**IDEA GROUP PUBLISHING** ITB11971

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 book, *Database Modeling for Industrial Data Management:Emerging Technologies and Applications* edited by Zongmin Ma © 2006, Idea Group Inc.

## Chapter VI

# Semantic-Based Dynamic Enterprise Information Integration

Jun Yuan, The Boeing Company, USA

## Abstract

*Decision support systems or mission control systems for network-centric operations typically need input from multiple heterogeneous information sources. While the number and size of information sources increase, information integration and its semantic interoperability are becoming a growing concern and a major challenge to information management. In this chapter, we will share our experience of enabling semantic-based dynamic information integration across multiple heterogeneous information sources. While data is physically stored in existing/legacy data systems across the networks, the information is integrated based upon its semantic meanings. Informally, it sounds like a virtual data warehousing technique without any physical data conversion required in advance. Ontology is used to describe the semantics of global information content, and semantic enhancement is achieved by mapping the local metadata onto the ontology. For better system reliability, a unique mechanism is introduced to perform appropriate adjustments upon detecting environmental changes.*

# Introduction

Information Integration has been a high priority in business for many years. For building applications such as enterprise-wide decision support systems or mission control systems in a network-centric environment, many companies have actually been struggling against integration of legacy/existing data systems for some time. With the recent explosion of Internet and Web-based resources, information integration and its semantic interoperability are becoming an even greater concern.

In this chapter, we propose a framework that enables dynamic information integration across multiple heterogeneous information sources, with special attention to semantic interoperability and adaptations to a dynamic environment. While data is physically stored in existing/legacy data systems across the networks, the information is integrated based upon its semantic equivalence. Ontology is used to explicitly describe the semantics of global information content. Such ontology is independent of any particular information sources, but only based upon domain experts' knowledge. Individual systems may have their own intended models, and there are likely various types of semantic heterogeneity in between those models. Semantic heterogeneity resolution methods are provided by mapping those local models onto the ontology with a unique mapping mechanism associated with the proposed framework. Our approach is leveraging with the state-of-the-art Semantic Web standards. For instance, the domain ontology as well as the mapping knowledge can be exported into RDF or OWL documents.

In our framework, each participating information source assumes full local autonomy. All existing software applications that were built on top of those information resources will remain to be functioning the same way as they did previously. There is no need to re-develop any of these pre-existing applications. The integration itself is thus a purely incremental enhancement.

A single information access point is provided. Users are able to access information as if it were a centralized information repository. More interestingly, it is a semantic information access point, which allows users to query against ontology, for example, concepts and relationships, directly. Users are thus able to retrieve information from multiple heterogeneous data systems without having to understand lower-level details such as physical distribution, semantic heterogeneity, and logical conflicts among information sources. In addition, query answers are also presented by an instantiation of the ontology, and the semantics can be easily captured and understood by both human beings and computer systems.