Creating Value Through Data Integrity

A Pragmatic Approach

Guy Gilliland, Sue Kolloru Barger, Vikrant Bhatia, and Ron Nicol

August 2011
Every company that works with large, complex, fast-changing sets of data requires a holistic data-integrity program in order to extract the most value from its data. The integrity of data determines whether or not a company can achieve cost savings, optimize investments, and gain competitive advantage, as well as how effectively its executives can make critical decisions.

**A Pragmatic Approach**
A successful data-integrity program must be premised on the fact that not all data are of equal value, and it must consider all data problems within a business-oriented context.

**Find, Fix, Prevent**
We take a three-step approach to improving data integrity. First, find the most critical problems. Second, fix the problems by analyzing their root causes. And third, prevent their recurrence by developing the company’s ability to ensure the integrity of its data.
Imagine an industry in which the typical company spends billions of dollars each year on the assets required to run the business. Then add a pace of change so rapid that revenues double every few years, new products are launched every two years, and a large percentage of the company’s assets must be reconfigured every year. Now imagine trying to operate such a business effectively when the information it depends on is only 50 percent accurate. Welcome to the world of the typical executive at many telecom operators and IT services companies, among many other types of business.

In the telecom sector, companies typically spend 20 to 25 percent of their annual revenues on capital investments in their networks. The typical mobile operator serves millions of subscribers, with thousands of circuits connecting cell sites to its network switches. Each cell site contains multiple network elements, and the switch sites and central offices require even more complex equipment. All these various elements are busy aggregating voice and data traffic onto the network.

Things get really scary when you consider that the average telecom network’s “data integrity” is only 40 percent—that is, the records of what actually takes place in the network are correct less than half the time. On top of that, the number of installation and change orders in the typical network exceeds what is typically forecast by 30 percent per year. For a telecom executive facing budget pressures to maintain or streamline operating costs, including manpower, while continuing to meet the regulatory requirements for both regular and 911 phone service, this is the definition of a nightmare.

IT services executives face an equally daunting challenge. The typical large data center contains more than 5,000 servers, 200 terabytes of storage, and potentially several mainframes. Add to that myriad load balancers, routers, switches, and firewalls. If the data center is client facing, then 60 to 70 percent of the equipment may be owned by clients and were likely built by many different OEMs.

Then there’s the software required to run every client’s enterprise applications effectively. Managing this infrastructure may require as many as eight people, all with different skill sets, to provision a single server, plus one person per 50 servers just for day-to-day management. This quite typical IT services company must maintain an accurate inventory not just of the equipment it has and where it is located, but also of how it is configured and what applications it is running. Now imagine all these challenges in an environment where all the servers are virtualized.
The Data Challenge

It isn’t just telecom and IT services companies that depend on successfully managing enormous and complex sets of assets—frequently with largely incomplete or inaccurate information about those assets. Every company with a large and complex set of assets that experiences a rapid rate of change must deal with these issues. It’s a perfect storm for the creation of chaos.

Companies in every industry have struggled—and sometimes failed outright—as a result of poor data access and management, an inability to translate data into valuable information, and poor data quality. Health care companies experience a huge amount of fraud, abuse, and waste, much of it preventable through improved data analysis. Financial services institutions are constantly rolling out new ways for customers to connect via all kinds of devices, yet they still know very little about those customers—their needs, how profitable they are, or how to target and service them as effectively as possible. The utilities industry can monitor power consumption at the level of the individual customer, yet it has few tools to help control demand or cost. Insurance companies, too, are only now starting to understand how to use data to create better models for claims management and fraud protection.

Yet there are companies that have built a competitive advantage specifically on their ability to manage this kind of information comprehensively, adroitly, and with speed—and many others that recognize the need to do likewise. Bharti’s network management and deployment based on traffic forecasts, eBay’s ratings system, Cisco’s daily sales reporting, BNSF’s logistics systems, Bloomberg’s company ratings, American Express’s credit-card-transaction reporting, Tesco’s supply-chain and customer-relationship-management systems, Gazprom’s pipeline system—all are completely dependent on the data these companies gather, analyze, and report on daily, and in some cases, instantaneously.

Every executive, therefore, needs to be able to answer two key data-management questions:

- Are data integrity problems inhibiting my ability to manage my business?
- If so, how can I address this problem effectively?

Lots of companies gather enormous amounts of data every day. But data gathered for data’s sake will never be of any real use. The key is to look at the problem of data integrity from the point of view of the value to be found in the data.

