DATA GOVERNANCE

Creating Value from Information Assets

Edited by Neera Bhansali

CRC Press
Taylor & Francis Group
AN AUERBACH BOOK

DATA GOVERNANCE

Creating Value from
Information Assets

DATA GOVERNANCE

Creating Value from Information Assets

Edited by Neera Bhansali

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
AN AUERBACH BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20130227

International Standard Book Number-13: 978-1-4398-7914-6 (eBook - PDF)

Contents

Preface

> As a rule, he or she who has the most information will have the greatest success in life.
>
> **Benjamin Disraeli**

In recent years, there has been enormous growth of data in organizations. This data has become the basis of competition, productivity, growth, and innovation. The rise in volume of data through the Internet, social media, and multimedia is adding new challenges and opportunities for harnessing the power of information. There is much discussion today on using organizational data to derive value. As organizations respond to strategic and operational challenges that demand high-quality data, data governance is emerging as an important area in enterprise information management.

As organizations deploy business intelligence and analytic systems to harness information and business value from their data assets, programs for the governance of data are gaining prominence. Data management issues have traditionally been assigned to and addressed by IT departments in organizations. However, organizational issues critical to successful data management require the implementation of enterprise-wide accountabilities and responsibilities. Data governance encompasses both the decision domains and the accountability for decision making. Effective data management requires a data governance structure and framework that emphasizes collaboration between business and IT to support organizational goals. It brings together diverse expectations and expertise from across the enterprise to achieve an agreed upon, consistent, and transparent set of processes that enable data-informed decision making. This book looks at the needs and processes for data governance to manage data effectively. It addresses the complete life cycle of effective data governance from metadata management to privacy and compliance. These also are highlighted through case studies.

The goal of this book is to assist others who are on the journey to derive value from informational assets using data governance. The book is a summation of experiences of experts and addresses an area that is of growing interest to the information systems and management community. Book chapters present how ideas have been adapted as techniques and policies

for practice in organizations in their journey to successful data governance. Case studies from healthcare and financial sectors, two industries that have successfully leveraged the potential of data-driven strategies, provide further insights into real-time practice.

> A popular government without popular information or the means of acquiring it is but a Prologue to Farce, or a Tragedy, or perhaps both. Knowledge will forever govern ignorance and a people who mean to be their own Governors, must arm themselves with the power which knowledge gives.
>
> **James Madison**

Acknowledgments

My happiest task is to thank my co-authors who have come together from around the globe to share their experiences and contribute to the book. Thanks Charlyn, Dasaratha, Michael, Christopher, Osamu, Julia, Gladys, and Fernando. Without you this book would not have been possible.

My thanks to Shekhar, Divya, and Shrenik for their unwavering support and patience during the writing of this book.

Contributors

Neera Bhansali
Executive Director, Biomedical Informatics
Herbert Wertheim College of Medicine
and
Faculty Director, MS HI&MS
Department of Decision Sciences and Information Systems
Florida International University
Miami, Florida

Fernando A. Faria
Analyst
Banco Central do Brasil
Brasilia, Brazil

K. Hasida
Social Intelligence Technology Research Laboratory
National Institute of Advance Industrial Science and Technology
Tokyo, Japan

Charlyn A. Hilliman
Chief Technology Officer
The Learningateway
New York, New York

N. Izumi
Social Intelligence Technology Research Laboratory
National Institute of Advance Industrial Science and Technology
Tokyo, Japan

K. Murata
School of Medicine
Kitasato University Hospital
Sagamihara, Japan

Dasaratha Rama
Professor
Department of Decision Sciences and Information Systems
Florida International University
Miami, Florida

Michael Schrader
Vice President, Data Architecture
Fidelity Investments
Evergreen, Colorado

Gladys E. Simpson
Visiting Faculty
Department of Decision Sciences and Information Systems
Florida International University
Miami, Florida

Christopher B. Sullivan
President and CEO
Image Research LLC
Miami, Florida

Osamu Takaki
School of Knowledge Science
Japan Advanced Institute of
Science and Technology
Ishikawa, Japan

Julia Zhang
Associate Director, Project
Management Strategy, Standard
and Architecture
R&D IS
Sanofi
Cambridge, Massachusetts

The Editor

Neera Bhansali, PhD, has more than 20 years of experience in information technology with an emphasis on strategic data warehousing. She is executive director of Biomedical Informatics at Herbert Wertheim College of Medicine (Florida International University (FIU), Miami). In Dr. Bhansali's previous roles, she was head of informatics at CTNeT, the Statewide Clinical Trials Network of Texas, and held leadership positions, including director of Data Quality and Standards at the H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida.

