INTRODUCTION

Data that is essential for a company’s successful businesses often resides in a variety of data sources. The reasons for this are manifold, e.g. load distribution or independent development of business processes. But data distribution can lead to inconsistent data which is a problem in the development of new businesses. Thus the consolidation of the spread data as well as giving applications a shared picture of all existing data is an important challenge. The integration of such distributed data is the task of Software AG’s “crossvision Information Integrator” one of the components in the crossvision SOA suite (crossvision).

Information Integrator is based on ontologies. On one hand, data source ontologies can be generated from metadata of underlying data sources. Currently, SQL databases, Software AG’s Adabas databases, and web services are supported types of data sources. On the other hand, more business oriented ontologies can be developed. These business ontologies make use of other business ontologies or can directly use data source ontologies. FLogic rules describe the information how objects in different ontologies are related to each other.

Using ontologies Information Integrator solves three major problems. First of all it provides all means to integrate different information systems. This means that comfortable tools are available to bring data from different systems together. This is partially already solved by systems like virtual or
federated databases (Batini et al. 1986). Information Integrator is more powerful compared to most of these systems as it not only supports databases but additional sources like web services, applications etc. The second problem which is solved is that Information Integrator allows reinterpretation of the contents of the information sources in business terms and thus makes these contents understandable by ordinary end users and not only by database administrators. Finally this semantic description of the business domain and the powerful mapping means from the data sources to the business ontology solves the semantic integration problem which is seen as the major problem in information integration. It maps the different semantics within the information sources to the shared conceptualization in the business ontology. As an intended side effect, once such an integrated semantic description is in place, it might be used by other semantic applications which do understand these semantic descriptions.

Within the next section will explain the conceptual ideas behind the Information Integrator and explicate the already sketched layered architecture for applications. Thereafter, we shortly describe the Information Integrator product itself. First of all we will present the product architecture. Then, we illustrate the tool set that shall allow even inexperienced and occasional users to create and modify semantic models. Within Software AG Information Integrator was used for a first project Customer Information Gateway (CIG) whose mission was to integrate data that on one side resides in a support information system and on the other side is stored in a customer information system. Experiences we gained from that first project are presented in the last major section before we conclude with a summary section.

2. CONCEPTUAL LAYERING

Conceptually Information Integrator arranges information and the access to information on four different layers (Figure 9-1):

- The bottom layer represents different data sources which contain or deliver the raw information which is semantically reinterpreted on an upper layer viz. ontologies. Currently relational databases, Adabas databases and web services are supported.
- The second layer assigns a so called “data-source ontology” to each of the data sources. These “data-source ontologies” reflect only database or WSDL schemas of the data sources in terms of ontologies and can be