

QDB's Important Dates

- June 4, 2011: Submission deadline
- July 15, 2011: Deadline reviews
- July 24, 2011: Notification of acceptance
- August 2, 2011: Deadline camera ready copy

Focus

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. These tasks usually assume that the data input conforms 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 effectively process the data for the application purposes. The Quality in Databases (QDB) workshop is focused on discussing various issues arising in detecting data anomalies and assessing, monitoring, improving, and maintaining the quality of information. Interesting topics include but are not limited to:

- Duplicate detection, entity resolution, and entity reconciliation
- Conflict resolution and data fusion
- Data quality models and algebra
- Quality of Linked Data
- Cleaning extremely large datasets
- Data Quality on the Web
- Privacy-preserving data quality
- Data quality benchmarks
- Data Quality on novel data management architectures (cloud, streaming data, ...)
- General techniques applicable across many domains with measurable improvements on specific domains.
- State of the art, recent progress, and prospects for the future in data quality
- Data scrubbing, data standardization, data cleaning techniques
- Quality-aware query languages and query processing techniques
- Quality-aware analytics solutions
- Data quality in data integration settings
- Role of metadata in quality measurement
- Data quality mining
- Quality of scientific, geographical, and multimedia databases
- Data quality assessment, measures and improvement methodologies

QDB's workshop organizers

Mourad Ouzzani, Purdue University **Paolo Papotti**, Università "Roma Tre" **Erhard Rahm**, University of Leipzig

Format and Paper Submission

For the QDB 2011 edition of the workshop, in addition to regular research papers, we would like to promote the presentations of research-in-progress papers and posters. For this reason, we propose three categories for submission:

- research papers (should not exceed 10 pages)
- overview/survey papers (should not exceed 10 pages)
- short papers (up to 6 pages)
- posters (up to 3 pages)

In line with our main intent for fruitful discussions and exchanges, in addition to previously unpublished material that can be submitted in any of the above formats, the program committee welcomes short paper submissions and posters that include ideas and material that may have appeared elsewhere, insofar as they are judged to be a sound and valuable contribution to the discussion. Papers to be included in the QDB workshop's research sessions will be reviewed by at least three people from the program committee. Depending on the number and quality of submissions, the best papers of the workshop may be recommended for a special issue in the Information Systems Journal.

The papers have to be formatted according the VDLB formatting guidelines (see <http://vldb.org/pvldb/format.html> but abide by the page limits set in this call and there is no 8+4 rule). All papers must be submitted in PDF. Please ensure that any special fonts used are included in the submitted documents. It is essential that the submitted papers print without difficulty on a variety of printers using Adobe Acrobat Reader. Accepted papers will be included in the VLDB proceedings as long as they follow the formatting guidelines. Paper submissions should be made on-line through QDB 2010's CMT paper submission site: <https://cmt.research.microsoft.com/QDB2011/>

In case of paper acceptance, the authors agree that at least one of them will register for the workshop and present the paper.

Invited Talks

We plan to include two invited talks in the program. Ahmed K. Elmagarmid (Qatar Computing Research Institute) and Wenfei Fan (University of Edinburgh), two prominent researchers in the area, have both accepted to present their research perspectives on data quality.

Program Committee

- Daniele Barone, University of Toronto, Canada
- Carlo Batini, University of Milano, Bicocca, Italy
- Ismael Caballero, University of Castilla, La Mancha, Spain
- Tiziana Catarci, Università di Roma, La Sapienza, Italy
- Peter Christen, Australian National University, Australia
- Eduard Dragut, Purdue University, USA
- Suzanne Embury, University of Manchester, UK
- Alvaro A. Fernandes, University of Manchester, UK
- Johann-Christoph Freytag, Humboldt, Universität zu Berlin, Germany
- Helena Galhardas, Technical University of Lisbon, Portugal
- Michael Gertz, University of Heidelberg, Germany
- Markus Helfert, Dublin City University, Ireland
- Melanie Herschel, Universität Tübingen, Germany
- Dmitri V. Kalashnikov, UC Irvine, USA
- Raghav Kaushik, Microsoft Research, USA
- Andy Koronios, University of South Australia, Australia
- Pei Li, University of Milano, Bicocca, Italy
- Paolo Missier, University of Manchester, UK
- Felix Naumann, Hasso, Plattner, Institut, Germany
- Mario Gerardo Piattini, University of Castilla, La Mancha, Spain
- Monica Scannapieco, ISTAT, Italy
- Divesh Srivastava, AT&T, Research, USA
- Panos Vassiliadis, University of Ioannina, Greece
- Vassilios Verykios, University of Thessaly, Volos, Greece
- William Winkler, U.S. Bureau of the Census, USA