



**Proceedings  
of the Fifth MIT  
Information Quality  
Industry Symposium**



Massachusetts Institute of Technology  
Cambridge, MA, USA

July 13-15, 2011

Edited by

WooYoung Chung

Forea Wang

Copyright © 2011 MIT Information Quality Program, Massachusetts Institute of Technology

All presentations included in these proceedings are published with explicit permission from their author or copyrights holder. These presentations are protected by the international and U.S. copyright law. Any copying, distribution, exhibition in public or private meetings, and/or pursuing any derivative work, by any third party, in part or as a whole, is strictly prohibited without explicit written permission of the original copyrights holder. Any third party who wishes to quote or refer to the information presented in these proceedings must comply with the fair use provisions of the U.S. copyright law 17 U.S.C. §107 and must properly cite the author and these proceedings as its source.

## FOREWORD

WELCOME to the *5th MIT Information Quality Industry Symposium (MIT 2011 IQIS)*. The last ten years have proven exciting for data quality practitioners, as more and more organizations have come to understand the critical importance of high-quality data. Data quality efforts in large enterprises often fail to deliver on the expectations of the business. As with any service provision, the key to avoiding such disappointment is to look at the problem from the beginning through the eyes of the business.

Many have made vast improvements serving as role models for others. The next ten years promise to be even more exciting. Leading organizations including the MIT Information Quality community are coming to realize that data, just like capital and people, are assets that must be explicitly managed, as any valuable corporate asset would be.

In order to produce an event that meets the needs of data and information quality practitioners, the MIT IQIS organizing committee has researched extensively with arrange of data experts in data management, data warehousing, business intelligence, customer information and information management, across a range of industries and sectors.

We are honored and privileged to have Mr. Steve Meurer, Senior Vice President, Comparative Data & Informatics, University HealthSystem Consortium to give the keynote address on Thursday, July 14.

This event is co-sponsored by the MIT Information Quality Program and in part, by [Acxiom Corporation](#), [Informatica](#), [QIBRAS](#), [Symbolic Systems, Inc.](#), and [Talend](#).

We would like to express our gratitude to Forea Wang, Rith Peou, Allen Moulton, and other MIT staff for their assistance in producing the Symposium proceedings, and managing the various aspects of the Symposium operations. The final Symposium program and other information regarding the Symposium are available at <http://mitiq.mit.edu/IQIS>.

## MIT 2011 Information Quality Industry Symposium Organizing Committee

### **Symposium Chair:**

Richard Wang, MIT Information Quality Program  
rwang@mit.edu, 617-304-3120

### **Symposium Vice Chairs:**

Willa Pickering, Lockheed Martin  
Steve Sarsfield, Talend

### **Track Chairs:**

Joseph Bugajski, Gartner Group (**Business & Finance**)

Suzanne Acar, Federal Data Architecture Subcommittee (**Government**)

Peter Anlyan, Anlyan Consulting (**Healthcare**)

Linda Kresl, Yahoo! (**Tutorial**)

Yang Lee, Northeastern University (**IQ Societies**)

### **Symposium Manager:**

Forea Wang, MIT Information Quality Program

### **Proceedings Chair:**

WooYoung Chung, Cleveland State University

### **Advisor Committee:**

Peter Aiken, VCU/Datablueprint  
Peter Benson, Electronic Commerce Code Mgt Association  
Bruce Davidson, Cedars-Sinai Health System  
Stuart Madnick, MIT Sloan School of Management  
Charles McKinney, Freddie Mac  
Kelly Morris, Gartner Group  
Michael Nix, Fletcher-Allen Health Care, Vermont  
Frank Ponzio, Symbolic Systems, Inc.