Dr. Bhansali is the author of *Strategic Data Warehousing: Alignment to Business* (CRC Press, 2009).

She earned a doctorate in business and a master's in business (information technology), both from the Royal Melbourne Institute of Technology, Melbourne, Australia.

1

The Role of Data Governance in an Organization

Neera Bhansali

CONTENTS

INTRODUCTION

In today's fast-paced global economy, changes in the environment result in new opportunities for wealth creation that decision makers use in strategy formulation and implementation. Organizations have grown global by not only expanding businesses and setting up branches overseas, but they have grown global in a different kind of way—leveraging on knowledge tapped globally rather than merely growing outward from a domestic base. In organizations, information from various departments and functions, as well as formal and informal sources of information, are bought

together to detect and interpret problem areas, identify opportunities, and implement strategic objectives. Strategies define the purpose of organizations, the competitive domain of firms, and the resource commitment these organizations make to achieve and sustain competitive advantage. Hence, developing an appropriate data strategy that fits the marketplace is one necessary ingredient for business success. Effective data governance reduces uncertainty and helps improve an organization's performance. An organization's ability to collect pertinent information and act on signals that others miss provides it a strategic advantage.

THE NEED FOR DATA GOVERNANCE

In the twenty-first century, organizational business is supported by data and information in many ways and forms. A large part of data generated today is in the form of electronic or digital data. For example, high-volume transaction processing systems (financial or accounting systems) generate reports, statements, and electronic checks; analytical systems produce documents, spreadsheets, forecasting models, visual images; marketing and planning systems generate emails, Web pages, XML from the Web or corporate intranet, and so forth. Data resides in databases, files, and other media. All this data needs data management to identify, categorize, preserve, and retrieve data when required for business purposes or regulatory needs. A rule of thumb often used is that corporate data volumes double every 18 months. However, mobile and online data are growing at an even faster rate. From cost containment, regulatory compliance, strategic initiatives, and beyond, the need for businesses to manage proliferating data is becoming even more urgent.

Maintaining an ever-increasing volume of data, processing it, and deriving information from it to meet the competitive needs of the market requires not only good data management practices, but a data governance process as well. Effective data governance improves data safety and security, improves data quality, and ensures compliance with data-focused regulations as well as helps an organization manage and use its data effectively. It increases data consistency, increases accountability for the organization's data, and improves decision making. Data governance must be a business-driven program that uses a data-governance maturity model to build a strategic road map.

Organizations in the public and private sectors continue their efforts to manage this enormous data and information inventory, regardless of medium. While there are many challenges associated with this effort, in today's highly complex business environment, one challenge that stands out is retaining and leveraging all knowledge assets. As we emerge slowly from the economic downturn of the past few years, growth continues against a backdrop of cost cutting. Data governance programs can help by delivering a unified view of the world. Organizations realize that to remain competitive, they must take advantage of what they know and what they are learning. A nonproactive approach to managing data leads to many pitfalls. Data governance process supports business strategies of operational excellence, cost reduction, and cost-effective regulatory and legal compliance while meeting the strategic objectives of the organization.

PLACE FOR DATA GOVERNANCE IN ORGANIZATIONAL STRATEGY

According to Porter (1980), successful strategy formulation involves the identification and exploitation of a firm's competitive advantage. There are five forces that affect a firm and its competitors in any given industry: threat of new entrants, threat of substitute products, jockeying for position among competitors, power of buyers, and power of suppliers. In Porter's framework, a firm's performance in the industry is a function of the environment and the firm's positioning in the market based upon the five forces that help establish competitive advantage over its rivals. Distinctive competencies and organizational effectiveness determine how a business performs. In today's data- and information-intensive age, good data governance strategies provide organizations with a distinctive competency in data and information management that enhances the effectiveness of business strategies.

