Exploiting tourism destinations’ knowledge in an RDF-based P2P network

Dimitris N. Kanellopoulos*, Alkiviadis A. Panagopoulos

Department of Tourism Management, Technological Educational Institute of Patras, Greece Meg. Alexandrou 1, Patras 26334, Greece

Received 8 March 2006; accepted 8 March 2006

Abstract

Destination Management Systems (DMS) is a perfect application area for Semantic Web and P2P technologies since tourism information dissemination and exchange are the key-backbones of tourism destination management. DMS should take advantage of P2P technologies and semantic web services, interoperability, ontologies and semantic annotation. RDF-based P2P networks allow complex and extendable descriptions of resources instead of fixed and limited ones, and they provide query facilities against these metadata instead of simple keyword-based searches. The layered adaptive semantic-based DMS (LA_DMS) and Peer-to-Peer (P2P) project aims at providing semantic-based tourism destination information by combining the P2P paradigm with Semantic Web technologies. In this paper, we propose a metadata model encoding semantic tourism destination information in an RDF-based P2P network architecture. The model combines ontological structures with information for tourism destinations and peers.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Destination Management Systems; P2P; Semantic Web; Distributed RDF repositories; Adaptability; Management; Design; Documentation; Standardization

1. Introduction

Peer-to-Peer (P2P) systems have been successful for special cases like exchanging music files (e.g. Napster), collaborative computing (e.g. SETI@home) and instant messaging...
(e.g. MSN, ICQ and Yahoo instant messengers). A P2P infrastructure provides a suitable architecture for distributed content management as it offers user-centered, data-centered and computing-centered models. P2P systems have shown to be scalable and provide low entry costs. The cost of entry refers to the amount of work related to the process of becoming a system consumer and/or provider. If the entry cost is too high, no one will use the system, thus making scalability rather irrelevant. Recent research in P2P systems focuses on providing techniques for evolving from basic P2P networks, supporting only file exchanges which use simple filenames as metadata, to more complex systems like schema-based P2P networks. These networks are capable of supporting explicit schemas to describe knowledge, usually using RDF and thematic ontologies as metadata (Nejdl et al., 2003). Schema-based P2P systems are the next step in distributed P2P networks with advanced data management and storage features. However, their architectures have to take into account the typical characteristics of P2P systems i.e. local control of data, dynamic addition and removal of peers, only local knowledge of available data and schemas, and self-organization and -optimization.

Through the use of metadata organized in numerous interrelated ontologies (Mizoguchi, 2004), tourism destination information can be tagged with descriptors that facilitate its retrieval, analysis, processing and reconfiguration. Tourism ontologies offer a promising infrastructure to cope with heterogeneous representations of tourism destination web resources. An ontology is a conceptualization of an application domain in a human-understandable and machine-readable form, and typically comprises the classes of entities, relations between entities and the axioms which apply to the entities that exist in that domain. The World Wide Web Consortium (W3C) has recently finalized the Web Ontology Language (OWL) as the standard format in which ontologies are represented online. With OWL it is possible to implement a semantic description of the tourism/travel domain by specifying its concepts and the relationships between these concepts. OWL (Dean and Schreiber, 2004) provides greater machine interpretability of web content than that supported by XML, RDF and RDF-schema.

In Destination Management Systems (DMS), data heterogeneity can be solved by providing semantic reconciliation between the different tourism information systems, with respect to a shared, conceptual reference schema: the tourism destination ontology. Core tourism ontologies contain knowledge about the domain of tourism and travel for developing intelligent tourism information systems. In the OnTour project, a working group at the Digital Enterprise Research Institute (DERI) deployed the e-Tourism ontology (Prantner, 2004) using OWL. This ontology describes the domain of tourism and it focuses on accommodation and activities (http://e-tourism.deri.at/ont/). It is based on an international standard: the Thesaurus on Tourism and Leisure Activities of the World Tourism Organization (World Tourism Organisation–WTO, 2002) (viz. a very extensive collection of terms related to the area of tourism).

The main issue in Web Tourism Information Systems (WTIS) design is the specification of the hypermedia generation process including the specification of which resources to use and how to map their data (integration) into the system. Vdovjak et al. (2003) presented their model-driven approach called Hera and introduced the transformation software that builds the heart of the hypermedia presentation process. In Hera framework, data is retrieved from the data repository composed of heterogeneous data sources distributed over the Web. During the “integration and data retrieval” phase, several autonomous sources are connected to a conceptual model (CM) by creating channels through which the
data will populate the concepts from a CM on request. Their CM provides a uniform semantic view over multiple-data sources and concept properties that together define the domain ontology. Their integration model (IM) is addressing the problem of relating concepts from the source ontologies to those from the CM. This problem can also be seen as the problem of merging or aligning ontologies. To automate the solution to this problem, they proposed the integration model ontology (IMO), which is a meta-ontology describing integration primitives that are used both for ranking the sources within a cluster and for specifying links between them and the CM. In their IMO, the main concepts are decoration and articulation. Decorations serve as a means to label “appropriateness” of different sources (and their concepts) grouped within one semantically close cluster, while articulations describe actual links between the CM and the source ontologies and clarify also the notion of the concept’s uniqueness which is necessary to perform join from several sources. From another perspective, many Semantic Web applications, such as Shared-HiKE, need a distributed RDF infrastructure in order to let their users to create, organize and share RDF data. For example, in Shared-Hierarchical Knowledge Editor (HiKE) (Cai et al., 2004), a collaborative knowledge editor lets users create, organize and share their RDF data. Each participant has his/her local hierarchical knowledge and also shares the external knowledge from other participants.