# MIT 2011 (5<sup>th</sup>) Information Quality Industry Symposium

## Location

Massachusetts Institute of Technology  
Tang Building (E51) in the MIT East Campus  
70 Memorial Drive, Cambridge, MA, USA  
Dress Code: Business casual

Date: July 13 – July 15, 2011

<http://mitiq.MIT.edu/IQIS/2011>  
Contact: Dr. Richard Wang, 617-304-3120, [rwang@mit.edu](mailto:rwang@mit.edu)  
MIT IQ Program Director

WEDNESDAY, July 13, 2011		
8:15-9:00 AM	CONTINENTAL BREAKFAST & REGISTRATION	Room
9:00-10:30 AM	<b>Session T1A</b> Financial Service Industry Enterprise Data Management Michael Atkin, EDM Council Executive Director	E51-325
Tutorial Sessions	<b>Session T1B</b> Improving Data Portability and Long Term Data Retention through ISO Standards 8000 and 22745 Gerald Radack, PiLog USA Inc.	E51-345 pg. 1
	<b>Session T1C</b> Building a Practical Data Quality Management Solution for an Academic Medical Center Bruce Davidson, Alein Chun, Ann Chu, and Thai Lam Data Quality Management Team Resource & Outcomes Management Department Cedars-Sinai Medical Center	E51-372 pg. 2
	<b>Session T1D</b> One-To-Many ...or Much Too Much? A tutorial about efficient, effective requirements analysis that promotes information quality by recycling some best practices, reconditioning some others, and debunking and discarding the rest Joe Maguire, O'Kelly Associates	E51-376 pg. 34
10:30-10:45 AM	COFFEE BREAK	
10:45-12:15 PM	<b>Session T2A</b> The Butterfly Effect of Data Quality Following the Impact of IQ on Revenue, Efficiency and Compliance Steve Sarsfield, Talend  Case Study: Catching Fraud in Alameda County, CA Earl Stevens, Alpine Consulting	E51-325 pg. 49



2:15- 2:30 PM	COFFEE BREAK	
2:30- 4:00 PM Parallel Sessions	<b>Session 1A</b> Making Data Valuable to the Enterprise Session Chair: Joe Bugajski	E51-325
	1. Data Quality Makes The Big Time: How We Got Here, and Why We'll Never Be Done..... Ted Friedman Vice President Gartner, Inc.	pg. 107
	2. Beyond Stakeholder Buy-In: Getting Programmers and Other IT Professionals To Embrace IQ Initiatives ..... Joe Maguire Principal Analyst and Consultant O'Kelly Associates.	pg. 116
	<b>Session 1B</b> Panel: Usability in Health Information Technology Session Chair: Bruce Davidson	E51-345
	<b>Data Quality as a Dimension of Usability for Health IT:          Update from a Working Group</b> .....  Wil Yu, Office of the National Coordinator for Health IT Bettijoyce Lide, National Institute of Standards and Technology Lana Lowry, National Institute of Standards and Technology Matt Quinn, Agency for Healthcare Research and Quality Emily Patterson, The Ohio State University Ellen Voorhees, National Institute of Standards and Technology	pg. 134
	<b>Session 1C</b> Civilian Government Agency Perspectives Session Chair: Suzanne Acar	E51-372
	1. Jump Starting Data Governance: A Program Manager's Story ..... Laura Duvall Senior IT Specialist, DHS	pg. 166
	2. Case Study: IRS Data Integrity ..... Daniel Chaddock, Director, Solution Engineering, IRS	pg. 177
4:00- 4:30 PM	COFFEE BREAK	
4:30- 6:00 PM Parallel Sessions	<b>Session 2A</b> Case Studies: IQ at Work Session Chair: Joe Bugajski	E51-325
	1. Case Study: How Data Quality Has Evolved at MathWorks ..... Kirk Amidon Marketing Data Analyst, MathWorks	pg. 178
	2. Integrating IPMAP with Information Life Cycle Management for the Upstream Oil & Gas Industry ..... Melanie Bell, Robert Elliott, Ed Moore, Yang Lee, Allen Moulton Chevron and MIT	pg. 200

