

## Chapter LIII

# Mediation and Ontology-Based Framework for Interoperability

**Leonid Stoimenov**  
*University of Nis, Serbia*

### INTRODUCTION

Research in information systems interoperability is motivated by the ever-increasing heterogeneity of the computer world. New generations of applications, such as geographic information systems (GISs), have much more demands in comparison to possibilities of legacy information systems and traditional database technology. The popularity of GIS in governmental and municipality institutions induce increasing amounts of available information (Stoimenov, Đorđević-Kajan, & Stojanovic, 2000). In a local community environment (city services, local offices, local telecom, public utilities, water and power supply services, etc.), different information systems deal with huge amounts of available information, where most data in databases are geo-referenced. GIS applications often have to process geo-data obtained from various geo-information communities. Also, information that exists in different spatial database may be useful for many other GIS applications. Numerous legacy systems should be coupled with GIS systems, which present additional difficulties in developing end-user applications.

But, information communities find it difficult to locate and retrieve data from other sources in a reliable and acceptable form. In such systems, reuse for geo-data is very often a difficult process because of poor documentation, obscure semantics, diversity of data sets, and the heterogeneity of existing systems in terms of data modeling concepts, data encoding techniques, and storage structures (Devogele, Parent, & Spaccapietra, 1998). Also, available information is always distributed and no one wants to share their own information in public without commitment.

The problem of bringing together heterogeneous and distributed information systems is known as interoperability problem. Today, research on interoperability solutions promises a way to move away the monolithic systems that dominate the GIS market (Sondheim et al., 1999). In that kind of environment, there often arises the problem of semantic heterogeneity and the correctness of the interpretation of data sets obtained from different geo-information communities.

Domain experts use the concepts and terminology specific for their respective field of expertise, and different parameters and different languages

to express their model of a concept. Therefore, very often different data sets can use different terms for the same kind of information. On the other hand, different data sets can use the same term for a completely different piece of information. Humans use their common sense, that is, their knowledge about the world, to translate the meaning of a foreign set of concepts and terms into their own terminology. Software systems usually do not have any knowledge about the world and have to explicitly be told how to translate one term into another. These problems can lead to serious conflicts during discovering and interpreting geographic data.

## **BACKGROUND**

### **Heterogeneity of Data Sources**

The realization of interoperable information systems is a weighty process involving two main system characteristics: distributed data sources and their heterogeneity. Information systems heterogeneity may be considered as structural (schematic heterogeneity), semantic (data heterogeneity), and syntactic heterogeneity (database heterogeneity; Bishr, 1998). Syntactic heterogeneity means that various database systems use different query languages (SQL [structured query language], OQL, etc.). Structural heterogeneity means that different information systems store their data in different structures. Semantic heterogeneity considers the content of an information item and its meaning. Semantic conflicts among information systems occur whenever information systems do not use the same interpretation of the information. Stuckenschmidt, Wache, Voge, and Vissar (2000) give an introduction to problems concerning the syntactic, structural, and semantic integration. This article also presents technologies for enabling interoperability.

### **Information Integration**

A number of proven and well-established methods exist that allow heterogeneous data sources (databases) to communicate on a technical level. Some of these methods include federated databases and schema integration (Larson, 1998), object-oriented approaches (Chawathe et al., 1994), data warehousing (Matousek, Mordacik, & Janku, 2001; Voisard & Juergens, 1998), and mediators and ontologies (Boucelma & Colonna, 2005; Fonseca & Egenhofer, 1999; Stoimenov et al., 2000).

The data warehousing approach (Voisard & Juergens) implies accumulation of spatial data in a few well-defined and tightly connected data stores, where information integration is precomputed. While efficient for a relatively small number of core spatial data sets, this approach is not readily extensible to a larger number of data sets with semistructured and ad hoc data. Mediator-based systems, alternatively, are constructed for a large number of relatively autonomous sources of data and services communicating with each other over a standard protocol and enabling on-demand information integration (Wiederhold, 1998). Structural and syntactic heterogeneity may be solved by mediation.

Information mediators are originally developed for integrating information in databases. Wiederhold defines a mediator and he is the first that points out the need for mediation for contemporary database systems. Wiederhold therefore sees a mediator as an intermediate abstraction layer between databases and applications that use them. Mediation is primarily an architectural concept, and so the precise implementation of a module is less important. The mediator architecture is therefore introduced as a three layer system: (a) applications, (b) mediators, and (c) DBMS. The three-level architecture of mediator-based systems is constructed of an application layer and a large number of information sources (heterogeneous data sources with wrappers) communicating with each other over a standard protocol.

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/mediation-ontology-based-framework-  
interoperability/20734](http://www.igi-global.com/chapter/mediation-ontology-based-framework- interoperability/20734)

## Related Content

---

### Human Factors Research on Data Modeling: A Review of Prior Research, An Extended Framework and Future Research Directions

Heikki Topi and V. Ramesh (2002). *Journal of Database Management* (pp. 3-19).

[www.irma-international.org/article/human-factors-research-data-modeling/3276](http://www.irma-international.org/article/human-factors-research-data-modeling/3276)

### Generic Model Management

Zinovy Diskin and Boris Kadish (2005). *Encyclopedia of Database Technologies and Applications* (pp. 258-265).

[www.irma-international.org/chapter/generic-model-management/11156](http://www.irma-international.org/chapter/generic-model-management/11156)

### Scalable QSF-Trees: Retrieving Regional Objects in High-Dimensional Spaces

Ratko Orlandic and Byunggu Yu (2004). *Journal of Database Management* (pp. 45-59).

[www.irma-international.org/article/scalable-qsf-trees/3315](http://www.irma-international.org/article/scalable-qsf-trees/3315)

### Dataveillance and Panoptic Marketspaces

Nikhilesh Dholakia, Detlev Zwick and Anil Pandya (2005). *Encyclopedia of Database Technologies and Applications* (pp. 170-175).

[www.irma-international.org/chapter/dataveillance-panoptic-marketspaces/11141](http://www.irma-international.org/chapter/dataveillance-panoptic-marketspaces/11141)

### Data Dissemination

Ludger Fiege (2005). *Encyclopedia of Database Technologies and Applications* (pp. 105-109).

[www.irma-international.org/chapter/data-dissemination/11130](http://www.irma-international.org/chapter/data-dissemination/11130)