Transforming enterprise information integrity

White Paper

Susanne Ruschka-Taylor, partner, IBM Business Consulting Services
Contributors: Christopher Evask, Piyush Malik, and Stephen Minsinger
Introduction

Say goodbye to the data age and hello to the information age, where leading-edge organizations treat their data as a critical resource and harness its power to generate information that enables them to stay ahead of the competition. In this era of on demand information, organizations generate, manipulate, manage and exchange more data than ever before. In addition, they continue to make significant investments in the development and deployment of technology that facilitates the collection, distribution and analysis of data. But where is the information? The old maxim — garbage in, garbage out — still applies. The data may cost more, be accessed faster and be better presented, but it's still garbage. The CEO sends back the numbers because they "look wrong;" the "spreadsheet jockeys" work through the weekend because nothing is ever as it appears; and somewhere a manager makes a bad decision — one that looks good on paper — and never realizes that things didn't work out. Sound familiar? Data sits at the bottom of the information hierarchy, yet without investing in its quality, no matter how good the tools and techniques, information will remain elusive, and the competition that has mastered its data foundations will surge ahead.
Understanding information integrity

Information integrity is a product — the result of people, processes and technology. See if the following hypothetical comments sound familiar:

• An IT director of a large corporation has this to say when his company was considering throwing out its enterprise resource planning (ERP) system and starting over because it wasn't working: “We converted bad data — I know I know, but that's what happened. Look, we are trying to run a business here, the implementation deadline loomed and no one wanted to hear that the data was bad, so we flagged it and converted anyway.”

• A project manager at a large service organization relays this story: “We spent US$50 million on 150 resources scrubbing the data full-time for four years and spent twice that amount on a new ERP [system] that was going to solve all of our information integrity problems. Two years later, it is even worse than when we started. How much more do we have to spend to make this problem go away?”

In IBM’s experience, over 50 percent of data corruption issues arise from sources other than technology. In fact, anything that causes an imbalance in the finely-tuned interrelationships between people, process and technology gives rise to poor information integrity. Furthermore, as anyone who has had to repeatedly correct the same data set can attest, achieving sustainable information integrity is not always as easy as it appears, particularly when an organization is large and complex. A company’s cultural approach to information integrity can also be a challenge. Consider the following account from a project manager at a large global organization:

“We know why our information integrity is poor — over the past five years, we have acquired a number of companies, and they don’t want to play by our rules. So by the time our data gets consolidated at the corporate level, we have apples, oranges and cucumbers. But how do we fix it? We can’t fire everyone.”
The moral of the story: fixing your data foundation requires a systematic, integrated approach that leverages the short-term gains of remediation where applicable, but looks beyond the quick-fixes to build an environment of people, processes and technology that will foster sustainable information integrity.

**What is information integrity?**

Information integrity is defined by how effectively data supports the transactions and decisions needed to meet an organization’s strategic objectives, as embodied in its ability to manage its assets and conduct its core operations. The level of information integrity required to effectively support operations will vary by product line or functional area, depending on their unique information requirements. For example, a finance department will require high-quality data, while a marketing database may be able to achieve its objectives with less stringent information integrity requirements.

Information integrity can be measured across multiple dimensions, each of which is used to gauge how well a data element meets a company’s information integrity goals, and ultimately, its information needs.

The degree to which the data conforms to its relevant dimensions dictates the level of quality achieved by that particular data element.
## Information integrity dimensions

