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, 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.

Building on the established tradition of nine previous international workshops on the topic of Data and Information Quality, namely IQIS 2004-2006, CleanDB 2006 and QDB 2007-2011, the Quality in Databases (QDB) workshop is a qualified forum for presenting and discussing novel ideas and solutions related to the problems of assessing, monitoring, improving, and maintaining the quality of data.

Website: www.cyber.purdue.edu/qdb2012

Topics of Interest
The Quality in Databases (QDB) workshop is a qualified forum for presenting and discussing novel ideas and solutions related to the problems of exploring, assessing, monitoring, improving, and maintaining the quality of data. Specific topics include, but are not limited to, the following.

- 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

**Important Dates**
- Deadline for submissions: May 30, 2012 10:00pm PDT
- Notification of acceptance: July 9, 2012
- Camera-ready versions: August 7, 2012

**Submission Instructions**
For the QDB 2012 edition of the workshop, we would like to promote the presentations of research-in-progress projects, including 1) new research ideas (even not yet fully explored), 2) experimental comparison of different existing approaches, and 3) industry experience of applying existing techniques on real-world data. All submissions are limited to six pages. The accepted papers will be presented orally and meanwhile as a poster in the program.

The papers have to be formatted according the VDLB formatting guidelines [www.vldb2012.org/call-for-contributions/preparation-and-formatting-guidelines/](http://www.vldb2012.org/call-for-contributions/preparation-and-formatting-guidelines/)

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.

Submissions should be uploaded through Microsoft's CMT submission site at:


All submissions will be handled electronically. Each submission will be reviewed by at least three members of 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.
Conference Organizers

**Workshop Co-Chairs:**
Xin Luna Dong, AT&T Labs-Research, USA
Eduard C. Dragut, Purdue University, USA

**Program Committee Members:**
Andrea Maurino, Università degli Studi di Milano-Bicocca, Italy
Beng Chin Ooi, National University of Singapore, Singapore
Carlo Batini, Università degli Studi di Milano-Bicocca, Italy
Daniele Barone, University of Toronto, Canada
Divesh Srivastava, AT&T Labs-Research, USA
Erhard Rahm, University of Leipzig, Germany
Fei Chiang, University of Toronto, Canada
Felix Naumann, Hasso-Plattner-Institut für Softwaresystemtechnik, Germany
Flip Korn, AT&T Labs-Research, USA
Hazem Elmeleegy, AT&T Labs-Research, USA
Laure Berti-Equille, IRD (Institute of Research for Development), France
Lise Getoor, University of Maryland, USA
Lukasz Golab, University of Waterloo, Canada
Mourad Ouzzani, Qatar Computing Research Institute, Qatar
Peter Christen, Australian National University, Australia
Tiziana Catarci, Università di Roma, Italy
Xiaofang Zhou, The University of Queensland, Australia
Weiyi Meng, SUNY at Binghamton, USA