	<p><b>Session 2B</b> Panel: DQ in Electronic Health Records Implementation Session Chair: Bruce Davidson</p> <p><b>Electronic Health Records (EHR): Benefits and Challenges for Data Quality</b> ..... pg. 202</p> <p>Michael Nix, Fletcher Allen Health Care, VT Alein Chun, Cedars-Sinai Medical Center, CA Elisa Horbatuk, Stony Brook University Medical Center, NY David Harriman, University of Chicago Medical Center, IL</p> <p>Physician Discussants: Dr. William Greene (Stony Brook Univ.), Dr. Sameer Badlani (Univ. of Chicago), Dr. Paul Rosenau (Fletcher Allen)</p>	E51-345
	<p><b>Session 2C</b> U.S. Air Force Information Quality Initiatives Session Chair: Willa Pickering</p> <p>1. The GIG is up! The Global Information Grid Changes Everything! ..... pg. 239 Harry Pape, Transformation Advisor Windmill International Inc. supporting SAF/AQX</p> <p>2. A Man With A Watch: The Convergence of Authoritative Data Sources &amp; Data Quality ..... pg. 240 Thomas Redman Navesink Consulting Group Wright Nodine Jr., Transformation Advisor, Windmill International Inc. supporting HQ USAF/A1X1</p> <p>3. Trusted Data Markets – The Role of Data Quality ..... pg. 252 David Becker, Principal Information Systems Engineer The MITRE Corporation</p>	E51-372
6:00-7:30 PM	SYMPOSIUM RECEPTION (ONLY) AT MIT FACULTY CLUB	



THURSDAY, July 14, 2011		
8:15-9:00 AM	CONTINENTAL BREAKFAST	
9:00-10:30 AM	<p><b>Session 3A</b> Advances in Information Quality Session Chair: Yang Lee</p> <p>1. Developing Standards for Financial Reporting Data ..... pg. 253            Yang Lee, Associate Professor            Northeastern University            Harry Zhu, Assistant Professor            Old Dominion University</p> <p>2. Update: IQ in China ..... pg. 255            Wei Huang, Executive Dean            Xi'an Jiaotong University</p> <p>3. Update: Chief Data Officer in the Office of the Secretary of Defense            Ron Bechtold, OSD CIO</p>	E51-325
	<p><b>Session 3B</b> DQ in Optimization of Electronic Health Records Session Chair: Bruce Davidson</p> <p>1. Human Factors Guidance to Prevent Healthcare Disparities with the Adoption of Electronic Health Records Panel ..... pg. 257            M. Chris Gibbons, Johns Hopkins Urban Health Institute            Lana Lowry, National Institute of Standards and Technology            Matt Quinn, Agency for Healthcare Research and Quality            Bettijoyce Lide, National Institute of Standards and Technology</p> <p>2. Re-Design of EHR Screens and Data Elements to Optimize Use by Providers ..... pg. 270            Paul Rosenau            Vermont Children's Hospital Director for Quality            Vermont Children's Hospital at Fletcher Allen Health Care            Michael Nix            Clinical/Operations Measurement Group Manager            James M. Jeffords Institute for Quality &amp; Operational Effectiveness            Fletcher Allen Health Care            Hannah Avarraschild            Measurement Support Associate            Clinical/Operations Measurement Group Manager            James M. Jeffords Institute for Quality &amp; Operational Effectiveness</p>	E51-345
	<p><b>Session 3C</b> National Defense Perspectives Panel Panel Chair: Willa Pickering</p> <p><b>National Defense Perspective</b> ..... pg. 290</p> <p>Panelists:            John Hohwald, Senior Principal Research Engineer, Lockheed Martin            Susmit Patel, MITRE Corporation            Frank Ponzio, President/CEO, Symbolic Systems, Inc.</p>	E51-372

10:30-11:00 AM	COFFEE BREAK	
11 AM-12:30 PM Parallel Sessions	<b>Session 4A</b> Measuring and Reporting Data Quality Session Chair: Joe Bugajski	E51-325
	1. The Value of Data Quality Solutions ..... Deepa Srinivasan IT Applications Manager National Instruments Christine McClary Internal Consultant, Business Intelligence Group National Instruments	pg. 292
	2. Visualizing Financial Information Quality Using Heat Maps and Semantic Data Quality Rating System ..... Ashu Bhatnagar, CEO Good Morning Research	pg. 308
	<b>Session 4B</b> HIT Trends and Your Organization Session Chairs: Peter Anlyan	E51-345
	1. HITECH Investments in HIT and What We're Seeing in the Field as a Result ..... Laura Adams, CEO Rhode Island Quality Institute (RIQI)  2. IT and Comparative Effectiveness Research Don Goldmann, Senior VP Institute for Healthcare Improvement (IHI)	pg. 324
	<b>Session 4C</b> Practical Approaches Session Chair: Bob Kenley, Program Manager, MIT Lean Advancement Initiative	E51-372
	1. Information Quality Approach at Colorado ..... Micheline Casey, Former Chief Data Officer Governor's Office of Information Technology  2. The "Tagged Data Authority Engine" — Assurance and Data Integrity for Government Agencies and Fusion Centers ..... Shawn P. McCarthy, Director of Research IDC Government Insights	pg. 326  pg. 347
12:30-2:00PM	LUNCH & KEYNOTE SPEECH	E51-345
	<b>Keynote Speech</b> Data, Quality, & Becoming a C.R.A.P. Detector .....  Steve Meurer Senior Vice President Comparative Data & Informatics University HealthSystem Consortium	pg. 365
2:00-2:30 PM	COFFEE BREAK	