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
<th>Example of poor information integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness</td>
<td>The data element is populated when required</td>
<td>Customer’s first name is missing</td>
</tr>
<tr>
<td>Validity</td>
<td>The data element contains a valid data format or domain value</td>
<td>State code is “ZZ”</td>
</tr>
<tr>
<td>Consistency</td>
<td>The data element is consistent with other related data or business rules</td>
<td>Client’s birth date is 01/01/1990 and retirement date is 01/01/2005</td>
</tr>
<tr>
<td>Uniqueness</td>
<td>The data element is unique — there are no duplicate values</td>
<td>Customer number used twice in information file</td>
</tr>
<tr>
<td>Accuracy (electronic)</td>
<td>The data element agrees with validated data from another source</td>
<td>Address does not match with post office database</td>
</tr>
<tr>
<td>Accuracy (real)</td>
<td>The data reflects the real-world object, event or transaction</td>
<td>Address does not match client’s application form</td>
</tr>
<tr>
<td>Precision</td>
<td>The data element holds data at sufficient granularity</td>
<td>Product costs stored as dollars, not dollars and cents</td>
</tr>
<tr>
<td>Accessibility</td>
<td>The data element is accessible when required</td>
<td>Birth date is in a source system, but not in the centralized customer information file</td>
</tr>
<tr>
<td>Timeliness</td>
<td>The data element is available when it is required</td>
<td>Data uploaded to data warehouse monthly, not daily</td>
</tr>
<tr>
<td>Clarity</td>
<td>The data element is clearly defined and understood</td>
<td>Customer record has free-text “order status” field, so the field cannot be analyzed</td>
</tr>
<tr>
<td>Sufficiency</td>
<td>The data element contents are sufficient without making assumptions</td>
<td>Order status “delayed” is used to identify out-of-stock items (because it is assumed that all delayed items are out of stock)</td>
</tr>
</tbody>
</table>
IBM experience indicates that information integrity issues come to the surface when new initiatives increase an organization's dependence on data. Imbalances in the interrelationship between people, process and technology are exposed by corporate initiatives such as:

- Customer relationship management (CRM) and ERP implementations
- Outsourcing
- New compliance requirements, such as Sarbanes-Oxley, Basel II and the USA PATRIOT Act
- Business intelligence and data warehousing
- Performance improvement
- New delivery channels (e.g., self-service)
- System upgrades
- Cost reduction
- Joint ventures, mergers, acquisitions and divestitures.

Without a foundation of information integrity best practices, such as those outlined in the information integrity framework described later in this paper, your organization may be ill equipped to handle these imbalances and find itself struggling to perform with the ensuing poor information integrity.
High-quality data optimizes effective, and by association, operational effectiveness. Tangible benefits of data quality are reflected in cost reductions and improved business metrics. A simple, yet powerful, example of the tangible costs of poor information integrity relates to direct marketing. If the customer data is cleansed and categorized by household, for example, costs of targeted direct marketing can be significantly reduced by sending out one mailer per household rather than multiple mailers to individual customers within the household. Poor-quality data creates costs in the forms of inefficient processes and costs to clean or resource data, fix applications and repeat analyses.
Direct costs related to poor information integrity can occur when an organization is not able to leverage its investments in technology. A classic example is the scenario in which a company makes an expensive investment in a new data warehouse only to scrap the project, because poor data quality makes it unusable. Another increasingly common occurrence is that of the accounts payable or accounts receivable department run amok — unable to collect from customers or pay invoices due to poor data. Finally, not to be discounted are the human capital and technology costs related to remediating data discrepancies.

Less-tangible and intangible costs

One reason that information integrity has not been top of mind for senior management is that many costs are not readily translated into hard dollars. They are less-tangible or intangible, but many of these costs — when left unattended — become measurable.

How do you quantify the cost of attrition in a customer base due to the inability of a database to communicate with a marketing database in order to relay address changes? How can you easily quantify the nature and extent of process inefficiencies that result from the poor morale that develops when managers operate blindly because nothing makes sense? Or because it takes too long to transform data into information? How do you put a price on customer dissatisfaction or the loss of faith in a new system? How can you quantify the cost of risk reduction or avoidance, or the impact of reduced time-to-market or inventory carrying costs? What is the true dollar impact of regulatory-penalty avoidance?
The less-tangible and intangible costs associated with poor information integrity are very real, but require more effort and diligence to identify, quantify and monitor. For a CRM system, poor information integrity may hamper the ability to accurately segment and predict aggregated patterns of customer behavior based on past trends. This can have serious repercussions when calculating customers’ lifetime value or delivering targeted offers to customers, and adds to costs as a lost opportunity.

