IDEA GROUP PUBLISHING **ITB13069**

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

This paper appears in the publication, *Advanced Topics in Database Research, vol. 5*
edited by Keng Siau © 2006, Idea Group Inc.

Chapter X

Reuse of a Repository of Conceptual Schemas in a Large Scale Project

Carlo Batini, University of Milano Bicocca, Italy
Manuel F. Garasi, Italy
Riccardo Grosso, CSI-Piemonte, Italy

ABSTRACT

This chapter describes a methodology and a tool for the reuse of a repository of conceptual schemas. Large amounts of data are managed by organizations, with heterogeneous representations and meanings. Since data are a fundamental resource for organizations, a comprehensive and integrated view is needed for it. The concept of data repository fulfils these requirements, since it contains the description of all types of data produced, retrieved, and exchanged in an organization. Data descriptions should be organized in a repository to enable all the users of the information system to understand the meaning of data and the relationships among them. The methodology described in the chapter is applied in a project where an existing repository of conceptual schemas, representing information of interest for central public administration, is used in order to produce the corresponding repository of the administrations located in a region. Several heuristics are described and experiments are reported.

INTRODUCTION

The goal of this chapter is to describe a methodology and a tool for the reuse of a repository of conceptual schemas. The methodology is applied in a large scale project related to the Italian Public Administration (PA); the goal of the project is to use the repository of conceptual schemas of the most relevant databases of the Italian central PA, developed several years ago, in order to build the corresponding repository of the local public administrations located in one of the 21 regions of Italy. Due to the limited amount of available resources, the methodology conceives and applies several approximate techniques, which allows for the rapid prototyping of the local repository. This is to be refined by domain expert, which results in a resource consumption one order of magnitude lower than by using a traditional process. We initially provide some details about the context in which the methodology has been investigated and developed.

In all countries, in the past few years, many projects have been set up to effectively use information and communication technologies (ICT) to improve the quality of services for citizens, by gradually improving on the services that are provided by information systems and databases of their administrations. In the following section, we focus in particular on the Italian experience.

In the past, the lack of cooperation between the administrations led to the establishment of heterogeneous and isolated systems. As a result, two main problems have arisen, namely, duplicated and inconsistent information and difficult data access. Moreover, the government efficiency depends on the sharing of information between administrations, due to the fact that many of them are often involved in the same procedures, but they are using different, overlapped, and heterogeneous databases.

Therefore, in the long term, a crucial aspect for the overall project is to design a cooperation architecture that allows both the central and the local administrations to share information in order to provide services to citizens and businesses on the basis of the "one-stop shopping" paradigm. A crucial aspect of such cooperation architecture is the data architecture: data have to be interchanged with an interoperable format; all the administrations have to assign the same meaning to the same data, achieving database integration in the long term. The database integration will provide for the spread of information within the government branches and will result in a more easily accessible working environment, in an increased quality of information management, and in an improved statewide decision-making process.

The long term goal of database integration has to be achieved in the complex organizational scenario of the Public Administration (PA). The structure of the Public Administration in Italy consists of central and local agencies that together offer a suite of services designed to help citizens and businesses to fulfill their obligations towards the PA. Central PAs are of two types: ministries, such as Ministry of the Interiors and Ministry of Revenues; and other central agencies, such as Social Security Agency, Accident Insurance Agency, and Chambers of Commerce. The main types of local administrations correspond to Regions (21), Provinces (about 100), and Municipalities (about 8,000).

To address this problem, the approach to cooperation among administrations followed in Italy is based on the concept of Cooperative Information Systems (CIS), that is, systems capable of interacting by exchanging services with each other. The general cooperative architecture for the Nationwide CIS network of the Italian PA is shown in Figure 1.