Chapter XXXIII
The Semantic Web in Tourism

Salvador Miranda Lima
Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Viana do Castelo, Portugal

José Moreira
IEETA / DETI, Universidade de Aveiro, Portugal

ABSTRACT

The emergence of the World Wide Web made available massive amounts of data. This data, created and disseminated from many different sources, is prepared and linked in a way that is well-suited for display purposes, but automation, integration, interoperability or context-oriented search can hardly be implemented. Hence, the Semantic Web aims at promoting global information integration and semantic interoperability, through the use of metadata, ontologies and inference mechanisms. This chapter presents a Semantic Model for Tourism (SeMoT), designed for building Semantic Web enabled applications for the planning and management of touristic itineraries, taking into account the new requirements of more demanding and culturally evolved tourists. It includes an introduction to relevant tourism concepts, an overview of current trends in Web Semantics research and a presentation of the architecture, main features and a selection of representative ontologies that compose the SeMoT.

INTRODUCTION

The concept of tourism emerged in the geographic area which is currently known as the European Union, and came out as a strongly European phenomenon in such a way that Europe is still the prime touristic destination as well as a spinning platform for intercontinental tourist flows, as confirmed by the statistics of the World Tourism Organization. In fact, tourism is the economical and social activity with the biggest impact and significance within the European Union, in such a way that it constitutes a strategic activity for sustained development that ensures the current
patterns of living in countries like Portugal, Spain, Austria or Greece.

Tourism demand in the following years or decades will be regulated by the a new social paradigm (considered in the current developments of Alvin Toffler’s “third wave” (Toffler, 1980)) centred in the multiplicity of “selves”; in other words, the new consumer will be profiled as having a great diversity of information which will be used as an instrument of demand, agile in the use of intelligent technologies and caring for environmental, cultural and ethnic problems.

In order to face the sustainability demands of European tourism development, as well as the new wave of more demanding and culturally evolved tourists, the development of models of touristic information in the light of the Semantic Web paradigm may play an important role.

Over more than a decade and a half after Tim Berners-Lee presented his original proposal for creating a universal platform of access and dissemination of information between the European Organisation for Nuclear Research (CERN) researchers, the Web became one of the artefacts that have metamorphosed all dimensions of society in an extraordinarily decisive, powerful and fast way. Such an artefact has also limitations that can be numbered as follows:

i. **Information search:** Current search engines ignore the context of the terms (restaurant, museum, handicraft, etc.) or the connections between them (a museum has a restaurant, a restaurant is classified as a space for smokers, etc.);

ii. **Extraction of information:** Obtaining decisive and influential information demands relying on human interpretation both for selection and filtering (“Margarida da Praça” can be both a restaurant or a cod-fish specialty);

iii. **Management:** Website content management can be complex, despite the simplicity of the documents’ organisation structure; in this way, the application of appropriate Web Engineering practices will play an important role in both the organisation and classification of touristic information in the Web;

iv. **Automatic document generation:** The inclusion of adaptability mechanisms will reduce the level of user dissatisfaction regarding Webpages which are adaptable to their profiles (museums with the most favourite painters, traditional fish-based gastronomies, etc.).

A touristic itinerary, in its essence, is an aggregator of touristic objects which semiotically have a significant and a meaning given by the resident communities, associating sufficiently attractive eclectic and singular values which promote or stimulate tourism. In this chapter, we contextualise such touristic objects in three kinds of information: spatial information (particularly, the geographical location defined by geographical coordinates and the type of location), temporal information (such as the working schedule and the duration of an event) and thematic information (such as intrinsic and extrinsic values of attractions and touristic descriptors).

Webpages contextualized to the touristic domain, in the current or syntactic Web, are prepared to be interpreted by humans and are, therefore, inadequate for “intelligent” processing by computing agents. Let us consider the example of a search for the restaurant “Margarida da Praça” in the Web. Current search engines are unable to differentiate our object of interest from other homonyms that occur in the Webpages retrieved by the search engines. This means that for the same object of interest we can identify the name of the restaurant, as well as a person’s name or a gastronomic specialty. To improve the effectiveness of search engines in what concerns our object of interest, we insert into the Webpages, besides the syntactic reference, the information of what “Margarida da Praça” means (metadata). So, we need semantic structures that represent