In addition, poor-quality data can lead to bad business decisions related, for example, to customer parameters such as insurability and credit-worthiness, resulting in the associated costs of increased claims, loan defaults and customer fraud. Many companies share data with their partners and suppliers. Sharing inaccurate data with suppliers might cause overstocking or shortages during important production cycles. Sharing inaccurate data with partners could impact sales by hampering the ability to identify and follow sales leads and could impact product development if the right teams don’t get accurate information from their partner channel.

These days, information is a corporate asset whose deteriorating quality will have an adverse impact on business, ultimately affecting the bottom line. The repercussions of poor information integrity are widespread and significant.
Benefits of good information integrity

How many times have you heard the adage, “You have to spend money in order to make money?” In the realm of data quality, nothing could be more true, and the benefits are very real.

Harvesting information integrity benefits requires an investment in an enterprise information integrity framework. The return on investment is high compared to the direct and indirect costs related to poor data quality. Information integrity projects commonly have a 12 to 24 month payback cycle. In addition, an initial investment in an enterprise information integrity framework can be leveraged across multiple systems, processes and organizational units, further compounding the initial return on investment.

Information integrity will be the critical factor underlying corporate success stories in the information age. Fixing your data foundation requires a systematic, integrated approach that leverages the short-term gains of remediation where applicable and looks beyond the quick fixes to build an environment of people, processes and technology that fosters sustainable data quality.
Roadmap to information integrity

Enterprise information integrity framework

The IBM enterprise information integrity framework recognizes that information integrity is not solely a technology issue, but that it arises in equal measures from process and organizational issues. It endeavors to achieve and sustain data quality by addressing organization, process and technology, and embodies both technical and organizational environments so that it can be applied to manage information integrity in a wide variety of situations.

The IBM enterprise information integrity methodology captures the process for designing and implementing the IBM enterprise information integrity framework. This enterprise information integrity program is based upon IBM’s experiences using an established methodology, applied to client data remediation and information integrity sustainability engagements in different functional areas and in organizations, such as yours, around the world.
The information integrity framework highlights the various areas that must be managed in order to achieve sustainable information integrity. The framework is used to describe what a current or future business environment looks like, but does not identify how the framework is developed or how an organization would make the transition to a new environment described in a framework. The information integrity methodology describes how you can tailor and implement the framework to help achieve and sustain information integrity.
Information integrity policy

An information integrity policy helps an enterprise to develop a strategic direction for the governance of its information assets, and directs, standardizes and safeguards the corporation’s information assets. An information integrity policy enables an organization to recognize that corporate data is a critical resource and should be managed as such. It helps an enterprise articulate corporate objectives with respect to information integrity and serves as a blueprint for what must be done, by whom, when, why and how it will be measured. At the enterprise level, information integrity policies focus on the best practices and strategy for each of the remaining information integrity framework elements, comprised of organization, architecture, administration, procedures, measurement, reporting and checkups. A clearly articulated enterprise-wide information integrity policy provides the framework for decision-making when various business groups, member companies and functions interpret their data world differently and hold conflicting views about information integrity.

Information integrity organization

Information integrity organization focuses on the combination of efforts of different people, at different levels, from different parts of the company, in such a way that the objectives of the information integrity program are achieved. Data organization focuses on empowering employees by verifying they are:

- Able — properly trained and provided with the necessary tools, assistance and information
- Willing — motivated to do things correctly and rewarded for a job well done
- Allowed — authorized to drive change and supported by management.
Components of information integrity organization are:

- Roles and responsibilities — Assignment of roles and responsibilities for performing particular activities for specific groups or individuals. Typical roles of an enterprise information integrity program include data steward, information integrity governance council, chief information integrity officer, data architect, data administrator, database administrator, business expert, business user and information integrity analyst

- Structure and work group design — Allocation of work activities, reporting relationships and the combination of resources needed to meet information integrity goals

- Job design — Assignment of groups of related tasks/activities and procedures to specific roles

- Skills and behavior development — Development of skill sets and behaviors required to support new processes at the individual and group levels

- Performance management — Development and tracking of key success indicators for individuals, groups and the overall business with respect to information integrity

- Communication — Education of everyone in the organization about the importance of information integrity, and creation of an awareness and understanding of the impact individuals have in the data value chain.
Transforming enterprise information integrity

