Making Data Quality a Way of Life

By adopting a rigorous operational and governance process for data quality, organizations can ensure the veracity and business impact of data big and small.

Executive Summary

Many organizations have contemplated, or are already on the journey to, making data quality a way of life. What many of them missed, however, are the necessary operational contexts of data quality and its interplay with a well-honed governance program that allows the functions which comprise data quality as a way of life.

This white paper explores the necessary operational activities that make data quality a way of life. It also examines why no data, including what is normally deemed unstructured data and big data, can be excluded from such a program and the necessary governance interplays that are required to make data quality efforts less of a chore and more normal operating procedure.

Data Quality

Data quality comprises the processes that allow business stewards to extract, capture and originate value from high-quality data by ensuring, through a repeatable process, that all data goes through a reasonability check. More specifically:

- Supplying the highest-quality data to those who use it in the form of knowledge, learned inferences, heard inferences and innovations as a vehicle to capture, originate or extract value for the organization.
- Providing a means for those who count on data being trustworthy to validate the data through an always-available audit trail that provides the lineage of transformations and references to reasonability exceptions.
- Holding the excellence of data quality available at the successful completion of a development project to the same standard long after the development effort is completed.

Done correctly, operational data quality provides a vehicle to ensure the trustworthiness of information, which will allow for more expedient and effective execution of actions. It enhances the value of data captured from the marketplace and thus reduces the risks associated with internal processes and disruptions originating outside the organization and optimizes the market disruptions associated with innovations introduced into the marketplace. The combined impact of value captured, risks mitigated and the assault of innovations is a material component of the value of the data assets of the organization.
To ensure that its data is highly relevant and accurate, an organization must:

- Engage a continuous process that certifies the accuracy of data made available for recurring and extraordinary business processes.
- Ensure that the data made available is current and just in time for recurring and extraordinary business processes. This means that data quality processes that hold up the certification of data negatively impact the value of data utilized for business processes.
- Ensure that the processes employed for certification of data made available for recurring and extraordinary business processes are repeatable and auditable.
- Encapsulate the operational data quality processes with proactive data governance processes that ensure that the data is relevant and that identified changes to the status quo (internal processes and system changes, market disruptions, etc.) are judiciously reviewed to ensure that they do not negatively impact the relevance, timeliness or quality of the data.

Why Operational Data Quality Is Different than Data Quality Projects

Operational data quality, while related to data quality efforts that occur during development efforts, requires a mind-set shift among data quality practitioners. Some processes that are prerequisites to enable operational data quality are often not enabled out of the box by the major data quality products. The baselining process, which we will detail later on in this white paper, is necessary for identifying data quality exceptions, which are then categorized in the operational data quality processes.

This concept of operational data quality is a departure from what many organizations use their data quality process for, which is often limited to creating a set of rules to elevate the quality of information consumed during a development project by enriching, de-duplicating or cleansing the information. However, the data quality efforts are commonly re-deployed after the completion of the development initiative to help with the next development initiative. The reasonability checks to ensure the quality of information such as messages, queries and files are rarely employed. Thus, the once pristine data quality associated with the
development project often dissipates into less than trustworthy information.

The vendors that publish data quality suites are mostly focused on assisting with ensuring pristine data during a delivery initiative. They rarely have complete solutions, and commonly have no capabilities for operational data quality controls.

In order to implement operational data quality capabilities, four major components are mandatory:

- **Translate the rules developed for data quality into a repeatable workflow that can be executed either prior to or in conjunction with extract, transform and load (ETL) processes.** Depending on the system, this can be done utilizing either a message/transaction process or a file (or query)/batch process. This workflow must be made operational to be useful. It is important to note that many data quality solutions available in the marketplace cannot be made operational and therefore cannot be used as an operational data quality engine.

- **Isolate a place to record exceptions identified during the operational data quality process.** This can be done by flagging exceptions in the normal data stream (e.g., if isolating exceptions would negatively impact the financial performance of the organization, such as not processing customer orders because of data quality issues) or holding them for cleansing and resubmitting.

- **Develop a series of baselines used to determine the reasonability of incoming data.**

- **Categorize the baseline exceptions and dispatch exceptions to the originators** (i.e., the suppliers of source data that is validated for reasonability using operational data quality controls) for exposed transformation errors and to compliance and management teams for process exceptions (market disruptions that require management attention, fraud, etc.).

