

# Rationalizing a Data Quality Dashboard Solution

Information Management Newsletters, January 4, 2010

[http://www.information-management.com/newsletters/data\\_quality\\_dashboard-10016811-1.html](http://www.information-management.com/newsletters/data_quality_dashboard-10016811-1.html)

Steve Ruff

When data quality is the topic of conversation, too often managers and execs are quick to sweep the discussion aside, or worse, sweep it under the rug. In my experience, data quality projects often have to demonstrate their resolve and overall worth early to establish their right to be implemented along with the multitude of other large projects that traditionally gobble up a company's IT budget. The usual suspects include data warehouse initiatives, enterprise resource planning and operational data store implementations and any number of industry-specific or operational systems, such as a policy administration or sales automation systems. While these usual suspects consume the bulk of the attention and budget from IT leadership, another smaller project has the potential to provide an array of benefits and payback for countless organizations: a data quality dashboard solution.

It is not always easy to articulate the key reasons why a data quality dashboard and reporting solution will provide true ROI. Numerous reasons can initially be cited:

- To ensure compliance with Sarbanes-Oxley and other regulatory requirements.
- To proactively increase awareness of the quality of an organization's data assets.
- To bolster confidence on the business side in the quality of the data upon which they rely for reporting and decision-making.
- To enable end users by providing them with tools that will help them be more efficient and productive in their day-to-day job, with fewer headaches along the way.
- To measure, trend and track data quality improvements over time, ensuring that investments in process improvements and technology infrastructure are protected.
- To audit, gauge and track the efficiency of systems integration exception handling and/or data stewardship processes.

If these reasons were presented to a seminar of consultants and mildly intrigued company managers, they would certainly make the ROI case for a data quality dashboard initiative. However, when an organization is truly considering putting pen to paper and allocating serious dollars to any prospective project, they need to understand how this will provide real value to their company; not just idealistic examples about future perceived, unrealized returns. They need to see hard ROI.

Then how can the justifications be articulated in a way that will not only inspire and invoke action and secure the funding that will ensure maximum benefit for an

organization? To best justify a data quality initiative, it is necessary to fully understand the individual organization's needs, both in the near-and- long term.

### **Crash Course**

It is helpful to clearly define a data quality dashboard and what it can provide to a business. First, let's break down the components of a generic dashboard report.

*"Dashboards provide visibility into key performance indicators through simple visual graphics such as gauges, charts and tables within a Web browser.*

*Dashboards are appealing because they:*

- *Present a wide number of different metrics in a single consolidated view.*
- *Roll up details into high-level summaries.*
- *Provide intuitive indicators, such as gauges and stoplights, that are instantly understandable for example, red bar means problem, green bar means everything is on plan.*
- *Display an at-a-glance view of the current operational state of the project."*

Now we need to modify this definition to be specific to a data quality dashboard:

*"A data quality dashboard is a visually intensive report that is created through collaboration between business and IT to define, measure and trend the consistency and quality of their data over time measured against business-defined KPIs. It presents numerous metrics and rollup details in a single consolidated view, utilizing intuitive indicators such as gauges, stoplights and charts to provide an operational snapshot of the current state of the organization."*

Now that definitions are cleared up, the next hurdle is to ensure that the data quality dashboard is adequately rationalized, ensuring maximum ROI for the individual organization.

### **Situation Assessment: Who Needs a Data Quality Dashboard Solution?**

In today's regulation-driven world, most companies are under a great deal of pressure to provide comprehensive oversight and management of their data assets. This pressure can mount from multiple angles, such as from industry or governmental regulatory bodies as well as internal leadership. Accountability is the resounding theme, but how can a data quality dashboard help?

#### **We have to first understand the individual organization.**

- Where is the company on the maturity roadmap?
- Do they already have a reliable data warehouse, and if so, do they fully utilize its capabilities for business intelligence, data mining and reporting?
- What other flagship IT programs are currently "in-flight" or have been successfully implemented in the past five years? Better still, what IT initiatives

begun within the last five years have stalled or failed?

- What reporting pain points does the business currently experience on a daily, weekly or monthly basis across the various business functions (such as finance, sales, marketing, support and services)?

These questions are geared at understanding the current maturity of the organization, the current IT landscape and the base reporting infrastructure in place. Once these current-state analysis questions (among others) are answered, the picture comes a bit more into focus, and we can begin to understand where a customized data quality dashboard can provide maximum value.

