

SEED, SEmantic E-tourism

Dynamic packaging

Jorge Cardoso and Jorge D. Fernandes
<http://seed.expedita.com.pt/>



Jorge Cardoso joined the University of Madeira (Portugal) in March 2003. He previously gave lectures at University of Georgia (USA) and at the Instituto Politécnico de Leiria (Portugal). Dr. Cardoso received his Ph.D. in Computer Science from the University of Georgia in 2002. In 1999, he worked at the Boeing Company on enterprise application integration. His research interests are in the areas of business processes, workflow management systems, semantic Web, and related fields.

Dr. Cardoso received his Ph.D. in Computer Science from the University of Georgia in 2002. In 1999, he worked at the Boeing Company on enterprise application integration. His research interests are in the areas of business processes, workflow management systems, semantic Web, and related fields.



Jorge Dias Fernandes works at Expedita (Portugal - Madeira), a tourism-specialized IT company, since 1999. He is responsible for research & development.

During various years he was an invited teacher at the University of Madeira. Previously, he participated in several D.A.I. and tourism-related research projects coordinated by the Instituto Superior Técnico (Lisbon) and the Faculdade de Ciências. Jorge received his MSc in Electrical and Computer Engineering in 1997 from Instituto Superior Técnico. His research interests are in the areas of business processes, tourism reservation systems and knowledge representation..

■ *Tourism is one of the sectors of the World economy with the best outlook. The World Tourism Organization predicts a 200% increase in tourist arrivals around the world by 2020. This is due to factors such as the increase in*

time for leisure activities and its social importance. According to Carl H. Marcussen from the Centre for Regional and Tourism Research of Denmark, the European online travel market size has increased from 2,5 Billion Euros in 2000 to 18,2 Billion Euros in 2004, and a further growth to 23,4 Billion Euros in 2005 and 28,5 Billion Euros in 2006 is expected.

The rapid growth of the Internet and the continual adoption of innovative technology have led to serious changes in the travel industry during. Due to the constantly changing business environment, one of the latest concepts of the tourist industry is "Dynamic Packaging". We can define dynamic packaging as the ability of a system to combine different tourism products in order to create a package. This ability will offer advantages for the tourist, because creating packages he can obtain products with lower prices. Another advantage is that the tourist only has to introduce his personal data, like name, address, credit card, etc, only one time to buy all the products in a package.

The SEED project was started with the objective of developing a new way to implement dynamic packaging systems. To create dynamic packages, systems must integrate different tourism data sources. This data sources can have a very different data formats and can be accessed by very different forms. To deal with the heterogeneity of the tourism data sources, we will use the Semantic Web technology. The use of Semantic Web technology will help us in the integration of the data sources. By creating a semantic model of the tourism domain and associating this model with each one of the data sources, we can more easily integrate them.

Funchal, Portugal, 25.02.2006

1. General Project Information

Objectives

The main objective of our project is helping the travel industry to take full advantage of the latest Internet technologies, i.e. Web services, Web processes, and semantics. In this context, we are implementing a semantic architecture for dynamic packaging applications. The architecture integrates tourism data systems in order to allow the creation of dynamic packages.

Main Deliverables

At the end of the project we expect to deliver the following items:

- Data Extraction System
- Data Model Mapping Application (Syntactic/Semantic)
- An Ontology for e-tourism
- Dynamic Packaging System
- Application for defining business rules for e-tourism

Scientific and Technical Coordinator

The project is coordinated by Jorge Cardoso, assistant professor at the University of Madeira, and by Jorge Fernandes, administrator of Expedita, an organization whose main activity is the development of systems for the tourism industry.

2. Partners Details

There are two main Partners in the project:

- The University of Madeira participates in the project through the Mathematics and Engineering Department. The PI is Prof. Jorge Cardoso. Currently, 3 master students (Miguel Gouveia, Jorge Sousa and Lisete Escórcio) and 3 senior students (Bruno Silva, Pedro Rosa and Toni Rodrigues) are involved in the project. Past contributors included Tobias Schmeing, from the University of Essen - Germany.
- Expedita - Arquitectura e Gestão de Sistemas de Informação Lda. is an IT

organization that develops several projects for the tourism industry. The company participates in the project through its administrator, Jorge Fernandes. He is the co-PI of the project.

3. WorkPlan Overview

The workplan for the project is divided into five sections. This division is necessary because five different teams are developing five different parts of the project.

Tourism Data Sources Integration System workplan:

- Web page development and study of ontologies and semantic (September 2005)
- Study extractor/wrappers technology and tools (October and November 2005)
- Architecture development (December 2005)
- Usage scenario description and study of related work (January 2006)
- Write a paper about Semantic Data Extraction for B2B Integration (February 2006)
- Implementation phase (March 2006)
- Testing and Optimization phase (May 2006)
- Integration and testing phase and write the final report (June 2006)

Data Model Mapping Application workplan:

- Getting familiar with several concepts related with the SEED project (September 2005)
- State of Art - Studying mapping and Semantic Web tools already available (October 2005)
- Preliminary technology evaluation (November 2005)
- Testing code as concept-proof for transforming XML to OWL using XSLT (November 2005)
- Requirement Analysis (December 2005)
- Design (January and February 2006)
- Implementation - Domain & Persistence (March and April 2006)
- Implementation - User Interface (April to June 2006)
- Final Report (June 2006)

