IBM

Getting started with data governance

Brett W. Gow
Associate Partner
IBM Global Business Services

Contents

Executive summary

Data governance is both a strategic and tactical approach to managing information. It involves political and technical organization to implement policies in business processes and control the quality of information design, collection, use and reporting. For the last three years, the IBM® Data Governance Council and IBM Global Business Services have been documenting data governance challenges and providing solutions. Today, your company can take advantage of this leadership and experience to shorten the learning process and get started with data governance.

This white paper explains the IBM model for achieving data governance success, including new IBM software and services solutions to help make governing data easier, and offers concrete actions you can take today to help increase revenue, cut costs and reduce the risks associated with managing information.

Data governance recognizes information as an enterprise asset

Data governance is the orchestration of people, processes and technology to enable an organization to leverage data as an enterprise asset. It helps make data more usable, accessible, consistent and trustworthy.

Driven by the need to have clear, concise, accurate data delivered to decision makers, organizations are now understanding the need for and the benefits of data governance.

According to a recent survey by The Data Warehousing Institute,[1] there are several reasons why data governance is becoming more important:

- *The current "age of accountability" demands compliance.*
- *Compliance and business intelligence demand high-quality, auditable data.*
- *Improving data quality is a cross-functional imperative.*
- *Data integration implementations cast an ever-widening net.*
- *Data governance helps reduce the risk incurred during business transformations.*

Despite the growing level of interest, the current state of data governance is hampered by organizations' failure to strategically and appropriately leverage their data. A survey from the International Association for Information and Data Quality (IAIDQ) and University of Arkansas at Little Rock Information Quality Program (UALR-IQ) reports that more than half of respondents recognize information as a strategic asset and manage it accordingly—yet "17% are neutral on this question and 25% feel their organizations do not recognize information as a strategic asset."[2]

Exploring the data governance lifecycle

Measuring data governance can be challenging. For that reason, the IBM Data Governance Council, an industry thought leader, and IBM Global Business Services suggest that companies use a Data Governance Maturity Model to address organizational performance (see Figure 1); track progress across a proposed data governance lifecycle and against a maturity model; and adopt methods and metrics for reporting program progress and performance.
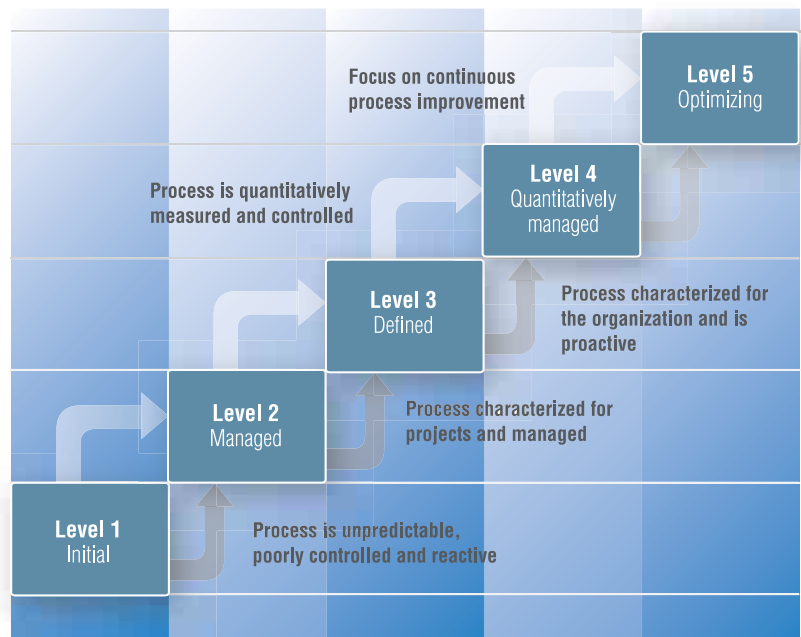