The combination of the IM model (Vdovjak et al., 2003) with RDF-based P2P networks constituted an inspiration for our work. As we shall consider later by combining Adaptive Hypermedia (AH), Semantic Web and P2P technologies in the destination management context, novel value-added tourism services can be provided.

In this paper we propose a metadata model, the layered adaptive semantic-based DMS model (LA_DMS), which combines ontological structures with information for tourism destinations and peers. In the LA_DMS framework, the destination information is adaptable according to the user’s personalization requirements and it is used for query processing. Taking into account the experience of the e-tourist and his/her personal needs we can avoid information overload and deliver knowledge at the right level of granularity as document repositories can be mined for interesting relations. Our work is strongly based on the work by Broekstra et al. (2003), which proposes a metadata model for semantic-based P2P systems. Their semantic model has several advantages over traditional ontologies in the context of P2P information exchange. In their semantic model, statements are not seen as being the truth as in most traditional models. They rather see their semantic model as a collection of opinions supported by different sources of information. Similarly, in the tourism destination context we see our proposed semantic model as a collection of Destination Managements Organizations (DMOs) opinions supported by different sources of destination information.

The rest of this paper is organized as follows: Section 2 discusses semantic-based DMS, and Section 3 describes the requirements imposed by the semantic-based DMS in the P2P context. Section 4 presents the LA_DMS metadata model and Section 5 describes its RDF-based P2P architecture. Section 6 outlines related work and Section 7 concludes with a discussion of open problems.

2. Semantic-based DMS

DMOs exploit tourism destinations (e.g. a village) developing marketing channels as tools to promote their destinations. Particularly, DMOs use DMS such as TISCover
(Haller et al., 2000), VisitScotland (Homepage of VisitScotland and Scotland) and Gulliver (Homepage of Gulliver and Ireland), to distribute their properties and to present their tourism destination as a holistic entity (Buhalis and Licata, 2002). DMS is the ICT infrastructure of the DMO, used for the collection, storage, manipulation and distribution of tourism information, as well as for the transaction of reservations and other commercial activities (Werthner and Klein, 1999). DMS provide complete and up-to-date information on a particular destination. They should be capable of handling both the pre-trip and post-arrival information, as well as of integrating availability and booking service too. A DMS is a full e-commerce system that consists of three basic modules:

- A **knowledge base** that is a subject-oriented, integrated, nonvolatile, time-variant collection of raw data (e.g. attractions, accommodation, travel information) that can be used to support destination management decision making.
- A **users’ database** containing the profiles for those users that access the tourism destination information.
- A **booking-reservation module**. Reservation of tourism products can be achieved using computer reservation systems (CRS) and global distribution systems (GDS) (Werthner and Klein, 1999). Dominant GDS (e.g. Sabre, Galileo) provide travel information services, such as real-time availability and price information for flights, hotels and car rental companies.

Current DMS could not interoperate with other systems and data sources, because they are heterogeneous. They are based on different hardware or software platforms, and they use different data representations for tourism services. To address the different data representation problems, the Travel Industry has formed the *Open Travel Alliance* (OTA) (http://www.opentravel.org), which adopted the *Web Services* technology, viz. a collection of standards that allows server DMS applications to “talk” to each other over the Internet. These standards are:

- **eXtensible Markup Language** (XML: http://www.w3.org/XML/) for driving DMS applications services, viz. XML schema is used in requests and replies (Dell’ Erba et al., 2002).
- **The Simple Object Access Protocol** (SOAP: http://www.w3.org/TR/soap) that provides a means of messaging between a service provider and a service requestor.
- **The Web Services Description Language** (WSDL) as the service description language (Christensen et al., 2001).
- **The Universal Description, Discovery and Integration** (UDDI: http://www.uddi.org/) as the service discovery protocol to find other applications (Bellwood et al., 2002).

Web services (abbr. WS) are interoperable software components that can be used in DMS application integration and component-based DMS application development, and they are accessible through a WS interface. As the demand for WS consumption is rising, a series of questions arise concerning the methods and procedures to integrate the most suitable of them with ontological knowledge.
2.1. Ontological description of DMS WS

Web Service Modelling Ontology (WSMO) (Roman et al., 2005) and OWL-S (The OWL Services Coalition, 2004) can provide the infrastructure for ontological description of DMS Web Services. In particular, WSMO can describe DMS semantic WS to solve the integration problem. The WSMO describes all relevant aspects related to DMS services with the ultimate goal of enabling the automation of the tasks involved in integration of DMS WS. These tasks are: discovery, selection, composition, mediation, execution, monitoring, etc. Novel techniques have been proposed to facilitate semantic discovery and interoperability of WS that manage and deliver web media content (Sakkopoulos et al., 2006). WSMO has its conceptual basis in the Web Service Modelling Framework (WSMF) (Fensel and Bussler, 2002) refining and extending this framework and developing a formal ontology and set of languages.

The OWL-S is an ontology for service description based on the OWL. It can facilitate the design of DMS semantic WS and can be considered as “a language for describing services, reflecting the fact that it provides a standard vocabulary that can be used together with the other aspects of the OWL description language to create service descriptions”. The OWL-S ontology consists of the following three parts: (1) a service profile for advertising and discovering services; (2) a process model that describes the operation of a service in detail; (3) the grounding that provides details on how to interoperate with a service via messages. The vocabulary defined by OWL-S may be used to provide semantic annotations of DMS services, and automatic agents may process this information.

