THE ADVANCE OF PROCESS

(Panel Description)

Panel Chair: Thomas C. Redman, Ph.D.

A business process is a "series of interrelated work activities, usually characterized by specific inputs and repeated value-adding steps, which produce a specific set of outputs." Business processes are the means by which organizations obtain (raw) input, complete their work, and deliver value to their customers.

The notions of process and process management are slowly (and painfully) entering the mainstream. Most organizations have not embraced process management. Some have done so only half-heartedly, with unsatisfactory results. Finally, those (relatively few) organizations that devoted serious attention to managing and/or reengineering their processes have reaped huge rewards. They have defined and implemented simpler, less costly, more robust, flexible, and faster processes that better meet customer needs (when it comes to process, you really can get it "better, faster, cheaper"). And process management has been especially effective for data quality.

This panel will explore "the ascent of process." It will address the following questions:
- What is process management and why is it so important?
- Is the continued ascent of process management assured or might it prove to be another management fad?
- What factors help/hinder the penetration of process management into organizations?
- What are the ramifications of process management on data/information quality and vice versa?

Panelists
Larry English, *President, Information Impact International*
A. Blanton Godfrey, *Dean and Joseph D. Moore Distinguished University Professor,*
North Carolina State University
Mark Parise, *Division President, Data Base Services, Information Resources, Inc.*
Thomas C. Redman, *President, Navesink Consulting Group*

MANAGING INFORMATION QUALITY –
A CRITICAL PROCESS FOR MOST ORGANIZATIONS
(Position Paper)

A. Blanton Godfrey
Joseph D. Moore Distinguished University Professor and
Dean, College of Textiles
North Carolina State University
Blanton_Godfrey@ncsu.edu

In the mid 1970s I inherited the supervision of developing and maintaining the Quality Information System for AT&T. It was then I first noticed how difficult it was to first define the quality of the information and second to assure this quality. Over the years first at AT&T and then with a wide range of companies while in Juran Institute, I constantly became more and more aware of the challenge of assuring the quality of the information used to make critical decisions each day. Few organizations had a clear definition of the quality needed, fewer still had any means of creating this quality.

Time after time information quality activities started only after a disaster. In AT&T we first learned of how much capacity in each central office was unused due to poor records (between 20-30%) only when a service crisis in New York City prompted us to try to find unused lines to reduce backlogs. Only when we introduced new 800 services did we find we had no practical means of detecting database errors and correcting these errors. It was years later that a quality improvement team leader thought to measure the waste in our direct marketing activities due to wrong addresses. It should not have surprised us to find we threw away over $10 million annually. What is not measured is rarely managed.

The same problems abound in all organizations. Products are shipped to wrong addresses. Bills are inaccurate, incorrectly addressed, and incomplete. Multiple bills are sent to the same address. Hospital medical records are notoriously incorrect. In one study over 80% had missing or incorrect critical information. Emergency vehicles are sent to wrong addresses. Medication errors are commonplace.

The good news is that information quality can be managed. The same tools and processes that we use so successfully to manage product and service quality work in managing information quality. Many organizations are starting to treat information as the critical resource it is and to actively manage the quality of this valuable resource.

PROCESS MANAGEMENT AND INFORMATION QUALITY: HOW IMPROVING INFORMATION PRODUCTION PROCESSES IMPROVES INFORMATION (PRODUCT) QUALITY

(Position Paper)

Larry P. English
Information Impact International, Inc.
Larry.English@infoimpact.com

PROCESS MANAGEMENT

Process management is a requirement for sustaining an environment that consistently produces quality information. Process management goes beyond process control, that is, consistently produces products whose variation is within defined specifications.

Process management requires a focus on the customers of the products produced by the process. It requires the definition of product specifications that meet end customer requirements, not just the specifications that may or may not meet customer needs.

Process as used here refers not just to single activity processes (input-process-output), but the entire value chain from the customer request (for a product) to the customer benefit (customer satisfaction with the product).

The automobile assembly line is a process or value chain. It is the entire assembly line as a whole that must be managed, not just each discrete activity (stand-alone process or activity). Individual activities may be controlled to meet its specification, but if the product produced does not fit with the components of the next processes or activities, process failure occurs and scrap and rework is required.

INFORMATION QUALITY AND PROCESS MANAGEMENT

Information quality practitioners often think of data or information quality as the focus of their attention, without thinking of the customers of that data, nor of the process activities (both automated and manual) that produce (create and maintain) it. Defective or nonquality data is simply the symptom of a defective process. To simply fix the defective data is only "information scrap and rework"—not true information quality management. Correcting data *only* to propagate to a downstream database or data warehouse without correcting the data at the source where it is still being used is *sub-optimized* and *defective* "information scrap and rework." This type of error correction is:

- *Sub-optimized* because it leaves defective data at the source databases allowing processes that use the source data to fail.
- *Defective* because it introduces a new information quality problem—inconsistency of the data in the target database or data warehouse and the source:
  - The redundant data cannot be reconciled

- o Aggregations of data cannot be drilled down to its originating source
- o Incremental updates to the target can cause subsequent contamination by the still remaining defective data at the source.

