Effective Data Governance in Banking - Failure is Not an Option
Effective Data Governance in Banking - Failure is Not an Option

In the movie Apollo 13, actor Tom Hanks, who plays astronaut Jim Lovell, delivers a line that almost instantly became an American idiom. “Houston,” he says, “we have a problem.” It may not be quite on par with the disaster faced by Lovell and his fellow crew members, but when it comes to effective data governance, many North American banks also have “a problem.”

Consumers are undefined, unknown, untracked and unmanaged. Lead representatives on client accounts are duplicative or non-existent. Accountability is assumed, yet unmanaged. Transaction dollar values are recognized, but cannot be accurately or consistently aggregated across districts, regions or national geographies. Clients move and the continuity of their addresses evaporates.

Combine that with complex customer profiles in which a consumer is, at the same time, an individual, a corporation, a borrower and an investor, and consider how difficult it is to establish and track risk exposure, employee activities and accountability. Factor in other critical elements required for maintaining capital requirements, providing protection for security and fraud, or accurately rendering consolidated post-acquisition financial statements, and the extent of the data challenge becomes obvious.

Data quality has gained much attention in the business press in relation to recent regulations and corporate governance requirements that mandate attention to improved corporate data quality. Unfortunately, most companies have been taking data quality for granted when it comes to effectively managing the business and preparing financial statements. But data quality issues can be costly. The Data Warehouse Institute estimates that data quality problems cost U.S. businesses more than $600 billion a year. In those terms, data quality issues represent approximately a five percent loss in productivity and nearly the GDP for Turkey - the world’s 17th-largest economy in 2005.
In addition to being costly, data quality problems are also widespread. According to The Data Warehouse Institute, the top technical challenges facing companies implementing these solutions are:

1) Data quality and security issues (55 percent)
2) Lack of business case and funding (45 percent)
3) Poor data integration infrastructure (38 percent)
4) Metadata management issues (36 percent)
5) Lack of IT data integration skills (33 percent)
6) Data transformation and aggregation (27 percent)

Gartner predicts that through 2007 more than 50 percent of data warehouse projects will have limited acceptance, or be outright failures, because of lack of attention to data quality issues. Furthermore, many organizations fail to recognize that they have an issue with data quality, and instead focus on identifying, extracting and loading data.

Federal regulators have enacted considerable changes with which even federal agencies must comply. The Data Quality Act of 2000 covers data quality policies and required corrective actions aimed at ensuring quality reporting. The Health Insurance Portability and Accountability Act of 1996 (HIPAA), for instance, encourages the electronic transmission of data, which implies higher levels of quality as well as security and performance improvements. The Gramm–Leach–Bliley Act of 1999, also known as the Financial Modernization Act of 1999, defines requirements for financial institutions, which includes banks, securities firms, insurance companies, and other businesses that provide financial products and services to consumers. These companies are now required to improve their protection of consumers’ personal financial information.

The Sarbanes-Oxley Act of 2002 introduced even more changes to financial practices and corporate governance regulation, including guidance “to protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws.”

A more recent requirement, the Basel II Accord, put in place at the end of 2005, imposes new regulatory methods for calculating capital for banking and non-banking subsidiaries of bank holding companies. Compliance with Basel II requires additional capital commitments from all banks, as well as
additional transparency and reporting to both the marketplace and regulators. Financial services firms are expected to rely heavily on new, revised or improved technology infrastructures to ensure compliance with the new regulations, and there likely will be a significant competitive advantage for banks meeting those guidelines. Their regulatory capital requirements will be reduced, which effectively increases their capital.

In the preliminary draft of their pending publication Voluntary vs. Mandatory Corporate Governance Regulation: Theory and Evidence, Anita Anand, Frank Milne and Lynnette Purda argue that there are considerable reasons for companies to embrace a governance strategy, regardless of governmental requirements, one of which is competition in the marketplace for scarce capital. Their study constructs a model of the interaction between incentives and firms’ propensity to adopt governance mechanisms, which identifies when voluntary adoption of governance mechanisms is a firm’s optimal choice.

