QDB 2010
8th International Workshop on Quality in Databases

The problem of poor data quality in databases, data warehousing, and information systems largely and indistinctly affects every application domain. Many data processing tasks (such as information integration, data sharing, information retrieval, and knowledge discovery from databases) require various forms of data preparation and consolidation with complex data processing techniques, because the data input to the algorithms is assumed to conform to nice data distributions, containing no missing, inconsistent or incorrect values. This leaves a large gap between the available "dirty" data and the available machinery to process the data for the application purposes.

In its long series of events, the Quality in Databases (QDB) workshop (formerly known as IQIS and CleanDB) has been a qualified forum for researchers and practitioners to discuss and get to know the most innovative and advanced experiences for detecting data anomalies and assessing, monitoring, improving, and maintaining the quality of information. In this year's edition, 8 papers have been selected for presentation. According to the topics they address, the papers can be grouped in three thematic areas. The first area includes contributions focused on data quality assessment. Specifically, this year's QDB has collected papers proposing new quality metrics for assessing the quality of social network tracing, the symbolic evaluation of metrics as functions of data set parameters with a view to their interactive usage and new methods to aggregate quality metrics for data sets. A second group of papers concerns system issues, covering a vast range of topics, such as the parallelization of entity matching via the appropriate partition of data, the construction of a toolkit for benchmarking the entity matching process and the situation-aware data quality management. Finally, a third category of papers covers data quality at the schema level, with two papers on conditional functional dependencies and conditional matching dependencies. Moreover, Xin (Luna) Dong will offer an invited talk on dependence and copying detection.

We would like to thank all the authors for contributing to the workshop with their high quality papers. We would also like to thank all the members of the Program Committee for their careful and dedicated work. Working in cooperation with them has been both a particular honor and a pleasure. Finally, we would like to express our gratitude to the members of the Organizing Committee of VLDB 2010 for their support in organizing this workshop.

The QDB 2010 Organization Committee:

Andrea Maurino, University of Milano-Bicocca, Italy

Cinzia Cappiello, Politecnico di Milano, Italy

Panos Vassiliadis, University of Ioannina, Greece

Kai-Uwe Sattler, Ilmenau Univ. of Technology, Germany