

Chapter 13

Semantic Integration Framework for Resources Identification

Miguel Pragosa

Polytechnic Institute of Leiria, Portugal

Vitor Basto-Fernandes

Polytechnic Institute of Leiria, Portugal

Lúisa Oliveira

National Health Institute Doutor Ricardo Jorge, Portugal

ABSTRACT

The global adoption of information technology systems throughout all activity domains lead to the existence of a huge amount of information systems unable to interoperate, first of all, because of different ways of identifying resources. Different transport and application level protocols for data exchange also raise some interoperability difficulties. Common interoperability scenarios rely on tightly controlled, specific communities of information technology islands. In this chapter we present and discuss the usage of lexical, syntactic and semantic lexical technologies to address interoperability problems at the product identification level, in the context of food consumption analysis.

INTRODUCTION

The case study presented here is focused in the integration of food point-of-sales (POS) systems, where sales data is available, and a public health care authority system, where food composition data is available. For the purpose of food consumption analysis, we need to solve the problem of matching each of the POS sales food description with a public health authority database (DB) food item description.

Nowadays several POS systems (e.g. WinRest, WinTouch, etc.) are widely used in restaurants, canteens, bars, however the adoption of international standards to identify the resources (food items) is very unlikely, and difficult to be applied to all ranges of food products (packed, unpacked, fresh, meals, etc.). Therefore, the main target of the study presented here, is the search for the best way to match the products description from a POS DB with the products description available in a national public food

DOI: 10.4018/978-1-4666-8368-6.ch013

composition DB. The matching process allows the food consumption analysis to include not only population consumption analysis but also nutritional intake, made available through the food composition data present in the national food composition DB.

In practice, this process will provide a unified record of the food consumption data. This will allow us to implement several different applications, for example to create consumption profiles based on the nutritional properties of the items sold in POS systems.

In order to do a proof of the concept, we developed a software application based on the Java programming language and MySQL (MySQL Test Labs, 2006) as the Database Management System (DMS).

This book chapter is structured according to the following outline. The semantic integration background and related work are presented in section two. Section three shows the components and structure of the proposed solution. Section four is focused on the datasets used for the integration case study. The attribute extraction from the food items description by the means of lexical, syntactical and semantic analysis is presented in section five. Section six presents the semantic processing and the scoring/relevance calculation algorithm. Finally some conclusions, limitations, contributions and future work are described.

BACKGROUND

The possibility of developing a computer application that allows automated alignment between food consumption/sales databases and food composition databases that contain their nutritional composition or other properties (e.g. contaminants, allergens) is seem of extreme value by public health care authorities.

Many efforts have been done in developing food classification and description systems, as well as the corresponding creation of controlled vocabularies, in order to allow food consumption analysis and related applications.

Although this represents a recurring issue/problem in this scientific area, no system or software application is known for enabling the automatic mapping of food descriptions between sales and food composition databases.

All known applications addressing these issues require the usage of pre-defined identifiers, pre-defined taxonomies and specific, controlled vocabularies for products identification, classification and description. One example of this type of approach is the FoodEx2 (European Food Safety Authority, 2011) specification done by the European Food Safety Authority. FoodEX2 aims to support food consumption, electronic transmission of chemical occurrence data and public health risk assessment in the framework of public health policies for the European Union countries.

In the context of our study, this type of initiatives is seen as feasible only for a small scale number of participants that are able to cope with the technical alignment and cost of adapting their systems.

Other type of approaches aiming for semantic integration of food related information are also worth to mention. Among those approaches, the semantic search engines and the semantic food repository systems are the most relevant in the context of our study.

The search engines could be highly enhanced with the aid of semantic tools. After that evolution we can expect that the search engines could be more intelligent, being able to understand, not only some predefined keywords, but also complex human-readable sentences.

The Knigine, the Hakia Search Engine and the DuckDuckGo are examples of search engines which already take advantage of Web semantic technologies. As in our work, this allows them to provide a broader and a more intelligent set of results.

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:
www.igi-global.com/chapter/semantic-integration-framework-for-resources-identification/133097?camid=4v1

Related Content

Designing Secure Data Warehouses

Rodolfo Villarroel, Eduardo Fernandez-Medina, Juan Trujillo and Mario Piattini (2006). *Enterprise Information Systems Assurance and System Security: Managerial and Technical Issues* (pp. 295-310).

www.igi-global.com/chapter/designing-secure-data-warehouses/18395?camid=4v1a

Evaluating the Influence of E-SCM on Business Integrity in SMEs

Hany Abdelghaffar and Nada Hassan (2013). *International Journal of Enterprise Information Systems* (pp. 1-11).

www.igi-global.com/article/evaluating-the-influence-of-e-scm-on-business-integrity-in-smes/100379?camid=4v1a

ERP and Beyond

Suresh Subramoniam, Mohamed Tounsi, Shehzad Khalid Ghani and K. V. Krishnankutty (2011). *Enterprise Information Systems: Concepts, Methodologies, Tools and Applications* (pp. 1960-1974).

www.igi-global.com/chapter/erp-beyond/48653?camid=4v1a

Assessing the Impact of System Quality, Information Quality, and Service Quality on Enterprise Resource Planning (ERP) Systems

Godwin Banafo Akrong, Yunfei Shao and Ebenezer Owusu (2021). *International Journal of Enterprise Information Systems* (pp. 69-84).

www.igi-global.com/article/assessing-the-impact-of-system-quality-information-quality-and-service-quality-on-enterprise-resource-planning-erp-systems/289845?camid=4v1a