2:30-4:00 PM Parallel Sessions	<b>Session 5A</b> Future Data: Drag the Enterprise into the 21 <sup>st</sup> Century Session Chair: Joe Bugajski  1. Tweeting & Blogging to World-Class Data Governance - A Tale or Reality? ..... Kimmo Kontra Chairman Datpro  2. Data Clouds: One Small Step for IT, One Giant Leap for Modern Enterprises ..... Lyn Robison Research VP and Director of Data Management Research Gartner, Inc.	E51-325  pg. 394  pg. 410
	<b>Session 5B</b> Data Aggregator's Dilemma: Partnerships for Managing Data Quality in a Shared Database Session Chair: Bruce Davidson  <b>Data Aggregator's Dilemma: Partnerships for Managing Data Quality in a Shared Database</b> .....  Panelists: Allen Juris University HealthSystem Consortium Martha Radford, MD, FACC, FAHA NYU, Chief Quality Officer Eric Hixson, PhD, MBA Cleveland Clinics, Director, Data Resources Management Allison Sabel, MD, PhD, MPH (invited) Denver Health, Director of Biostatistics & Clinical Data Warehouse	E51-345  pg. 427
	<b>Session 5C</b> NATO Perspectives Session Chair: Willa Pickering  1. The DNA of Modern Logistics -- NATO Codification ..... Ian Smith, NCS Transformation Steering Group  2. The NATO Codification System: Implementing the Standard and Continuous Improvement of Cost Effectiveness and Data Quality ..... Henriette Schromm, Senior Project Manager, ESG  3. Faster, Better and Lower Cost Access to Authoritative Supplier and Manufacturers Data ..... Peter R. Benson, Executive Director, ECCMA  4. Cataloging at Source in Practice ..... Salomon de Jager, Executive Director, PiLog	E51-372  pg. 450  pg. 460  pg. 478  pg. 491
4:00-4:30 PM	COFFEE BREAK	

4:30-6:00 PM Parallel Sessions	<b>Session 6A</b> Models and Metadata: Session Chair: Joe Bugajski  1. Repository Redux: Lessons Learned from 40 Years of Metadata Repositories ..... David Eddy, General Manager David Eddy & Associates  2. Semantic Metadata for Data Quality: Impact on Data Integration ..... David S. Frankel, Lead Standards Architect – Model Driven Systems SAP Labs	E51-325  pg. 492  pg. 500
	<b>Session 6B</b> The Business Value of Data Quality Session Chair: Anne Marie Smith, LiquidHub, Inc.  1. Making It Real – The Business Value of Data Quality at Sallie Mae ..... Michele Koch, Director, Enterprise Data Management Sallie Mae  2. How to Assess the Business Value of Data Quality ..... Danette McGilvray, President and Principal Granite Falls Consulting, Inc.	E51-345  pg. 502  pg. 522
	<b>Session 6C</b> IQ Perspectives Session Chair: John Talburt  1. Bringing in the Emerging <i>Realized</i> Information Age 3.0 ..... Larry English, President Information Impact International, Inc.  2. Update: IQ Certification Program – UALR/IAIDQ ..... John Talburt, UALR	E51-372  pg. 545  pg. 574
6:00-7:30 PM	SYMPOSIUM RECEPTION ONLY AT MIT FACULTY CLUB	
<b>FRIDAY, July 15, 2011</b>		
8:15-9:00 AM	CONTINENTAL BREAKFAST	

