Preface

The international workshop of Web Information Systems Modeling (WISM) targets the recent issues involved in modeling Web Information Systems (WIS). The sixth edition of this workshop follows five successful editions organized in Barcelona (2008), Trondheim (2007), Luxembourg (2006), Sydney (2005), and Riga (2004). Taking in consideration the current developments of the Web, i.e., the Semantic Web, the workshop has as a special focus of interest on how WIS modeling can benefit from Semantic Web technologies.

The first paper by Marinette Savonnet et al. presents a top-down methodology based on business patterns for the development of WIS. The methodology has three steps: development of the business architecture, mapping activity patterns onto the business architecture, and defining the abstract Web services that correspond to the previously proposed activities. The activity flows that cross pattern boundaries are identified as points of risk and are implemented by compensation Web services that are able to guarantee the non-functional properties of the system.

The second paper by Sonia Guehis proposes a semi-automatic methodology for determining the semantics of existing WIS. It is assumed that the relational schema of the underlying database is given as well as the right to execute the application code on instances of this schema. Such an approach is useful for example when the source code of the application is difficult to understand or is no longer available. The solution is based on the execution of the program with canonical instances of the database schema and the analysis of the generated dynamic documents. The dynamic documents are scrutinized for distinguishing the static parts from the dynamic ones. From the dynamic parts one can define the program interpretation as a mapping between document structure and the corresponding database schema.

The third paper by Mark Roantree focuses on the optimization of database and streaming data used in a WIS for sport scientists. The proposed approach is based on the DataSpace architecture where the subject data is central, and the sensor data is stored in raw as well as in XML format. XML is useful for representing the integrated data but the conversion process of raw data to XML is rather slow and thus not acceptable for health specialists. For this purpose, the author proposes the use of a hybrid approach that allows the mix of XPath, SQL, and Java for querying the data sources.

The fourth paper by Hadjouni et al. presents an approach for personalizing the information retrieval step in WIS. The proposed solution is based on an implicit and multidimensional user modeling. The user model has four components: textual, i.e., keywords used for search by the user, spatial, i.e., positions of the user, navigational, i.e., entities visited by the user, and user profile, i.e., the user interests computed based on the previous three dimensions. The personalization is based on the spatial and semantic distances between the location of the user

and the objects of interest (deduced from the user position on a map or search keywords).

The fifth paper by Ben Mustapha et al. explores how the processes of semantic search and ontology learning from text can support each other in a WIS. Information retrieval tasks can benefit from domain ontologies by refining the user queries, and ontology learning can enrich an ontology with new constructs extracted from retrieved Web documents. The approach is based on the notion of case bases that store search queries and their corresponding retrieved Web documents. The ontology learning step is based on OntoCoSemWeb, a methodology that uses lexico-syntactic patterns for information extraction from Web documents.

The sixth paper by Gielissen and Marx specifies the requirements and the functional design of PoliDocs, a WIS for distributing parliamentary data in the Netherlands. For requirements definition it uses Sargasso, a collective politically oriented Weblog, TheyWorkForYou.com, the UK WIS for disclosing parliamentary information, and conversations with prospective users/professionals. The functional design describes the WIS back-end, which relies heavily on the Extract-Transfer-Load process to harvest, integrate, transform, and store the data. The functional design also specifies the WIS front-end, which uses state-of-the-art techniques (e.g., faceted search, result aggregation/summarization, etc.) to meet the previously defined requirements.

The seventh paper by Capellari et al., presents a keyword search paradigm for RDF graph models. Differently than other approaches which focus on the best answer trees, the proposed algorithm computes all matching RDF subgraphs for a keyword-based query. First, the computation starts by sorting all the paths that end with a keyword using the path length. Then, these informative paths are clustered based on the template they share and sorted using exhaustivity (based on the number of matches with keywords). The clusters are sorted based on the average exhaustivity of the contained informative paths. The final step is combining the informative paths so that all subgraphs matching the keyword-based query are obtained.

We do hope that the above topics match the readers interest and invite them to a have closer look at the papers gathered in these proceedings. Last, but not least, we would like to thank all the authors, reviewers, and participants for their contributions making thus the organization of this workshop possible.

June 2009

Flavius Frasincar
Geert-Jan Houben
Philippe Thiran
Workshop Co-Chairs
WISM 2009

Organization

WISM 2009 is organized by the Econometric Institute, Erasmus University Rotterdam, the department of Computer Science, Delft University of Technology, and the department of Business Administration, Namur University.

Executive Committee

Workshop Co-Chairs & Organizers:	Flavius Frasincar (Erasmus University Rotterdam, NL)
	Geert-Jan Houben (Delft University of Technology & Eindhoven University of Technology, NL)
	Philippe Thiran (Namur University, BE)
Local Organizer:	Peter Barna (TOPIC, NL)

Program Committee

Djamal Benslimane (University of Lyon 1, FR)
Sven Casteleyn (Vrije Universiteit Brussel, BE)
Virginia Dignum (Utrecht University, NL)
Martin Dzbor (The Open University, UK)
Flavius Frasincar (Erasmus University Rotterdam, NL)
Martin Gaedke (Chemnitz University of Technology, DE)
Irene Garrigos (Universidad de Alicante, ES)
Hyoil Han (Drexel University, US)
Geert-Jan Houben (Delft University of Technology & Eindhoven University of Technology, NL)
Manfred Jeusfeld (Tilburg University, NL)
David Lowe(University of Technology Sydney, AU)
Zakaria Maamar (Zayed University, AE)
Michael Mrissa (Namur University, BE)
Tommaso Di Noia (Technical University of Bari, IT)
Moira Norrie (ETH Zurich, CH)
Jose Palazzo Moreira de Oliveira (UFRGS, BR)
Philippe Thiran (Namur University, BE)
Hajo Reijers (Eindhoven University of Technology, NL)
Stefan Schlobach (Vrije Universiteit Amsterdam, NL)
Riccardo Torlone (Universita di Roma Tre, IT)
Lorna Uden (Staffordshire University, UK)