Defective, nonquality data is a symptom of a "defective" process. There are of course many precipitating and root causes of defective data, such as:

- Mis-entering data at the source
- Not having appropriately defined update processes that capture updates to volatile information subject to *information quality decay*
- Environmental causes such as lack of training, inappropriate performance measures, no accountability in management for information quality

From a process management standpoint, there are other causes, such as:

- Lack of robust information product specifications
- Defining information product specifications from the perspective of only the immediate (business area) requirements, missing potential requirements from all categories of downstream knowledge workers who require the data

INFORMATION PROCESS MANAGEMENT

To manage information quality, one must manage the processes across the information value chain from origination information producer to its most remote information customers.

Information process management includes:

- Identifying an information group, such as "Customer," "Product," "Order" or "Insurance Claim"
- Identifying and defining the information value/cost chain[1]
- Defining the information product specifications based on *all* information stakeholders (producers and customers)
- Defining the data movement specifications (when data must be moved from one data store to another)
- Managing the processes that create, update, store the data by the information producers
- Managing the processes that extract, transform and load or propagate the data
- Managing the processes that retrieve, format and present the data to the knowledge workers
- Providing feedback across the value chain for continuous improvement

Information process management requires:

- Process identification. What is our enterprise mission, what objectives must we accomplish, what strategies shall we take, what processes, then, must we perform?
- Customer identification. Who are the information customers? These are the knowledge workers—both internal and external, such as end-customers and regulatory reporting authorities whose requirements must be understood and met.
- Supplier identification. Who are the information producers? These are the people who are the natural create point of knowledge about objects and events because they perform activities in which data becomes known. These individuals must have training, and processes designed in which to error-proof the collection, gathering, or creating of the data. These individuals must also know who their information customers are and their requirements, so they understand the importance of the information they create, but may not require themselves to perform their jobs. Information producers require the resources necessary to perform their jobs properly to produce quality information.

- Customer product requirement specification. Define the information requirements, i.e., what data must/should we know to accomplish our mission and meet our objectives. Then define the information quality requirement characteristics (completeness, accuracy, timeliness, etc.) that meet the most rigorous expectations based on risk and cost of process failure. This information requirement is the same team approach that Deming describes in Quality Point 4, when he says purchasing should be a team effort, consisting of "product engineer and representatives of manufacturing, purchasing, sales, or whatever other departments will be involved with the product."[2] Defining data requires input from all information stakeholders, not just the immediate beneficiaries of an application being developed.

- Process definition. Before you can put a process into control, let alone optimize it, the process must be defined so that it is repeatable. To define a process the first time, Masaaki Imai describes the "Standardize, Do, Check, Act" cycle as a requirement before processes can effectively be improved.[3]

- Process improvement. To continually meet the ever-increasing quality expectations of knowledge workers and end customers, a Plan-Do-Check-Act process is required. This is required to improve processes to the point of optimization.

- Process optimization. Processes can be in control, yet not optimized. An organization can have superbly managed business area activities, everyone doing a good job with no one having problems, yet fail, because the business areas activities are sub-optimizing its own work by not working together as a team across the value chain.[4] To optimize a process, one must evaluate the activities that are cost-adding and eliminate or minimize them. Creating redundant databases or files is cost-adding if data can be shared from a single or replicated database. The interfaces and redundant databases are cost-adding only. This is "muda" (waste) of processing and should be eliminated, so that "at every step in which a work piece or a piece of information is worked on" value is added, not cost.[5]

- Management accountability. Managers of activities must be held accountable for the quality of the information created by the information producers who report to them. Manager's accountability is not just to its up-line management, but also horizontally to its information customer peer managers who depend on the information produced by their business area. This is called managerial information stewardship.[6]

CONCLUSION

Information quality is outcome of managed processes. Defective data is the symptom of broken processes. Unnecessary redundant data and transforming interfaces reflect sub-optimized processes that contain waste and are sources where errors may be introduced.

The goal of information quality is to eliminate waste of information scrap and rework, unnecessary processing, and to increase business effectiveness by increasing customer satisfaction (both in the organizations products and services, and in the information it provides to internal and external information customers).

By managing the processes across the information value chains, one eliminates unnecessary costs, increases customer satisfaction and with it customer lifetime value.

GLOSSARY:

Business value chain: "An end-to-end set of activities that begins with a request from a customer and ends with specific benefit or benefits for a customer, either internal or external."[7] The terms "process" and "business process" are sometimes used interchangeably.

Information product specifications : The set of information resource data (meta data) characteristics that define all characteristics for a process and creating/updating applications can produce quality information. Information product specification characteristics include: data name, definition, domain or data value set (code values or ranges) and the business rules that identify policies and constraints on the potential values. These specifications must be understandable to the information producers who create and maintain the data and the knowledge workers who apply the data in their work.[8]

Information quality : Consistently meeting all knowledge worker and end-customer expectations in all the characteristics of the information products and services they deem important. The degree to which information consistently meets the requirements and expectations of all knowledge workers who require it to perform their processes.[9]

Information value/cost chain: "An end-to-end set of processes and data stores, electronic and otherwise, involved in creating, updating, interfacing, and propagating data of a specific type from its origination to its ultimate data store, including independent data entry processes, if any."[10]

Process management: The process of ensuring that a process is defined, controlled to consistently produce products that meet defined quality standards, improved as required to meet or exceed all customer expectations and optimized to eliminate waste and non-value adding.[11]

---

[1] See L. English, *Improving Data Warehouse and Business Information Quality (IDB&BIQ)*, NY: John Wiley & Sons, 1999, pp. 160-163 for a discussion of IQ Assessment, Step 3 "Identify the Information Value and Cost Chain."

