4–6, 6–7, &gt;7</td>
</tr>
<tr>
<td>GPT.QoS/Operator</td>
<td>This is the service quality the traveller prefers from the GPT operator.</td>
</tr>
<tr>
<td></td>
<td>The service quality has three values: low, medium, high</td>
</tr>
<tr>
<td>GPT.Safety</td>
<td>This attribute can take three values: low, medium, and high</td>
</tr>
<tr>
<td>GPT.Mode</td>
<td>The mode of a GPT that the traveller prefers. The mode can be: air, cruise, rail or bus</td>
</tr>
<tr>
<td>GPT.DepartureDate</td>
<td>This is the departure date that the traveller prefers for his/her package tour</td>
</tr>
<tr>
<td>GPT.DeparturePlace</td>
<td>It indicates the departure place</td>
</tr>
<tr>
<td>Max price</td>
<td>It is the maximum price the traveller is willing to pay for a package tour. The search engine to match GPT offers against suitable travellers uses this information.</td>
</tr>
<tr>
<td>GPT.Accommodation</td>
<td>This is the accommodation (e.g., hotel, motel, condominium, other) offered during a package tour</td>
</tr>
<tr>
<td>GPT.Environment</td>
<td>It is the environment (sun and sand, scenic beauty, weather/climate, peace and quite) of the travel package that the traveller prefers</td>
</tr>
<tr>
<td>GPT.DestinationCountry</td>
<td>It is the European country of the GPT that the traveller prefers</td>
</tr>
<tr>
<td>GPT.NonSportActivities</td>
<td>These are the non-sports activities: educational/historical, cultural sites, sightseeing, naturism, spas/hot springs, shopping, dining, nightlife that the GPT may include</td>
</tr>
<tr>
<td>GPT.SportActivities</td>
<td>These are the sports activities: surfing, white-water rafting, water skiing, snorkelling/diving, Other water sports, golf, tennis, biking, hiking, equestrian, hunting, angling, snow skiing that the GPT should include</td>
</tr>
<tr>
<td>GPT.Events</td>
<td>The traveller may desire: sporting events, cultural events or natural events</td>
</tr>
<tr>
<td>Announcement of GPT availability</td>
<td>It models the availability of package tours</td>
</tr>
<tr>
<td>Name</td>
<td>This attribute is inherited from the concept ‘Offer’ and can be seen as a short description of the offer</td>
</tr>
<tr>
<td>Number of seats</td>
<td>It represents the number of individuals for this specific announcement</td>
</tr>
<tr>
<td>Expiration date</td>
<td>It is the date the offer is cancelled</td>
</tr>
<tr>
<td>Offer</td>
<td>It models all the elements a travel agency can offer in a GPT</td>
</tr>
<tr>
<td>Profile</td>
<td>This concept is composed of several types of GPT’s characteristics such as: GPT.NoofDays, GPT.QoS/Operator, GPT.Kind, GPT.Safety, GPT.Mode, GPT.Accommodation, GPT.Cost, GPT.Departure date, GPT.Departure place, GPT.Accommodation, GPT.Environment, GPT.Destination_country, GPT.Non-sportActivities, GPT.SportActivities, GPT.Events</td>
</tr>
<tr>
<td>Travel agency</td>
<td>It represents the entity that makes the offer</td>
</tr>
<tr>
<td>E-mail</td>
<td>It is the e-mail address of the entity</td>
</tr>
<tr>
<td>Password</td>
<td>It is a privacy code established by the entity</td>
</tr>
</tbody>
</table>
An ontology-based system for intelligent matching of travellers’ needs

