

5th International Workshop on Quality in Databases (QDB) 2007

at the VLDB 2007 conference in Vienna, Austria

Data and information quality has become an increasingly important and interesting topic for the database community. Solutions to measure and improve the quality of data stored in databases are relevant for many areas, including data warehouses, scientific databases, and customer relationship management. The QDB 2007 workshop focuses on practical methods for data quality assessment and data quality improvement. QDB 2007 continues and combines the successful three IQIS workshops held at SIGMOD 2004-2006 and the CleanDB workshop held at VLDB 2006.

We are happy to host invited talks by Renee Miller (University of Toronto) and AnHai Doan (University of Wisconsin), and a great program of six research papers with authors from seven different countries.

Our thanks are due to Microsoft both for hosting the conference management tool and in particular for generous sponsorship of the participation fee.

The editors

Venkatesh Ganti (Microsoft Research)

Felix Naumann (Hasso Plattner Institute, Potsdam)

Program, Sunday, September 23

Introduction

9:15 - 9:30 Felix Naumann and Venkatesh Ganti

Invited talk: Renee Miller

9:30 - 10:30

Management of Inconsistent and Uncertain Data

by Renee J. Miller (University of Toronto)

Coffee break

10:30 - 11:00

Research Session 1

11:00 - 13:00

Accuracy of Approximate String Joins Using Grams

Oktie Hassanzadeh, Mohammad Sadoghi, and Renee J. Miller

QoS: Quality Driven Data Abstraction Generation For Large Databases

Charudatta V. Wad, Elke A. Rundensteiner and Matthew O. Ward

Quality-Driven Mediation for Geographic Data

Yassine Lassoued, Mehdi Essid, Omar Boucelma, and Mohamed Quafafou

On the performance of one-to-many data transformations

Paulo Carreira, Helena Galhardas, Joao Pereira, Fernando Martins, and Mario J. Silva

Lunch

13:00 - 14:00

Invited talk: AnHai Doan

14:00 - 15:00

Data Quality Challenges in Community Systems

by AnHai Doan (University of Wisconsin)

Research Session 2

15:00 - 16:00

Towards a Benchmark for ETL Workflows

Panos Vassiliadis, Anastasios Karagiannis, Vasiliki Tziouvara, and Alkis Simitsis

Information Quality Measurement in Data Integration Schemas

Maria da Conceição Moraes Batista, Ana Carolina Salgado

Coffee break & end of workshop

16:00 - 16:30

