Editors' Comments
Where the JDIQ Articles Come From:
Incubating Research in an Emerging Field

The *ACM Journal of Data and Information Quality* (*ACM JDIQ*) has become the leading publication forum for the data and information quality field and now marks the publication of its 3rd issue. We actively encourage researchers and practitioners to use this journal as a research forum in which new ideas and discoveries may flow freely. A question of no small importance that the readers might ask is: where do the articles that appear in this issue come from?

Prior to the debut of *ACM JDIQ*, many high-quality and relevant articles had been published under data or information quality headings or keywords in other more broadly-focused journals as profiled in the inaugural issue of this journal (available free for download from the ACM Digital Library through December 2010). Since the inaugural issue, the number of submissions to *ACM JDIQ* has increased as more prospective authors learned about the journal. We have also maintained contacts with program chairs of relevant conferences and workshops to identify high-quality papers as candidates for a fast-track review process at the journal. Thus far, we have received many high-quality papers from various conferences, including those presented at the *International Conference on Information Quality (ICIQ)*, *Americas' Conference on Information systems (AMCIS)*, *International Conferences on Information Systems (ICIS)*, and at the *Quality Database (QDB) Workshop*. We continue to encourage the program committees of other such conferences and workshops to contact us with their recommendations.

In this Issue

We present four articles in this issue. Three are extended works based on previously presented papers at the *ICIQ*. These three articles share a common concern for assessing various quality aspects of data stored in a given database. The fourth article addresses the preservation and authentication of the original quality of content information. All four articles develop new solutions, employing simulations and experiments to demonstrate proof of concept and the efficacy of the suggested techniques, methods, and models.

In Article 14 "Towards a Method for Data Accuracy Assessment Utilizing a Bayesian Network Learning Algorithm" by Valerie Sessions and Marco Valtorta, the authors present a new algorithm, the Accuracy Assessment Algorithm (AAA), for estimating the accuracy of a dataset. The authors derive

this new algorithm by explicitly incorporating accuracy metrics into the algorithm development process. By applying Bayesian learning algorithms stemming from the artificial intelligence perspective, Sessions and Valtorta develop and test their new algorithm using three different sets of data from the medical area. The results provide fertile ground for future research in developing fine-grained algorithms for estimating the accuracy of a dataset. More broadly speaking, this article exemplifies the application of perspectives from artificial intelligence for refining assessment techniques and associated algorithms to a data quality problem. In practice, contingency plans and tools for data assessment could emerge from articulating different levels and utility of data quality for various types of data. This is a good read particularly for data quality researchers and practitioners who are interested in new assessment techniques and tools beyond those conventionally available in the areas of data mining, artificial intelligence, and knowledge discovery.

This article went through three revisions by the authors during the review process. The original manuscript was an extended research of their initial study presented at the *11th International Conference on Information Quality* at MIT in 2006 and received the Madnick Best Paper Award with a USD $1,000 cash award.

Article 15 "Dual Assessment of Data Quality in Customer Databases" by Adir Even and Ganesan Shankaranarayanan, provides two useful complementary ways for assessing data quality, what the authors call "impartial" and "contextual" data quality. The authors provide these assessment techniques using an example dataset in a Customer Relationship Management database for managing alumni relations. The authors also describe the techniques involved as a way for comparing the two assessment methods, "impartial" and "contextual" assessments, in order for a manager to prioritize the improvement solutions for a given dataset.

This article is a good read particularly for researchers who are interested in furthering the work of developing a cost-benefit model of data quality improvement and associated assessment techniques and methods that are anchored on economic utility of existing or potential use of data. For practitioners, this work can be applied in gaining insights into the differentiated benefit from managing data quality in their organizations.

This article went through five revisions by the authors during the review process. The original manuscript was an extended work based on their initial study which was presented at the *12th International Conference on the Information Quality* at MIT in 2007 and received the Madnick Best Paper Award with a USD $1,000 cash award.