Their results show that in Canada, firms voluntarily implemented best-practice standards between the years 1999-2003, and that after 2002, non-cross-listed firms adopted Sarbanes-Oxley-like reforms, voluntarily. The suggestion is that mandatory governance legislation may not be necessary, at least where certain incentives exist.

Obviously, firms need to act swiftly and surely to ensure systems that render quality data. Board members, accountable executive officers, chief executives and chief financial officers have taken notice, but many may not truly understand what is required to achieve success. It is not enough to decentralize the planning and execution of such efforts, for example, since very few internal lines of business (LOBs) have the understanding, ability or objectivity to do it well.

IT organizations have an enterprise-wide view and have hands-on knowledge of data, but often lack business knowledge, which is required to understand data quality and accuracy. It is important to be sure a company’s business information lifecycle is supported by the inclusion of processes to transform data into information, information into knowledge, and knowledge into business value. To do that, a comprehensive corporate data strategy is required.
Data governance, while complex and certainly important, is relatively intuitive. A corporation must understand definitions, locations, transformations, uses, users and timing, among other requirements. While businesses have long understood the need for data governance and built large and layered organizations to manage it, the need for it is no longer optional and must meet numerous, specific criteria.
Data Governance Defined

IBM Global Business Services has defined data governance as “the orchestration of people, process and technology to enable an organization to leverage data as an enterprise asset.” As a corporate asset, data should be managed at an enterprise-wide level, including all lines of business, geographies and all functional areas. Given the recent spate of regulations demanding transparency and accountability and the stiff market and non-market penalties associated with non-compliance, business and government organizations would be well served to take the matter seriously.

All infrastructural areas, including people, process and technology, should be addressed. A data governance foundation, including policies, procedures, organization, roles and responsibilities, along with associated communication and training, are required to design, develop and provide ongoing support for the effort. New and important organizational structures and processes, including data stewardship, must be defined and operationalized within a set model or framework and executed consistently for maximum benefit.

Execution at the data-element and higher levels requires substantial work defining elements at source and ancillary downstream systems. Managing the movement of data across and through the organization requires an understanding of each element, each transformation, each migration and each source and target system. The use of common corporate definitions must become familiar in the corporate culture and reporting requirements built on common aggregation and disaggregation will reduce the complexity of understanding and managing the environment.
Among a number of other requirements, businesses need to understand:

- What data exists in the environment
- In what form it exists and how it is transformed from system to system across the enterprise.
- Where it comes from
- How one data element relates to another
- Where it resides
- How business terms and data elements should commonly be defined and used across the enterprise
- How it got there
- How and why a data element does or does not vary from the accepted and common corporate definition (if one exists)

It is not a question of whether data quality is poor; it is matter of degree, either in scope (the range of errors across the company and the number of instances) or in the impact of poor data quality on an organization. Variances occur in both the breadth and depth of issues, and in considering how to address those issues, actions can be taken to stop data from being corrupted further. Remediation activities will address pre-existing data quality issues, but addressing known issues is only part of the effort; investigative actions to address unknown issues can yield surprising results.
There are many partners and tools available to assist, but data governance requires a well-defined corporate discipline. A certain level of understanding and expertise is required, but subjectivity may be the most important challenge. Be aware that the effort surrounding issue identification, root cause analysis and remediation design and execution can become an exaggerated fault-finding exercise. Companies are facing serious considerations, after all, when they address federal regulatory actions based on improper governance.

Rather than limiting the effort to identifying issues and remediating problems, some players get tied up in trying to figure out who did what and assigning blame. While there are instances (think Enron, WorldCom) in which data errors, element definition and reporting differences are not benign, organizations need to focus on configuring new organizational elements, new roles and responsibilities, and new processes.

IBM Global Business Services uses the following approach to designing, implementing, managing and maintaining an effective data governance support structure:

1. Design an Operational/Organization Model that provides a high-level data governance organization and structure, including an overview of roles and responsibilities and processes required to fulfill those responsibilities.
2. Develop a Data Governance Management Manual (DGM) that outlines the phases and activities (description, objectives, roles, tools/aids, and tasks) required to execute the DGM methodology. The DGM methodology encompasses processes such as the enforcement of data governance policies and procedures, issue escalation and metadata management.
3. Create a Data Quality Management User Manual (DQM) that outlines the phases and activities (description, objectives, roles, tools/aids, and tasks) required to execute the DQM methodology. The DQM methodology encompasses processes such as data domain definition, scorecard construction and data quality issue identification and remediation.
The primary workstream of a Data Quality effort are shown above. There is significant detail behind each of these annotated elements. The detail includes flow charts, definitions, templates, RASIC charts, and more to provide for consistent deployment and learning. Descriptions of these workstreams are listed below.