Sources of Data Integrity Problems

Our experience has shown that the size, scope, and complexity of a company’s databases, and the velocity at which the data change, are the key drivers of difficulty in managing data. Companies with a large number of product SKUs, detailed billing information, significant product customization, or rapid changes in the location, configuration, and capacity of their assets are likely to suffer from data integrity problems.
In addition, most competitive organizations are under intense daily pressure. Often, too much is going on too quickly, and poorly defined data-management processes are not functioning as required. Just as often, organizational structures and lines of accountability are not well defined, and employees in different parts of the organization can become confused as to their roles and responsibilities regarding the collection and use of data. Moreover, a lack of trust in the data's integrity can lead to the inconsistent use of data and data systems. Finally, the data systems themselves may be the cause, owing to overly flexible data definitions, lack of standardization, or the lack of a proper audit trail for keeping track of data changes.

The impact of these data-integrity issues can be felt throughout an organization. Product development and engineering groups may find themselves with unutilized or underutilized network capacity and network elements. Product manufacturing and service delivery groups may face delays, avoidable order issuance, and unnecessary penalties. Service maintenance units may have to deal with increased work orders and jobs, unnecessary work, longer time to dispatch for work, and delays in resolving problems.

**Best Practices**

The value of business data is greater in some circumstances than in others. We have identified five critical areas in which the quality, completeness, and accuracy of data can have a significant effect on how businesses operate. (See Exhibit 1.)

<table>
<thead>
<tr>
<th>EXHIBIT 1</th>
<th>Data Integrity Is Critical to a Company’s Operations</th>
</tr>
</thead>
</table>
| **Products** | • Concept development  
  • Competitive positioning  
  • New-product launch  
  • Is this a concept worth pursuing?  
  • Do our products have a competitive advantage?  
  • Should we launch this product in this way? |
| **Channel/go-to-market plan** | • Demand  
  • Pricing  
  • Services  
  • How will demand for a product change in the future?  
  • How should a specific product be priced?  
  • Do we have the appropriate service offerings to support our products? |
| **Assets** | • Management  
  • Deployment (capital expenditures)  
  • Can we streamline our cost structure?  
  • Can we redeploy assets? Streamline capital expenditures? |
| **Supply chain** | • Product customization  
  • Should we meet our customers’ request to customize this product? What is the cost of complexity? |
| **Human resources** | • Resource deployment  
  • Should we outsource management of this product?  
  • Are we deploying the right people in the right place/with the right projects at the right time? |

Source: BCG analysis.
integrity of the data will determine whether or not a company can achieve cost savings, optimize investments, and gain a competitive advantage, as well as how effectively its executives can make critical decisions.

Companies that excel at managing their data tackle these areas by thinking in terms of the value of the data in each context and then instilling the ability to understand and manage data in terms of that value. This is best accomplished in a three-stage effort:

1. Establish clear process definitions and create a detailed data model that is closely linked to the business model and underlying processes.

2. Make clear who owns and who takes care of the data, and devise processes for measuring, monitoring, and reporting on the data's accuracy and completeness.

3. Create feedback mechanisms to rectify errors that inevitably creep into every data system and process.

As critical as these best practices are, they cannot be implemented until all data-integrity issues have been addressed. There are six key dimensions of data integrity, each of which brings with it challenges that every company faces in improving the quality and completeness of its data. (See Exhibit 2.)

### EXHIBIT 2 | The Six Dimensions of Data Integrity

<table>
<thead>
<tr>
<th></th>
<th>Acid-test questions</th>
<th>Common challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accessibility</strong></td>
<td>Are the data easily accessible?</td>
<td>Unstructured/unreadable data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inaccessible source systems</td>
</tr>
<tr>
<td><strong>Completeness</strong></td>
<td>Are the data complete?</td>
<td>Inconsistent definitions/business rules</td>
</tr>
<tr>
<td></td>
<td>Are critical data fields populated?</td>
<td>Poor validation</td>
</tr>
<tr>
<td><strong>Correctness</strong></td>
<td>Does the format correspond to what is expected?</td>
<td>Inaccurate source systems</td>
</tr>
<tr>
<td></td>
<td>Is the format consistent (similar fields, different systems)?</td>
<td>Cross-reference tables not well maintained</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rounding errors, poor validation</td>
</tr>
<tr>
<td><strong>Timeliness</strong></td>
<td>Are all elements up to date?</td>
<td>Infrequent source systems synchronization</td>
</tr>
<tr>
<td></td>
<td>Are the data being produced and delivered at the right time?</td>
<td>No transaction time-stamping</td>
</tr>
<tr>
<td></td>
<td>Are the data compared and aggregated?</td>
<td>Long calculation times</td>
</tr>
<tr>
<td><strong>Synchronization</strong></td>
<td>Do proper data links exist?</td>
<td>Inconsistent or missing links</td>
</tr>
<tr>
<td></td>
<td>Are the data relationships consistent?</td>
<td>Data independence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of parent/child corporate links</td>
</tr>
<tr>
<td><strong>Usage</strong></td>
<td>Are data effectively used?</td>
<td>Unused data becoming stale</td>
</tr>
<tr>
<td></td>
<td>Are data still relevant?</td>
<td>Irrelevant data a source of errors</td>
</tr>
</tbody>
</table>

Source: BCG analysis.
A Pragmatic Approach

In our work with clients on data integrity projects, we have found it critical to avoid attempting a one-off, “grand IT” solution or beginning with a full-scale data clean-up. These approaches rarely work and are often out of date by the time they are deployed. In the context of data integrity, trying to solve world hunger all at once is a fool’s errand. Instead, companies are most successful when they adopt a practical approach to improving their data-integrity and data practices—an approach that is focused on value.