<p>9:00-10:30 AM Parallel Sessions</p>	<p><b>Session 7A</b> The Data Economy Session Chair: Joe Bugajski</p> <p>1. Infonomics: The Economics of Information and Principles of Information Asset Management ..... pg. 590 Doug Laney Eminence Lead, Deloitte Analytics Institute Deloitte Consulting LLP</p> <p>2. The Decision Model: An Introduction and a Case Study at Freddie Mac ..... pg. 604 Barbara von Halle and Larry Goldberg Managing Partners Knowledge Partners International, LLC (KPI) Tom Schweikert Business Rules Manager, Information &amp; Decision Mgmt Freddie Mac</p>	<p>E51-325</p>
	<p><b>Session 7B</b> Key Information Quality Issues in Healthcare IT Session Chair: Peter Anlyan</p> <p>1. Meaningful Use in Five Years ..... pg. 623 Michele Madison Morris, Manning and Martin, LLP</p> <p>2. Data Quality and Interoperability of Identity Data in the Veterans Health Administration ..... pg. 625 Sara Temnitz &amp; Alice Cave Veterans Health Administration</p>	<p>E51-345</p>
	<p><b>Session 7C</b> Perspectives from Military and Intelligence Activities Session Chair: TBD</p> <p>1. Lessons Learned in Implementing Metadata and XML Standards to Improve Quality and Discoverability of Intelligence ..... pg. 639 Doug Whall, Alanthus Associates</p> <p>2. A Data Quality Strategy for Large Organizations – How the German Armed Forces manage master data quality in the logistics domain ..... pg. 650 Joachim Sobota and Christian Fürber, Joint Support Command, German Armed Forces</p>	<p>E51-372</p>
<p>10:30-10:45 AM</p>	<p>COFFEE BREAK</p>	
<p>10:45 - 12:00 PM Plenary Session</p>	<p><b>CLOSING PLENARY SESSION</b> <b>Open Discussion &amp; Closing Remarks</b> Moderator: Stuart Madnick, MIT Sloan School of Management</p>	<p>E51-345</p>
	<p><b>END OF THE SYMPOSIUM</b></p>	

**Pre-Symposium Event I: The Organizing Committee of MITIQIS is offering a workshop on Tuesday, July 12, 2011 (0900-1600), E51-372 at MIT. The registration is free for Symposium attendees. Facilitators are Peter Anlyan and Earl Hadden and, in part, by Richard Wang.**

“Leading Through Information Quality” -  
Part of the PPI – Strategic Breakthrough Methodology

The confluence of development of people, process and information is crucial to reaching the mature stage of data quality. In many organizations, an important challenge within this triumvirate is the ability of people to navigate the political waters and conflict inherent in teaming with cross-functional teams and gaining sponsorship at the top. Successful efforts are dependent on basic management, leadership and communication skills. [Register here](#). For more information, contact Peter Anlyan at [peter@anlyanconsulting.com](mailto:peter@anlyanconsulting.com) or 919-669-0052.

**Post-Symposium Event I: Experiences Achieving Data Quality in a High Volume DoD Operational Environment, Friday, July 15, 2011 (1300-1700), E51-325, MIT. Frank Ponzio, Jr. (Symbolic Systems, Inc.)**

The forum discussions will be based on seven years (and counting) experiences that have resulted from implementing and now maintaining data quality processes while building tailored data products that rely on authoritative data sources for input. This high volume production environment develops data products for military systems that are built on a demanding schedule.

To continuously improve the data products’ data quality requires an agile development environment, with enduring capabilities and certain critical skills. The workshop will focus on the risks, threats, and opportunities that the team has had to address in order to achieve and stay successful despite unreliable authoritative data sources. It will also include a discussion of experiences and challenges developing and maintaining data quality processes, tools, and staff.

**Post-Symposium Event II: EDM Council Meeting by Michael Atkin, Friday, July 15, 2011 (1300-1700), E51-345, MIT. The event is independent of MIT2011IQIS Symposium. Contact [atkin@edmcouncil.org](mailto:atkin@edmcouncil.org).**