21.1 Introduction to the Notion of Governance ........................................ 21-1
   Scope of Data (Information) Governance in the Overall Corporate Governance

21.2 Data Governance ................................................................. 21-4
   Introduction • Definition

21.3 Designing Effective Data Governance.................................... 21-5
   Goals of Data Governance • Focus of Data Governance • Data Governance Principles • Benefits of Data Governance • Key Data Governance Components

21.4 State of the Art in Data Governance.................................... 21-9
   Academic Research on Data Governance • Data Governance from Practitioners

21.5 Data Governance Journey: Application and Business Cases with the Data Excellence Framework.................................................. 21-14
   Data Excellence Framework: Govern by Value • Managing Data as a Company Asset • Data Excellence Framework Methodology • Business Case • Data Excellence Process • Governance Model • Maturity Model

21.6 Conclusion and Future Research on Data Governance.......... 21-29
   References.................................................................................. 21-30

21.1 Introduction to the Notion of Governance

Before diving into the specific case of data governance (DG), it is important to outline the notion of governance in the broader sense because DG is but one form of governance. A governance program is used to define the decision-making processes and authority around a specific domain. In other words, a governance program is about deciding how to decide in order to be able to handle complex situations or issues in the future in the most controlled and efficient way possible.*

Over the past decade, corporate governance has been defined as a set of relationships between a company’s management, its board, its shareholders, and other stakeholders that provide a structure for determining organizational objectives and monitoring performance, thereby ensuring that corporate objectives are attained (Khatri and Brown, 2010). Under this broad umbrella that sets enterprise-wide policies and defined standards, it is necessary to implement different governance areas to achieve the goal of a good and responsible corporate governance structure. Weill and Ross propose a framework
for linking corporate governance with six key assets of a company, which are the basis of generating business value and for realizing the company’s strategies (Weill and Ross, 2004). Figure 21.1 is based on Weill and Ross’ understanding of corporate and key asset governance and depicts the dependencies between these areas.

Since information technology (IT) is an essential part within a company’s structure, quality management has to ensure corporate governance compliance of its IT systems. Thus, the discipline IT governance has been established and concentrates its activities on handling of technology and its usage within the company. For Weill and Ross, the IT governance has the goals in specification of decision rights and accountability frameworks to encourage a desirable behavior in the IT usage. In addressing the following three questions, effective IT governance will be set up:

- What decisions regarding the use of IT must be made?
- Who should make these decisions?
- What are the characteristics of the decision and monitoring processes?

In this context, IT governance relates to the management of IT life cycles and IT portfolio management. This scope does not include the information itself, its value, and its life cycle management. The necessary examination of the relevance and the value chain of information should be part of an information and DG. Unlike Weill and Ross, many researchers emphasize that data and information governance are not a subset of IT governance. Accomplishing corporate data quality requires close collaboration among IT and business professionals who understand the data and its business purpose; in addition, both have to follow corporate governance principles. DG broadens the scope of IT governance by considering data quality aspects and processes especially defined to clarify the life cycle of data. According to Khatri and Brown, IT assets refers to technologies (computers, communication, and databases) that help support the automation of well-defined tasks, while information assets (or data) are defined as facts having value or potential value that are documented. Data Governance Institute (DGI)* makes the difference between IT and DG by analogy with the plumbing system: IT is

---

* [http://www.datagovernance.com](http://www.datagovernance.com)
like the pipes and pumps and storage tanks in a plumbing system. Data (information) are like the water flowing through those pipes.

21.1.1 Scope of Data (Information) Governance in the Overall Corporate Governance

DG is an essential part of an overall corporate governance strategy as an equal subdiscipline alongside IT governance. DG is a cross-organization and organization-wide initiative that requires that any barriers between IT and business are brought down and that they are replaced with well-defined roles and responsibilities for both the business areas and the technical areas of the organization. The “who does what and when” question is much more important than the “where” question. Its goal is to govern and manage data and information by organizing the proper usage of data and defining the responsibilities for data and respective data quality.

Furthermore, DG should be clearly distinguished from data management and data quality management (DQM) (Brown and Grant, 2005) because the most common definitional mistake companies make is to use DG synonymously with data management. DG is the decision rights and policy making for corporate data, while data management is the tactical execution of those policies. In other words, DG is not about what specific decisions are made. That is management. Rather, it is about systematically determining who makes each type of decision (a decision right), who has input to a decision (an input right), and how these people (or groups) are held accountable for their role. DG requires a profound cultural change demanding leadership, authority, control, and allocation of resources. DG is the responsibility of the board of directors and executive management and more focused on corporate environment and strategic directions, including other areas beyond DQM, such as Data Security and Privacy and Information Life-Cycle Management (Lucas, 2010). With DG, organizations are able to implement corporate-wide accountabilities for DQM, encompassing professionals from both business and IT units. Figure 21.2, extended from Brown and Grant (2005), illustrates the scope of DG in the corporate governance and the position between governance and management.

In this chapter, in line with most data or information governance publications, we have used synonymously the terms “data” and “information” although they may have different meanings in some contexts.
21.2 Data Governance

21.2.1 Introduction

Nowadays, volumes of data are steadily growing up, and organizations are seeking new ways to exploit their data as enterprise assets in order to fuel their growth. However, data of high quality are a prerequisite for fulfilling these objectives. Moreover, understanding how the data are being used throughout the business is vital for ensuring that legal, ethical, and policy-based requirements are met. Over the past few years, many organizations adopted DQM focusing on collecting, organizing, storing, processing, and presenting high-quality data. The organizational issues that must be addressed, such as defining accountabilities, managing expectation and compliance, avoiding scope creep, and handling political issues, push DQM into an area of conflict between business and IT. On the one hand, DQM has to provide many stakeholders (e.g., CEOs, sales, controlling, procurement, IT, business units, customers, public authorities) with high-quality corporate data. On the other hand, these stakeholders have different interests (e.g., local/regional/company-wide differences), which do not necessarily accumulate to the best achievable result for the company as a whole. Because of these particularities of DQM, large multi-business companies are likely to have difficulties with institutionalizing DQM, that is, defining accountabilities, assigning people accountable for DQM within the organizational structure, and enforcing DQM mandates throughout the company.

DG addresses these particular issues by implementing corporate-wide accountabilities for DQM that encompass professionals from both business and IT. It specifies the framework for decision rights and accountabilities to encourage desirable behavior in the use of data. It establishes organization-wide guidelines and standards for DQM and assures compliance with corporate strategy and laws governing data.