In the semantic WS area, other major initiatives are METEOR-S, and IRS-II. METEOR-S (http://lsdis.cs.uga.edu/Projects/METEOR-S/) aims at integrating WS technologies, such as Business Process Execution Language for Web services (BPEL4WS) (Andrews et al., 2003), WSDL and UDDI with Semantic Web technologies, in order to automate the tasks of publication, discovery, description and control flow of WS. The Internet Reasoning Service II (IRS-II) (Motta et al., 2003) is a semantic WS framework which allows applications to semantically describe and execute WS. Compared to IRS-II, WSMO focuses more on the description elements that are needed to deal with semantic WS. Conceptually, WSMO and IRS-II are not too different in the sense that both have common roots in UPML (Fensel et al., 2003a). IRS-II and WSMO are expected to converge as future versions of IRS which are planned to be WSMO compliant.

In this project we have been working with existing WS technologies, which have been semantically annotated to facilitate discovery and matching. Undoubtedly, the Semantic Web (Berners-Lee et al., 2001) enables better machine processing of DMS information on the Web, by structuring web destination documents in such a way that they become understandable by machines. It allows the destination content to become aware of it, and this awareness allows users and software agents to query and infer knowledge from DMS information quickly and automatically. Semantic Web technologies can provide interoperability, reusability and shareability among modular and service-oriented DMS (Kanellopoulos et al., 2004). A semantic-based DMS application can allow consumers or travel agents to create, manage and update itineraries. Moreover, it permits the customer (user) to specify a set of preferences for a vacation and query a set of information sources to find components such as air fares, car rental and leisure activities in real-time. Semantic-based DMS include: ontology-driven subject domain, repository of tourism destinations, destination presentation, adaptation and personalization. They can be adaptive to the user’s
personalization interests (e.g. information about transportation, restaurants, accommodation, services, weather, events, itinerary tips, shopping, nightlife, daily excursion, car rental, sport activities, etc.) as ontologies and associated meta-data deliver semantic information about the user behaviour and user interests. A prerequisite for this adaptation is to attach semantic metadata to tourism destination elements. Expertly, tourism destination tasks must be annotated in terms of tourism destination concepts and some instructional relationships between the involved concepts. Semantics can be used in the discovery and composition of DMS WS by applying service registries, and semantics can also be used in the discovery of DMS WS registries.

3. The requirements of semantic-based DMS in P2P context

We see different DMOs as one or many independent operating nodes within a “tourism destination knowledge” network. Nodes can join or disconnect from the network at any moment. A node acts as a peer in the network and access destination knowledge sources (e.g. a users’ file system, a local database). It can communicate with other nodes to achieve its goals. A node must be designed to meet the following requirements that arise from the task of sharing tourism destination information from the external resources with other peers:

- Distribution of tourism destination information within the network.
- Support for query answering and routing.
- Multiple views on available destination information.
- Multiple sources of destination information.
- Adaptation of destination information view according to the user’s personalization requirements.

The proposed metadata model (LA_DMS) needs to reflect these requirements.

Integration: Each piece of destination knowledge (e.g. a hotel’s location) requires metadata about its origin. To retrieve external information, we need to capture information about where the piece of information was obtained. This information will allow identifying a peer and locating resources in its repositories.

Information heterogeneity: As tourism destination information is added from a variety of peers, inconsistencies may occur in a local repository. For example, two DMOs have different information about a tourism destination’s element and inconsistencies occur in the local repository of a DMO. Destination information needs to be assigned such a confidence rating that the system will be able to handle heterogeneity and provide useful information. In the same way, a level of trust can be assigned to DMO peers to model their reliability. The level of trust (that the trusting peer has in the trusted peer at a given time slot with a given type of association relationship) is depicted by the trustworthiness that is computed in Hussain et al. (2004) and Dillon et al. (2004). Nejdl and Olmedilla (2004) considered the problem of passing credentials in P2P systems. More specifically, they recommended sharing credentials with third parties, if trust negotiation strategies allow this (Yu et al., 2003). In the PeerTrust project (Basney et al., 2004), trust negotiation in Semantic Web and P2P environments is investigated. Within this project, digital credentials can be signed XML or RDF statements that express peer properties, and policies are expressed as logic programs that tie resource access to required credentials.
Besides, the semantic interoperability is addressed through ontology mapping. This is a process whereby two tourism destination ontologies are semantically related at conceptual level, and the source ontology instances are transformed into the target ontology entities according to those semantic relations. An interesting approach to ontology mapping has been taken in the GLUE system (Doan et al., 2002). Furthermore, as each peer uses its own local ontology, mapping may be required to overcome the heterogeneous labelling of the same objects (e.g. hotels).

Security: Some tourism destination information may be of private nature and should not be visible to others peers (DMOs). Other information may be restricted to a specific set of peers. The proposed model (LA_DMS) needs to provide means to express these security policies. For example, in the case of Freenet a high level of security is maintained in order to support anonymity and freedom of speech (Sandberg and Clarke, 2002).

Caching: The availability of other peers is not always guaranteed within P2P systems. In particular, some peers may have better connectivity, in terms of bandwidth, to the rest of the network than other peers. Therefore, to improve network efficiency, caching of information can be useful. The caching mechanisms need to be transparent to the user, but must be captured by the proposed metadata model.

Adaptability: Destination concepts (e.g. hotel’s availability, museum’s location, attraction’s category, etc.) can also be used as a basis for implementing the adaptive behaviour of a DMS. Consequently, context-specific configuration of tourism destination elements and their adaptation according to the user’s personalization requirements can be enabled as semantic metadata are attached to tourism destination elements. For achieving this, ontologies being used must be aligned with the ontologies defining the context and user profile.

