Chapter V

Data Integration Through Service-Based Mediation for Web-Enabled Information Systems

Yaoling Zhu
Dublin City University, Ireland

Claus Pahl
Dublin City University, Ireland

ABSTRACT

The Web and its underlying platform technologies have often been used to integrate existing software and information systems. Traditional techniques for data representation and transformations between documents are not sufficient to support a flexible and maintainable data integration solution that meets the requirements of modern complex Web-enabled software and information systems. The difficulty arises from the high degree of complexity of data structures, for example in business and technology applications, and from the constant change of data and its representation. In the Web context, where the Web platform is used to integrate different organisations or software systems, additionally the problem of heterogeneity arises. We introduce a specific data integration solution for Web applications such as Web-enabled information systems. Our contribution is an integration technology framework for Web-enabled information systems comprising, firstly, a data integration technique based on the declarative specification of transformation rules and the construction of connectors that handle the integration and, secondly, a mediator architecture based on information services and the constructed connectors to handle the integration process.
INTRODUCTION

The Web and its underlying platform technologies have often been used to integrate existing software and information systems. Information and data integration is a central issue in this context. Basic techniques based on XML for data representation and XSLT for transformations between XML documents are not sufficient to support a flexible and maintainable data integration solution that meets the requirements of modern complex Web-enabled software and information systems. The difficulty arises from the high degree of complexity of data structures, for example, in business and technology applications, and from the constant change of data and its representation. In the Web context, where the Web platform is used to integrate different organisations or software systems, the problem of heterogeneity arises also. This calls for a specific data integration solution for Web applications such as Web-enabled information systems.

The advent of Web services and service-oriented architecture (SOA) has provided a unified way to expose the data and functionality of an information system. Web services are provided as-is at certain location and can be discovered and invoked using Web languages and protocols. SOA is a service-based approach to software application integration. The use of standard technologies reduces heterogeneity and is, therefore, central to facilitating application integration. The Web services platform is considered an ideal infrastructure to solve the problems in the data integration domain such as heterogeneity and interoperability (Orriens et al., 2003; Haller et al., 2005; Zhu et al., 2004). We propose a two-pronged approach to address this aim: firstly, data integration and adaptivity through declarative, rule-based service adaptor definition and construction; and, secondly, a mediator architecture that enables adaptive information service integration based on the adaptive service connectors. Abstraction has been used successfully to address flexibility problems in data processing; database query languages are a good example here.

XML as a markup language for document and data structuring has been the basis of many Web technologies. XML-based transformation languages like XSLT, the XML Stylesheet Transformation Language, XML-based data can be translated between formats. With recent advances in abstract, declarative XML-based data query and transformation languages beyond the procedural XSLT, this technology is ready to be utilised in the Web application context. The combination of declarative abstract specification and automated support of the architecture implementation achieves the necessary flexibility to deal with complexity and the maintainability of constantly changing data and system specifications.

Our objective is to explore and illustrate solutions to compose a set of data integration services. The data integration services deliver a unified data model built on top of individual data models in dynamic, heterogeneous, and open environments. The presentation of this technology framework aims to investigate the practical implications of current research findings in Web information systems technology.

A lightweight mediated architecture for Web services composition shall be at the centre of our solution. Data integration is a central architectural composition aspect. The flexibility of the architecture to enable information integration is essential in order to separate the business process rules from the rest of the application logic. Therefore, the data transformation rules are best expressed at the abstract model level. We apply our solution to the Web Services platform in the context of information technology services management in the application service providers ASP (on demand) business area. We focus on this context to illustrate problems and solutions. Portals, provided by ASPs, are classical examples where data might come from different sources that motivate our research. In order to consume the information, the data models and representation...