Ontology for e-tourism workplan:

- Study the domains in which ontologies and semantic are being used (September 2005)
- Study the use of ontologies and semantic (October 2005)
- Study RDF, RDFS and OWL Languages (November 2005)
- Define the ontology vocabulary (November 2005)
- Install Protégé and use it to build a small ontology for e-tourism (November 2005)
- Study the different methodologies to build ontologies (December 2005)
- Write a survey describing the different ontology tools (January 2006)
- Identify ontology instances (February 2006)
- Building and develop a new methodology to build ontologies (March 2006)
- Build an ontology according to the proposed methodology (April 2006)
- Write the thesis (July 2006)

Semantic Rules workplan:

- Study First Order Logic (September 2005)
- Study rules reasoner (September 2005)
- Study the use of RuleML and SWRL rule languages (October 2005)
- Study Bossam Language (November 2005)
- Select the best rule language for the project (December 2005)
- Study the different kind of business rules (January 2006)
- Identify the different categories of business rules that can be applied to the tourism domain (February 2006)
- Identify the different rules in each categories (February 2006)
- Specify the rules in SWRL (March 2006)
- Write the Thesis (July 2006)

Dynamic Package System workplan:

- Study of data sources systems and extractor systems that can be integrated to the dynamic package system (September and October 2005)
- Define the integration of the data sources with the dynamic package system (November and December

2005)

- Analyze the possible query languages to use in the system (January 2006)
- Define the main modules of the architecture (February 2006)
- Detailed each one of the modules (March 2006)
- Implementation of the defined architecture (April to June 2006)
- Integration of all the modules and testing the system (July 2006)
- Write the Thesis (August and September 2006)

4. Deliverables

The architecture will be composed by components that can be grouped into packages to possibly be used and installed individually:

- Data Extractor: Component that allows the dynamic integration of different information sources and channel data. Besides handling different data types, the component is able to manage data sources, connect to remote and local data sources, knows how, and where, to get specific information, respond to data retrieval queries and handle queries and exceptions. In the global architecture, it has the responsibility to retrieve the information that is in the tourism systems and that is requested to build the dynamic packages.

- Data Model Mapping: Implements a user-friendly interactive mapping tool that allows users to map syntactic data in a XML format to concepts of an OWL ontology. In the context of the global architecture, this component must generate mapping rules that, once defined, allow converting any XML structure to instances of a global shared data ontological model.

- Ontology for e-tourism: An ontology that defines the concepts that are present in the e-tourism domain. The ontology will be developed by taking in consideration the traveler's point of view. It defines the global shared data model for the architecture and will be used in

the integration of the different information sources and in the definition of packaging rules.

- **Semantic Rules:** A component that allows the definition of rules based on a specific ontology. It will include an engine to validate and manage rules. It is used in the dynamic package architecture to define the rules for the creation of dynamic packages. Using this component, we can separate dynamic packaging logic from dynamic packaging applications.

- **Dynamic package system:** The final system, composed by all the components that are necessary to create the dynamic packaging system. It has the responsibility of integrating all the other components and guarantee that processes are executed in order to build the dynamic packages.

5. Workshops/sponsored Events

IWDDS 2006. International Workshop on Dynamic Distributed Systems (IWDDS 2006), In conjunction with the ICDCS 2006, The 26th International Conference on Distributed Computing Systems July 4-7, 2006 - Lisbon, Portugal.

SDWP 2006. The 3rd International on Semantic and Dynamic Web Processes (SDWP 2006), In conjunction with the 2005 IEEE International Conference on Web Services (ICWS 2006), September 18-22, 2006, Chicago, USA.

6. Open Tools for public

In the end of the project we intend to provide the following tools and modules:

- **Data Extractor:** A tool to extract information for different data sources and integrate them through XML Schemas and OWL representing a specific domain model. It can respond to queries based in the XML Schemas and returns information respecting the same data schemas.

- **Data Model Mapping:** A tool for

mapping XML data to OWL data. Returns a XSD document that defines the transformation of XML data to OWL data. This tool allows storing the mapping rules for future editing.

- **Ontology for e-tourism:** An ontology that defines the concepts presented in the e-tourism domain.

- **Semantic Rules:** A tool to define and manage semantic rules.

- **Dynamic Package Application:** A system for creating dynamic packages. The packages are created using the information of data sources integrated in the system, by an ontology for the e-tourism and by specific rules defined using the ontology concepts.

All these tools will be freely available in the SEED Web site:

<http://seed.expedita.com.pt/>

Besides these tools, the SEED Web site contains a lot of useful information related to the project, such as work plans, the technology that is being used, analyses of similar systems, problems and rational about the decisions taken, etc.

7. Contact details

Jorge Cardoso
Department of Mathematics and Engineering
University of Madeira
9050-078 Funchal - Portugal
jcardoso@uma.pt
Tel: +351 291 705 156
Fax: +351 291 705 309

Jorge Fernandes
Caminho da Penteada
Polo Científico e Tecnológico da Madeira, Piso 2
9000-390 Funchal - Portugal
jfernandes@expedita.com
Tel: + 351 291 723 410
Fax: +351 291 723 411