Figure 1: The five stages of data governance maturity

Focus on continuous process improvement — **Level 5** Optimizing

Process is quantitatively measured and controlled — **Level 4** Quantitatively managed

Level 3 Defined — Process characterized for the organization and is proactive

Level 2 Managed — Process characterized for projects and managed

Level 1 Initial — Process is unpredictable, poorly controlled and reactive

Initially, companies can embrace data governance by following this path:

1. Recognize the need, value or imperative of data governance
2. Assess the organization's current state using objective assessment, specifically considering elements of policy, data quality and value creation
3. Clarify the value or organizational imperative using the results of the objective assessment
4. Identify leadership, responsibility and accountability for data governance activities and accomplishments
5. Initiate data governance programs and projects
6. Define and deploy management processes and incorporate them into "business-as-usual" processes
7. Evolve data governance capabilities as the organization changes or matures

An objective assessment or health check can provide an organization with an informed, documented perspective on strengths and weaknesses. It often confirms known truths and validates, invalidates or modifies commonly held assumptions. It also brings new truths forward and can be used to establish a foundation for strategies and tactics, including ownership and timelines.

IBM offers Data Governance Assessments that are designed to be comprehensive, objective and quantifiable to help develop dialogue and build consensus across the organization. Based on IBM's considerable experience and a library of more than 150 questions, these enterprise-wide assessments are well defined and can be delivered quickly—often in six weeks or less—to help effectively manage costs.

Moving up the data governance maturity model

To move up the data governance maturity curve, organizations must consider developing or acquiring several common elements or leading practices. Four of the most important are data policies and guidelines; organizational change management; data architectures; and data definitions or metadata.

Data policies and guidelines

Developing policies and guidelines for enterprise architectures is an important step in defining and building conformance. A distributed structure can lead to redundancy and a lack of consistent information, breaking down process efficiency and effectiveness while increasing exposure to risk and regulatory noncompliance.

Creating policies and initiatives to standardize architectures helps uncover the redundant elements and structures that are often root causes of data quality issues. Corollary metadata initiatives assist in this process by identifying where similar data is housed and managed.

Organizational change management

Clear leadership roles and responsibilities are required as organizations plan, align and mobilize data governance activities. Appropriate leadership and management of the human element of change are critical to the success of a transformation project, and organizations should not underestimate organizational change management requirements and challenges. Plus, data governance needs will change over time as the company grows and markets evolve.

To ensure appropriate planning and support for these changes, the IBM Data Governance Council and IBM Global Business Services recommend the following framework to understand how the current organization is structured and behaves, and what enablers exist to motivate behaviors and support the structure (see Figure 2).
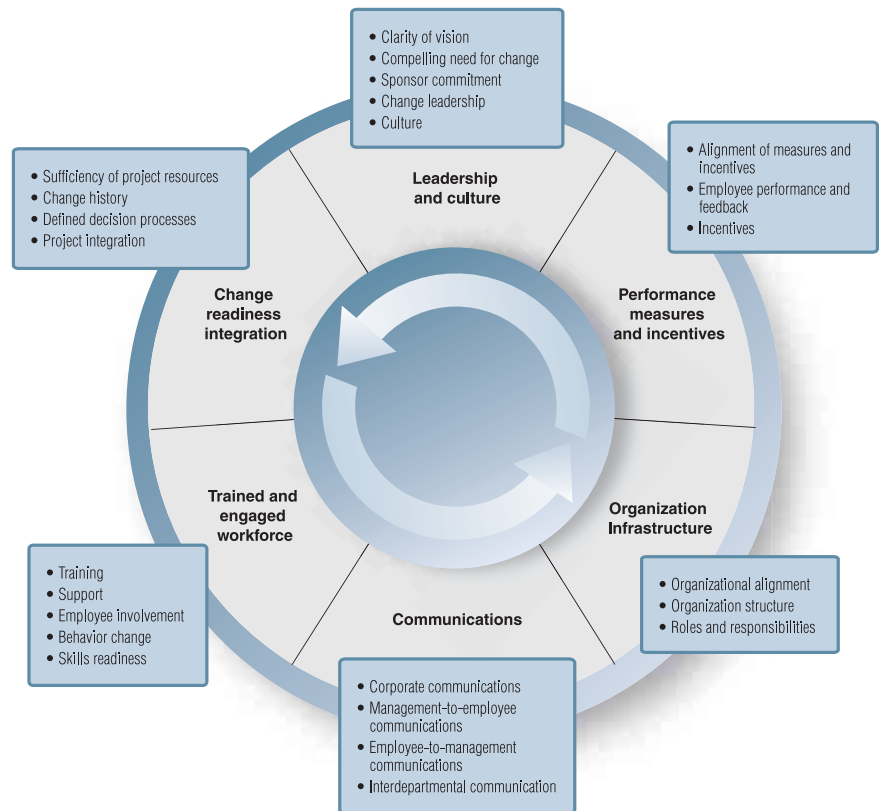