It is important to build or transform the information integrity organization in a way that leverages existing organizational structures, roles and communication channels, as opposed to creating a new layer of bureaucracy or excessive complexity. Augmenting existing responsibilities or realigning existing resources and reporting lines to support tighter coordination is typically required. As for essential leadership roles, changes most companies face include establishing a senior governance council and chief information integrity officer, which together drive the information integrity strategy and performance targets, and align the efforts of all individuals engaged in the information integrity program. The chief information integrity officer, who, depending on scope, may be a dedicated program manager, also develops and manages the methods and tools that verify information integrity practices are consistent across the enterprise. Overall, the primary organizational changes include formalizing these and other key information integrity roles and building community, skills and linkages among people in those roles.

Information integrity data administration

Poor information integrity often arises as a result of an unclear understanding within an organization as to the meaning and usage of data elements, as well as a lack of standardization of those elements. Information integrity data administration involves envisioning a data administration framework that is applied consistently across the organization, verifying that data is standardized and that all stakeholders can readily access information about the current meaning and use of the data. Although current technological developments such as metadata repositories may serve as an enabler, the key to successful information integrity administration is the design of organizational structure and processes which help ensure that business events which cause changes to reference data can be proactively captured, documented and distributed. The key elements of information integrity administration are the metadata model, business rules and the data model.
The key elements of information integrity administration are the metadata model, business rules and the data model.

- Metadata — Documenting definitions for data elements helps build a solid foundation for the enterprisewide metadata strategy. To develop an enterprise data model, it is important to capture business entities and their relationships, as well as data elements. Valid domains for each key data element that identify the value ranges and types must also be defined.

- Business rules — Business rules are policies governing business actions. A business rule describes how to make transition from one state to another or how to prohibit such a transition. Business rules govern updating, creating or deleting objects and generally result in constraints or integrity parameters on data relationships or values. Since they tend to be volatile or subject to frequent change, there must be a single source where business rules are documented and distribution is controlled. This helps prevent different business areas or application development teams from creating inconsistent business rules that govern the same activity.

- Data model — Data design calls for understanding the business requirements and translating them into a data model. A data model can be conceptual, logical or physical. Experience proves that source or target systems with suboptimal designs can create data errors. As companies rush to deploy new systems, developers often skirt fundamental design and modelling principles, which can lead to future data integrity problems. While a conceptual data model captures the business at a very high level, it provides an overview of the business entities in the scope of any system. A logical data model captures the entities and their relationships and cardinality, as well as rules governing the establishment of relationships. The business rules recorded in the logical data model and supporting documentation need to be incorporated into the physical designs (e.g., physical data model and the physical database design). Lack of attention to the business rules results in an inability to meet business requirements and could potentially cause information integrity issues. The physical data model addresses the alignment of the data with the metadata model by implementing data type constraints, domain rules and business rules within the data administration framework.
Information integrity architecture

The information integrity architecture component of the framework is composed of data architecture, technology and infrastructure, and is closely linked to information integrity administration. It includes key information integrity best practices, such as verifying that:

• Data is electronically validated at all entry points
• Systems sharing common data elements are linked so that they always share the most current view of those data elements
• All data is held in a staging area where information integrity processes can be applied to it before it is populated into operating systems
• Data redundancies are eliminated or streamlined through good architecture
• Systems development methodologies incorporate data quality best practices.

Strong data architecture results from robust data design practices and enforcement of standards. This contributes significantly to the overall improvement of information integrity across the enterprise, meeting the core objective of the framework. As with information integrity administration, it is vital to link information integrity architecture directly to the change management process.

Information integrity process

Information integrity process focuses on activities and tools used across the enterprise to support the consistent management and modification of data as an asset across the information value chain. Information integrity processes are aligned with, and integrated into, the organization’s other business processes and technologies to ascertain information integrity controls are in place. The information value chain maps out the generation, management, manipulation, assessment and transfer of data (both electronic and paper based) to achieve a business outcome.
A complete information value chain is made up of the following procedural components:

• Delivery processes — The activities and tools required to achieve a quality business output
• Quality flagging processes — The activities and tools required to record the quality of the business output
• Remediation processes — The activities and tools required to correct the business output in order to achieve sufficient quality.