21.2.2 Definition

DG is an emerging discipline with an evolving definition. There are varying definitions of DG because each practitioner tailors DG to its needs and abilities. Here are some definitions that cover almost all the components and goals of DG:

Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.

*The Data Governance Institute*

Data governance is a quality control discipline for adding new rigor and discipline to the process of managing, using, improving and protecting organizational information. Effective data governance can enhance the quality, availability and integrity of an organization’s data by fostering cross-organizational collaboration and structured policy-making.

[IBM, 2007] IBM†

Data governance encompasses the people, processes, and information technology required to create a consistent and proper handling of an organization’s data across the business enterprise.

Wikipedia‡

‡ http://en.wikipedia.org/wiki/Data_governance

© 2014 by Taylor & Francis Group, LLC
The formal orchestration of people, processes, and technology to enable an organization to leverage data as an enterprise asset.

Customer Data Integration Institute*

All these definitions help us understand the aim and principle of DG, which is to help maximize the value of an organization’s data and information assets. In the absence of academic definitions of DG, the working definition used in this chapter was adapted from research on IT governance definition (Weill and Ross, 2004). Thus, DG specifies the framework for decision rights and accountabilities to encourage desirable behavior in the use of data; it encompasses people, policies, procedures, and technologies to enable an organization to leverage data as an enterprise asset.

21.3 Designing Effective Data Governance

The scope of a DG program can be very wide or specific due to the wide variety of backgrounds, motivations, and expectations of DG practitioners. However, a good DG framework ensures that organizational data are formally managed throughout the enterprise as an asset and provides efficient help to the organization to achieve one of the following business drivers:

- Increasing revenue and value
- Lowering costs
- Reducing risks

The following are few descriptions of goals, focus areas, and principles of a DG program. Due to the limited academic research, most of sources come from practitioners on DG, namely, reports by analysts, white papers by consulting companies, vendors, or software manufacturers such as IBM, GDE, Gartner, and Oracle.

21.3.1 Goals of Data Governance

DG is not meant to solve all business or IT problems in an organization. Its goal is to continually produce high-quality data while lowering cost and complexity and supporting risk management and regulatory compliance. The main goals and objectives of DG include the following points:

- Ensure data meets the business needs
- Protect, manage, and develop data as a valued enterprise asset
- Lower the costs of managing data and increase the revenue
- Enable decision making and reduce operational friction
- Train management and staff to adapt common approaches
- Build and enforce conformance to data policies, standards, architecture, and procedures
- Oversee the delivery of data management projects and services
- Manage and resolve data-related issues

21.3.2 Focus of Data Governance

DG frameworks can have a focus on the following areas:

- Policy, standards, and strategy
  The focus on data policies, data standards, and overall data strategies are usually the first step when an organization initiates a DG program. The DG program contributes to the definition of a data strategy and will ensure that consistent policies and standards are used within the

* http://www.tcdii.com
organization. It will steer the creation of new policies and standards and work on the alignment of the existing ones. It will finally monitor that those policies and standards are correctly used with the organization. The main activities of a DG program with a focus on policy, standards, and strategy include the following:

- Establish, review, approve, and monitor policy
- Collect, choose, approve, and monitor standards
- Establish enterprise data strategies
- Align sets of policies and standards
- Contribute to business rules
- Data quality

The desire to improve the quality and usability of the data is the main driver for this type of DG program. It will set the direction for the data quality program. It will help the organization to define and monitor the quality of its data and help to establish ownership, decision rights, and accountabilities to ensure the quality, integrity, and usability of the data.

- Privacy, compliance, and security

Depending on the type of the organization, this area might be more or less critical. Companies within the financial services and health-care industries are main adopters of these programs for apparent reasons. In any case, the DG program will help to protect sensitive data and enforce the legal, contractual, or architectural compliance requirements. This will help the organization to control and limit the risks on its data.

- Architecture and integration

In this domain, the DG program will mainly have a support role by bringing cross-functional attention to integration challenges. It will support initiatives like metadata or MDM programs and help in the definition of architectural policies and standards in order to improve systems integration.

- Data warehousing and business intelligence

The DG program will help to enforce rules and standards in the various source systems of the organization, which will affect the data warehouses and business intelligence data by providing better quality data. It will also help to clarify and promote the value of the data-related projects in general.

- Management alignment

The DG program will promote and measure the value of any data-related projects and initiatives toward management. It will help management to align the different initiatives and monitor for them the data-related projects.

Independently of the area of focus, the main activity should be to identify stakeholders, establish decision rights, and clarify accountabilities.

### 21.3.3 Data Governance Principles

Here are some DG guiding principles that help stakeholders come together to resolve the types of data-related conflicts that are inherent in companies:

- Integrity

DG participants will practice integrity with their dealings with each other; they will be truthful and forthcoming when discussing drivers, constraints, options, and impacts for data-related decisions.

- Transparency

DG and stewardship processes will exhibit transparency; it should be clear to all participants and auditors how and when data-related decisions and controls were introduced into the processes.
Governance of Organizational Data and Information

- **Audit ability**
  Data-related decisions, processes, and controls subject to DG will be auditable; they will be accompanied by documentation to support compliance-based and operational auditing requirements.

- **Accountability**
  DG will define accountabilities for cross-functional data-related decisions, processes, and controls.

- **Stewardship**
  DG will define accountabilities for stewardship activities that are the responsibilities of individual contributors, as well as accountabilities for groups of data stewards.

- **Checks-and-balances**
  DG will define accountabilities in a manner that introduces checks-and-balances between business and technology teams as well as between those who create/collect information, those who manage it, those who use it, and those who introduce standards and compliance requirements.

- **Standardization**
  DG will introduce and support standardization of enterprise data.

- **Change management**
  DG will support proactive and reactive change management activities for reference data values and the structure/use of master data and metadata.

### 21.3.4 Benefits of Data Governance

The scope of a DG program can be very wide and needs to be carefully defined in order to provide the benefits the organization is looking for. It also requires specific expertise to implement DG initiatives in a successful way in order to transform the data in the organization into a real asset. Effective DG does not come together all at once, but it is an ongoing process. Before adopting an approach, it is important to assess the current state maturity of the DG capability. Usually, DG program appears in an organization when it reaches a certain maturity level. In that case, it will be necessary to align the existing rules and processes of different groups by using the following iterative approach:

- **Creates and aligns rules**: the governance program will help in the definition of global rules by coordinating and balancing the need of all the stakeholders. It will define the decision process as well as the roles and responsibilities of all the stakeholders involved.