There is a movement to manage data as a non-expiring asset of the organization, making this a primary function of the newly-created position of chief data officer. If information is trustworthy, it can be wielded with an understanding that the underlying data is relevant and accurate. Furthermore, it means that the information's derivation can be reviewed at any point in time, should its derivation be questioned. Highly relevant and accurate information that can be wielded just in time is just worth more to the organization because it can be used to generate greater degrees of value. Measuring the value of the data asset is the next great challenge for the business community. Peter Aiken’s Data Blueprint describes (through data strategy and other means) the processes to ensure the accuracy and relevance of data and to measure the value of the information exposed to the data blueprint processes.

When the focus of data quality is on a project or program either consuming data from or supplying data assets to other programs, the interplay

---

**Gauging Trustworthiness**

- What is the lineage of the data?
- What was done to it before I received it?
- Do I have the most current data?
- Is there more relevant data I should be using?
- What is the context of the data?
- Does the data mean what I think it means?
- Will my analyses and decisions be tainted by bad data?
- Can I trust the data?

Figure 2
between the organization’s data governance and data quality functions comes into focus. At organizations that treat their data as a key asset, the data governance function is accountable for prioritizing the means to elevate the value of data assets, one of which is data quality. The symbiosis between data quality and data governance is especially strong and will be explored later in this white paper.

The Secret to Making Data Quality a Way of Life

Data assets used to derive value for a business entity can and should be mapped to a business model to assess the organizational risks associated with the business model artifacts that consume or supply data to the business. These exposures will be one of two types: originated data sources or transformed data sources that should be carefully scrutinized as part of a well-honed data trustworthiness assurance process. Only then can a business measure and ensure the creation, capture and origination of value of data assets used throughout the enterprise (see Figure 3).

The value of data can be measured by the incremental benefits it provides to the organization – through fine-tuning processes, shepherding innovations for market disruption, capitalizing on opportunities presented in the marketplace or mitigating risks sourced within the organization (i.e., process misfires, fraud, etc.) or outside the organization (i.e., competitors cherry-picking customer bases, competitive market disrupters, etc.).

Operational data quality taken to its natural end state is the process for determining which components generated from within the organization or

---

### Figure 3

#### Data Sources

<table>
<thead>
<tr>
<th>Internal Data Sources</th>
<th>Acquired Data Sources</th>
<th>Log Sourced</th>
<th>Other Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Data</td>
<td>Reference Data</td>
<td>Web Logs</td>
<td>Customer Provided</td>
</tr>
<tr>
<td>Change Data Capture</td>
<td>Market Data</td>
<td>Mobile Logs</td>
<td>Internet Accessible</td>
</tr>
<tr>
<td>Transaction Data</td>
<td>Crowd Sourced Data</td>
<td>Call Center Logs</td>
<td>RFID Provided</td>
</tr>
<tr>
<td>Reference Data</td>
<td>Social Media Data</td>
<td>Workflows</td>
<td>Acquired Inferences</td>
</tr>
<tr>
<td>Master Data</td>
<td></td>
<td></td>
<td>Heard Inferences</td>
</tr>
</tbody>
</table>

#### Data Origination Trustworthiness Assurance

<table>
<thead>
<tr>
<th>Offering</th>
<th>Interactions</th>
<th>Channel</th>
<th>Monetize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information/Knowledge</td>
<td>Customer Facing</td>
<td>Brick &amp; Mortar</td>
<td>Licensing</td>
</tr>
<tr>
<td>Capabilities</td>
<td>Supplier Facing</td>
<td>Brick &amp; Mortar</td>
<td>Usage Fee</td>
</tr>
<tr>
<td>Non-Tangible Goods</td>
<td>Employee Facing</td>
<td>Brick &amp; Mortar</td>
<td>Sale</td>
</tr>
<tr>
<td>Tangible Goods</td>
<td>Financier Facing</td>
<td>Digital</td>
<td>Leasing</td>
</tr>
<tr>
<td>Services</td>
<td>Agent Facing</td>
<td>Digital</td>
<td>Subscription</td>
</tr>
<tr>
<td></td>
<td>Media Facing</td>
<td>C:C</td>
<td>Advertising</td>
</tr>
<tr>
<td></td>
<td>Regulator Facing</td>
<td>Digital</td>
<td>Promotion</td>
</tr>
</tbody>
</table>