In open networked systems the following features affect collaboration, and need to be addressed by appropriate techniques (Castano et al., 2005): (1) dynamism of the system, regards the fact that the DMOs are allowed to join and leave the networked organization at any moment; (2) autonomy of DMO, in that each DMO is responsible for its own information resource management and representation; (3) absence of a priori agreement, about ontology specification vocabulary and language to be used for knowledge specification; (4) equality of responsibilities, no centralized entities with coordinating tasks are recognized and each DMO enforces interaction facilities with other DMOs for resource sharing and collaboration. Peers are not usually created equal but have different characteristics with respect to their capabilities (e.g. bandwidth, storage space or processing power). As discussed in Yang and Garcia-Molina (2002), exploiting the different capabilities in a P2P network can lead to an efficient network architecture, where a small subset of peers, called super-peers, takes over specific responsibilities for peer aggregation, query routing and possibly mediation.

4. The LA_DMS metadata model

To meet the requirements on a destination knowledge node, we designed a general infrastructure in the JXTA P2P framework (Gong, 2001). JXTA is a set of XML-based protocols to cover typical P2P functionality. It is a platform independent set of open protocols to facilitate P2P communication and application development. JXTA provides a Java binding offering a layered approach for creating P2P applications. It provides P2P protocols and services, including peer discovery, peer groups, peer pipes and peer
monitoring. The LA_DMS environment aims at providing a general view on the destination knowledge each peer has. It should facilitate the access to different destination information sources and enable the user to take advantage from other peers’ knowledge. It allows caching information to make the entire network more efficient.

*Internal destination knowledge sources:* Peers may have local sources of destination information, which represent the peer’s body of knowledge (e.g. local file systems, bookmark lists).

*Individual views:* Peers provide different views on the destination information they have in their local sources. The views can be implemented using different visualization techniques (topic hierarchies, thematic maps, etc.).

*Local node repository:* Each peer maintains an internal working model, which provides the following functionality (Broekstra et al., 2003):

- Mediate between views and stored tourism destination information.
- Support query formulation and processing.
- Specify the peer’s interface to the network.
- Provide the basis for peer ranking and selection.

### 4.1. The RDFS classes of the LA_DMS model

The LA_DMS model (as shown in Fig. 1) was defined in RDFS (Brickley and Guha, 2003; Decker et al., 2000) and consists of two RDFS classes namely the “Destination”-class and a “Peer”-class. Several properties support these objects:

![Fig. 1. The LA_DMS metadata model.](image-url)
**Destination:** Each piece of destination information links to a “destination”-object. This object contains the metadata or links to it. Kanellopoulos et al. (2005) proposed an ontology for tourism destinations, which is used for the annotation of unstructured web pages at the destination web sites. This destination ontology is able to answer four types of questions that can be asked. These questions involve mainly the predicates: *what, where, when* and *how.*

\[
\text{Destination\_id} = \{\text{What services are provided, Where these services are provided, When these services are provided, How these services are consumed}\}
\]

*What:* These are the available activities and attractions involved in a tourism destination.

*Where:* These are the places where the attractions and activities are located.

*When:* These are time-marks in which the tourists must visit this destination and its points of interest.

*How:* These are the ways that the tourist can get to this destination to participate in an activity.

*For\_travellers:* This information specifies for what travellers the destination is addressed.

*views:* This information includes possible views for the destination’s presentation.

**Peer:** Each destination information is originated by a peer and the local repository stores different information about each peer. The “Destination”-object links to the corresponding peer.

*peerID:* Each peer has a unique ID to be identified. This is the first attribute stored within the peer object, and it is the JXTA UID.

*peerLabel:* It is the peer label, which is a human readable and understandable description of the peer. An example could be: “AlteMar Central Computer”.

*peerTrust:* It is a measure to include trust. Some peers might be more reliable than others. To control the peerTrust it is defined to have a value between 0 and 1, with zero meaning having no reliability at all and one the peer being very trustworthy.

*URI:* To keep track of the origin of the information and its primary universal resource indicator (URI) this is explicitly saved among the metadata. Across the network an object can be addressed with *uri.*

**Location:** To access a document not only do we need the document-object in the destination ontology but also the address where the document can be accessed (e.g. file:///c:/work/myfile.txt). Only the peer where the information is physically stored needs to be able to interpret the expression. The location information is also required when doing updates of the local repository.

**Label:** The label saves how the specific destination information is called on the peer it originates from. The label-attribute is formulated in natural language (e.g. *Parthenon, hotel A, museum B, castle C,*…). The label is human readable and is added to each “Destination”-object and not only to the original object.

**Confidence:** Trust is used to measure the reliability of a specific peer. The confidence attribute returns a figure describing a specific statement. Again it is stored as a number between 0 and 1.
additionDate: This attribute keeps track of the date it was added to the local repository. This could be used to determine confidence; old information might become less reliable. The main reason for this will be to make the right updates.

Security: Some access control is required to ensure proper usage of the destination information.

Visibility: Some objects will have to be hidden, instead of being completely removed.

Cache: Caching of information is necessary to increase the network efficiency.

4.2. The LA_DMS model components

The LA_DMS model was built on the idea of separation of concerns, and thus advocates the separation of tourism destination information from the authoring perspective, as well as from the storage point of view. The conceptualization of the LA_DMS model was based on the LAOS model (Layered AHS-Authoring Model and Operators) (Cristea and Mooij, 2003) that is designed for the Educational Hypermedia. The LA_DMS model consists of five layers (or components):

1. **Domain model (DM):** containing a collection of linked tourism destination resources. It contains the specifications of tourism destination concepts and relations. The DM is a hierarchical model, in which hierarchy can be considered as the hierarchy of attractions (e.g. monasteries, castles, caves, hotels …) in a tourism destination as shown in Table 1. In this hierarchy, each component is a concept that has a semantic unity and it is labelled appropriately, and reused. The domain attributes represent different aspects (views) about the same concept. For example, a destination location is one of the aspects of a tourism destination.