Information integrity procedures can be augmented by information integrity tools, which fall into two general categories:

• Information integrity assessment tools — Provide an organization with a quantified view into the current state of its data. Without this information, it is difficult to target specific areas for data correction and accurately forecast the duration and amount of resources required for the effort. Once the extent of the information integrity problem is quantified, it becomes the benchmark for information integrity improvement. Information integrity assessments are often performed as a precursor to system integrations, where data is migrated from one system to another. They are also performed when a company wants to create an effective, targeted action plan to correct known or suspected problems with transaction or reporting system data.

Information integrity assessment tools assist with this process by analyzing data in a company’s internal system. Data quality is determined by comparing the analysis results to the company’s quality requirements. There are several, well-established assessment tools on the market that contain sophisticated, out-of-the-box reports that would take a significant amount of time to custom develop internally. The amount, duration and complexity of analysis that is required will determine whether an automated tool should be purchased.
• Data cleansing/validation tools — Analyze incoming data and standardize, correct and validate the data against predefined business rules and databases (e.g., postal directories). Similar to information integrity assessment tools, there are several, well-established data cleansing/validation tool vendors with experience in global data. Many of these vendors, however, evolved from the customer name and address cleansing world, and thus offer limited functionality for assessing other types of data. These tools can be broadly classified into first- and second-generation tools. First-generation tools require specific configuration for existing business rules, and second-generation tools utilize “fuzzy logic” to derive business rules from data sets.

Information integrity validation

An enterprise's ability to measure the quality of its data is critical to the success of its information integrity initiatives; it is a means through which an organization can monitor and execute its information integrity policy. Validation standards allow monitoring over time to verify ongoing group and staff performance, helping to create an effective data organization that can execute information integrity policy. Monitoring performance against validation standards and intervention (as required) also verifies that information integrity levels continue to support business requirements.

Successful enterprise information integrity programs identify specific validation components that are reviewed regularly to foster a high level of information integrity within an organization.

Information integrity validation results in the creation of metadata related to the following areas:

• Criticality of data — Critical data elements, those that have a significant impact on the most important business processes, can be identified and prioritized. Prioritization helps ensure that the most important data elements/data objects are being assessed and that implemented changes are effective, efficient and produce significant value.

• Dimensions of quality — Data quality is measured across several dimensions, each representing a facet of the data element that could be managed by the enterprise.
Standards and diagnostics — Once the dimensions of data quality have been determined, they can be applied to all data elements. Documentation of the expectations for each data element helps create standards that can be reused throughout the organization.

Data quality metrics — Within an organization, there may be thousands of data elements. Even when the critical list of data elements are identified, the list of all applicable diagnostics can still be overwhelming. Diagnostics and standards must be aggregated into metrics, which can be defined at many levels, to enable decision making and drive performance. At the lowest level, the quality of individual data elements can be measured to verify that all of the information integrity categories are being met. At a higher level, aggregation of elemental measures can be used to determine the quality of a specific data object, data table or data source.

Communication is vital to educating users enterprise-wide, highlighting individuals’ impact on information integrity and relaying measures and metrics.

Information integrity communication

All information integrity stakeholders should be told of the current state of an organization’s information integrity. Communications should focus on both the quality of results and the improvements made to the framework. Information should be tailored to each target audience and should be meaningful and viable. To accomplish this, reports can be categorized by business area, stakeholder group or information integrity responsibility.

The communication component of the information integrity framework refers to communication with employees and various groups and units in an organization.

The purpose is to:

- Educate and inform everyone in an organization about the importance of information integrity
- Create an awareness and understanding of the impact individuals have on information integrity
- Communicate information integrity measures and metrics to people who can react to them in a timely manner.
Successful information integrity programs create a communication component that includes transition and update information, knowledge management and performance reporting. The following examples relate reporting requirements to support roles and responsibilities:

- **Information integrity stewardship council** — Requires reports on quality measures; for example, measures related to application data
- **Data steward** — Needs reports on the quality of specific databases; for example, the central application database
- **Information integrity analyst** — Uses reports on quality aspects from specific processes; for example, processing of online applications into the central application database
- **Business user** — Consumes reports on business outputs created by specific staff and processes; for example, processing online applications for a specific line of business.