#### Data Transformation Trustworthiness Assurance

<table>
<thead>
<tr>
<th>Data Generation</th>
<th>Operations</th>
<th>Analytics</th>
<th>Reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combining Known</td>
<td>Data Creation</td>
<td>Descriptive</td>
<td>Statutory</td>
</tr>
<tr>
<td>Log Preparation</td>
<td>Processing</td>
<td>Predictive</td>
<td>Management</td>
</tr>
<tr>
<td>CrowdSourcing</td>
<td>Execute</td>
<td>Prescriptive</td>
<td>Regulatory</td>
</tr>
<tr>
<td>Transaction</td>
<td>Track</td>
<td>Visualization</td>
<td>Financier</td>
</tr>
<tr>
<td>Master</td>
<td>Annotate</td>
<td>Disseminate</td>
<td>Customer</td>
</tr>
<tr>
<td>Priorities</td>
<td>Collaborate</td>
<td>Collaborate</td>
<td>Media</td>
</tr>
<tr>
<td>Plans</td>
<td>Innovate</td>
<td>Detect &amp; Alert</td>
<td>Employee</td>
</tr>
<tr>
<td></td>
<td>Manage</td>
<td>Business Model</td>
<td>Investor</td>
</tr>
</tbody>
</table>

#### Key Activities

<table>
<thead>
<tr>
<th>Business Model Artifacts</th>
<th>Data Governance Interfaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Origination Trustworthiness Assurance</td>
<td>Data Transformation Trustworthiness Assurance</td>
</tr>
<tr>
<td>Offering</td>
<td>Interactions</td>
</tr>
<tr>
<td>Information/Knowledge</td>
<td>Customer Facing</td>
</tr>
<tr>
<td>Capabilities</td>
<td>Supplier Facing</td>
</tr>
<tr>
<td>Non-Tangible Goods</td>
<td>Employee Facing</td>
</tr>
<tr>
<td>Tangible Goods</td>
<td>Financier Facing</td>
</tr>
<tr>
<td>Services</td>
<td>Agent Facing</td>
</tr>
<tr>
<td></td>
<td>Media Facing</td>
</tr>
<tr>
<td></td>
<td>Regulator Facing</td>
</tr>
</tbody>
</table>

---

cognizant 20-20 insights 4
collected from external data sources are outside the expected tolerances. Statistically known as an outlier, this point can be either accurate good or bad news or something that requires research. It does not mean that the data point is bad; it simply means that what arrived through the operational data quality processes was an outlier and presents either a data quality remediation opportunity or something requiring management action. Something that requires management action is in fact the gateway to data-driven risk management, or a detector for something that isn’t meeting expected data tolerances.

There are four data stores or databases required to perform operational data quality. It is important to note that in most cases the data stores provided by third-party vendors do not provide out-of-the-box capabilities to execute operational data quality; in most cases, they are designed to augment delivery projects.

- **Data stores:**
  - To validate the reasonability of incoming data, the total population of previously processed data is required. This will be used to compute the variance from expected values (typically two to three standard deviations from the mean for analyzable information such as metrics, numeric data, etc.) and the alignment to employed taxonomies, ontologies and reference data for non-tabular and textual data.
  - A place to house the acceptability criteria or baseline criteria. Criteria are used at runtime to determine if incoming data is reasonable. Examples of baselines include sales with a variance greater than one sigma, or a variance of more than one standard deviation from the mean of sales; such instances should be held for review. It doesn’t mean that this information is bad; it just means that it requires further investigation.
  - A place to house data that is deemed worthy of further investigation prior to committing it to the intended target.
  - A place to house baseline statistics, or the measures used to orchestrate operational data quality. An example of a baseline statistic is that a +/-5% tolerance is acceptable for day-over-day valuation of a managed asset within a trading desk portfolio; thus, day-over-day variations greater than +/-5% would be recorded as an operational data

---

**Making Sense of Data Reasonability**

![Figure 4](image-url)
quality exception. It is important to note that the operational data quality team will be accountable for reviewing the number of exceptions triggered for each measure as a way to determine if the rules used to trigger operational data quality exceptions were too lax or stringent. Accordingly, in coordination with the data governance team, they would make the necessary changes to the baseline statistics.

