

17th International Conference on Information Quality

(ICIQ 2012)

**Paris, France
16-17 November 2012**

Editors:

**Laure Berti-Equille
Monica Scannapieco**

Isabelle Comyn-Wattiau

ISBN: 978-1-62748-396-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2012) by the MIT Information Quality Program
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact the MIT Information Quality Program
at the address below.

MIT Information Quality Program
Attn: Dr. Richard Wang
225 Crafts Road
Chestnut Hill MA 02467

Phone: (617) 304-3120

rwang@mit.edu

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

MAIN TRACK RESEARCH PAPERS

Session 1 – IQ and Organizations	<i>page</i>
— <i>Riikka Vilminko-Heikkinen, Samuli Pekkola</i> , Organizational Issues in Establishing Master Data Management Function.....	1
— <i>Elizabeth Pierce, C. Lwanga Yonke, Piyush Malik, Chitra Kagathur Nargaraj</i> , The State of Information and Data Quality Efforts in Today’s Organizations.....	14
— <i>Angélica Caro, Alfonso Rodriguez, Cinzia Cappiello, Ismael Caballero</i> , Designing Business Processes Able to Satisfy Data Quality Requirements.....	31
Session 2 – IQ and Knowledge	
— <i>Meredith Nahm</i> , Knowledge Acquisition from and Semantic Variability in Schizophrenia Clinical Trial Data.....	46
— <i>Umair ul Hassan, Sean O’Riain, Edward Curry</i> , Towards Expertise Modelling for Routing Data Cleaning Tasks within a Community of Knowledge Workers.....	58
— <i>Sarah Ayad, Samira Si-said Cherfi</i> , Domain Knowledge Based Quality for Business Process Models.....	70
Session 3 – Information Accuracy	
— <i>Sami Laine</i> , APC-SIMULATOR: Demonstrating the Effects of Technical and Semantic Errors in the Accuracy of Hospital Reporting.....	85
— <i>Alisa Wechsler, Adir Even</i> , Assessing Accuracy Degradation over Time with a Markov-Chain Model.....	99
— <i>Meredith Nahm, Joseph Bonner, Philip L. Reed, Kit Howard</i> , Determinants of Accuracy in the Context of Clinical Study Data.....	111
Parallel Sessions: Session 4 – IQ Improvement	
— <i>Philip Woodall, Alexander Borek, Ajith Kumar Parlikad</i> , Customized Data Quality Improvement.....	126
— <i>Hendrik Decker</i> , Checking and Repairing the Quality of Information in Databases by Inconsistency Metrics.....	139
— <i>William Rybolt, Leo Pipino</i> , Introducing Data and Information Quality Principles in Today’s College Curriculum via an Introductory Probability and Statistics Course.....	151
— <i>Mario Mezzanzanica, Mirko Cesarini, Fabio Mercorio, Roberto Boselli</i> , Towards the Use of Model Checking for Performing Data Consistency Evaluation and Cleansing.....	163

Parallel Sessions: Session 5 – IQ Dimensions

— <i>Phyllis Illari, Luciano Floridi</i> , IQ : Purpose and Dimensions	178
— <i>Philip Woodall, Andy Koronios, Jing Gao, Ajith Kumar Parlikad, Elaine George</i> , An Investigation into Data Quality Root Cause Analysis.....	193
— <i>Roman Lukyanenko, Jeffrey Parsons</i> , Impact of Conceptual Modeling Approaches on Information Quality: Theory and Empirical Evidence.....	206
— <i>Carlo Batini, Matteo Palmonari, Giuanluigi Viscusi</i> , The Many Faces of Information and their Impact on Information Quality.....	212

Parallel Sessions: Session 6 – Measurement of IQ

— <i>Michael Feldman, Adir Even, Yisrael Parmet</i> , The Effect of Missing Data on Classification Quality.....	229
— <i>Abrar Haider, Snag Hyun Lee</i> , Information Quality Assessment in Korean Asset Managing Organization – Using a Product Perspective.....	B#5
— <i>Yonelbys Iznaga, César Guerra, Ismael Caballero</i> , CALYDAT: A Methodology for Evaluating Data Quality Dimensions based on Data Profiling Techniques.....	260
— <i>Fabian Panse, Wolfram Wingerath, Steffen Friedrich, Norbert Ritter</i> , Key-based Blocking of Duplicates in Entity-Independent Probabilistic Data.....	278

Parallel Sessions: Session 7 – IQ and Social Media

— <i>Yasuhiro Tanaka, Akihisa Kodate</i> , Research on the Role of Social Media and Motivation to Use in the Local Community – Index of Information Quality and Private Space Function.....	297
— <i>Ganesan Shankaranarayanan, Bala Iyer, Donna Stoddard</i> , Quality of Social Media Data and Implications of Social Media for Data Quality.....	311
— <i>Olivier Blattmann, Patrick Kaltenrieder, Patrizia Haupt, Thomas Myrach</i> , Measuring Information Quality on the Internet – A User Perspective.....	326