- **Enforces rules and resolves conflicts**: the governance program will then make sure that those rules and processes are used and applied correctly by all the stakeholders. If necessary, they will be enforced. The governance program will also help to prevent or resolve conflicts by providing ongoing services.

- **Provide ongoing services**: the governance program must support the stakeholders in the application of the processes and rules and by identifying any new opportunities to create new rules or to align or adapt existing ones.

Thus, DG program can provide the following benefits to the organization:

- **Reduce costs and complexity**: reduce duplicate data and management processes and reduce likelihood of errors due to the lack of understanding of data or poor-quality data

- **Support compliance**: quicker access to authoritative data in order to achieve compliance goals and avoid cost of noncompliance penalties

- **Support impact analysis**: increase ability to do useful impact analysis (by providing authoritative business rules, system of record information, and data lineage metadata) and provide a capability to assess cross-functional impacts of data-related decisions
• Help align efforts: assist business teams to articulate their data-related business rules and requirements to IT, architecture, and data management teams; consider requirements and controls in an integrated fashion; craft cross-functional accountabilities; and develop common data definitions to be used across various systems increasing integration and cross-flow of data

• Improve data repositories: provide accountability and support for improving the quality of data in the repository so it can become an authoritative source of information, reduce likelihood of architectural decisions that limit the organization’s ability to analyze its information, and increase ability to find authoritative information quickly

• Improve confidence in data: increase confidence in data-related decisions and output reports, increase ability to make timely data-related decisions, and increase confidence in data strategy by providing a cross-functional team to weigh in on key decisions

21.3.5 Key Data Governance Components

DG is a cross-organization and organization-wide initiative where people work together to establish standards and enforce policies (or rules) over data processes. Figure 21.3 illustrates that an effective DG involves four key components:

Standards and policies: A key function of DG is to establish standards and enforce policies that guide DG program implementation. These include vision and goals of the organization regarding data standards, data management processes, decision-making jurisdiction, responsibilities, enforcement, and controls. The DG approach should be consistent with the organization’s overall mission and stakeholders’ expectations. Program goals should be clearly stated, and it should be made clear how these goals address data content needs, what outcomes will be considered a success, and how the progress will be measured. Furthermore, an organization should evaluate the resources required for the long-term sustainability of the program to ensure it can sustain necessary levels of data quality and security over the entire data lifecycle.

Processes: Establishing and enforcing processes around the creation, development, and management of data is the foundation of an effective DG practice. Companies need to define data and data-related business rules, control access to and delivery of data, establish ongoing monitoring and measurement mechanisms, and manage changes to data.

People: Arguably the most important issue that companies must address when launching a DG initiative is how to manage people in order to design the organizational structure. Companies need to define the
roles and responsibilities within the organization that are accountable for the data. The organization may include several different roles at different levels, involving both business and IT personnel from executive councils to day-to-day implementers, such as data stewards and data analysts. Addressing training and organizational change management issues is also critical if DG programs are to succeed.

Technology: Hypothetically, companies could embark on a DG initiative without an underlying technology infrastructure. Indeed, many organizations launch their initial DG programs using manual tools (spreadsheets, Word documents, etc.) to capture data definitions and document processes. However, most quickly realize that this kind of manual approach is severely limited. It is difficult to ensure high data quality and availability, security is at risk given the ad hoc nature of the approach, and maintaining detailed documentation is an almost insurmountable task. Indeed, it is nearly impossible to achieve the ultimate goals of DG using a manual approach. Thus, technologies can help automate and scale the development and enforcement of DG standards, policies, and processes.

21.4 State of the Art in Data Governance

Both researchers and practitioners consider DG as a promising approach for companies to improve and maintain the quality of corporate data, which is seen as critical for being able to meet strategic business requirements. It receives increasing interest and has recently been given prominence in many leading conferences, such as The Data Warehousing Institute (TDWI) World Conference, the Excellence Quality Information* Association, the Data Management Association (DAMA) International Symposium, the Data Governance Annual Conference, and the MDM Summit. Recently, Gartner has published the 2012 list of Cool Vendors in Information Governance and MDM (Bill et al. 2012). A recent survey conducted by TDWI highlights the increasing number of organizations investing in formal DG initiatives. Eight percent of the 750 responding organizations have deployed a DG initiative; 17% were in the design or implementation phase and 33% are considering it (Russom, 2006). The results also make clear that DG is an emerging practice area, most governance efforts are in the early phases, and many organizations are still working to build the business case for DG. However, only few scientific findings have been produced so far (Otto, 2010) and academic research on DG is still in its infancy. Research on IT governance is more advanced with the first publication released in 1997 (Wende, 2007) and constitutes the starting point for research in DG.

IT governance follows a more flexible approach for the assignment of accountabilities. Early research distinguished two IT governance models: in centralized models, corporate IT performs all IT functions, whereas in decentralized models, business units’ IT performs these tasks. Subsequent research specified more precise IT governance models, acknowledging several IT functions and more than one organizational level. Finally, Weill and Ross (2004) proposed five IT functions, three organizational units, and a distinction between decision and input rights. The combination of these three dimensions resulted in six feasible IT governance models. In conclusion, IT governance research proposes three elements that compose an IT governance model: roles, major decisions areas, and assignment of accountabilities. This idea of IT governance models has been the starting point for research on DG models.

21.4.1 Academic Research on Data Governance

In the academic research area, Wende has proposed a model for DG similar to IT governance models to document the company-specific decision-making framework of DQM (Wende, 2007). His model, shown in Table 21.1, outlines the three components of such a framework, namely, DQM roles, decision areas, and responsibilities. For the components, he identifies typical data quality roles and decision areas and proposes a method to assign responsibilities. The approach respects the fact that each company needs a specific DG configuration.

