Chapter V

XML-Based Toolkit for Interoperability of Web Information Systems

Christophe Nicolle and Kokou Yétongnon
University of Bourgogne, France

ABSTRACT

This chapter presents a methodology and tools framework, called X-TIME, to support data integration via web sites. It is a data model translator toolkit based on a metamodel and XML technology that is aimed at facilitating the design of wrappers, semantic reconciliators or mediators. X-TIME is an adaptable semantics-oriented metamodel approach that takes into account important characteristics of interoperable information systems, including extensibility and composability. Extensibility requires a translation scheme that can easily integrate new data models while composability, which is the ability of an application to define dynamically the subset of data sources it needs, requires on-demand translation among a subset of data models. To meet these requirements, X-TIME is based on an extensible metamodel that defines a set of metatypes for representing meta-level semantic descriptors of data models that can be found in web sites. The metatypes are organized in a generalization hierarchy to capture semantic similarities between modeling concepts and correlate constituent data models of interoperable systems.

INTRODUCTION

With the emergence of web-oriented technologies, the amount of information accessible to users or applications is growing rapidly. Information system
Interoperation has emerged as a central design issue when data and services must be shared among heterogeneous systems.

Data diversity and heterogeneity stemming from the large number of data format or models used to represent and store information on the web is a major obstacle to information systems interoperability. These data models range from the structured data models (network, relational, OO) found in traditional databases to flat files and the emerging semi-structured data models found on the web. The main objective of interoperability is to support the amalgamation of a collection of autonomous heterogeneous systems into an integrated system to provide access to information from multiple disparate sources. Database integration-related issues, such as heterogeneous data conflict resolution have been extensively studied in the past. Several approaches have been proposed to bridge semantic gaps between heterogeneous information systems, including distributed systems, federations, language-based multidatabase and mediation. As new data models are developed for web-based information systems, there is a need to extend the interoperability solutions to take into account the new models and their requirements. For instance, XML has emerged recently as an important model to describe and share web-based data. This importance stems from two major factors. First, XML is becoming a de facto data standard supported by many software vendors and applications developers. Second, XML is based on a relatively simple structure that is both user and machine readable and that can be used by many non-professional users who are not expert in database administration. Existing web technologies are not intended to address some of the issues involved in database integration. For instance the web-browsing paradigm is efficient for looking up data in a large environment, but it cannot adequately support data integration from multiple information sources. To use this paradigm to locate and fuse data requires costly applications that are often tailored to specific integration needs. New challenges have arisen from the development of web-based information systems. One of the challenges is the need to develop a web-oriented tool to support information integration to allow access to local as well as remote information sources. In this chapter we propose a toolkit to support the interoperability of web-based information systems.

**Background**

Many approaches have been proposed for achieving information system interoperability. They include database translation, standardization, federation, language-based multidatabase, ontology and mediation:

- The *database translation* interoperability approach is a point-to-point solution that uses direct data mapping to resolve data heterogeneity between pairs of databases (Andersson, 1994; Cluet, Delobel, Siméon & Smaga,
A Survey of Web Service Discovery Systems

[www.igi-global.com/article/survey-web-service-discovery-systems/2627?camid=4v1a](www.igi-global.com/article/survey-web-service-discovery-systems/2627?camid=4v1a)