2. **Goal and constraints model (GM):** containing goal-related information, such as instructional information about the tourism destination resource. Therefore, the GM model can be called Travel Model, because (by using this) e-tourists can set their travelling goals concerning a tourism destination, and organize their travel. The GM is a constrained version of the DM, as from one DM, several presentations can be produced or multiple GM versions can be generated. The GM represents the filtering steps towards putting together and organizing the tourism destination material for satisfying the e-tourist’s personalization requirements. Note that the division of the contents into DM and GM does not restrict the user view.

3. **User model (UM):** containing user-related information, such as information about the e-tourists. The UM is a concept map, allowing explicit relations between the variables within the UM.

4. **Adaptation model (AM):** containing the behaviour and dynamics, such as, a navigation style related strategy. The AM determines the whole dynamics of the adaptive

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hierarchy of attractions in a tourism destination</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tourism destination</th>
<th>Hotel</th>
<th>Monastery</th>
<th>Museum</th>
<th>Cave</th>
<th>...X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Castle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subcategory</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subcategory1</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subcategory2</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
hypermedia system and is composed of IF-THEN rules (e.g. IF concept.knowledge<threshold THEN concept.hide). The IF-THEN rules form a low-level assembly adaptation language. The UM can generate automatically adaptation rules by applying templates for specific UM attributes. For example, to illustrate a pure usage of UM elements to generate an AM rule, we consider the same state of ‘interest’ about a user concept (e.g. hotel_X). We want a rule that displays everything in the user concept, if this concept is of interest to the user:

\[
\text{IF (hotel}_X\text{.interest}>\text{threshold})\
\text{THEN } \{\text{hotel}_X\text{.name.available = ‘true’; hotel}_X\text{.contents.available = ‘true’}\}
\]

Such rule is a generic one, which can be applied on all user concepts in a user map, therefore drastically reducing the workload. Adaptation strategies represent the only dynamic part of the LA_DMS model. They are the ones instructing the delivery engine about how to handle the static data generated by the other layers.

5. Machine or presentation model (MM): containing display and machine-related information, such as the foreground-background colour scheme for the destination presentation. It addresses the specifications of views for DMS websites, including navigational structures and detailed user interfaces. The views can be implemented using different visualization techniques (topic hierarchies, thematic maps, etc.). There are some predefined views that correspond one-to-one to the different sources of destination information (e.g. hotels, museums, etc.).

The LA_DMS model has five RDF schema elements: ‘domainMap’, ‘goalAndConstraintsMap’, ‘userMap’, ‘adaptationStrategy’ and ‘presentationMap’, which are defined as being of a type of the same name. The key-idea of the LA_DMS model is that it allows flexible re-composition of its elements, according to the user’s personalization requirements. For each layer type (e.g. domain, user, machine ...) the LA_DMS model can contain an unbound number of maps. This means, for instance, that several DM maps corresponding to several kinds of tourism destinations (e.g. hotels, museums, monasteries, castles, caves, etc.) can be described within this model. Therefore, in the DM there are hotel, museum, castle, caves maps and other tourism domain maps. Similarly, different tourists’ goals will result in transforming the same domain map, for instance, into many different goal and constraints maps. Clearly, more than one users (e-tourists) can be defined with this model. Although individual user maps can also be defined with LA_DMS, the idea is to define either stereotypes, or group of users, so that the same basic model can be reused. For example, during the actual interaction of an e-tourist, with the delivery system, the basic model will get updated and will generate several individual versions. This can happen, if all e-tourists are interesting in museums, but some of them have been visiting (navigating) archaeological museums. Their knowledge level will be accordingly updated by the delivery system and will be different for each user, although they all belong to the same basic category: “users interested in museums”.

Additionally, different presentation maps can be defined, giving the parameters, for instance, for a desktop presentation, or a palm presentation (of a tourism destination). Finally, the tourism destination material can be presented according to one or more adaptation strategies that can correspond to instructional strategies. For example, an instructional strategy can be implemented by displaying the destination concepts
preferentially at the same depth of the conceptual tree of either a goal and constraints map or domain map. Hereafter we give generic definitions of the LA_DMS layers.

A **domain map** (DMap) of the LA_DMS is determined by the tuple \( < C, L, Attr > \); where \( C \) a set of concepts; \( L \) a set of links and \( Attr \) a set of DMap attributes. A domain concept \( c \) that belongs to \( DMap_i \) is defined by the tuple \( < A, C > \); where \( A \) is a set of domain map attributes; \( C \) a set of domain map sub-concepts (e.g. “archaeological museum” is a sub-concept of the domain concept “museum”).

A **goal and constraints map** (GMap) is a tuple \( < G, GL, Gattr > \); \( G \) represents a set of goal and constraints concepts; \( GL \) is a set of goal and constraints links and \( Gattr \) is a set of goal and constraints attributes. A goal and constraints concept \( g \) (e.g. hotels of class A, located in Athens) is defined by the tuple \( < GA, G, DMap_j.c.a > \); \( GA \) is a set of attributes; \( G \) a set of sub-concepts; \( DMap_j.c \) is the ancestor domain map concept and \( DMap_j.c.a \) is an attribute of that concept. Each goal and constraints concept \( g \) must be involved in at least one special link \( gl \), called prerequisite link.

A user concept \( u \) is defined by the tuple: \( < AU, U, GMap_i(g.(a))/DMap_j.c.(a) > \) \( AU \) is a set of UM attributes; \( U \) a set of UMap sub-concepts; \( GMap_i(g.(a))/DMap_j.c.(a) \) is the ancestor goal and constraints map.