**Data Domain Definition** – This phase focuses on the development of common business definitions for a data domain. It begins by defining concepts contained within a data domain and bringing different SMEs and stakeholders across involved LOBs to an agreement on a meaning and data elements for each concept area. (For example, a customer has a name, an address and a Social Security number.) A Data Quality team develops standards about the meaning and format of each critical data element, which eventually become part of the way data quality is measured in the enterprise. Throughout the process, stewards are identified to support and enforce program policies and standards.

**System Mapping** – This phase identifies systems that contain data for the domain defined in Data Domain Definition. Documentation helps to detail the data stored in each system, the criticality and uniqueness of the data, and its use. Then each critical element as defined through Data Domain Definition is mapped to the actual field instance in the system, so appropriate standards can be applied.

**Gap Analysis** – This phase focuses on the collection and discovery of data quality issues. Issues that are identified in Data Domain Definition and System
Mapping are collected and formally logged for the data governance organization. Steps required to profile data elements also are detailed.

Data Quality Remediation Strategy – This phase provides the steps and tools necessary to research a data quality issue, develop options to remediate it, and select the best remediation plan for the organization. At completion of the phase, the data governance organization will have an approved remediation plan to address specific data quality issues.

Remediate – The goal of the Remediate Phase is to prioritize, obtain funding and execute approved remediation plans. At the conclusion of this phase, data quality issues should be resolved to the level required by the business.

Monitor and Control – Once remediation plans are executed, data quality must be monitored and assessed continually. This phase examines the effectiveness of remediation plans on data quality and the data quality of important data elements.

As an enterprise-wide effort, any data governance project may be centrally designed, managed and funded as a portion of allocated overhead. If individual business units design independent strategies, the enterprise effort may be negatively impacted. And if management and funding are decentralized, the fate will be the same.

Figure 3: Budget Impact Assessment

<table>
<thead>
<tr>
<th>Critical Resources</th>
<th>Quality Ratio (Column A / Column B)</th>
<th>Annual $\ to Operate (Column A)</th>
<th>Annual $\ to Improve (Column B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employees</td>
<td>%</td>
<td>Salary, Benefits, Bonuses, etc.</td>
<td>Training</td>
</tr>
<tr>
<td>Products</td>
<td>%</td>
<td>Manufacturing, Distribution</td>
<td>Research &amp; Development</td>
</tr>
<tr>
<td>Customers</td>
<td>%</td>
<td>Sales, Marketing, Advertising</td>
<td>Customer Service</td>
</tr>
<tr>
<td>Equipment Systems</td>
<td>%</td>
<td>Capital Equipment Budget</td>
<td>Capital Maintenance Budget</td>
</tr>
<tr>
<td>Applications</td>
<td>%</td>
<td>Software, Development Staff</td>
<td>Quality Assurance Staff, Testing Software</td>
</tr>
<tr>
<td>Data</td>
<td>%</td>
<td>Information System less Capital Expense</td>
<td>Data Quality Programs</td>
</tr>
</tbody>
</table>

Source: The Data Warehousing Institute (TDWI)
No one is fond of hearing “We’re from the corporate office and we’re here to help you.” Divisional, department and LOB owners will find it difficult to allow external access to and deep inspection of (think audit) their data and processes. There also may be a general misunderstanding of the roles of information management or management information services, so responsibilities need to be stated clearly. As defined earlier, information is a corporate asset and needs to be addressed as such.

Consider an anecdotal alternative-management message that supports centralized funding with decentralized management:

“We know you have near-term business performance expectations and don’t have the specialized resources (regulatory, technology, process skill sets) required, but you need to transform the way your group gathers, defines and processes information, and make sure it aligns with all other business units and functional areas by next year. Use this special general ledger account for any expense to pay for the work.”