- **Program code used to orchestrate operational data quality:**

  - This is a means of interrogating incoming transactions. It can be done one of three ways:

    » A batch cycle that interrogates data reasonableness after the fact through regularly scheduled routines. This means of validating requires the ability to flag data written into the intended target with some designation indicating that the data is suspect and under investigation. Batch operations do not have the stringent performance constraints of online processing environments. Therefore, while they are important, the impact of data quality processes on the overall performance characteristics of production operations will require less scrutiny, thereby allowing additional in-line or parallel processing to identify and flag or set aside data suspected of significant data quality anomalies. The flagging of data (setting a data quality flag or setting data aside) is the primary means of identifying suspect data that needs further investigation.

    » Generally, when target environments are complex – such as a big data lake or a complex data environment such as SAP, PeopleSoft or a federated data warehouse – it is less manageable to flag data designated for reasonableness validation. This is because such complex environments house data in ways that require program code to locate the data (e.g., ABAP code for accessing data in SAP, or NOSQL interfaces for big data). The batch approach to reasonableness does not impair performance characteristics of production operations, but will allow decisions to be made on suspect data until such time that the reasonableness flag can be set.

    » A listener-based technology, which grabs a copy of data for the target environment while it is being processed to validate its reasonability. It then either sets a flag or writes data to a new data store as close to the writing of data to the intended target as practicable. While many data quality tools do not possess this capability out of the box, it can be implemented with minimal effort. This approach does not impair the performance characteristics of production operations and minimizes the timeframe in which there is a risk that decisions can be made using suspect data.

    » Using stream data quality routines to validate the quality of information as part of the job stream and either set a flag to designate data that needs validation or write data to a suspense data store. This approach eliminates the timeframe in which decisions can be made using suspect data, but introduces a risk that the performance characteristics of production operations will be inconsistent due to in-stream data quality checks.

    » A means to send out data quality alerts concerning data items worthy of review. It is important to prioritize the alerts so that those who need every data alert get them, while managers who need to review the alerts (i.e., the CFO for financial alerts) receive only material alerts and summary statistics.

    » A means to ensure that data items flagged for review are actually reviewed and dispensed. Typically, the data quality team accountable for operational data quality produces aging reports of the exceptions and publishes the categorized aging through a dashboard or scorecard.

A new class of software has emerged to perform data quality checks at every data touch-point (e.g., when data was combined with other data, moved to another state in the job flow or otherwise transformed). The rule base, which typically serves as the means to judge the tolerances for numeric data variances and the reasonableness of transactions, is used to score the incoming data and determine whether it should be reviewed prior to committing it as an executed transaction or transmission file.
It should be expected that there will be a large amount of data that is created by the data lineage process, and that understanding what systems are touched by a data lineage system throughout its data flow will be quite difficult in advance without some level of listener technology (i.e., technology that tracks the movement of data rather than explicitly pre-recording the hops the data is expected to traverse).

Such oversight is required for new regulatory processes in the financial brokerage (i.e., Dodd-Frank) and insurance spaces (i.e., Solvency II). But there are additional outside data touch-points in other industries, such as trading exchanges, where there is an inherent trust that data is transmitted properly, or royalty-based solutions and other transaction environments where the lineage of data needs to be understood (i.e., fraud detection, validated systems, Basel III, etc.).

Enabling Operational Data Quality

To effect operational data quality, it is critical that data is pre-validated to assure reasonability before being made available for reporting and analytics. This requires operational interfaces to assure the reasonability of data prior to when it arrives in staging environments, data warehouses and reporting marts.

Such a process will become unwieldy if the business lacks an organization that is tasked with orchestrating data quality validation processes. This organization should be accountable for:

- Completeness of the baselines (baselines are written for information used in key decision-making processes).
- Assurance that processes used to ensure information reasonableness are operating and that they are not negatively impacting the stability and performance of the business's operational procedures. Without this, it is unlikely that the business will embark on the operational data quality journey.
- Validation that information flagged as questionable is dispensed with. Aging of the reasonableness flags or information held in suspense will be used as one of the primary vehicles.
- A process to review information flagged for reasonability and bucket exceptions as origination issues, transformation issues, procedural issues, procedural exceptions (i.e., fraud) and market disruptions (new innovations from competitors, etc.). Note that this requires an understanding of the business represented by the data exceptions needed to make this call. The bucketed exception summary and highlight lists will be provided to the governance and/or compliance functions. Governance will utilize the list to prioritize initiatives while compliance will use the information to fine-tune compliance rules and potentially build surveillance cases.