* http://www.exqi.asso.fr/excellence-qualite-information/
The three components are arranged in a matrix, the columns of the matrix indicate the roles in DQM, the rows of the matrix identify the key decision areas and main activities, and the cells of the matrix are filled with the responsibilities, that is, they specify degrees of authority between roles and decision areas. A company outlines its individual DG configuration by defining data quality roles and decision areas and responsibilities and by subsequently arranging the components into the model. This configuration is unique for each company. However, this model only focused on the DQM accountability aspect of DG and did not address other areas of DG such as compliance and privacy. Similarly, Khatri and Brown (2010) proposed a framework, which inherits an existing IT governance framework with a set of five data decision domains. For each decision domain, they create a DG matrix, which can be used by practitioners to design their DG model. The framework places special interests in data principles and data life cycle in addition to data quality, data access, and metadata. Boris Otto has developed a morphology of DG organization for answering the question: What are the aspects that need to be considered in order to capture the entirety of DG organization? Based on deductive analysis, the morphology constitutes an analytic theory, that is, it identifies and structures the basic concepts of DG organization (Otto, 2010).

Figure 21.4 describes the morphology of DG organization developed in Otto’s research and identifies 28 organizational dimensions.

The first organizational dimension (1) relates to an organization’s goals divided into formal goals (1A) and functional goals (1B). Whereas the former measure an organization’s performance and maintain or raise the value of a company’s data assets, the latter refer to the tasks an organization has to fulfill represented by the decision rights defined. The first set of formal goals specified by the morphology comprises a number of business goals or focus areas of DG. The second set of formal goals specified by the morphology is a two-piece set of information system (IS)/IT-related goals: “increase data quality” and “support IS/IT integration.” Functional goals exclusively relate to the decision areas for which DG specifies certain rights and responsibilities.

The second organizational dimension (2) is the organizational structure. It comprises three aspects, namely, the positioning of decision-making power within the hierarchical structure of an organization.

<table>
<thead>
<tr>
<th>Decision Areas</th>
<th>Executive Sponsor</th>
<th>Data Governance Council</th>
<th>Chief Steward</th>
<th>Business Data Steward</th>
<th>Technical Data Steward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan data quality initiatives</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Establish a data quality review process</td>
<td>I</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Define data producing processes</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Define roles and responsibilities</td>
<td>A</td>
<td>R</td>
<td>C</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Establish policies, procedures and standards for data quality</td>
<td>A</td>
<td>R</td>
<td>R</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Create a business data dictionary</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>Define information systems support</td>
<td>I</td>
<td>A</td>
<td>C</td>
<td>R</td>
<td></td>
</tr>
</tbody>
</table>


R: Responsible; A: Accountable; C: Consulted; I: Informed

© 2014 by Taylor & Francis Group, LLC
Governance of Organizational Data and Information

(2A, “locus of control”), the division of labor and the organizational form (2B) resulting from this, and the allocation of tasks to roles and committees (2C).

21.4.2 Data Governance from Practitioners

Numerous DG frameworks and maturity models have been proposed by different organizations with an attempt to provide guidance in designing and developing effective DG approaches. The DG maturity model is a tool to assess your organization’s current state of DG awareness and effectiveness. These instruments need to be used judiciously and with regard to particular organizational environments (Chartres, 2012). Each organization will set its own goals and methods based on individual business scenarios, plans, and needs. This leads different organizations to have their own focus on specific aspects of DG. Despite this, they still share some common decision domains, such as data quality at least. The following are a few examples (not comprehensive) of DG frameworks from DG organizations.

21.4.2.1 IBM Data Governance Council Maturity Model

The IBM Data Governance Council is an organization formed by IBM consisting of companies, institutions, and technology solution providers with the stated objective to build consistency and quality control in governance, which will help companies better protect critical data. This common forum, where practitioners explore challenges and solutions, has been instrumental in developing benchmarks, best practices, and guides to successful DG. This is known as the Data Governance Council Maturity Model, shown in Table 21.2. The model has 11 categories, where each category is both a

---

**FIGURE 21.4** Morphology of DG organization.
starting place for change and a component in a larger plan. The categories are grouped into four major
DG domains (IBM, 2007):

1. **Outcomes**: data risk management and compliance; value creation
2. **Enablers**: organizational structures and awareness; policy; stewardship
3. **Core disciplines**: DQM; information life cycle management; information security and privacy
4. **Supporting disciplines**: data architecture; classification and metadata; audit information logging and reporting

### 21.4.2.2 Data Governance Institute Framework

The DGI framework* provided by the DGI is a logical structure for classifying, organizing, and communicating complex activities involved in making decisions about and taking action on enterprise data. **Figure 21.5** describes the architecture of the framework with the 10 universal components:

1. Mission and vision
2. Goals, governance metrics and success measures, and funding strategies

* [http://www.datagovernance.com/fw_the_DGI_data_governance_framework.html](http://www.datagovernance.com/fw_the_DGI_data_governance_framework.html)
21.4.2.3 Data Excellence Framework

The Data Excellence Framework (DEF) provided by Global Data Excellence (GDE)* describes the methodology, processes, and roles required to generate business value while improving business processes using data quality and business rules (El Abed, 2009b). The framework supports the creation of a new cultural shift focused on data excellence, motivating the broader team and supporting collaboration between the stakeholders. DEF is more than improving data quality and reducing costs; it is about taking a sustainable approach to an organization’s data assets to drive value from the data. A key difference about the DEF is that it is focused on generating value in comparison to most initiatives relating to data, which are only focused on reducing costs. The key success of DEF is that it is grounded on both academic research (Feuto and El Abed, 2012) and practical experiences.

* http://www.globaldataexcellence.com
Thus, the data excellence maturity model helps organizations worldwide successfully move from chaos to data excellence stage.

Innovative capabilities in maximizing the value of enterprise data through more effective ways in

1. Aligning and linking business objectives to data management processes via business rules and data quality
2. Measuring and visualizing the business impact of data quality compliance and the value generation of high data quality
3. Organizing and executing sustainable DG processes based on a proven DEF methodology supported by a product solution called the data excellence management system (DEMS)

earned GDE a spot on Gartner’s Cool Vendors in Information Governance and MDM for 2012 list (Bill et al. 2012).*

Through the Cool Vendors list, Gartner shines light on companies that use information to enable business outcomes and improve and enrich the meaning and use of master data.

### 21.5 Data Governance Journey: Application and Business Cases with the Data Excellence Framework

#### 21.5.1 Data Excellence Framework: Govern by Value

As already mentioned in the previous section, DEF describes the methodology, processes, and roles required to measure and govern the business impact of noncompliant data on business activities and the business value generated by fixing the data (El Abed, 2009a,b). The leading concept is the governance by value in order to enable the new paradigm shift as illustrated in Figure 21.6. It supports the creation of a new cultural shift focused on new concept called data excellence, for example, elevating data to

* [http://www.gartner.com/id=1984515](http://www.gartner.com/id=1984515)

**FIGURE 21.6**  Govern by value: the new paradigm shift.
DEF is based on four value pillars (shown in Figure 21.7) that are essential to the survival of any organization or enterprise in the information age: agility, trust, intelligence, and transparency.