Sub-concepts (e.g. archaeological museum) within a concept (e.g. museum) can be OR-connected (therefore becoming concept alternatives, from which the appropriate one will be selected according to user map variable settings) or AND-connected (meaning that a user has to view all sub-concepts regardless). To facilitate this, a concept contains a concept attribute, which in its turn contains a holder for OR-connected sub-concepts or a holder for AND-connected sub-concepts. The holder contains the actual sub-concepts in a specified order. A concept attribute contains, besides the sub-concept holders, one or more pointers to domain concept attributes. This is the link with the concept domain. The idea is that a concept puts pieces of information that are stored in the concept attributes together in a suitable way for presentation to an e-tourist. A sub-concept which has no sub-concepts (e.g. is a leaf in the sub-concept hierarchy) corresponds to a (one) concept attribute. By using predefined algebraic operators (as those used in the LAOS model) we can create, delete, repeat, view, list … certain elements of the different LA_DMS model’s layers.

Destination elements information (e.g. hotel’s availability) is changed frequently, and it has to malleable sufficiently so that it can be reused in different settings. This kind of change can focus on the new issues of a tourism destination and refine the old ones. Therefore, a layered approach with appropriate semantic labelling for tourism destinations is necessary (Kanellopoulos et al., 2005). The layers of such an architecture should reflect a higher-level semantics, such as DM (e.g. hotel’s class), user’s characteristics (e.g. e-tourist’s profile) and machine characteristics. The LA_DMS model architecture is depicted in Fig. 2 for the hotel context. In the DM at a lower level, the semantics have to be applied all the way to the lowest level of modification (and thus of reuse). For example, the hotel’s availability is changed over and over again (i.e. highest level of reuse) and it should be appropriately labelled in order to be easily retrieved according to its semantics (Kanellopoulos et al., 2005).

The use of the LA_DMS system is not restricted to LA_DMS peers which have used the LA_DMS-common namespace. It represents the peer’s destination knowledge about the network. As long as the other peer understands the query language it can participate as an information provider and seeker.
Fig. 2. The LA_DMS model architecture in the hotel context.

The example shows a statement with its corresponding “Destination” information:

```xml
<!-- peer1 -->
<rdf:Statement
    rdf:about = "http://LA_DMS://1234567890.jxta#statement01">
    <rdf:subject rdf:resource = "http://LA_DMS://1234567890.jxta#project"/>
</rdf:Statement>
```

```xml
<!-- peer1 -->
<rdf:Statement
    rdf:about = "http://LA_DMS://1234567890.jxta#statement01">
    <rdf:subject rdf:resource = "http://LA_DMS://1234567890.jxta#project"/>
</rdf:Statement>
```
4.3. Applications of the LA_DMS model

By using the LA_DMS model:

(a) An e-tourist (user) can express a goal concept and its constraints in the global destination context and find the relevant tourism material. For example, a user wants to find and view “hotels of class A located in Athens” (via the Domain Map of hotels). This goal and constraints concept $g = \text{"hotels of class A located in Athens"}$ is involved in at least one special link $gl$. The goal concept and its constraints can be more complicated. However, the predefined constraints for a tourism destination are based on the predicates: what, where, when and how, which are related to tourism destination ontology.

(b) An e-tourist can view destination presentations (e.g. a hotel’s presentation) according to his/her requirements or the characteristics of his/her computer.

(c) An e-tourist interested in specific destination elements (e.g. museums) can view automatically destination presentations for these elements according to his/her profile that was generated in past navigation sessions. Using ontology to model user profiles, the semantics of user preferences can be captured so that the LA_DMS system to be more accurate in providing personalized services to the users. Using ontological inference, implicit interests can be derived based on the explicit preference specifications (Middleton et al., 2004). If an e-tourist is interested in a particular DMS service of a class $C$ (e.g. archaeological museums), he/she might also like services of the superclass of $C$ (museums) to some degree. Adaptation strategies determine what tourism destination elements the user should see, if the user belongs to a category of e-tourists (e.g. users interested in museums).
The LA_DMS model can constitute an ideal platform for dynamic packaging applications as the Web has permanently changed the manner in which vacation packages can be created. Consumers can now acquire packages from a diversity of websites including online agencies and airlines. The creation of dynamic vacation travel packages is vital for the travel industry. The objective of dynamic packaging is to pack all the components chosen by a traveller to create one reservation. Regardless of where the inventory originates, the package that is created is handled seamlessly as one transaction, and requires only one payment from the consumer. Dynamic packaging systems create customized tourism packages for the consumers. Cardoso (2005) proposes a platform to enable dynamic packaging using Semantic Web Technologies. A dynamic packaging application allows consumers or travel agents to bundle trip components. The range of products and services to be bundled is too large: guided tour, entertainment, event/festival, shopping, activity, accommodation, transportation, food and beverage, etc. Dynamic packages can be created and booked effortlessly with private and published air, car hire, hotels, attractions and insurance rates.

Furthermore, in the LA_DMS model the Cluster Map technique (Fluit et al., 2002) is adopted, which supports several user tasks such as analysis, query and navigation of semantic structures. Inspecting the same data source from different perspectives and visualizing it according to different views, we achieve semantic structure analysis. For example, views can obtain all the sub-concepts of the “geographic-location” concept and result in the visualization of the tourism destination reservations according to their geographical distribution.

5. The RDF-based P2P architecture

As an example we consider the destination management organization DMO1, which provides information resources related to the tourism destination domain. As shown in Fig. 3, the DMO1 ontology contains the Accommodation, Attraction, Reservation, and Travel_document concepts; the Reservation concept is a specialization (i.e., subclassOf) of the Travel_document concept. Furthermore, the value of the property Accommodation of the Reservation concept refers to the Accommodation concept. In particular, strong property constraints are defined with a property restriction by setting the min-Cardinality clause to the value 1 (e.g. reservation_code). Each DMO in the system acts both as a client and a server in the networked organization interacting with other DMOs directly, by submitting queries containing a request for one or more resources. To receive the right answers without flooding the peer network with queries, queries are directed to the ‘right’ peers via ontology-based peer selection mechanisms which exploit similarity of ontologies for this purpose.

