# Mediation-based Web Services fed Data Warehouse

John Samuel*, Christophe Rey*
Franck Martin **, Lionel Peyron**

*LIMOS, CNRS, Université Blaise Pascal, Aubière, France
samuel@isima.fr, christophe.rey@univ-bpclermont.fr
**Rootsystem, Vichy, France
{franck.martin,lionel.peyron}@rootsystem.fr

**Abstract.** We present our prototype DaWeS, Data Warehouse fed with Web Services. Its ETL process is grounded on a mediation approach usually used in data integration. This enables DaWeS (i) to be fully configurable in a declarative manner only (XML, XSLT, SQL, datalog) and (ii) to make the warehouse schema dynamic so it can be easily updated. (i) and (ii) aim at making DaWeS scalable and adaptable to smoothly face the ever-changing and growing web services offer. We point out the fact that this also enables DaWeS to be used with the vast majority of actual web services defined with basic technologies only (HTTP, REST, XML,and JSON) and not with more advanced standards (WSDL, WADL, hRESTS or SAWSDL).

## 1   Introduction

Everyday we use a large number of web services and applications catering to our various requirements. This trend has now caught up even among the enterprises, especially the small and medium scale enterprises, resulting in a situation where their own business data is spread across multiple data centers spanning even across continents, managed and controlled by numerous web service providers. Enterprises need an integrated view of their data in the form of performance dashboards to track the overall growth of their companies and business units. Traditionally enterprises may use a data warehouse (cf Kimball and Ross (2002)) to perform data analysis and compute business performance measures. But with the advent of multiple heterogeneous, autonomous and ever-evolving web services, the task of data integration has become a challenging one. The purpose of our work is to aid enterprises using the web services as their data source with a data warehouse service to track their performance measures. We built DaWeS (Data warehouse fed with Web Services) (cf Samuel and Rey (2014); Samuel (2014)), a multi-enterprise data warehouse service able to fetch interesting data from various web services and expose them in a manner so that the end users can compute their own interesting business measures in a transparent manner hiding from them the various intricacies of the underlying web service application programming interface (API). Recent web services and semantic web technologies (like WSDL, WADL, hRESTS or SAWSDL) allow to solve many integration issues as for instance the automatic generation of wrappers or the automatic discovery of web services. But these technologies are constraining: developers must learn the associated languages, the integrated systems must be fully and precisely described and the