Article 16 "An Accuracy Metric: Percentages, Randomness and Probabilities" by Craig Fisher, Eitel Lauria, and Carolyn Matheus, investigates the complexity involved in assessing accuracy of data. Using two simulation studies, the authors identify the accuracy metric to include two complementary measures: a randomness measure (based on Lempel-Ziv complexity measure) and a probability distribution (Poisson) value of accuracy. This article is a good read for researchers who are interested in developing assessment techniques for various aspects of data problems. For practitioners, this article can be a

useful input for developing automated tools, as well as discussing contingency plans for an organization's competing data improvement projects.

This article went through four revisions by the authors during the manuscript review process. The original manuscript was an extended work based on their initial study that was presented at the *12th International Conference on the Information Quality* at MIT in 2007. Their conference paper was chosen as one of the finalists for the Madnick Best Paper Award.

Article 17 is "Compensated Signature Embedding for Multimedia Content Authentification", by Surfyan Ababneh, Rashid Ansari, and Ashfaq Khokhar. This article deals with the problem of guaranteeing the original quality of information in multimedia. Ababneh, Ansari and Khokhar provide improved compensated signature-embedding techniques and demonstrate their efficacy of performance using simulations of example image authentication. They propose a system which uses the combination of robust watermarking and fragile content-based signatures to achieve effective authentication without adding additional information. This article is a good read for researchers who are interested in devising new techniques in security and privacy protection for various information products. For practitioners, this article should be useful for establishing requirements for exchanging or sharing content information in the supply-chain or in various types of collaboration situation. This article can also be used for developing an implementation plan for a new tool for image signatures or, more broadly still, in the security and privacy area.

The article went through three revisions by the authors during the manuscript review process prior to an acceptance for its publication.

The Reviewer Process

We would like to share some information and observations of the review process. Each submitted manuscript goes though a double-blind review process. Neither reviewers nor authors are known to the other party, and typically, an Associate Editor and at least three reviewers are involved in reviewing each manuscript.

We found most reviewers and AEs to be much more thorough than we had expected (which is wonderful news!). Among the 12 articles accepted for the first three issues, none was accepted based on its initial submission; they all required minor or major revisions. The breakdown was:

—accepted after first revision: 1 manuscript,
—accepted after second revision: 3 manuscripts,
—accepted after third revision: 6 manuscripts,
—accepted after fourth revision: 1 manuscript,
—accepted after fifth revision: 1 manuscript.

We are reducing the review cycle time, particularly for the first-round review, while improving the quality of the manuscripts with help from associate editors, reviewers, and authors.

Acknowledgements: Associate Editors and Reviewers

As is evident from the multiple rounds of reviews by the review teams, many individuals have worked diligently behind the scenes to bring this issue to fruition. Since establishing *JDIQ*, many reviewers, associate editors, guest editors for special issues, and invited ad hoc associate editors have worked tirelessly to go through the manuscript review process.

We appreciate our devoted Associate Editors who have led the review process and worked with the reviewers and those who have worked on various activities for establishing the journal and identifying quality research in the past year. They are recognized on the inside cover page of the journal. We also recognize the ad hoc associate editors and reviewers (through September 2009) who devoted their time to provide insightful recommendations for each review process at the end of this editors' comments section. We especially thank the authors for their tremendous work and their invaluable contribution to our dialogue on data and information quality.

We continue to welcome new and innovative research papers. While there is often value in addressing the same or similar questions found in the existing literature, we look forward to publishing articles that ask new and interesting questions. We also welcome research that applies theories from diverse fields to data and information quality issues, and we do not discriminate against any particular research method.

Enjoy the read, and we look forward to new issues in 2010.