Where possible, communications should be localized, so users understand and have a sense of ownership of the results. This is especially true in the case of data stewards who are responsible for specific data. When dealing with data at an enterprise level, the diagnostic and transactional results captured in information integrity measurement should be summarized to highlight key performance areas. A scorecard approach to reporting helps communicate current performance and identify quality trends, and focuses on improvements, not the overall quality level. The scorecard can provide motivation at the beginning of a project, because it does not highlight substandard quality performance levels. Communication of framework improvements can be as important to the stakeholders as the actual information integrity results. The completed improvements suggest the future state of the data and where the changes in information integrity will occur.
Compliance with an information integrity framework is critical in order to maintain high data quality

Information integrity framework compliance

The interrelationships between organization, process and technology are continually evolving, and can cause the information integrity framework components to become outdated if a strong change management process is not present. For example, new errors can be caused by changes in data interfaces, staff, new business policies and regulatory requirements. If existing processes and systems do not detect the errors, an enterprise will not be aware of the loss in data. To mitigate this risk, the framework compliance component of the information integrity framework assesses the information integrity program within an organization by evaluating the effectiveness of the integrated framework in sustaining data quality. The results of the framework compliance can then be used to certify systems, procedures and business units to meet the standards set out in the information integrity policies and to identify areas where the framework elements need to be strengthened due to a change in circumstances. The process for executing information integrity framework compliance represents a condensed version of the methodology used to design and implement an information integrity assessment.
The IBM information integrity methodology describes how to implement the framework in order to improve and sustain information integrity.

Information integrity methodology: applying the information integrity framework
The information integrity framework highlights the various areas that need to be managed in order to achieve and sustain information integrity. The framework is used to describe what a current or future business environment looks like, but it does not identify how the framework is developed or how an organization makes the transition to the new environment described in a framework. The five phases of the information integrity methodology describe how to implement the framework in order to improve and sustain information integrity.

![Diagram of the IBM information integrity methodology]

- **Initiate**
  - Identify scope of information integrity project
  - Select roadmap for information integrity project
  - Create information integrity workplan
  - Estimate and obtain resources
  - Assess risks, readiness, and quick wins

- **Define**
  - Define areas of focus
  - Identify known data issues & current initiatives
  - Profile existing data*
  - Identify critical data elements
  - Define information integrity criteria

- **Assess**
  - Assess framework elements
  - Assess critical data*
  - Perform root cause analysis

- **Cleanse**
  - Envision remediation environment
  - Implement remediation environment*
  - Cleanse and connect data*
  - Assess results*

- **Assure**
  - Envision new environment
  - Implement new environment
  - Transition to new environment
  - Assess results

* Tool supported
The information integrity methodology can be implemented on a top-down, enterprisewide basis or from the bottom up, starting with a business area that requires analysis and improvement, and then iterating the process throughout the enterprise. In light of the significant change management that accompanies the design and implementation of an information integrity framework, adopting the latter approach helps reduce the risk of failure of the program and should be seriously considered in at least a pilot capacity. Each phase of the methodology for applying the information integrity framework has a specific objective.

**Initiate**

The initiate phase is used to obtain an overview of the entire information integrity project. This initial planning stage is required to obtain the optimum results in the least amount of time and with limited effort. Planning also develops a common understanding of the project, fostering cooperation, availability of people and systems, and a framework for communication, monitoring and control. The initiate phase includes:

- Identifying the information integrity project scope — Defining project boundaries to prescriptively establish what is in scope and the interfaces
- Creating an information integrity workplan — Providing the step-by-step plan for executing the information integrity project, measuring its progress and communicating tasks and project status
- Estimating, obtaining and scheduling resources — Estimating the skills required and the number of people needed within each skill category
- Assessing risks, readiness and quick wins — Analyzing the information integrity project situational elements to determine the conditions or factors that could cause potential loss, and identifying benefits that can help support the project.
Define

During the define phase, the data quality scope is determined, the critical data elements are identified and the data quality requirements are defined. This information provides the foundation for the remaining project phases. The define phase includes the following processes:

- Defining focus area — Highlighting the people, processes and systems related to the initiative
- Identifying known data issues and current initiatives — Identifying current issues and planned improvements
- Profiling existing data — Examining the data initially to confirm expectations related to the data, and identifying additional issues
- Identifying critical data elements — Identifying data elements that have a significant impact on business processes and systems
- Defining information integrity criteria — Documenting the quality criteria that represent the rules to which the data must conform.