These characteristics are fundamental value pillars to enable business sustainability and support economic growth:

- **Agility** is needed to react to external and internal changes and ensure prompt and successful integration that supports fast business: how fast can you change a process, introduce a new product, or merge a new business.
- **Trust** is associated with the integrity of the data, which induces confidence in enterprise image, brands, and products; empowers business partners, employees, and stakeholders; and ensures legal and compliance traceability.
- **Intelligence** at all levels of the enterprise leads to better and flawless execution, operational efficiency, and accurate financial consolidation based on just-in-time quality data from reporting systems and applications.
- **Transparency** is critical to the organization’s performance as it increases visibility and collaboration across and outside the enterprise: not about showing everything but controlling what should be communicated, for example, social responsibility.

These fundamental value pillars result in new ways of working and will lead to maximizing the value and lowering the cost of using data.

### 21.5.2 Managing Data as a Company Asset

A key difference about the DEF is that it is focused on generating value in comparison to most initiatives related to data that are only focused on reducing costs. Data are an asset, and an asset has a value. The main difference between a tangible asset and the data asset is that the value of data is dynamic and will depend on the context of its usage. Out of the context of usage, the value of data is zero. Moreover, it has a cost, and when the value is not demonstrated, the organization tends to cut this cost.
Price, Cost, and Value

Price and cost are very familiar notions, but value is less understood as a concept. However, value is present in our daily life: We attach sentimental value to our belongings especially when we have to separate from them; we buy goods according to the use or the amount of pleasure they will provide to us. Intuitively, we don’t attach value to what is of no use to us although it may have a price or a cost. If you have a fridge filling your needs, it is of no use to buy another one even if there is a bargain (e.g., a low price) at the local discount store: therefore, the item has no value for you (except if you think to sell it at a higher price). The price is the monetary amount required to buy an item. It is fixed by the seller and is generally the same for all buyers. It can be negotiated depending on certain factors (volume, delivery mode, etc.). Cost is linked to cost price, which is the total amount of ownership required for the producing and the utilization of an item. Taking the fridge case, its price is the posted price to buy it and cost is the price plus the total expense to run it through its use: electricity, repair, maintenance, etc. Cost is also objective (same for all users).

However, value is subjective and contextual depending on the use we make of it. A bottle of water in the desert can be worth your life (i.e., a lot of money if your life depends on this water), but the same bottle has little value in your daily life as you have plenty of water on the tap. Hence, the value is contextual and instantiate itself at transaction time. Value is therefore subjective as it can vary for different individuals and vary in time for the same person.

Since Frederick Taylor invented his theory of scientific management to improve economic efficiency in the industrial age, enterprises have separated the value generation and the cost management. The former was given to the “business” and the latter to the support functions whose mission is to rationalize and optimize processes (the means). This approach has proved efficient and effective in increasing volumes and reducing costs in an industrial world. However, since we have entered the information age, this separation has disconnected the value creation from the means supporting it and has, in a way, “devalued” the support functions like IT or data management.

21.5.2.1 What Are the Three Barriers to Manage Data as an Enterprise Asset?
The main barriers are as follows (Figure 21.8):

- **Insufficient alignment between business managers, data management, and IT**: data are noncompliant and are not fit for purpose, and as a result, transactions fail. Despite investing millions in infrastructure and architecture, organizations have not invested in the data. However, data are a key differentiator for organizations when all organizations have similar ERP systems implemented by the same consulting firms. The business, the data management team, and IT are not aligned, resulting in problematic decision outcomes.
- **Inability to demonstrate the business impact of bad data management**: nobody knows the cost of poor data quality and nobody is able to quantify the potential benefit of fixing the data. Worse, nobody knows which data are critical and what the impact of this bad data on the business is. Building a business case to invest in data quality is always difficult, and as a result, the problem still exists. People have lots of anecdotes but find it difficult to quantify their thoughts.
• **Unclear roles and responsibilities to govern data as an asset**: most organizations are struggling to maintain momentum to achieve sustainable business process improvement and flawless execution through data quality improvement processes. This is because nobody feels accountable or responsible for data quality. There is no understanding of the roles or processes needed to sustain high-quality data to fulfill the goals of a DG process. There is no easy way to motivate or sustain the data management team or IT to support the business goals. The DEF provides a method to lower each of these three barriers.

### 21.5.3 Data Excellence Framework Methodology

The DEF’s methodology emphasizes a value-driven approach for enabling business excellence through data excellence. Data excellence becomes an imperative to unlock the enterprise potential and enable sustainable value generation. Thus, data are made visible so that they can become resources for the enterprise to make informed decisions.

In order to lower the three barriers to manage data as a company asset and to maximize the business value of enterprise data, DEF proposes a three-step approach presented at Figure 21.9.

#### 21.5.3.1 Align and Link: Business Excellence Requirements

The purpose here is to align and link business objectives to data management processes via business rules and data quality (see example in Figure 21.10). We introduce the concept of business excellence requirements (BERs): the BER is a prerequisite, business rule, standard, policy, or best practice that business processes, transactions, and data should comply with in order to have the business goals flawlessly executed to generate value. Defining BERs is defining compliance on data quality. The role of the business rules concept implementation enables to action the data quality key performance indicators (KPIs) at the record level and to address the root cause (El Abed, 2008a).

#### 21.5.3.2 Measure and Visualize: DEI and KVI

We measure and visualize the business impact of data quality compliance and the value generation of high data quality. To do so, we use a data excellence index (DEI) measuring the ratio of compliant data and a key value indicator (KVI) measuring the value of compliant data and the value of noncompliant
data (the impact). In compliant and noncompliant data DEIs and KVIs, value and impact can be visualized through different angles (geography, business unit, process, business rule, role, etc.) using the contextual polarization. This helps organizations in

- Providing global visibility of data quality to the whole organization at all levels.
- Publishing the data KPIs monthly on the operation site on the intranet at all levels.

