Decision Support for Agricultural Consultants With Semantic Data Federation

Mohammad Sadnan Al Manir, Department of Computer Science, University of New Brunswick, Saint John, Canada
Bruce Spencer, University of New Brunswick, Fredericton, Canada
Christopher J.O. Baker, IPSNP Computing Inc., Canada & Department of Computer Science University of New Brunswick, Saint John, Canada

ABSTRACT

Informational needs of agricultural consultants are increasingly complex. Advising farmers on the appropriate measures for optimizing cropping yields demands access to custom data archives and analytics tools. In line with the increasing number of archives, the expertise required of consultants goes beyond the capabilities of these non-technical agri-specialists. These end users have diverse ad-hoc query needs and require tools that provide simple access to distributed data silos and easy ways to integrate relevant information. In this article, the authors report on a pilot deployment of Semantic Automated Discovery and Integration (SADI) Web services for the federation and computation of agricultural data. A registry of 9 SADI Web services was deployed to expose data from a variety of different data resources in support of a defined set of query needs. The authors demonstrate that the deployment of these services facilitates the ad-hoc creation and execution of mission critical workflows targeting use cases in agricultural operations management. Using HYDRA, a semantic query engine for SADI Web services with a custom built graphical user interface, agricultural consultants can identify optimal crop varieties, and compute profit margins of each variety using a complex cost model.

KEYWORDS

Agricultural Consultants, Decision Support, Semantic Automated Discovery, Semantic Data Federation

1. INTRODUCTION

Large not for profit, intergovernmental organizations such as CABI (Cameron, Somachandra, Curry, Jenner, & Hobbs, 2016) provide information and scientific expertise to help solve problems in the agriculture and environment sectors. They have significant structured data sources providing detailed information in the form of fact sheets and images about crops and pests on a per country basis. These online resources, such as the Plantwise Online Management System (POMS) (Finegold et al., 2015) are published as an access-controlled system allowing authorized partners with permission from local authorities to view summary reports of the data. Data sets can also be downloaded as spreadsheets and the off line Plantwise Knowledge Bank is available on USB cards to Plantwise partners. These and other decision support resources are also targeted to providing educational information on specific

DOI: 10.4018/IJAEIS.2018070106

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
protocols by way of Smartphone Apps\textsuperscript{1,2,3} where users can look up, in a local language, images or audio that can assist end users with tasks such as composting, bee-keeping, pest control, and soil health.

The World Soil Information Service (WoSIS) provides the Soilgrids REST API\textsuperscript{4} where developers can access their interface, query data and download files for geo-referenced data points. This form of data access enables developers to build applications that retrieve data and combine it with other available data providing value added services such as advanced online maps or mashups showing custom views that are informative to specific user communities. While clearly useful, not all data providers offer APIs and the development of mashups is not without challenges, as developers often have to manually navigate complex data descriptions in order to understand how to gain access. In a typical scenario the time frame for new mashups to be coded and become operational may be on the order of months, which is costly, and does not offer consultants ad-hoc access to multiple data sources according to their clients’ needs and objectives.

To improve the effectiveness of Decision Support Systems (DSS) in delivering relevant, reliable and accurate information to a user on-demand a number of challenges must be addressed. These can be itemized as follows; (1) timely discovery of appropriate new data resources (2) assessment of data ownership and licenses, (3) provision of web access to data stores, (4) rapid assessment of the suitability of data for reuse, (5) integration of data sets with community adopted metadata standards to ensure interoperability, and (6) provision of an effective means of querying of distributed data.

In a recent study (Blomqvist, 2014) it was noted that technologies for decision support have increasingly leveraged Semantic Web methodologies to address data aggregation, abstraction, and analysis, and the scalable management of big and dynamic datasets. The Semantic Web approach\textsuperscript{5} seeks to address challenges of semantic heterogeneity where information resources use different ways to identify the same object or concept, thereby impeding data integration. It advocates for common standards to allow data to be shared and reused across applications, enterprise, and community boundaries. Specifically, the approach proposes the use of common vocabularies to annotate data and services offering easier integration and system interoperability. For data the W3C Standard RDF data model is mandated and domain ontologies scripted in W3C Standard OWL are used as a reference for semantic relationships. All data published in RDF using the same reference ontology can be semantically integrated without further formatting or manipulation.

For provision of web access to data stores Web services are of particular interest because they offer secure access to data from distribute sources. They can also facilitate remote computation or transformation of data. Whereas researchers have explored the use of Web services (Vinoksi, 2003; Zhu et al., 2004) for data integration, services alone do not support desirable features upon which intelligent decision support systems can be built, namely automated discovery, dynamic composition or enactment of workflows. The Semantic Web approach however introduces the addition of semantic annotations to Web services (Wilkinson, Vandervalk, & McCarthy, 2011), and ensures that the operational requirements (I/O) and functions of the services are intelligible, so that they can be discovered, composed into workflows, and invoked automatically by client softwares.

By default, Decision Support Systems must also be able to provide users simple ways to build complex queries to fetch data, however advanced expertise in specific query languages is typically necessary. One solution is for developers to implement, on behalf of end users, a form of semantic querying (Bizer & Seaborne, 2004; Calvanese et al., 2011) over the data where domain knowledge written in the form of axioms or rules is automatically applied to the data resources. Specifically, axioms are used to map concrete data into virtual domain models based on semantic knowledge representation standards RDF\textsuperscript{6} and OWL\textsuperscript{7}, such that databases can be queried as RDF graphs. For end users it means they can formulate their queries using the terminologies of their domain, without any knowledge of how the underlying data is structured. The most basic form of semantic querying is querying semantic data, e.g., RDF, with a language like SPARQL\textsuperscript{8}. Engines that permit semantic querying are increasingly suitable for consultants seeking to gain real time access to diverse target data. Graphical interfaces now allow non-technical users to compose queries as graphs using familiar
Towards Self Energy-Management and Sustainable Citizens’ Engagement in Local Energy Efficiency Agenda

CloudGanga: Cloud Computing Based SDI Model for Ganga River Basin Management in India
www.igi-global.com/article/cloudganga/188646?camid=4v1a