Any organization undertaking one of the large, usual suspects projects previously mentioned is a prime candidate for data quality standardization, cleansing/scrubbing and governance. In 2008, Gartner Inc. published “Nine Fatal Flaws in BI Implementations.” Flaw number three specifically addresses data quality issues:

*“Flaw No. 3: “Data quality problem? What data quality problem?” Data quality issues are almost ubiquitous and the impact on BI is significant people won’t use BI applications that are founded on irrelevant, incomplete or questionable data. To avoid this, firms should establish a process or set of automated controls to identify data quality issues in incoming data and block low-quality data from entering the data warehouse or BI platform.”*

Clearly, data quality can either positively or negatively impact any large system implementation, as well as the day-to-day reporting and customer interactions that the business must navigate. A data quality dashboard designed to monitor and measure critical KPIs is the answer to identifying and tracking these issues. It provides the necessary visibility throughout the organization that will help ensure data quality issues do not doom your program objectives. It also enables the business and data governance functions by providing reliable and consistent access to data quality information and exceptions.

### **Utilizing Data Quality Scorecards**

Scorecards, like dashboards, are a highly effective way of highlighting an organization’s data issues. They can quickly and efficiently underscore data anomalies and discrepancies and visually represent these metrics in a way that is easily consumed by even the most nontechnical people. Specifically, data quality scorecards can be geared toward an organization’s senior business folks and leadership team. These are the decision-makers, and it is essential to present a technical issue such as data quality in a form that can be implicitly explained and quickly understood. By assigning a “grade” or confidence level to an organization’s data assets, a data scorecard quickly proves its point to its audience. From the highly technical database administrator to the senior VP, a data scorecard communicates results in a way that virtually anyone can identify with.

In my experience, the difference between a dashboard and a scorecard is blurry. Usually the most recognizable difference between the two is the mechanism in which they are presented. A data quality dashboard is usually built within a dynamic reporting toolset, while a data quality scorecard can be found in something as rudimentary as a spreadsheet.

However, there are some definitional differences between the two. Wayne Eckerson of The Data Warehouse Institute provides a solid distinction: “The primary difference is that dashboards tend to monitor the performance of operational processes, whereas scorecards tend to chart the progress of tactical and strategic goals. Dashboards also tend to display charts and tables with conditional formatting whereas scorecards use graphical symbols and icons to represent the status and trends of key metrics.”

In practice, data quality scorecards are a great starting point in driving out the metrics and KPIs that can be directly input into a data quality dashboard. They are a great tool for understanding the data quality pain points of an organization at a granular level, and should be considered as part of any data quality dashboard implementation.

### **Making the Case: Realizing True ROI**

Through all of this, we can't lose sight of the ROI factor. Maximizing the benefit to the business is the ultimate goal. This benefit can be realized by quantifying the process improvements and efficiency gains that will be the direct result of a one-stop-shop for data quality monitoring and reporting. To present true ROI in a way that people can understand, it has to be quantified. Ask these questions, in no particular order, to help you come up with some hard numbers:

What daily, weekly and monthly reports or processes require overly manual work to produce/complete on the part of the organization's valuable resources?

- How many reports rely on manual data pulls in the form of spreadsheets or Access databases that require time and effort on the part of numerous IT and businesspeople? How frequently are these reports run, and how much time is spent weekly/monthly verifying manual data feeds across systems?
- Do business functions agree on common definitions of shared data elements, such as revenue or expense? If not, how much time is spent translating and communicating cross-functional reports across the enterprise? How much time is wasted recalculating and reconciling reports when an issue is discovered at the eleventh hour?
- How poor is the data that resides today in the source systems and is utilized for reporting and decision-making? Is it possible to quantify the resources involved and amount of hours necessary to manage and remediate these issues on an ad hoc basis?
- What IT exception management reporting is in place today? What data steward processes are in place to handle these exceptions? How many resources are involved to manage, monitor and manually track and remediate exceptions as

they occur? How is the efficiency of this process tracked, if at all?

Many times, it's all of these superfluous hours spent by numerous folks across IT and business that go unrecognized and untracked. The effort, hours, resources and subsequent frustration and headaches cost time and money, and money is certainly quantifiable. Answer these questions, and you will be able to derive the true cost to the organization and justify a data quality dashboard solution that will ease the burden and produce measurable returns.

*Steve Ruff is a senior consultant at Attevo, a global business and information technology consulting firm. He may be reached at [sruff@attevo.com](mailto:sruff@attevo.com).*

For more information on related topics, visit the following channels:

- Data Quality
- Scorecards and Dashboards

©2010 Information Management and SourceMedia, Inc. All rights reserved.

SourceMedia is an Investcorp company.

Use, duplication, or sale of this service, or data contained herein, is strictly prohibited.