**FIGURE 21.9** The three steps to maximize the business value of enterprise data.

- **Scope:**
  - All vehicles (e.g., one million)
- **Rule**
  - if then “Stop”
  - else then “Prepare to Stop”
  - else then “Go”
- **KPI results**
  - 93.4% for consistency
- **Business impact**
  - 6.6% of indecision (inconsistent with the rule), may lead to crashes, i.e., 66,000 instances
- **Business value**
  - Cost of each crash is 15k. Preventing 66,000 instances may save 990 million

**FIGURE 21.10** Example of how a business rule links to business impact and business value.
• Presenting the data KPIs at the top management level and at the market management level each month (El Abed, 2008b).
• Measuring the business value through KVIs (aggregated sum and number for real business impact) enables the business users to visualize how DG can help them get there fast (Michellod and El Abed, 2010).

21.5.3.3 Organize and Execute: Accountability and Responsibility

In step 3, the business impact of poor-quality data must be reduced by fixing the wrong data and identifying and analyzing the root causes. In order to manage it efficiently and effectively, new DG must be established including the business, the data management, and IT. The business managers who are accountable and responsible of the value generated in the enterprise will take accountability for the BERs and the associated data, DEIs and KVIs. They will decide which data should be corrected to generate more value by reducing the impact. They will also define accountability and responsibility around the business rules KVIs with business stakeholders (Orazi and Gessler, 2010). Data managers, in charge of the data management of the BER-associated data, will take responsibility for these data. They will correct data and optimize processes upon business order. IT will provide and support the systems necessary to manage these data. The Global DEF is modifying the role of the MDM manager, giving the role higher visibility and strategic positioning (Delez, 2010). The higher the person is in the hierarchy, the faster he or she understands how powerful and useful the DEF to achieve their objectives, and people can no longer question the importance of data errors—with the KVIs, the business impact and value generation are now obvious (Michellod and El Abed, 2010).

In order to facilitate the implementation of this three-step approach, a process including eight questions is provided as shown in Figure 21.11. A simplified view of DEMS, which is a collaborative and evolutionary system to fully support the implementation of DEF capabilities, is presented in Figure 21.12.

FIGURE 21.11 The eight critical questions linking your business and your data.
Figure 21.12 shows the general architecture of DEMS. However, it does not show how DEMS could be integrated in the enterprise environment. Thus, Figure 21.13 provides a detailed functional architecture of DEMS in the enterprise environment. It makes a clear difference between the master data and the transactional data and presents the data sharing sphere and how it can be implemented and governed.
21.5.4 Business Case

We are now going to illustrate the application of the DEF through a simple business case in a financial institution chosen from many possibilities: the allocation of loans by a bank to its customers.

**Business Case: Allocation of Loans by a Bank to Its Customers**

SmartBank is a retail bank who has two brands: BankA and BankB both operating in two regions, West and East.

Strategic goal: to govern the risk linked to the distribution of loans and maximize their value

Scope: all loan applications

Data: customer data

21.5.4.1 Scope

The scope is the set of data objects and business processes that are defined before starting the exercise. This scope is dynamic and can be extended or reduced according to the business critical requirements.

In this example, the scope is the set of loan applications at BankA and BankB in the East and West regions. This scope could be reduced to BankA in the East region or extended to another brand of SmartBank (BankC) or to all regions.

The business process is the loan validation process and data objects are the loans (transactions) and customer data.

21.5.4.2 Business Excellence Requirements

The BERs, a unique value proposition representing the backbone of the DEF, are a prerequisite business rule, standard, policy, or best practice that business processes, transactions, and data should comply with in order to have the business goals flawlessly executed to generate value.

21.5.4.2.1 Example

In this example, the SmartBank loan business manager has defined three BERs for the loan allocation process:

BER 1: “a customer record of default must be provided for any loan application if the amount is over or equal to 10000€.” This record states whether the customer had defaulted on a loan repayment in the past.

BER 2: “no loan is granted to minor customers (e.g., under 18).”

BER 3: “each application must be accompanied with a payslip.”

It is the responsibility of each branch account manager to make sure that any loan application is compliant with these three BERs before they are examined by the SmartBank loan committee for approval.
21.5.4.3 Data Excellence Index

21.5.4.3.1 Definition and Role

The DEI is the instrument used to measure the compliance of data records with the BERs. It consists of the following:

- The percentage of data records compliant with a collection of BERs
- The list of compliant and noncompliant records

Each data record impacts the DEI (positively or negatively) only once regardless of the number of BER applied to this record.

The DEI results are used to evaluate the value and impact of data on business processes and transactions. The BER concept makes the DEI actionable at the record level allowing finest root cause analysis and surgical DG.

21.5.4.3.2 Example

In this example, the three BERs have to be applied to the two data objects: the loans and the clients included in the case scope (see preceding text). The extraction of data from BankA and BankB ISs yields the following data:

<table>
<thead>
<tr>
<th>Records</th>
<th>Région</th>
<th>Bank</th>
<th>Account Manager</th>
<th>Customer Name</th>
<th>Customer First Name</th>
<th>Age</th>
<th>Income</th>
<th>Loan Amount</th>
<th>Pay slip Provided</th>
<th>Default Record Provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record 1</td>
<td>East</td>
<td>BankA</td>
<td>Claus French</td>
<td>Dupont Jean</td>
<td>Jean</td>
<td>75</td>
<td>30,000</td>
<td>9,000</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Record 2</td>
<td>East</td>
<td>BankB</td>
<td>Claus French</td>
<td>Proust Marcel</td>
<td>Marcel</td>
<td>50</td>
<td>40,000</td>
<td>11,000</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Record 3</td>
<td>East</td>
<td>BankA</td>
<td>Julian Wood</td>
<td>Flaubert Gustave</td>
<td>35</td>
<td>25,000</td>
<td>11,000</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Record 4</td>
<td>East</td>
<td>BankB</td>
<td>Gilbert Sullivan</td>
<td>Monet Claude</td>
<td>22,000</td>
<td>5,000</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 5</td>
<td>East</td>
<td>BankA</td>
<td>Claus French</td>
<td>Courbet Gustave</td>
<td>57</td>
<td>60,000</td>
<td>25,000</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Record 6</td>
<td>East</td>
<td>BankA</td>
<td>Julian Wood</td>
<td>Dupont Jean</td>
<td>75</td>
<td>31,000</td>
<td>5,000</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Record 7</td>
<td>West</td>
<td>BankA</td>
<td>Julian Wood</td>
<td>Hugo Victor</td>
<td>52</td>
<td>58,000</td>
<td>17,000</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 8</td>
<td>West</td>
<td>BankB</td>
<td>Gilbert Sullivan</td>
<td>Cezanne Paul</td>
<td>40</td>
<td>1,500</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Record 9</td>
<td>West</td>
<td>BankA</td>
<td>Claus French</td>
<td>Duchamp Marcel</td>
<td>32</td>
<td>15,000</td>
<td>2,500</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 10</td>
<td>West</td>
<td>BankB</td>
<td>Gilbert Sullivan</td>
<td>Lupin Arsene</td>
<td>XXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td></td>
</tr>
<tr>
<td>Record 11</td>
<td>West</td>
<td>BankA</td>
<td>Julian Wood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are two records with specific features:

