Abstract of the PhD Thesis

Foundations of Information Integration
Analyzing, Managing and Reasoning about Schema Mappings

Emanuel Sallinger

Reviewers: Prof. Dr. Reinhard Pichler
Prof. Dr. Phokion G. Kolaitis

Today, we are witnessing a rapid growth of information available in databases and on the web or, as some like to call it, we are the beginning of the era of “big data”. The sources of this information are manifold, among them companies and private organizations, universities and governments. Access to data integrated from many different sources is becoming increasingly important for decision makers at all levels – individual persons, companies, and governments alike. It is thus not surprising that information integration has been on top of the agenda of database research over the past decade, and continues to be a key issue.

The database community has recently made remarkable progress in the area of information integration. Strong formal foundations have been laid for many aspects of information integration through the concept of schema mappings, which are high-level descriptions of the relationship between databases. For building effective systems based on schema mappings, analyzing and optimizing schema mappings is an essential task. Beyond that, with the number and complexity of schema mappings growing, there is a strong need for managing schema mappings. Taken together, the areas of analyzing, optimizing and managing schema mappings form the core of the diverse field of reasoning about schema mappings.

In this thesis, we embark on a systematical study of analyzing, managing and reasoning about schema mappings. We advance the state of the art, completing the picture for many areas of classical schema mappings. We also introduce a number of new concepts that go beyond classical schema mappings, opening up new lines of research. In doing this, we show that schema mappings offer a solid foundation for information integration and promise to be one of the keys for taking advantage of the rapidly growing amount of information in the future.