[2] Mary Walton, *The Deming Management Method*, New York, NY: Perigee Books, 1986, p. 64.

[3] M. Imai, *Gemba Kaizen: A Commonsense, Low Cost Approach to Management*, New York: McGraw-Hill, 1997, p. 5.

[4] W. E. Deming, *Out of the Crisis*, Cambridge: Massachusetts Institute of Technology Center for Advanced Engineering Study, 1986, p. 63.

[5] M. Imai, p. 79.

[6] *IDB&BIQ*, p. 406-407.

[7] *Ibid.*, p. 495.

[8] L. English, *Glossary of Terms (InfoImpact Glossary),* Brentwood, TN: Information Impact International, pp. 23-24.

[9] *Ibid.*, p. 12.

[10] *IDB&BIQ*, p. 480.

[11] *InfoImpact Glossary*, p. 18.

DATA QUALITY MANAGEMENT – SIMPLY MANAGING THE PROCESS

(Position Paper)

Mark S. Parise
Division President, Data Base Services
Information Resources, Inc.
mark.parise@infores.com

BACKGROUND ON IRI AND THE ENVIRONMENT

IRI Has Experience in Managing Data and Data Quality

- Leading CPG Provider - Market Tracking, Volumetric Measurement, Panel Based Purchase Data, and Product Testing.
- 120+ tera bytes of data
- 17 million UPC's
- 90 billion UPC records
- 1.6 million data sets
- 13 million batch jobs per year
- 45, 000 UPC's accessed per second

Marketing Leading Companies are our Customers - Examples
Very Tough Business Environment is Requiring Clients To Make Even Tougher Decisions. They are Looking for Data To Develop and Support Those Decisions, Especially When The Stakes are High

This is Causing Clients to Push Data to Previously Unknown Levels of Granular and Frequency

Trolling For Completive Opportunities - Micro Levels…

- Marketing and Sales Investment
- Rationalizing Products
- Integrating Market Data w/ Financial Data for Performance Tracking
- Support M&A Activity

Speed and Access are No Longer The Primary Requirement. - Driven by Improvements in Technology, The Decreasing Cost of Hardware/Software and The Flexibility of Tools

Data Quality Has Been Pushed To the Forefront with Granularity and Consistency With Other Sources Being Paramount.

Wrong Data Regularly Effects Multi Million Dollar Decisions Everyday - Example Using Hypothetical Client Situation
However, The Same Economic Environmental Pressures Also Require That This Improvement Not Come at an Incremental Cost.

Variation Exists in All Data - As you increase granularity you increase variation. As you increase frequency and cycle time you also increase the opportunity for variation -. Example

It's All About Understanding the Variation Threshold and Managing Variation - Variation in the Process that Produce the Data, the Raw Data Itself, the Manipulation of the Data and the Final Output and Expectations for What the Data Should Be

Data Quality is Mainly Driven By the Process of Gathering, Transforming and Delivering Information and Insights to Clients - Process Map
- Getting It -Capturing and Gathering Raw Data
- Storing It - Transmitting, Loading, Storing, Normalizing
- Transforming It- Value Functions
- Delivering It – Packaging it and Transmitting
- Disseminating and Communicating – Formatting It, Presenting It To Clients, Clients Presenting It/Using It

Enabling Processes
- Responding to Inquiries and Reactive Error Correction
- Proactive Policy Management and Values Function Management
- Developing New and Enhancing Existing Products
- After the Fact Quality Checking, Validation and Audit

Each Sub Process and Handoff Produces Variation and Error That Effects Client Decision Making – Overlay Error Chart

Because the Process of Creating and Delivering the Data Have the Primary Impact of the Quality of the Data, the Approaches Used to Efficiently Manage Classic Process Quality and Improvements Also Apply to Data Quality Management
- Ongoing Customer Requirement Understanding and Measurement
- Focus on the Entire Process From the Point of Capture to Client Action
- Proactive Process Management and Tools are Critical To Controlling and Reducing Error
- Critical; Measurement and Dashboards of Big Y's and X's
- Process Owner Performance Management Tied to Measures
- Process Improvement Tools
- Efficiency Management

Implemented a Process Management Centered Data Quality Management Approach Is not Difficult - There are Relatively Low Cost Quality Tools, Process and Sources That are Available Today.
- Six Sigma - IRI's Approach
- TQM
- ISO
- "Light" Versions and any combination of the above.

CONCLUSION

The key to improving data quality, efficiently, is implementing a process management centered data quality management approach focused on managing down variation.