In today's digital world, technology makes it easy to create, transmit, store, access, and use information that is becoming the basis for business operations, customer service, and government relations. However, technology itself is so ubiquitous that by itself it no longer provides a distinguishing competitive advantage in business. It is rather people's creative use of information that counts rather than the technology. A distinguishing feature of sound strategic management is its flexibility, responsiveness to change, and ability to respond to new challenges. Information and data governance

supports an organization's strategic approach involving planning, choosing, and sometimes improvising or shifting approaches dynamically based on the competitive environment. Data governance supports the systematic management and use of information to achieve objectives that are clearly aligned with and contribute to the organization's objectives.

To compete in the global arena, organizations adopt two broad strategies types: either Adaptive/Defender or Proactive/Prospector strategies. Organizations pursuing the Defender strategy type respond to competitive forces with better management of costs and reliance on internal strengths for their positioning in the marketplace. Organizations pursuing Prospector strategy type utilize marketing strengths in response to competitive pressures. In both the cases, data and information are utilized to assess, evaluate, and formulate these business strategies. Data and information internal to the organizations are used in Defender strategies to lower operating costs and achieve efficiencies to better position the organization in the competitive marketplace. Similarly, data and information of external environment and marketplace play a strong role in the formulation of proactive strategies. Organizations with Prospector strategy are å and marketing-oriented, focusing on identifying and satisfying the needs of customers with value, quality, and product offerings based upon their knowledge of customer and industry trends. Good data governance policies and procedures are invaluable in the management of data and information in organizations of both types of strategies, and contribute to building distinctive competencies for the organization.

DATA GOVERNANCE ACROSS INTRAFIRM NETWORKS

"Intrafirm networks" are a set of formal and/or informal relationships among business units of the same legal entity (Achrol and Kotler, 1999). Each business unit has a sufficient degree of freedom to make its own resource allocation and data governance decisions while still working in close cooperation with its affiliated business units. Business units obtain and utilize data and information from within their formal boundaries and from other business units. In order to acquire this data, business units within the organization must follow some procedures and policies.

Data governance across the business units in an organization facilitates easy and timely access to each other's data and to data from businesses

outside the organization. Data governance also helps the business units to be better prepared for information sharing with the other units within the organization. Data governance structures promote the development of intrafirm networks within the organization in order for data and other resources to be transferred or exchanged and distributed efficiently throughout the organization to gain competitive advantage. Strong intrafirm networks enhance reciprocity, cohesiveness, and connectivity among business units (Rindfleisch and Moorman, 2001).

Globalization of the economy, dealing with trusting business partners at great distances and ensuring that information is properly used and has agreed-upon qualities and limits, emphasizes the need of a data governance structure. These needs are met by data governance policies that oversee the flow of information and design information products to meet new values criteria. Managements have to make data governance decisions as to what information should be shared, what quality controls should be in place to assure others of its validity, who should have access to data and information, in what forms, and at what levels of access,

DATA GOVERNANCE CHARACTERISTICS OF ORGANIZATIONS THROUGH THEIR LIFE CYCLE

Organizations face many challenges in this era of high growth and competition. It encompasses the numerous demands of new product innovation, increasing market shares, and customer satisfaction. To address time and resource constraints, companies formulate and disseminate an elaborate, structured policy of data governance based on the guidelines of accountability, responsibility, internal controls, and audit procedures. A well-developed system of corporate governance, along with data governance, adds to the needed synergy for growth. Data governance encompasses features such as ethics policy, reporting transparency, and corporate citizenship. It is a mechanism to maximize data and information value.

Good governance systems lead to better access to data and information, higher quality data, and reduction of risk due to inaccurate data. However, implementing effective governance systems also comes at a cost. Resource constraints, lack of business understanding of the system, and the cost of implementing and communicating data governance policies throughout the organization are crucial barriers that many organizations face. Once a

successful working environment is established, a well-oiled system of data governance will be highly rewarding. However, in the beginning, the tendency of management will be to focus resources on revenue increase and the value chain that links data and customers through the organization.