Figure 2: The IBM Organization Framework provides a view of the organization's structure and supporting activities.

Data architectures

Shared or common enterprise data architectures aid in accelerating technology initiatives while maintaining quality and data governance standards. Consistently leveraging these architectures will also help prevent duplication of effort across multiple projects. Once a model is validated, it helps increase the consistency and performance of business solution specification, design, delivery and use with minimized risk.

A common architecture model helps organizations achieve streamlined, consistent business processes; reduce IT complexity; improve business processes internally and across external business partners; achieve consistent processes across products and channels; and achieve straight-through end-to-end processing. Moreover, it facilitates consistent provision of products through multiple channels and defines workflows and processes independent of line of business, product, channel, organization structure and technology.

Companies can also use a common data model and reference architecture to address compliance issues, promoting conformity of key performance and risk indicators and reporting summary levels across the enterprise. Appropriate designs can also support detailed financial reporting, down to the transaction level of internal business processes across the organization.

Data definitions or metadata

Data governance tools and processes enable visibility into the impact of business change on the IT infrastructure. To do this, they employ a comprehensive and consistent dictionary to describe business issues, applications and components, helping to define the organizational systems and the impact of new initiatives.

Metadata plays a key role in business glossary efforts. It describes aspects of data used within and across an organization, including its definition, source, stewardship or ownership, location and its state and usage specifications. Metadata management helps support a company's business information lifecycle by capturing the transformation of data as it turns it into information, as information becomes knowledge and as it is finally translated into business value.

- *Initial stage: Data*—*The organization's raw source data has few quality constraints, no quality checks and inconsistent structures and definitions for the same or similar data across multiple repositories and databases.*
- *Second stage: Information*—*Data is consistently conformed through data structures and common business definitions, which are provided and maintained through shared processes. They often use tools and technology as accelerators, as well as to increase the efficiency and consistency of maintenance and management activities.*
- *Third stage: Knowledge*—*Information is available for sharing consistently and coherently, both vertically and across the organization, in support of analysis, fact-based decision making and consolidated enterprise reporting and forecasting.*
- *Fourth stage: Business performance*—*The organization uses analysis and fact-based decision making in the marketplace on behalf of individual business and functional areas to the benefit of the entire organization. Few, if any, decisions of magnitude are made without leveraging common enterprise-level tools and decision methodologies.*

Technology and tools

IBM offers a wide range of tools that can help accelerate the speed and success of data governance initiatives, including initial adoption, process and training conformation and value realization.

Policy tools

IBM Rational® RequisitePro collects and tracks all of the documents that are relevant to a software development project, especially those representing software requirements and use cases. It is both a document repository and a database of project information, tracking the origination and change history of requirements as they evolve.

RequisitePro enables project managers to classify and track software requirements against a wide variety of project-specific criteria such as completion status, degree of difficulty and size. Criteria can be tracked over time using the tool's metadata features.