Our RDF-based P2P infrastructure is based on the following building blocks (Nejdl et al., 2003):

- The RDF-S language, to define and use the five schema elements of the LA_DMS, which specify the kind of destination data available in our P2P network.
- The query language, to retrieve the data stored in the P2P network.
- A network topology combined with an appropriate query routing algorithm, to allow efficient queries.
- Facilities to integrate heterogeneous destination information stored in the P2P network.
5.1. The RDF-schema language

Using RDF-S we represent schemas based on classes, properties and property constraints, and we define the vocabulary used for describing our destination resources. RDF triple \( \langle \text{subject}, \text{property}, \text{value} \rangle \) represent specific annotations, where subject identifies the resource we want to describe (using a URI), property specifies what property we use, and value the specific value, expressed as primitive data type or another URI. RDF schemas (Lassila and Swick, 1999) are flexible and extensible and can be evolved over time (e.g. by extending them with additional properties). We can use any properties defined in the schemas we use, possibly mix different schemas, and relate different resources to each other, when we want to express interdependencies between these resources, hierarchical relationships, or others. In the LA_DMS framework, the five RDF schema elements: ‘domainMap’, ‘goalAndConstraintsMap’, ‘userMap’, ‘adaptationStrategy’ and ‘presentationMap’ have been defined as being of a type of the same name. We used the RDFS Reasoner of Jena2 (a Hewlett-Packard’s Semantic Web toolkit for Java programmers: \text{http://sourceforge.net/projects/jena}) to solve heterogeneity at the ontological level. The RDFS Reasoner provides an implementation of the RDFS closure rules, and in RDFS Reasoner each domain, range, sub-property, and sub-class declaration is eagerly translated into a single query rewrite rule.

5.2. Network topologies

In our RDF-based P2P network, data is not owned by any participant (DMO) while each participant is responsible for supporting the community. In particular peers provide
and use explicit RDF descriptions of their content. Our network can adopt or use two different approaches:

- The **super-peer topology** (Nejdl et al., 2004), in which the super-peer backbone can be responsible for message routing and integration/mediation of tourism destination metadata. This super-peer based network can support our RDF-based network, with different metadata schemas and ontologies. Moreover, it can provide better scalability than broadcast-based networks. Based on metadata routing indices (stated in RDF), the super-peer network can support sophisticated routing and distribution strategies. Our super-peers network can be arranged in a hypercube topology, according to the HyperCuP protocol (Schlosser et al., 2002). Using this protocol, super-peers connect to each other. Different kinds of super-peer indices describe the data and schema characteristics of the peers connected to the “destination” network. Super-peer indices exploit the RDF ability to uniquely identify schemas, schema attributes and ontologies, and provide a necessary ingredient for our schema-aware P2P data management infrastructure.

- The **content-based networking model** (Carzaniga et al., 2001), in which a message is transmitted from a DMO sender to one or more DMO receivers without the sender having to address the message to a specific receiver. Receivers DMOs would express interests in the kind of messages they would like to receive, and the network delivers to the receivers any and only messages matching those interests. DMOs receivers express interests through **predicate advertisements**. In this receiver-driven style of communication, our network is responsible for efficiently applying predicates to the content of messages so as to minimize the computational and communication costs of the network.

6. Related work

In the travel domain there have been efforts in defining semantics, such as the Harmonise and Satine projects. The Harmonize project allows participating tourism organizations to keep their proprietary data format and use ontology mediation while exchanging information (Dell’ Erba, 2004; Missikoff et al., 2003). Satine project developed a secure semantics-based interoperability framework for exploiting WS platforms in conjunction with P2P networks in tourism industry (Dogac et al., 2004). The EU-IST project SWAP (http://swap.semanticweb.org/) demonstrated that the power of P2P computing and the Semantic Web could actually be combined to share and find “knowledge” easily with low administration efforts, while participants can maintain individual views of the world. In the travel domain, the advantages of Web Semantics and P2P computing for service interoperation and discovery have been analyzed by Maedche and Staab (2003). In integrating P2P and ontologies, a big challenge is to design ontology-based **peer selection services** in order to avoid flooding the peer network with queries (viz. one must ask the ‘right’ peers to receive the right answers). These peer selection services will be relied on mechanisms, which will exploit similarity of ontologies for this purpose (Fensel et al., 2003b).

An overview of development methodology and applications for tourism ontologies is given in (Jakkilinki et al., 2005). Ontologies are created using ontology development tools, such as **Protege 2000** (Protegé-2000 tool, 2000), viz. a Java-based ontology editor with **OWL Plugin**: that means that it allows ontology implementation as an applet on the Web. This permits multiple users to share the ontology. Mondeca’s tourism ontology (http://www.mondeca.com) defines tourism concepts based on the WTO thesaurus (World
Tourism Organisation–WTO, 2002). These concepts include terms for tourism object profiling, tourism and cultural objects, tourism packages and tourism multimedia content. The development of another comprehensive and precise reference ontology named Comprehensive Ontology for the Travel Industry (COTRIN) is the major result of another research group (Cardoso, 2004). The objective of COTRIN is the implementation of the semantic XML-based OTA specifications. Major airlines, hoteliers, car rental companies, leisure suppliers, travel agencies and others will use COTRIN to bring together autonomous and heterogeneous tourism WS, web processes, applications, data, and components residing in distributed environments.