We believe that a successful data-integrity program has the following four overarching characteristics:

1. There is a recognition that not all data are of equal value to the entire company, but that different pieces of data may be critical to the work of individual employees.

2. There is a recognition that the impact of “bad data” must be understood in a business context, in terms of wasted operating and capital expenses, unproductive labor, service interruptions, and impact on customers.

3. The effort takes a programmatic, business-case-driven approach in which the impact on value is always the key driver.

4. It is a holistic effort to sustain improvements in data quality that is synchronized with the data cleanup effort.

Our approach, which we call “find, fix, and prevent,” stems from our focus on business value in data management. It involves finding the data problems that have the greatest effect on a company’s ability to reap the benefits of its data, developing a fix based on an assessment of the root causes of the problems, and implementing a robust solution and capability to prevent a recurrence.

Finding the root causes of data integrity problems is a two-step process. It involves, first, assessing the owners, creators, editors, and users of critical data fields and, second, identifying the likely root causes of any data errors that are occurring. Then we develop a “surgical” plan to prevent the data integrity problems from recurring, followed by a pilot program and ultimately by a broad program rollout.

Our approach is strictly pragmatic. Every data-integrity program we execute is based on a well-defined business case. We carefully track the ongoing success of the business case so that all captured savings can be “booked.” The savings that each program generates are computed on the basis of a clear rationale linked to the specifics of the program.

The benefit of this approach is that it generates value in two ways. First, it creates value directly by targeting the most critical data. Second, it builds the organizational capability to build and use a more robust and accurate set of data to run the business and build competitive differentiation.
We identify the errant data and the correct source of data, upload the correct data into the production database, and develop a “prevent” solution to inhibit the recurrence of the data error. Our prevent solution is based on a surgical set of changes that address the root causes of bad data. The solution is based on six pillars, each of which must be defined and implemented to sustain a differentiated data-integrity capability. (See Exhibit 3.)

Data Integrity in the Telecom Industry

The physical and logical networks of telecom operators are enormous, and the databases underpinning them are equally immense. Some operators we have worked with run key databases consisting of more than 100 million records and more than 1,500 data fields spread across hundreds of tables—although the number of truly critical data fields is typically just 100 or so.

Given this level of complexity, it comes as no surprise that a significant subset of telecom service outages are a direct result of problems with data integrity, the majority of which are “self-inflicted.” Moreover, such outages take 2.5 times as long to resolve, on average, than those not related to data issues. Reducing these outages can have a major impact on operations: fewer outages directly correlate with reduced churn and increased customer satisfaction. (See Exhibit 4.)

Another common problem among telecom companies can occur when a large number of stakeholders from all over the company have the ability to input and change data. A critical step in ensuring the accuracy of the data is, first, to

---

**EXHIBIT 3 | BCG’s “Prevent” Solution Is Based on Six Pillars**

<table>
<thead>
<tr>
<th>Data ownership</th>
<th>Data accountability</th>
<th>IT system solutions</th>
<th>Process reengineering</th>
<th>Metrics and scorecards</th>
<th>Organizational update</th>
</tr>
</thead>
</table>
| • Field prioritization  
• Key field owners and responsibility mapping  
• Master data for fields | • Owners/stakeholders responsible for field integrity and accountable for process adherence | • System and other automation enhancements | • Root cause issues  
• Key process-reengineering solutions | • Process for measuring and monitoring, with scorecards and metrics  
• Organizational roles and responsibility for data management | Cross-functional effort across the organization |

Source: BCG analysis.
identify both the stakeholders who originate data in key data fields and those
who are affected by that data and, second, to consolidate the ownership rights.
(See Exhibit 5.)

We have worked with numerous telecom clients to solve just these kinds of prob-
lems. In one case, a U.S. telecom provider had undergone significant network
growth and rapid expansion over the previous few years. But that growth resulted
in a number of immature processes:

- Too much was going on too quickly, and employees were overwhelmed.
- The databases were being used inconsistently, and employees were losing trust
  in the competency of the data.
- Many end-to-end processes were not clearly defined, and others were entirely
dysfunctional.
- The organizational structure was weak and the lines of accountability unclear,
  which led to confused roles and responsibilities.
- Problems with the system were arising as a result of poor system structure,
lack of standardization, and the lack of a clear audit trail when changes were
made.

