Workshop on Semantic Data Management

Reto Krummenacher, Atanas Kiryakov, Karl Aberer, and Rajaraman Kanagasabai

Semantic Data Management Initiative
www.semdata.org

The Semantic Web represents the next generation Web of Data, where data is published and interlinked in order to facilitate the exploitation of its structure and meaning. Semantic applications require database management systems for the handling of structured data, taking into consideration the models used to represent semantics. To foster the realization of the Semantic Web, the World Wide Web Consortium (W3C) developed a set of metadata models, ontology models, and query languages. Most of the Semantic Web repositories are database engines, which store data represented in RDF, support SPARQL queries, and can interpret ontologies represented in RDFS and OWL. We are at the point where the adoption of semantic technologies is growing. However, these technologies often appear to be immature, and tend to be too expensive or risky to deploy in real business. Solid data management concepts, architectures, and tools are important for the semantic ecosystem, and creating them requires a strong community, with a critical mass of involvement.

Semantic data management refers to a range of techniques for the manipulation and usage of data based on its meaning. It enables sustainable solutions for a range of IT settings, where the usage of mainstream technology is inefficient or entirely unfeasible: enterprise data integration, life science research, data sharing in SaaS architectures, querying linked data on the Web. In a nutshell, semantic data management fosters the economy of knowledge, facilitating more comprehensive usage of larger scale and more complex datasets at lower cost.

The SemData workshop provided a platform for the investigation of various aspects related to semantic databases and data management, such as semantic repositories, their virtualization and distribution, and interoperability with related database solutions. Many of the semantic data management challenges cumulate in the need for scalable database solutions for semantic data, a building block that runs largely behind comparable non-semantic solutions. In order to make semantic technologies take on the market, it is indispensable that technological progress allows semantic repositories to reach near performance parity with some of the best RDBMS solutions without having to omit the advantages of a higher query expressivity compared to basic key-value stores, or the higher schema flexibility compared to the relational model. It is time that one must no longer pay a heavy price in terms of longer run times or more expensive equipment for profiting from the flexibility of the generic physical model underlying the semantic graph-based structures of RDF. We recognize that there will always be a burden with more flexibility, hence, the goal is to minimize the drawbacks and maximize the advantages of the semantic repositories.