---

**Data Integrity Scorecard**

<table>
<thead>
<tr>
<th>Data Integrity Functionality</th>
<th>Data Quality</th>
<th>ETL</th>
<th>Data Lineage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validate data conformance to standards and business rules during loads, inserts, audits, into a system.</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Identify errors as data moves between systems.</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Recognize in situ errors caused during manual data updates.</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Identify false positives (data conforms, but isn't actually valid).</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Spot data that is missing.</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Spot data that is abnormal (not in line with norms/trending).</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Maintain system lineage, history and context for each data value event across all systems for the term of the data (who, what, where, when).</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Consolidate data value consistency and conformance across all data portfolio sources for each LOB in a dashboard view.</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Provide an enterprise data integrity analysis repository.</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Provide troubleshooting ability to crawl backward/forward along the value chain and recreate a point-in-time view.</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Certifies data's conformance to standards and rules.</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
The process, which accommodates both message-based and file/query-based operations, has several important components, which will be evaluated section by section.

**Operational Data Quality Step 1: Ingestion**

The data quality process will never have, nor should it have, the performance characteristics of operational processes. Therefore, it should never be in line to operational processes because they will upset the notion of consistently favorable performance characteristics of operational systems.

The ingestion process is either to attach to a backup or read-only copy of data or to ingest messages and files into the DQ environment for analysis.

**Operational Data Quality Step 2: Processing**

From the data cache, data profiling rules will be executed using baselining rules that will identify and dispense with exceptions, or data that fails reasonability tests.

Data that passes all reasonability tests will be ignored by the operational data quality processes.

**Operational Data Quality Step 3: Categorizing**

The categorization process will segregate exceptions into:

- Market disruptions: items requiring management attention for potential strategy refinement.
- Procedural exceptions: items requiring compliance attention for potential fraud and other exceptions.
- Procedural issues.
- Transformation errors, requiring the attention of ETL teams.
- Origination errors, requiring the attention of development teams.

**Operational Data Quality Step 4: Communicating**

Once categorized, the exceptions are sent off:

- To the supplier of the source exposed to operational data quality processes.
To the governance counsel for addressing changes to the priorities that impact the valuation of the data assets of the organization.

To the compliance team for dispensation of fraud and other issues.

Scorecards are also constructed for measuring the continued improvement of data quality and aging of operational data quality exceptions.

The supporting data stores to the operational data quality process are:

• The rules used for operational data quality.
• The source exposed to the scrutiny of operational data quality.
• The identified baselining exceptions.
• The reference data and standardization rules utilized.
• The quality mart used to measure increased data quality.

Data Quality Workflow

Figure 7

Figure 8

Figure 9

Operational Data Quality: Supporting Data Stores

The supporting data stores to the operational data quality process are:

• The rules used for operational data quality.
• The source exposed to the scrutiny of operational data quality.
• The identified baselining exceptions.
• The reference data and standardization rules utilized.
• The quality mart used to measure increased data quality.
Getting on the Road to Operational Data Quality

There are several ways to build a well-honed operational data quality program — one devised to make data quality a way of life. Very few of the components are technical in nature but those that are have critical specific roles.

Leadership

While leadership is a subset of people, it is an important component of the operational data quality roadmap. It is leadership of the data quality organization that will shepherd the organized transformation from a project data quality focus (largely a development effort) to an operational data quality focus (largely a support, not a maintenance, effort). Leadership also will justify the necessary financing for the execution of operational data quality controls.

Without proper leadership, ensuring buy-in to a potentially disruptive process will become an issue for those who have the highest need for these operational data quality controls. This is because there will be additional work for them to dispense with the exceptions identified by the operational data controls. (The process appears disruptive because it creates work for the greatest sources of suspect data). These exceptions have in the past been lumped into a general pool of challenges to data trustworthiness.