We supported the client by developing and executing an actionable program
designed to have a major impact on business operations. First, we established a
Creating Value through Data Integrity

baseline understanding of known problems, assessed the business impact of data errors, and identified the key “pain point” data fields. Then we created programs to address these problems, prioritizing them on the basis of several critical factors and then executing them. Finally, we ensured that the data would stay clean by executing our prevent solution and helping the client acquire the tools and capabilities it would need to maintain its newly won data integrity.

Overall, we helped the client create significant value in five critical ways:

1. Reduced operating costs by eliminating wasteful spending (typically 3 to 5 percent of the cost of access to third-party and external carriers) and by increasing and releasing unused capacity in the client’s networks (typically by 5 to 10 percent).

2. Deferred capital investments through increased transparency into existing network inventory (typical savings of 2 to 10 percent).

3. Decreased network outages (typically by around 20 percent), dispatches of service trucks, time to dispatch resources, time to resolve outages, and time to resolve data integrity problems (typically by more than 50 percent).

4. Increased workforce productivity by streamlining work processes and improving the accuracy of the data required to get work done (typically by around 15 percent).

EXHIBIT 5 | Most Data-Field Ownership Rights Should Be Consolidated

Source: BCG analysis.
5. Most important, increased customer retention and satisfaction (churn reduction, the impact of which is critical but difficult to measure).

Our client captured the added business value from the program quickly—much of the payback came within 12 to 18 months or less, depending on the type of savings. (See Exhibit 6.)

Success in improving data integrity depends on taking a holistic and practical approach to the effort, one that is focused on people, processes, and systems. It is critical to focus first on value creation and then on data integrity, prioritizing each stage of the project according to the data integrity issues that will have the greatest impact on value. The effort to capture the full value of a company's data depends not just on strong analytics capabilities but also on a carefully planned change-management program, including the ability to handle cross-functional issues and organizations and to navigate political tensions that can inhibit true long-term data-integrity solutions.

The data that companies collect, maintain, and use are—and increasingly will be—vital to producing business value and creating real competitive differentiation. Only by maintaining the accuracy, completeness, and timeliness of data—that is, overall data integrity—can companies sustain the critical advantage that a strong data program can provide.

**EXHIBIT 6 | A Data Integrity Program Can Capture Substantial Savings**

**Business case estimate**

- **Increased customer retention and satisfaction (churn reduction)**
  - Improved network quality
  - Reduced outages
  - Reduced dropped calls
  - Increased customer satisfaction
  - Terminating or altering cost of access circuits, roaming charges
  - Optimizing and changing circuit costs
  - Deferral of network equipment capital expenditure
  - Redevelopment or reuse of existing element ports

- **Workforce productivity improvement**
  - Reduced frequency and duration of trouble tickets, truck rolls, outages
  - Ancillary impact on productivity, revenues, and churn
  - Improved workforce efficiency
  - Reduced overtime

- **Reduction in network outages, mean time to resolve, and mean time to dispatch**

- **Deferred capital spending on equipment**

- **Reduction in access costs/roaming and release of unused capacity**

**Source:** BCG client example.
About the Authors
Guy Gilliland is a partner and managing director in the Dallas office of The Boston Consulting Group. You may contact him by e-mail at gilliland.guy@bcg.com.

Sue Kolloru Barger is a principal in the firm’s Atlanta office. You may contact her by e-mail at kolloru.sue@bcg.com.

Vikrant Bhatia is a partner and managing director in BCG’s Dallas office. You may contact him by e-mail at bhatia.vik@bcg.com.

Ron Nicol is a senior partner and managing director in the firm’s Dallas office and the global leader of the Technology, Media & Telecommunications practice. You may contact him by e-mail at nicol.ron@bcg.com.

Acknowledgments
The authors would like to thank Rob Trollinger, Howard Wilkins, Kirsten Allerdt, and Edward H. Baker for their contributions to this report, as well as BCG’s editorial and production team, including Kim Friedman, Gina Goldstein, and Kirsten Leshko.

The Boston Consulting Group (BCG) is a global management consulting firm and the world’s leading advisor on business strategy. We partner with clients in all sectors and regions to identify their highest-value opportunities, address their most critical challenges, and transform their businesses. Our customized approach combines deep insight into the dynamics of companies and markets with close collaboration at all levels of the client organization. This ensures that our clients achieve sustainable competitive advantage, build more capable organizations, and secure lasting results. Founded in 1963, BCG is a private company with 74 offices in 42 countries. For more information, please visit www.bcg.com.

For a complete list of BCG publications and information about how to obtain copies, please visit our website at www.bcg.com/publications.

To receive future publications in electronic form about this topic or others, please visit our subscription website at www.bcg.com/subscribe.

© The Boston Consulting Group, Inc. 2011. All rights reserved.
8/11