

ACM Journal of Data and Information Quality

Article 14
(2 pages)

L. Raschid

Editorial

Article 15
(17 pages)

F. Wijnhoven
C. Amrit
P. Dietz

Value-Based File Retention:
File Attributes as File Value and
Information Waste Indicators

Article 16
(38 pages)

W. Fan
S. Ma
N. Tang
W. Yu

Interaction between Record Matching
and Data Repairing

Article 17
(40 pages)

N. Martin
A. Poulovassilis
J. Wang

A Methodology and Architecture Embedding Quality
Assessment in Data Integration



Association for
Computing Machinery

Advancing Computing as a Science & Profession

ACM Journal of Data and Information Quality

ACM
2 Penn Plaza, Suite 701
New York, NY 10121-0701
Tel.: (212) 869-7440
Fax: (212) 869-0481
<http://www.acm.org>

Home Page: <http://jdiq.acm.org>
JDIQ email: acm-jdiq@mit.edu

Editor-in-Chief
Louiqa Raschid

University of Maryland, College Park, USA / <http://www.umiacs.umd.edu/people/louiqa>

Senior Associate Editor
Felix Naumann

Hasso-Plattner-Institute, Germany /
http://www.hpi.uni-potsdam.de/naumann/people/felix_naumann.html

Associate Editors

Irit Askira Gelman
Laure Berti-Équille

DQIQ Research and Solutions, USA / email: askirai@dqiqsolutions.com
Institut de Recherche pour le Développement (IRD), France /
<http://pageperso.lif.univ-mrs.fr/~laure.berth/>

Cinzia Cappiello
InduShobha Chengalur-Smith

Politecnico di Milano, Italy / <http://home.dei.polimi.it/cappiell/>
SUNY, Albany, USA /
<http://www.albany.edu/business/about/faculty/chengalur-smith.html>

Tamraparni Dasu

AT&T Labs-Research, USA /
http://www.research.att.com/people/Dasu_Tamraparni/index.html

Michael Gertz

Institute of Computer Science / University of Heidelberg, Germany /
<http://www.informatik.uni-heidelberg.de>

Markus Helfert

Dublin City University, Ireland / <http://www.computing.dcu.ie/~mhelfert/>

Panos Ipeirotis

New York University, USA / <http://people.stern.nyu.edu/panos/>

Barbara Klein

University of Michigan, Dearborn, USA / <http://www.som.umd.umich.edu/3363/>

Xiaobai (Bob) Li

University of Massachusetts, Lowell, USA /
<http://www.uml.edu/msb/faculty/li-xiaobai.aspx>

Paolo Missier

Newcastle University, UK / <http://www.ncl.ac.uk/computing/people/profile/paolo.missier>

Paolo Papotti

Qatar Computing Research Institute (QCRI), Qatar /
<http://qcri.org.qa/our-people/bio?pid=34&par=acc&name=PaoloPapotti>

Elizabeth Pierce

University of Arkansas at Little Rock, USA / <http://technologize.ualr.edu/msiq/>

Monica Scannapieco

Istat, Italy / email: scannapi@istat.it

Ganesan Shankaranarayanan

Babson College, USA /
<http://www.babson.edu/faculty/profiles/Pages/shankaranarayanan.aspx>

Maria-Esther Vidal

Universidad Simón Bolívar, Venezuela / <http://ldc.usb.ve/~mvidal/>

Hongwei (Harry) Zhu

University of Massachusetts, Lowell, USA /
<http://www.uml.edu/MSB/faculty/Zhu-Harry.aspx>

Senior Advisors

Robert Austin

Harvard University, USA / <http://hbswk.hbs.edu/faculty/raustin.html>

Peter Chen

Louisiana State University, USA / <http://csc.lsu.edu/~chen/display.html>

Jay Nunamaker

University of Arizona, Tucson, USA /

<http://mis.eller.arizona.edu/cgi-bin/web/faculty/detail.cgi?id=34>

M. Tamer Özsu

University of Waterloo, Canada / <http://db.uwaterloo.ca/~tozsu/>

Edgar Sibley

George Mason University, USA / http://policy.gmu.edu/faculty/bio/facbio_sibley.html

Richard Wang

Massachusetts Institute of Technology, USA / <http://mitiq.mit.edu/>

Andrew Whinston

University of Texas, Austin, USA /

<http://www.mcombs.utexas.edu/dept/irom/faculty/profiles/index-vita.asp?addtarget=78>