In tourism, new offers and requests typically come in by the minute and late vacancies of rooms, flights or lodging can be lost easily. Therefore, there is a need for a fast match between providers and requestors. In e-markets that exploit semantic descriptions, semantic-based matching of products and requirements is made fast between tourism providers and requesters, while a large volume of transactions is executed. Antoniou et al. (2005) considered the brokering and matchmaking problem in the tourism domain, that is, how a requester’s requirements and preferences can be matched against a set of offers collected by a broker. Their proposed solution uses RDF to represent the offerings, and a deductive logical language for expressing the requirements and preferences. Sycara et al. (1999) described a comprehensive software agent framework that allows the set up of semantic-based e-markets. In semantic-based e-markets, intelligent software agents can exploit semantics on the Web. The Semantic Web can utilize a variety of traveller, hotel, museum and other software agents to enhance the tourism marketing and management reservation processes (Hendler, 2001). For example, a hotel software agent might undertake many of the routine administrative tasks that currently consume large amounts of a hotel manager’s time. On the other hand, traveler software agents can assist travellers in finding sources of tourism products and services and in documenting and archiving them. An additional capacity of the Semantic Web is realized, when software agents extract information from one DMS application and subsequently utilize the data as input for further applications. Thus, software agents create greater capacity for large scale automated collection, processing and selective dissemination of tourism data.

The Edutella project (Nejdl et al., 2002) relies on the W3C metadata standards RDF and RDF-S to describe distributed learning resources, and uses basic P2P primitives provided as part of the JXTA framework (Gong, 2001). It implements an RDF-based infrastructure for P2P networks based on the JXTA framework. An Edutella successor (Nejdl et al., 2004) provides better scalability by introducing super-peers and schema-based routing; however, it requires up-front definition of schemas and designation of super peers. In Broekstra et al. (2003) an interesting metadata model was proposed that combines features of traditional ontologies with rich metadata about the origin of information and the reliability of sources. Also, methods for creating, assessing and accessing semantics and metadata were introduced.

Cai et al. (2004) proposed a scalable P2P RDF repository (RDFPeers) which stores each triple in a multi-attribute addressable network by applying globally known hash functions. In their framework, queries can be efficiently routed to the nodes that store matching triples. Zhou et al. (2002) focused on searching in P2P based hypermedia systems. Their search algorithm is based on a notion of semantic equivalent data, and uses this knowledge to rearrange the network topology in order to cluster semantic related entities. This is achieved by letting each peer publish a topic list. These lists are used by other peers to
maintain information of neighbours, which hold semantically equivalent topics. Bernstein et al. (2002) proposed the local relational model (LRM) enabling general queries to be translated into local queries with respect to the schema supported at the respective peer, using the concept of local translation/coordination formulas to translate between different schemas. Finally, Aberer and Hauswirth (2002) proposed schema-based peers and local translations to accommodate more sophisticated information providers by a Gnutella-like P2P topology.

7. Conclusions

In this paper, we propose a metadata model (LA_DMS) encoding semantic tourism destination information in an RDF-based P2P network architecture. The LA_DMS model combines ontological structures with information about tourism destinations and peers. The LA_DMS project aims at providing semantic-based tourism destination information by combining the P2P paradigm with Semantic Web technologies and AH. The model is adaptive to the user’s personalization requirements (e.g. information concerning transportation, restaurants, accommodation, etc.) and it is a collection of opinions supported by different sources of tourism destination information. In the LA_DMS framework, the semantic models of different peers (DMOs) are being aligned dynamically. The LA_DMS model supports mainly:

- **Flexibility**: Different semantically meaningful combinations can be generated by automatically populating the different layers of the LA_DMS model, based on previous ones. This processing is done automatically, because the data in the original layers is semantically well labelled.
- **Expressivity**: The semantics of the LA_DMS model’s elements are machine understandable and also easy to grasp for humans. Therefore, an author of a tourism destination can understand what data and metadata he/she is creating.
- **Reusability**: The model enables reuse of all aspects of the AH.
- **Interoperability**: The LA_DMS framework is generic enough for the authoring of adaptive DMS to be easily converted into material for different DMS delivery platforms.
- **Standardization**: The LA_DMS framework describes and extracts patterns at the different levels of granularity, starting with the above five layers and detailing each layer separately. These patterns should be able to feedback into extant standards and provide information for enriching them according to the needs of adaptivity and the e-tourists’ demands. Besides, the feedback from e-tourists can be used to extend the adaptation language.

DMOs demand to manage tourism destinations effectively utilizing Semantic Web technologies, which will re-engineer their role as planners, coordinators, regulators and promoters of tourism destinations. DMS should take advantage of semantic services, interoperability, ontologies and semantic annotation. Semantic Web could offer more flexibility in DMS through use of new emergent RDF-based P2P network services.

Nonetheless, **proof** and **trust** is emerging as another central issue. The **proof layer** (of the Semantic Web architecture) involves the actual deductive process, as well as the representation of proofs in web languages and proof validation. **Trust** emerges through the use of digital signatures, and other kind of knowledge, based on recommendations by software agents we trust, or rating and certification agencies and consumer bodies. **But,**
how do we know that what our intelligent software agent has discovered through its trawl of the Semantic Web can be trusted? Even in the case of ontologies how should we decide whose tourism ontology to trust? This is especially important where the two ontologies may conflict with one another. The current design of the LA_DMS model does not answer satisfactory the above questions. Similarly, we are faced with the difficulties of ensuring and maintaining semantic integrity and a lack of methods for testing its presence.

At the present time, the requirements of intelligent tourism information systems raise several technical research issues, such as: (1) semantic interoperability and mediated architectures; (2) e-business frameworks supporting processes across DMOs; (3) mobility and embedded intelligence; (4) personalization and context-based services and (5) information-to-knowledge transformations—data mining and knowledge management.

References


Homepage of Gulliver, Ireland: http://www.gulliver.ie/.