Enterprise architecture

The IBM Industry Models are a comprehensive set of industry-specific business models that represent best practices for the given industry. The models form an efficient communication bridge between business and technology communities. They are designed to be readily accessible to business users and focus on issues in areas such as customer insight, multichannel transformation, core systems and risk and compliance.

The IBM Industry Models comprise:

- *Information Models: Industry-specific data content to enable capabilities such as enterprise-wide views of information*
- *Process Models: Industry-specific business process content to address areas such as business process re-engineering*
- *Integration Models: Business services content to address areas such as Service Oriented Architectures (SOAs)*

IBM InfoSphere™ Master Data Management (MDM) Server delivers a rich, extensible enterprise repository and a robust set of business services for master data. Master data is the high-value information (such as data about customers, suppliers, partners, products, materials and employees) that is critical for running a business, yet is typically scattered among heterogeneous application silos across the enterprise. With IBM InfoSphere MDM Server, master data can be brought into an application-independent repository, yielding a complete, 360-degree view of the master data as well as a common platform for maintaining the integrity and quality of master data across the enterprise.

Data quality and metadata

The IBM InfoSphere Information Server platform helps organizations derive more value from the complex, heterogeneous data within their information ecosystems. It enables organizations to integrate disparate data and deliver trusted information wherever and whenever needed, in line and in context, to specific people, applications and processes.

IBM InfoSphere Information Server also helps business and IT personnel collaborate to understand the meaning, structure and content of any type of information across any type of source and provides capabilities for cleansing, transforming and delivering this information consistently and securely throughout the enterprise.

Information lifecycle management

IBM Optim software is a powerful information lifecycle management (ILM) solution that offers a central point to implement ILM policies across heterogeneous structured data sources. It can extract, archive, port and privacy-mask valuable application data from its creation through to its eventual decommissioning—across all enterprise databases.

IBM Optim provides the following core functionalities to support data governance efforts:

- *Test data management: IBM Optim helps streamline creation and management of test environments; subset and migrate data to build realistic and right-sized test databases; and eliminate the expense and effort of maintaining multiple database clones.*
- *Data privacy: Protecting your sensitive data does not stop at your production system. This data is commonly replicated in multiple test environments across your organization, as well as in extract files and staging tables. IBM Optim provides automatic data transformation capabilities to mask personal information and confidential information to protect privacy.*
- *Archive: IBM Optim Data Growth provides proven database archiving capabilities, empowering organizations to segregate historical from current data and to store it securely and cost-effectively while maintaining universal access.*

Value creation and performance measurement
IBM Cognos software features business intelligence and performance management capabilities to help organizations:

- *Integrate disparate data into a clear, consistent and accurate view of performance at every level of your organization*
- *Deliver business-critical information to decision makers when, how and where it is needed*
- *Identify and leverage "sweet spots" of information in core functional areas for competitive advantage*
- *Plan, understand, manage and improve financial and operational performance*
- *Answer the three questions that drive better performance: How are we doing? Why? What should we be doing?*

IBM expertise helps support data governance success

C-level executives face an onslaught of information-related processes, from addressing regulatory compliance and reporting requirements to ensuring that all business units are using the correct data to track customer preferences. IBM continues to define, innovate and deliver the software, hardware and services required to make even the most daunting data governance undertaking successful.

Data governance is not new and neither are the challenges—but IBM tools, techniques and support can help business decision makers at every level be more effective, more consistent and more successful.

For more information

For more information on IBM data governance products and services, please contact your IBM marketing representative or IBM Business Partner, or visit

ibm.com/software/data/information/trust-governance.html

[1] "Data Governance Strategies: Helping your Organization Comply, Transform, and Integrate." The Data Warehousing Institute. April 2008.

[2] "The State of Data Governance." International Association for Information and Data Quality (IAIDQ) and University of Arkansas at Little Rock Information Quality Program (UALR-IQ). April 2008.