Leadership from the data quality organization will also be accountable for spotlighting recurring data quality themes so that the data governance organization can prioritize and sponsor the necessary development efforts to eliminate these data quality challenges.

Leadership from both the data quality and governance programs will be continuously required to measure and communicate the value of data assets and their impact in the valuation of data assets.

People

Trained experts must be available to manage the codification of the baseline and categorization processes in order to execute an operational data quality program. These individuals will be involved in several components of the operational data quality program:

- Facilitator and caretaker for the baseline rules used for operational data quality.

---

**Data Quality Program Flow**

---

**Inputs on Data Quality Management**

---

**Figure 10**

(largely a development effort) to an operational data quality focus (largely a support, not a maintenance, effort). Leadership also will justify the necessary financing for the execution of operational data quality controls.
• Process owner for the categorization of operational data quality exceptions.
• Process owner for the computation of operational data quality metrics, including those that measure the positive incremental value of data assets associated with the operational data quality program.
• Process owner for the operational data quality scorecard.
• Process owner for the aging of operational data quality exceptions.
• Facilitator for the processes associated with the publication of alerts to source systems, governance and compliance.

Metrics
The metrics are the message that the operational data quality program is functioning as expected and that it is having the intended effects on the organization's data assets. Key asset categories that should be monitored include:
• **Usage metrics**, such as data quality sessions, volume of data accessed by data quality facilities, number of fixes made to data through cleansing algorithms, number of records enriched through standardization and other routines, etc. Usage metrics should track upward over time.
• **Impact metrics**, such as the improvement in customer information, increased revenue associated with better quality, cost reductions due to reduced data validation expenditures, etc. Impact metrics should track upward over time.
• **Effort metrics**, such as the hours of training provided for data quality, the number of presentations made about data quality, the effort in obtaining data that requires cleansing, etc. Effort metrics should track downward per initiative over time.
• **Cost metrics**, such as the cost of software, hardware, acquired data and manpower required to perform the data quality function. Both direct and indirect costs should be tracked, and these should track downward per data quality initiative over time.
• **Operational effectiveness metrics**, which measure the success of managing data as an organizational asset and the reduction of operational data quality exceptions caught per interfaced application over time.

Process
The processes of data quality will be directed by the governance council. A key requirement is that the process employed for operational data quality fits the organization rather than trying to change the organization's culture and operational processes to fit the operational data quality program.

Technology
There are minimal technology components required to orchestrate an operational data quality program. These include:
• It can have direct access to files and messages used in the company's operations without adversely impacting the consistently acceptable performance of production operations. Many data quality facilities do not have the capability of analyzing incoming data without placing it in an environment augmented for data quality operations. Such a requirement adds an unacceptable delay on making data available just in time for regularly occurring and extraordinary business processes.
• It can be integrated as the primary interface or interfaces to an operational data quality program without significant uplift in data quality effort.
• It can be operated without manual intervention. Such manual intervention would eliminate the benefits achieved from an operational data quality program.

Data
Data is the raw material for an operational data quality program. Without data that leverages the operational data quality program, the process is simply a test case and will have minimal if any impact on incrementally improving the value of the organization's data assets.

Looking Ahead
One closing remark on the roadmap to operational data quality as a vehicle to making data quality a way of life is that all of the items on the roadmap do not happen at once. Also, there will be organizational strife that hopefully can be dispersed by the governance council. Without a functional, active data governance council, reducing organizational strife associated with initiating an operational data quality roadmap must be the responsibility of the leadership of the data quality program.
An Operational Data Quality Scorecard

A scorecard (as depicted in Figure 12) is traditionally the communication vehicle of choice to report on the activity and success of the operational data quality program and its effectiveness as a major cog in the management of data as a non-expiring asset of the organization and making data quality a way of life.

About the Author

Mark Albala is the Lead for Cognizant’s Data Quality and Governance Service Line. This service line provides the expertise in data quality and governance solutions to the portfolio of activities tackled by the Enterprise Information Management Business Unit. A graduate of Syracuse University, Mark has held senior thought leadership, advanced technical and trusted advisory roles for organizations focused on the disciplines of information management for over 20 years. He can be reached at Mark.Albala@cognizant.com.

About Cognizant

Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world’s leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 75 development and delivery centers worldwide and approximately 199,700 employees as of September 30, 2014, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.