<div align="right">

Stuart E. Madnick
Yang W. Lee
Editors-in-Chief
http://jdiq.acm.org

</div>

List of Ad Hoc Associate Editors* and Reviewers

Mr.	Sufyan	Ababneh	Dr.	Martin	Michalowski
Dr.	Elainie	Abebe	Dr.	Diego	Milano
Prof.	Latif	Al-Hakim	Dr.	Nora	Misiolek
Dr.	Irit	Askira Gelman*	Dr.	Paolo	Missier
Prof.	Donald	Ballou	Dr.	Bamshad	Mobasher
Dr.	Philip	Bernstein	Prof.	Yang-Sae	Moon
Dr.	Laure	Berti-Equille	Mr.	Heiko	Mueller
Dr.	Indrajit	Bhattacharya	Mr.	Sham	Navathe
Dr.	Jit	Biswas	Prof.	Pam	Neely
Dr.	Sandra	Boughal	Dr.	David	Olson
Prof.	Matthew	Bovee	Dr.	Mourad	Ouzzani
Prof.	Paul	Bowen	Prof.	Themis	Palpanas
Prof.	Anthony	Bryant	Prof.	Hee Sun	Park
Prof.	Ismael	Caballero	Mrs.	Claudia	Perlich
Dr.	Toon	Calders	Dr.	Elizabeth	Pierce
Dr.	Cinzia	Cappiello	Prof.	David	Pingry
Dr.	Monica	Chiarini Tremblay	Dr.	Shekhar	Pradhan
Dr.	Peter	Christen	Prof.	Ken	Pu
Dr.	Sarah	Cohen-Boulakia	Prof.	Daphne	Raban
Dr.	Yasha	Crnkovic	Prof.	Amnon	Rapoport
Dr.	Nick	Dedeke	Dr.	Eliot	Rich
Prof.	Marlon	Dumas	Prof.	Blake	Roger
Mr.	Mohammed	Eltabakh	Prof.	Kai-Uwe	Sattler
Dr.	Suzanne	Embury	Dr.	Monica	Scannapieco
Dr.	Adir	Even*	Dr.	Thomas	Schweiger
Dr.	Craig	Fisher*	Dr.	Valerie	Sessions
Prof.	Helena	Galhardas	Prof.	G.	Shankaranarayanan
Prof.	Monica	Garfield	Dr.	John	Slone
Prof.	Sanjay	Goel	Prof.	Diane	Strong
Prof.	Janis	Gogan	Prof.	Besiki	Stvilia
Prof.	P. Y.	Hsu	Dr.	Sherry	Sun
Dr.	Melanie	Herschel	Dr.	Peter	Tarasewich
Prof.	H V	Jagadish	Prof.	Singwhat	Tee
Dr.	Matthew	Jensen	Prof.	P.	Thangavel
Dr.	Zhengrui	Jiang	Dr.	Nikola	Trcka
Prof.	Beverly	Kahn	Mr.	Cihan	Varol
Prof.	Dmitri V.	Kalashnikov	Prof.	Suresh	Venkatasubramnaian
Dr.	Ralph	Kimball	Prof.	Vassilios	Verykios
Miss	Anja	Klein	Mrs.	Kristin	Weber
Dr.	Barbara	Klein	Prof.	Andreas	Wichert
Prof.	Andy	Koronios	Dr.	Ningning	Wu
Prof.	Nick	Koudas	Dr.	Moe	Wynn
Prof.	Eitel	Lauria*	Prof.	Hui	Xiong
Prof.	Guy	Lebanon	Mr.	Guanshuo	Xu
Prof.	Dongwon	Lee	Dr.	Peng	Xu
Prof.	Hans	Lehmann	Mr.	Mohammed	Yakout
Mr.	Wolfgang	Lehner	Dr.	Xiaoxin	Yin
Prof.	Hans-J.	Lenz	Dr.	Xiaohui	Yu
Prof.	Chen	Li	Dr.	Chrisy	Yao
Dr.	Xiaobai	Li	Dr.	Fan	Zhang
Prof.	Herbert	Longenecker	Dr.	Dong	Zheng
Prof.	Bertram	Ludaescher	Prof.	Bin	Zhu
Dr.	Andrea	Maurino	Dr.	Hongwei	Zhu
Mr.	Chris	Mayfield			