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>It indicates the entity official name</td>
</tr>
<tr>
<td>TAC</td>
<td>It comes from the ‘Travel Agency Code’ and represents a code by which an entity is uniquely identified</td>
</tr>
<tr>
<td>GPT Kind</td>
<td>This attribute indicates the GPT class in which the entity is involved</td>
</tr>
<tr>
<td>Service quality</td>
<td>low, medium or high</td>
</tr>
<tr>
<td>Phone number</td>
<td>It is the entity contact phone number</td>
</tr>
<tr>
<td>Province</td>
<td>It establishes the place (province) where the entity has its headquarters</td>
</tr>
<tr>
<td>City</td>
<td>It establishes the place (city) where the entity has its headquarters</td>
</tr>
<tr>
<td>Address</td>
<td>It establishes the place (address) where the entity has its headquarters</td>
</tr>
<tr>
<td>Postal code</td>
<td>It is the postal code of the area where the headquarters are located</td>
</tr>
<tr>
<td>Contact person data</td>
<td>It contains all data related to the person who represents the travel agent</td>
</tr>
</tbody>
</table>

**Figure A1** The domain ontology (see online version for colours)

**Figure A2** GPT activity taxonomy (see online version for colours)
A.2 Axioms

This section presents the constraints, restrictions and axioms defined for the GPT ontology. Some of these axioms are general but others are exclusive for the GPT domain. In order to explain these axioms, some general variables and functions have been used. These variables are shown in italics:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>thisDate</td>
<td>It represents the current date</td>
</tr>
<tr>
<td>this year</td>
<td>It represents the current year</td>
</tr>
<tr>
<td>thisMonth</td>
<td>It represents the current month</td>
</tr>
<tr>
<td>thisDay</td>
<td>It represents the current day</td>
</tr>
<tr>
<td>allOfficialEuropeanCountries</td>
<td>It represents the set of all the European countries officially recognised</td>
</tr>
<tr>
<td>All_European_Provinces</td>
<td>It represents the set of all the provinces within Europe.</td>
</tr>
<tr>
<td>IN</td>
<td>It is the set of natural numbers.</td>
</tr>
<tr>
<td>Length(attribute)</td>
<td>This function receives as parameter an attribute and returns its length (it can be applied to strings of characters or numbers indiscriminately)</td>
</tr>
</tbody>
</table>

A.2.1 Format constraints

- Traveller.e-mail: ‘____@____.__’
- Length(Traveller.password) >= 6
- 9 ≤ Length(Traveller.phoneNumber) ≤ 12
- 9 ≤ Length(Traveller.cellularPhone) ≤ 12
- Traveller.birthday: dd/mm/yyyy
Announcement of GPT availability.expirationDate: dd/mm/yyyy
Travel Agent.e-mail: ‘_____@____.__’
Length(Travel Agent.password) ≥ 6
Travel agency.tac: cnnnnnnn, c is a character, 0 ≤ n ≤ 9, where n ∈ IN
9 ≤ Length(Travel Agency.phoneNumber) ≤ 12
Length(Travel Agency.postalCode) = 5

A.3 Relations

Hereafter we refer to the relationships established by the concepts in the domain ontology. It has been necessary to use just two types of relations: the relation ‘IS-A’ (it implies a taxonomic categorisation) and the relation ‘Part-of’ (it forms a mereology). At this point, some aspects should be pointed out. First, a profile is composed of preferences. GPT destination country, GPT mode, desired departure date, GPT environment, GPT activities are considered as preferences. In this way, a profile may be composed of several partial preferences. On the other hand, travellers and announcement for GPT availability will be partially composed of profiles. Given that an offer is established by a travel agency, it will form part of the agency itself. Finally, two important characteristics of this ontology are put forward: on the one hand, the relation between the concept ‘Traveller’, secondly, the concept ‘Announcement for GPT availability’ through the concept ‘Profile’. The fact that the same profile can form part of both travellers and announcement for GPT availability makes us sure that the matching when a traveller searches for a desired package tour will be almost perfect. So, as the search process of announcement for GPT availability executed by the system is based on both, that is, the traveller’s profile and the announcement for GPT availability’ profile, the system is likely to find the appropriate set of announcements for a concrete traveller.