Alternatively, consider a management approach using decentralized funding and decentralized management:

“We know you have near-term business performance expectations and didn’t plan for this. We know your staff already has been reduced to meet competitive pricing pressures and you don’t have the specialized resources required, but we want you to transform the way your group gathers, defines and processes information, and make sure it aligns with all other business units and functional areas by next year. By the way, don’t miss your performance metrics.”
Clearly, those are sub-optimal approaches to a complex issue, but there is a positive aspect, which is the considerable value of managing and maintaining high quality data and enterprise-wide information services. Significant gains may be available to marketing, especially in areas of product development, product pricing and sales and service delivery. Considerable savings also may be available in operations and information management. Most important may be management’s new-found ability to truly understand the business, have insight into the accurate yield of various strategic options and the impact on specific organizational areas, customers and clients.

When considering data quality (complete, valid, timely, unique, accurate, precise and integrity), organizations may be struck by the element of integrity, which is defined as “ensuring consistency of and relevancy in the definition of a data element across systems or within the context of the environment or practice.” It also raises the issue of honesty. In The Naked Corporation, authors, Don Tapscott and David Ticoll suggest that stakeholders now expect nothing less than a relationship based on a trust bound by four values: honesty, accountability, consideration and transparency. Those elements provide the context for what Tapscott and Ticoll call “the New Business Integrity and the foundation of the open enterprise.” How then, can companies meet that new and admittedly higher level of acceptability for themselves, their employees, their customers and their shareholders?
The need for change is evident, but it is not enough for an organization to build systems, establish policies, create task teams and standing review committees. Change and change management is at the core of effective data governance. The practice will deliver a general impact, but it also requires specific and individual change. Employees need to do things differently, and accomplishing that on a corporate-wide scale is no small feat. Do not assume that everyone will understand or that everyone will comply. Build new expectations into job descriptions to provide for minimum compliance. Providing incentives for exceeding minimum compliance levels will enhance the speed and ease with which an organization embraces and benefits from effective data governance.

Beyond the benefits of meeting regulatory compliance requirements, effective data governance also provides a solid foundation for enterprise business intelligence and business performance management. Follow the implementation of data governance with enterprise-wide strategic planning, locating new pools of previously unknown profits and tracking performance with enterprise-wide dashboards and scorecards. Leveraging established tools and practices such as traditional balanced scorecards and Six Sigma methodologies can turn a company’s run-of-the-mill, previously non-compliant, capital-restricted company into a Wall Street darling.

Considering the options, it may be worthwhile to hearken back to Apollo 13 and consider the lead flight director’s announcement to Mission Control: “Failure is not an option.”
About the Author

Brett W. Gow is the financial services data governance leader for IBM Global Business Services. His email address is bgow@us.ibm.com.

Contributing Authors:
Andrea Castellow, castellow@us.ibm.com
Scott Kimbleton, kimbleton@us.ibm.com
Doris Saad, DorisSaad@us.ibm.com
Rajesh Srivastava, srajesh@us.ibm.com
Paul Tomko, ptomko@us.ibm.com
Bruce Tyler, bruce.tyler@us.ibm.com
Bao Vuong, baovuong@us.ibm.com

End Notes

1 Wayne W. Eckerson, The Data Warehouse Institute, Quality and the Bottom Line, The Data Warehouse Institute (TDWI) Report Series, 2002
3 According to the World Bank, the Gross Domestic Product for the U.S. in 2005 was about $12.4 trillion according to the World Bank. The Gross Domestic Product for the U.S. in 2005 was about $12.4 trillion according to the World Bank; reference http://siteresources.worldbank.org/DATASTATISTICS/Resources/GDP_PPP.pdf
4 The Data Warehouse Institute, Data Integration: Using ETL, EAI, and EII Tools to Create an Integrated Enterprise, 2005
6 Quote “An Act To protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws, and for other purposes.” Sarbanes Oxley, Public Law 107-204; reference http://frwebgate.access.gpo.gov/cgi-bin/useftp.cgi?fIPaddress=162.140.64.88&filename=publ204.pdf&directory=\diskc\wa\data\107_cong_public_laws