Chapter 5

Using the Rhizomer Platform for Semantic Decision Support Systems Development

Roberto García
Universitat of Lleida, Spain

ABSTRACT

Decision support systems get more useful as they manage to make decisions more informed. However, the cost of information and of combining and making it available in the appropriate context make this a tricky trade-off. Fortunately, Semantic Web technologies make it possible to easily publish and reuse data. But this is not simple data, it is semantic data, which makes it easier to query, browse and combine it. Apart from semantic data, it is also important a user interface that carries all this potential to the user. Rhizomer is a framework for semantic data publishing and user interaction that facilitates building semantic dashboards. It is possible, for instance, to build a simple dashboard on top of semantic data generated from financial reports and incorporate web services that provide specialised ways to interact with semantic data, like showing geo-located resources in a map or events in a timeline.

INTRODUCTION

One of the challenges of Decision Support Systems (DSS) in the context of increasingly virtual and distributed enterprises is data integration. Decision must be drawn from the combination of heterogeneous data coming from distributed and third party sources, whether it comes from structured relational databases or semi-structured content sources like documents or web pages. It’s from the combination of all these sources that it is possible to attain unprecedented levels of information access, sharing and collaboration. However, most current business intelligence implementations lack a technological foundation that facilitates the integration of data coming from a broader non-relational domain of data, which additionally might be distributed and outside enterprise boundaries and control. Semantic web technologies can help here, thanks to their ability to model and interlink data.

One additional benefit is that once in the Semantic Web, data is formalised in a way that logical inference can operate on it. Inference is guided by the knowledge captured by ontologies, which can model enterprise structure, business...
rules, etc. Moreover, web ontologies can be also distributed so it is easier to reuse both data and ontologies.

The original objective of the semantic web is to enable the description of web content using domain-specific data models (called ontologies) that make it possible for applications to locate and reason over the Web’s resources. Similarly, ontologies can feasibly be adapted to model the structure of a data warehouse’s schema.

It allows decision support systems queries to make semantic inferences over an extended range of external and distributed data that is not necessarily stored in the data warehouse; rather it is referenced through an ontology. The analysis is also enriched by a deeper semantic understanding of data relationships within the data warehouse itself, and is not restricted to the rigid relationships pre-coded in DB schemas.

Information often lacks a meaningful context. The knowledge extracted from unstructured sources must be enriched with links to ontologies that capture their context and facilitate their integration and exploitation together with other more structured sources. There is some early work being done on semantically driven data integration with the semantic web equivalent of the SQL query language for accessing relational databases, SPARQL (Prud’hommeaux & Seaborne, 2008).

But semantic data is not enough if decision-makers cannot easily access and manipulate it. Semantic Dashboards challenge information and interaction design as well as information architecture. These challenges come from integrating data from different sources in an open environment like that made possible by the Semantic Web and that enables entities like virtual world wide enterprises. Creating a dashboard architecture as a set of independent components whose configuration is driven by the available data.

This paper presents a the Rhizomer framework for semantic data publishing and user interaction that can be used in order to develop decision support systems on top of semantic data. The platform is presented in Section 2. Then, in Section 3, a use scenario is presented. In that scenario, Rhizomer is used in order to publish, mix, query and browse semantic data resulting from the SEC’s EDGAR program for financial reporting based on the XBRL standard.

In that scenario it is illustrated how, thanks to semantic technologies, it is possible to present and integrated view on XBRL plus data coming from other sources. On top of this integrated view it is possible to offer some basic user tasks that allow them interacting with data and make decision based on a more integrated and flexible data set. The current set of simple user task are the basis for future developments, in this and other application contexts, based on Rhizomer as a framework for semantic decision support systems.

Introduction to the Semantic Web

Looking at Semantic Web technologies and methodologies more relevant from the point of view of decision support systems, it is important to first note that it is rooted on a data model, a way to represent data, geared towards interoperability. It is based on a directed graph, i.e. a set of nodes connected by edges with a direction, from node A to node B. This graph model constitutes the first building block for semantic interoperability because a graph can be used to represent many other kinds of data structures.

For instance, it is easy to model a tree using a graph—it is just a graph without cycles or a table—each row is represented by a node that is connected to the different row values by edges labelled after each column name. This makes it easier to integrate data coming from XML documents or relational databases into the Semantic Web. Moreover, it is easier to mash-up data from disparate sources into a graph because the result is always a graph.

The Semantic Web graph model is named RDF, Resource Description Framework (Tauberer, 2008). However, this is not enough. We can put it