Stanford University, USA / <http://infolab.stanford.edu/~widom/>

Founding Editors-in-Chief

Yang W. Lee

Northeastern University, USA / <http://www.cba.neu.edu/ylee>

Stuart E. Madnick

Massachusetts Institute of Technology, USA / <http://web.mit.edu/smadnick/www/home.html>

Headquarters Journals Staff

Laura Lander

Journal Manager

Irma Strolia

Editorial Assistant

Media Content Marketing

Production

The *ACM Journal of Data and Information Quality (JDIQ)* (ISSN: 1936-1955) is published quarterly in Winter, Spring, Summer, Fall by the Association for Computing Machinery (ACM), 2 Penn Plaza, Suite 701, New York, NY 10121-0701. Printed in the U.S.A. POSTMASTER: Send address changes to *ACM Journal of Data and Information Quality*, ACM, 2 Penn Plaza, Suite 701, New York, NY 10121-0701.

Copyright © 2012 by the Association for Computing Machinery (ACM). Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: Publications Department, ACM, Fax +1 212-869-0481 or email.permissions@acm.org.



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

For other copying of articles that carry a code at the bottom of the first or last page or screen display, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

ACM Journal of Data and Information Quality

<http://jdiq.acm.org>

Guide to Manuscript Submission

Submission to the *ACM Journal of Data and Information Quality* is done electronically through <http://acm.manuscriptcentral.com>. Once you are at that site, you can create an account and password with which you can enter the ACM Manuscript Central manuscript review tracking system. From a drop-down list of journals, choose *ACM Journal of Data and Information Quality* and proceed to the Author Center to submit your manuscript and your accompanying files.

You will be asked to create an abstract that will be used throughout the system as a synopsis of your paper. You will also be asked to classify your submission using the ACM Computing Classification System through a link provided at the Author Center. For completeness, please select at least one primary-level classification followed by two secondary-level classifications. To make the process easier, you may cut and paste from the list. Remember, you, the author, know best which area and sub-areas are covered by your paper; in addition to clarifying the area where your paper belongs, classification often helps in quickly identifying suitable reviewers for your paper. So it is important that you provide as thorough a classification of your paper as possible.

The ACM Production Department prefers that your manuscript be prepared in either LaTeX or Ms Word format. Style files for manuscript preparation can be obtained at the following location: <http://www.acm.org/publications/submissions>. For editorial review, the manuscript should be submitted as a PDF or Postscript file. Accompanying material can be in any number of text or image formats, as well as software/documentation bundles in zip or tar-gzipped formats.

Questions regarding editorial review process should be directed to the Editor-in-Chief. Questions regarding the post-acceptance production process should be addressed to the Journal Manager, Laura Lander, at lander@hq.acm.org.

Subscription, Single Copy, and Membership Information.

Send orders to:

ACM Member Services Dept.
General Post Office
PO Box 30777
New York, NY 10087-0777

For information, contact:

Mail: ACM Member Services Dept.
2 Penn Plaza, Suite 701
New York, NY 10121-0701
Phone: +1-212-626-0500
Fax: +1-212-944-1318
Email: acmhelp@acm.org
Catalog: <http://www.acm.org/catalog>

Subscription rates for *ACM Journal of Data and Information Quality* are \$53 per year for ACM members, \$48 for students, and \$165 for nonmembers. Single copies are \$18 each for ACM members and \$40 for nonmembers. Your subscription expiration date is coded in four digits at the top of your mailing label; the first two digits show the year, the last two show the month of expiration.

About ACM. ACM is the world's largest educational and scientific computing society, uniting educators, researchers and professionals to inspire dialogue, share resources and address the field's challenges. ACM strengthens the computing profession's collective voice through strong leadership, promotion of the highest standards, and recognition of technical excellence. ACM supports the professional growth of its members by providing opportunities for life-long learning, career development, and professional networking.

Visit ACM's Website: <http://www.acm.org>.

Change of Address Notification. To notify ACM of a change of address, use the addresses above or send an email to coa@acm.org.

Please allow 6–8 weeks for new membership or change of name and address to become effective. Send your old label with your new address notification. To avoid interruption of service, notify your local post office before change of residence. For a fee, the post office will forward 2nd- and 3rd-class periodicals.