Organizations move through various stages in their life cycle from start-up phase, growth phase to maturity phase, renewal, and, finally, decline phases. Different stages in an organization's life necessitate different roles for data governance. In the early phases of the life cycle, data governance is a regulatory requirement and not a competitive tool. It consists of policies, control procedures, guidelines, and mechanisms to ensure accountability. A data governance structure specifies the distribution of rights and responsibilities amount different participants in the data governance process. It spells out the rules and procedures for making decisions on data affairs so that the organization's data and information needs and objectives are met.

In the early stages of an organizations growth, the management hierarchy usually remains flat rather than layered. Management executives of such organizations are still trying to find an internal organizational structure that works well and, therefore, will not have the time or the opportunity to set up a well-defined data governance system. An effective system of data governance has many features that involve board structure, planning and monitoring, risk management, audit committees, internal control, ethics, and transparency. An organization in its early stages of growth may not have the resources to develop every facet of an effective data governance policy.

The growth of an organization usually implies constant changes to products, processes, and organizational and managerial practices. This also requires continuous adaptation to the changing business environment, and developing sustainable data governance processes. A structure of rules and regulations to effectively govern data and information is established to encourage the efficient use of resources and to require accountability of those resources.

Mature organizations would not have "flat" management systems, but well-developed hierarchical structures that require a good system of data governance for effectiveness and transparency. In later stages of an organization's life cycle, a good system of data governance enhances returns, provides for better risk management, improves customer satisfaction, and increases the organization's reputation. Costs involved are generally increased managerial and supervisory time.

DATA GOVERNANCE PROGRAM

A data governance program provides an opportunity of taking an enterprise-wide view rather than focusing on a particular department, involving business heads in planning, devising budgets that meet the needs of the enterprise, and being aware of the powerful effect data and information have on organizational competitiveness. Increasing awareness at the executive level of the importance of managing data and information strategically is an integral part of increasing the recognized value of the data governance program.

To begin a successful data governance program in an organization, aligning the program with the business strategy is important. A multiyear-phased data governance program should focus on critical business scenarios rather than invest into the technology fostered by vendors selling compliance wares. From an assessment of the business use cases and scenarios, a business case should be developed that draws a distinction between strategic business and operational needs. The governance plan should present an end-to-end view of both strategic business and operational needs and requirements with established priorities.

A data governance program requires significant investment from the organization in terms of time and resources. However, these costs are offset by the business value it delivers. The data governance solution should be enterprise-wide and the enabling components should fit and function within the strategic direction of the organization. With a defined business case, clearly defined business drivers, and executive sponsorship, an enterprise-wide data governance program can be successfully launched to provide significant returns.

An enterprise data governance program impacts and touches upon the organization, its processes, people, and enabling technologies comprised of hardware and software. An assessment of the business use cases and scenarios presents not only an understanding of the issues, but of the current business context, the processes followed, the technologies employed, and the interactions among people, groups, and departments. It presents an understanding of the organizational practices, challenges and issues, and priorities and opportunities for data governance. Business use cases and scenarios present a description of current processes and practices, help identify improvement opportunities, and help identify

the business and technology operation's roles required. It helps define a common vocabulary in managing both structured and unstructured data.

Understanding an organization's current processes and issues is not enough to build an effective data governance program. To gather business, functional, and technical requirements, understanding the future vision of the business or organization is important. This is followed with the development of a visual prototype or logical model, independent of products or technology, to demonstrate the data governance process. This business-driven model results in a definition of enterprise-wide data governance based on key standards and processes. These processes are independent of the applications and of the tools and technologies required to implement them. The business and functional requirements, the discovery of business processes, along with the prototype or model, provide an impetus to address the "hard" issues in the data governance process.

Building an effective enterprise-wide data governance program, aligned with the strategic direction, maintains a business focus rather than a product or technology focus. A business architecture should be developed identifying the domains of data, its owners, users and custodians, and their roles and responsibilities. To establish a common vocabulary, standards should be adopted for naming, metadata, and records management definitions. The data governance process should include structured data and unstructured data, data extraction, data retention, data sharing, legacy data, and data archiving processes.

