Chapter III
A Mediation Architecture for Global Comparison Services

Hongwei Zhu
Old Dominion University, USA

Stuart E. Madnick
Massachusetts Institute of Technology, USA

ABSTRACT

Global comparison services facilitate easy comparison of product offerings around the world. To offer such services, one has to address the semantic heterogeneity problems that often arise when data is collected from sources around the world. In this chapter, the authors use examples to illustrate three types of semantic heterogeneity problems that a global comparison service may encounter. Then they present a mediation architecture as a solution to addressing these problems. The feasibility of using the architecture to enable global comparison is demonstrated with a prototype application. An evaluation of the solution shows that it is scalable due to its capability of automatically generating necessary conversions from a small set of predefined ones.

With the global reach of the Internet, every business can be a global business. In this environment, it is often desirable to know how things are selling in various parts of the world – vendors need to price their products strategically and consumers want to find the best deals around the world. This need can be facilitated by a global price comparison service. Although there have been numerous regional comparison services which compare prices of vendors in a particular country, there have not been many global comparison services. This is partially due to the problems of semantic heterogeneity: what seems to be the same data (e.g., price data) has different meanings in different sources (e.g., prices in USD without taxes on one web site vs. prices in Swedish krona with 25%
A Mediation Architecture for Global Comparison Services

taxes included on another). A global comparison service has to make data from worldwide sources meaningful to diverse end users. For example, for users in South Korea, all price data should be in thousands of South Korean won with taxes and shipping charges excluded, regardless of how prices are reported in the original sources.

Comparison service providers are also known as comparison aggregators for their capability of transparently aggregating information from multiple web sources (Madnick and Siegel, 2000). They are also called price comparison agents, shopping agents, and shopping robots (or shopbots for short) elsewhere; we will use these terms interchangeably in the rest of the chapter.

In this chapter, we use examples to illustrate the semantic heterogeneity problems. Then we present a mediation architecture for solving these problems. We show the feasibility of the architecture using a prototype implementation. Although we use price comparison as an example, the architecture is applicable for comparison of other characteristics, such as product dimensions and weights.

**MOTIVATING EXAMPLE**

Imagine for the moment you are from Sweden and interested in buying a pocket sized digital camcorder. After some research on the Web you decide to buy a SONY MICROMV DCR-IP5, which records video in MPEG format for easy editing on computers and weighs only 0.336 kilograms (i.e., 12 ounces). You use your favorite comparison service Kelkoo (at www.kelkoo.se) to find the best deals and it returns information as shown in Figure 1.

Among the vendors found, 18,082 Swedish krona (SEK) is the lowest total price (see the Totalpris column in Figure 1). Is this the best deal, or is there a substantially better deal, on a global basis? If you plan to use the camcorder while on an upcoming trip to several countries, is it better to buy it in Sweden before the trip or buy it at the first stop (say, the U.S.) on your arrival? Without a global comparison service, this can only be done manually by visiting numerous regional comparison aggregators available in other countries. Our manual exercise found one vendor in the U.S. sells...