1. Record 10 could not be retrieved for security reasons (no access rights).
2. Record 11 was not found, although account manager, Julian Wood, reported it in the CRM.
A simple check of each record against each BER yields the following results:

<table>
<thead>
<tr>
<th>Records</th>
<th>Customer Name</th>
<th>Customer First Name</th>
<th>Age</th>
<th>Income</th>
<th>Loan Amount</th>
<th>Pay Slip Provided</th>
<th>Default Record Provided</th>
<th>BER 1</th>
<th>BER 2</th>
<th>BER 3</th>
<th>All BERs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record 1</td>
<td>Dupont</td>
<td>Jean</td>
<td>75</td>
<td>30,000</td>
<td>9,000</td>
<td>No</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Record 2</td>
<td>Proust</td>
<td>Marcel</td>
<td>50</td>
<td>40,000</td>
<td>11,000</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Record 3</td>
<td>Flaubert</td>
<td>Gustave</td>
<td>35</td>
<td>25,000</td>
<td>11,000</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Record 4</td>
<td>Monet</td>
<td>Claude</td>
<td>22,000</td>
<td>5,000</td>
<td>Yes</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Record 5</td>
<td>Courbet</td>
<td>Gustave</td>
<td>57</td>
<td>60,000</td>
<td>25,000</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Record 6</td>
<td>Dupont</td>
<td>Jean</td>
<td>75</td>
<td>31,000</td>
<td>5,000</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Record 7</td>
<td>Hugo</td>
<td>Victor</td>
<td>52</td>
<td>58,000</td>
<td>17,000</td>
<td>Yes</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Record 8</td>
<td>Cezanne</td>
<td>Paul</td>
<td>40</td>
<td>1,500</td>
<td>Yes</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Record 9</td>
<td>Duchamp</td>
<td>Marcel</td>
<td>32</td>
<td>15,000</td>
<td>2,500</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Record 10</td>
<td>Lupin</td>
<td>Arsene</td>
<td>XXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Record 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

The DEI is the ratio of compliant records against the total number of records. Noncompliant records for all three BERs are records #1, #4, #7, #10, and #11. Action in the further steps will be focused on these records (see subsequent sections).

It should be noted that

- Although each BER have a DEI of 72.7%, the global DEI is lower (54.5%), which shows that this model is different from the traditional aggregation models (OLAP).
- Each record counts only for one in the final calculation of DEI (records 10 and 11) although they can be noncompliant to several BERs.

At this stage, we have a level of compliance (DEI) similar to a KPI. However, we have no information regarding the impact in value of noncompliant transactions (records).

21.5.4.3.3 Views

Using the same method, other DEIs can be calculated according to different views: per region, per bank, per account manager, per customer, etc. It is easy to check that the DEI by region yields the following results:

<table>
<thead>
<tr>
<th>Regions</th>
<th>BER 1</th>
<th>BER 2</th>
<th>BER 3</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>East</td>
<td>100.00%</td>
<td>83.33%</td>
<td>83.33%</td>
<td>66.67%</td>
</tr>
<tr>
<td>West</td>
<td>40.00%</td>
<td>60.00%</td>
<td>60.00%</td>
<td>40.00%</td>
</tr>
<tr>
<td>All regions</td>
<td>72.73%</td>
<td>72.73%</td>
<td>72.73%</td>
<td>54.55%</td>
</tr>
</tbody>
</table>

Although DEIs by regions (and for other views) are different, the global DEI stays unchanged as the whole set of records is identical and noncompliant records remain the same.

21.5.4.4 Data Excellence Dimensions

Another view called data quality view according to data quality dimensions (called data excellence dimensions (DEDs) in the framework) can be calculated. DEDs are as follows:

- **Uniqueness**: The uniqueness dimension is to allow the identification in a deterministic way of an entity, a relationship, or an event instance within a specific context to execute a specific business process (e.g., the business process must be able to identify the unique address for sending the invoice for a specific customer). In the example, Records #1 and #6 raise an issue regarding the uniqueness dimension.
21 Data, Information, and Knowledge Management

- **Completeness**: The completeness dimension is to ensure that the data required for a successful execution of a process in a specific domain and context are present in the database (e.g., to pay your supplier, you need its bank account number). In the example, records #4 and #8 are not complete.
- **Accuracy**: The accuracy dimension is to ensure that the data reflect a real-world view within a context and a specific process. In the example, either record #1 or #6 may not be accurate.
- **Non-obsolescence**: The non-obsolescence dimension ensures that the data required to execute a specific process in a specific context are up to date. Either record #1 or #6 may be obsolete.
- **Consistency**: The consistency dimension is to ensure that the data values are delivered consistently across all the databases and systems for the execution of a specific business process in a specific context. In the example, records #1 and #6 may not be consistent.
- **Timeliness**: The timeliness dimension is to enable first-time right delivery of data required to enable flawless execution of business processes and fulfilling the service-level agreements. In the example, record #11 has not been delivered on time.
- **Accessibility**: The accessibility dimension is to ensure that people, systems, or processes have access to data according to their roles and responsibilities. In the example, record #10 is not accessible.

Each DED provides a DEI, and the DEI resulting from all DED is called the GDE index.

### 21.5.4.5 KVIs

#### 21.5.4.5.1 Definition and Role

A KVI is a measurement of the value and impact of the DEI on the business operations. It consists of

- **Value**: related to the collection of data elements of the DEI that comply with the BERs
- **Impact**: related to the collection of defective data elements of the DEI that do not comply with the BERs

A fundamental KVI represents the key business value and impact affected by a specific BER. A BER could impact multiple KVIs and a specific KVI can be affected by multiple BERs.

