In this special issue, we take a look at a number of new trends in the management of heterogeneous data. I have titled this issue "Integration Management" because the development of robust, comprehensive management and maintenance solutions is a central theme underlying these trends.

We begin with two papers from integration practitioners that present several important research challenges. The first by Seligman *et al* presents a survey of how effort is distributed among different integration tasks. This survey highlights several time-consuming tasks that are ripe for enhanced automation. A few of these tasks that are considered within this issue include: obtaining source knowledge (including the understanding and representation of structure and data semantics), the creation of schema correspondences (or schema matches), the creation of data combination rules, the creation of logical mappings, and the creation of data translation programs. Narayanan and Subramanian present a different look at integration, but again with a practitioner's eye. They identify several open problems including the management of partial or incomplete data. In addition, some important issues are raised with respect to materializing or replicating integrated data. Finally, Narayanan and Subramanian consider the need to manage and integrate not only data, but also the queries (and query capabilities) and the context or environment of the data. In effect, the authors are proposing more robust and extensive knowledge management facilities that manage the full integration problem, not just the data.

After the practitioners' call-to-arms, we move on to three specific systems designed to manage heterogeneous data. Buneman *et al* present a system for integrating vertically partitioned databases. The main goal of the system is to provide fast, flexible data transformation operators. The emphasis is on data integration in environments where the integrated data may be materialized and this materialized data must be kept updated. The system is used in the very dynamic environment of Wall Street trading. Our next two systems are the result of strong industry-university collaborations. The first is a collaboration between the University of Washington and Microsoft Research on *generic model management*. In the paper presented here, Pottinger and Bernstein consider a system for creating integrated schemas that is *generic*. The emphasis is on abstracting out the data-model-independent portions of the schema integration task. The algorithm makes use of schema correspondences to produce an integrated schema. While these correspondences do not contain data semantics, this work suggests that the merge result may be used to constrain the set of possible mappings. The next system is a collaboration between the University of Toronto and IBM Almaden Research Center on *schema management*. In the paper presented in this issue, Andritsos *et al* focus on the management of schemas and mappings. They describe solutions for creating and managing: schemas; correspondences (matches) between schemas; mappings between schemas (which provide a semantics for transforming data); and translation programs or queries. The authors consider the use of mappings for both data integration and data exchange.

We conclude with two papers related to the management of incompleteness and inconsistency that may arise from the integration or transformation of data. These two papers lay a foundation for identifying, representing, and reasoning about incompleteness and inconsistency within the context of data transformation. Calì *et al* address the problem of data integration where a unified (virtual) view is provided over a set of heterogeneous sources. They argue that integrity constraints on the view are required to model the semantics of the integrated data and to correctly query the integrated data. In the presence of such constraints, they show that query answering in data integration raises the issue of dealing with both incompleteness and inconsistency. Grahne also considers the relationship of querying integrated data and querying incomplete information. He surveys recent result on query rewriting for incomplete information and shows how they may be applied in data integration.

It is notable that the importance of *data semantics* in managing heterogeneous data is highlighted throughout the issue.