Assess

The goal of the assess phase is to determine the current level of data quality and identify any information integrity issues. An assessment of the data and the business environment is performed, including analysis of data quality metrics, data issues and root causes, processes, systems and organizational structure. The information related to the existing environment is then used as a baseline for the subsequent cleanse and assure phases. The assess phase includes:

- Assessing framework elements — Reviewing all the areas within scope using the information integrity framework
- Assessing critical data — Reviewing the data directly to identify errors in critical data elements based upon the criteria established in the define phase. The results are metrics for each element that pinpoint the various errors that can exist for the element
- Performing root cause analysis — Identifying specific errors and why they are occurring. This process converges the two assessments into issues and root causes.
During the cleanse phase, the remediation environment is envisioned and implemented, data is cleansed and corrected, and the results are assessed. At phase completion, data issues should be resolved to address any significant business impacts. The cleanse phase includes:

- Envisioning a remediation environment — Targeting specific focus areas, including identifying fields to be improved, prioritizing cleanup and identifying where, when and how to improve, as well as any training requirements.
- Implementing remediation environment — Implementing a solution based on the strategy developed in the envision phase. A data remediation group is assembled, required tools are developed and implemented, data is obtained and training completed.
- Cleansing and correcting data — Cleansing and correcting data utilizing the implemented remediation environment.
- Assessing results — Reassessing the data quality to measure improvement. Any data quality exceptions that still exist are identified, and any required improvements can be planned.
The objective of the assure phase is to design, develop and implement solutions that address information integrity issues. The assure phase identifies the root causes of the data quality issues identified in the assess phase and creates solutions to ensure sustained data quality. The support phase includes:

- Envisioning a new environment — Reviewing the issues identified in the data assessment and the root causes of these issues, and developing solutions by leveraging the information integrity framework best practices
- Implementing a new environment — Developing methods, training staff and implementing changes based on the strategy developed in the envisioning phase
- Transitioning to a new environment — Making the business transition to the new environment created by the implementation of changes
- Assessing results — Reassessing data quality to measure improvement. Any data quality issues that still exist are identified, and any required improvements can be planned.
Conclusion

Leveraging data is critical to competing effectively in the information age. To capitalize on the value of their information, businesses must invest in the quality of their data and achieve information integrity. The result of the interaction of people, processes and technology, information integrity is achieved through sophisticated interrelationships. It is defined by how effectively data supports the transactions and decisions needed to meet an organization’s strategic objectives. Achieving information integrity means fixing the data foundation through a systematic, integrated approach that leverages the short-term gains of remediation where applicable, but looks beyond the quick-fixes to build an effective supporting environment. It also involves addressing tangible, less-tangible and intangible costs to protect and enhance customer, supplier and partner relationships.

IBM Business Consulting Services can help organizations achieve information integrity through established best practices, proven through numerous successful customer engagements. The IBM enterprise information integrity framework highlights areas that must be managed to achieve information integrity. And the IBM enterprise information integrity methodology captures the five-phase process for designing and implementing the framework — all of which is designed to help organizations foster sustainable data quality and yield a high return on investment.
About IBM Business Consulting Services
With consultants and professional staff in more than 160 countries globally, IBM Business Consulting Services is one of the world's largest consulting services organizations. IBM Business Consulting Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues and the ability to design, build and run those solutions in a way that has the potential to deliver bottom-line business value.

For more information
To learn more about how IBM Business Consulting Services can help you transform your enterprise information integrity, contact your IBM sales representative, or visit:

ibm.com/bcs