The need for a comprehensive strategy cannot be underestimated. An organization is constantly undergoing change. Although a data governance program could be on track and progressing, constant justification and rejustification have to be provided to the sponsors of the program. Toward this end, education, meetings, governance functions, and presentations become significant aspects of the program. Establishing a data governance committee early in the process provides the necessary direction and momentum. A diverse set of skills are required to bring data governance into practice, including application developers, subject matter experts, process control leaders, and business users. Data governance committees usually include a broad spectrum of stakeholders: users, sponsors, business department leaders, IT staff, and consultants.

Key to good data governance is co-ordination and communication. Through constant communication, all members and stakeholders are involved in the process and informed of what is being developed and implemented. Through this practice, the contrarians surface early and

resistance to change is addressed. Communication also fosters a sense of ownership and contribution. Feedback and an iterative process help manage change and maintain the momentum for an effort that takes time to show tangible results. Recruiting ambassadors throughout the organization at various levels of the organization further strengthens the data governance process.

Actively engaging executive sponsors is vital for a successful governance program. Continuous involvement and contribution of key executive sponsors maintains the focus of the program. Business leader support helps maintain resources needed for the program. Consistency among the leaders involved, a sound business proposition, and a value-focused enterprise approach provides the best strategic direction for data governance programs.

One of the challenges in this information age is the superabundance of information. It overwhelms an organization's ability to sift through, organize, and act on it. Information has varying degrees of timeliness, pertinence, and importance. Governance structures provide opportunities to share or withhold information. Informal or formal meetings where individuals from different departments meet and have the opportunity to informally share information face-to-face is still one of the most effective means of sharing relevant and current information.

Data governance is the systematic management of information to achieve objectives that are clearly aligned with and contribute to the organization's objectives. There is a growing recognition that it is the information rather than the information technology that really counts. An enterprise-wide data governance program should be flexible, responsiveness to change, and have the ability to respond to new challenges. As discussed above, a strategic approach to data governance involves planning, choosing, trading off, improvising, and shifting approaches to meet the changing needs of the organization.

BENEFITS FROM DATA GOVERNANCE

One of the main benefits of data governance is recognizing the value inherent in data and treating data as a valuable and manageable organizational asset. Information governance strategy and process establishes the necessary framework to turn data into business value. An effective

data governance framework can help organizations manage data more efficiently. It provides consistent definition, establishes enterprise data management, and measures and tracks the quality of transactional and analytical data used across the organization. It also improves coordination between different functions of business and provides broader insights into data across the products and business units. The data steward groups, part of the data governance framework, help create, implement, and establish measures of the standards across the enterprise.

An effective data governance framework leads to a lowering of information costs. With the reduction of duplicative data stores throughout the organization, iterative data cleansing costs can be reduced via better quality source data. Through the application of standard processes across the business, substantial information cost reduction is achieved as well. It also leads to a higher data quality, greater trust in the data leading to greater insight and better decision making aligned to business goals. With proper data governance over their lifetime, organizations are better equipped to deliver competitive offerings to the market faster and support business goals with less risk.

Effective data governance helps improve compliance and control efforts. Today, organizations are information-aggressive. To conduct business, they routinely collect, analyze, and use information in key areas relating to customers, products, changes in the business environment, and other areas. Unexpected and unintended disclosure of data and information negatively impacts the business and reinforces the need for data governance. Many businesses lose sensitive information because of a lack of proper data governance policies and a lack of understanding and appropriate use of data.

With effective data governance, data standards facilitate high-quality data. Data standards applying uniformly across business functions and lines of business create a uniform transactional and analytical environment for compliance monitoring. Also, with effective data governance, data stewardship is an organization-wide effort, which reduces risk of noncompliance with regulatory and statutory requirements.

DATA QUALITY AND DATA GOVERNANCE PROCESS

Research by Information Difference (Waddington, 2010) suggests that the top six main drivers for implementation of data governance are to support