The KVI is a fundamental deliverable of the DEF: Through the KVI, the DEF enables multifocal governance linking business management and data management. The business transaction is, therefore, considered as a key component that enables managing data as a company asset. The DEI becomes the pivot that links the BER, the KVI, and the data elements.

#### 21.5.4.5.2 KVI's Guiding Principles

In order to be efficient and effective, the KVI must fill the following requirements:

- **Tangible value**: a KVI must have tangible value to be considered as an asset linked to the data through the DEI.
- **Stewardship**: each KVI must have a clear steward in the business function and at all organizational levels. The steward from the function must practice their accountability striving for the improvement of the said KVI.
- **BERs**: each KVI must be based on defined BERs.
- **Adequate**: the KVI should reflect real and important business factors to enable control or analysis/evaluation of business operations and to enable sharp decision support.
- **Simple and intuitive**: the KVIs should convey their meaning in a compelling way. Their impact on business should be clearly understandable using common sense.
- **Precise and unambiguous**: the measures should provide no room for misinterpretation or result negotiation.
- **Comprehensive**: they should describe all aspects that are relevant to achieving expected results.
• **Manageable:** calculation of indicators and their presentation to interested parties (reporting) must be straightforward and must not involve personnel interaction.

• **Limited in number:** the whole set of measures should be looked at when evaluating performance.

### 21.5.4.5.3 Example

Back to our loan application management example, the loan business manager of SmartBank can decide to define the KVI as the amount of the loan for each application. The rationale behind it is that the risk for the bank is the amount of money lent to its customers; it is also linked to the risk to lose a transaction if some data are missing for a loan application. More widely, a specific KVI linked to the loss of customer derived to a bad loan application management could be considered.

From the DEI calculation, we have derived that records #1, #4, #7, #10, and #11 are noncompliant: the following table shows the KVI and impact according to the KVI definition for each of the three BERs, all BERs.

The value and impact per BER are calculated by summing, respectively, the loan amount of all compliant records and noncompliant records. For example, for BER1 only, record #7 is noncompliant. The impact is the loan amount of record #7, that is, 17,000€. The KVI (value) is the sum of the loan amount of the compliant records, that is, 70,000€. The total value is 87,000€. If the ratio KVI/total value of 80.9% is slightly better that the DEI (72.7%), the impact can be zeroed by fixing record #7 by providing the information related to the customer default record. The value impact % column in the table is the value impact distribution across noncompliant records. It can be seen that fixing record #7 and record #1 (surgical approach to data quality) would cut the impact by 84%, which may be sufficient for the business.

<table>
<thead>
<tr>
<th>Records</th>
<th>Customer Name</th>
<th>Customer First Name</th>
<th>Age</th>
<th>Income</th>
<th>Loan Amount</th>
<th>Pay Slip Provided</th>
<th>Default Record Provided</th>
<th>BER 1</th>
<th>BER 2</th>
<th>BER 3</th>
<th>All BERs</th>
<th>Value</th>
<th>Impact %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record 1</td>
<td>Dupont</td>
<td>Jean</td>
<td>75</td>
<td>30,000</td>
<td>9,000</td>
<td>No</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>29.03%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 2</td>
<td>Proust</td>
<td>Marcel</td>
<td>50</td>
<td>40,000</td>
<td>11,000</td>
<td>Yes</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 3</td>
<td>Flaubert</td>
<td>Gustave</td>
<td>35</td>
<td>25,000</td>
<td>11,000</td>
<td>Yes</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 4</td>
<td>Monet</td>
<td>Claude</td>
<td>22,000</td>
<td>5,000</td>
<td>Yes</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>16.13%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 5</td>
<td>Courbet</td>
<td>Gustave</td>
<td>57</td>
<td>60,000</td>
<td>25,000</td>
<td>Yes</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 6</td>
<td>Dupont</td>
<td>Jean</td>
<td>75</td>
<td>31,000</td>
<td>5,000</td>
<td>Yes</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 7</td>
<td>Hugo</td>
<td>Victor</td>
<td>52</td>
<td>58,000</td>
<td>17,000</td>
<td>Yes</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>54.84%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 8</td>
<td>Cezanne</td>
<td>Paul</td>
<td>40</td>
<td>1,500</td>
<td>Yes</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 9</td>
<td>Duchamp</td>
<td>Marcel</td>
<td>32</td>
<td>15,000</td>
<td>2,500</td>
<td>Yes</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 10</td>
<td>Lupin</td>
<td>Arsene</td>
<td>XXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that if DEIs for the three different BERs are identical (72.7%), the KVIs are different because each record does not represent the same business value. This is the main reason why the management by KPIs (DEIs) of the industrial age is inadequate as a good KPI may hide a terrible KVI leading to a severe impact for the business. Alternatively, a bad KPI could generate expensive process improvements, which could prove useless with regards to high KVIs.

In the information age, enterprises must shift from a KPI-driven organization to a KVI-driven one. For emphasizing the point, Patrick Kabac has concluded in Givaudan MDM governance initiative that: "Business understands facts, not theory" (Kabac, 2010).
21.5.5 Data Excellence Process

Data and information governance mission: The establishment of accountability and responsibility at all levels to maximize the business value of enterprise data

Accountability
The willingness and commitment of business executives to be accountable for the definition and management of the BERs of their area and related KVI targets

Responsibility
The willingness and commitment of data managers to be responsible for individual data records assuring the compliance with BER

DG can be operationalized through collaborative governance networks where data stewards can play a pivotal role between business, data management, and IT.

21.5.5.1 Roles

We have defined the key roles at the center of the DEF:

- **Data accountables**: they have BER responsibility. They are part of line of business managers responsible for processes and applications. They define, approve BER and KVI, follow the results, and drive priorities of actions. There is a tree of data accountability called the governance tree. In our example, the loan account manager is data accountable for the loan application process of SmartBank. He or she has responsibility over the local account managers in BankA and BankB in all regions for the loan application process: this defines the governance tree.

- **Data responsibles**: they are the data managers. Part of the local data management, they are responsible for individual source data records for each BER and defect management.

In our example, we suppose that for each bank (BankA and BankB), there is a data responsible per region for customer and loan data. Each data responsible in each bank reports to the bank’s global data responsible reporting to the SmartBank’s data responsible: this defines the responsibility tree. Note that there could be different data responsibility for customer data and loan application data.

- **Data stewards**: link the business, data management, and IT. They coordinate the setup of the DG structure, DEI, KVIs, and BERs with data accountable. They collaborate with data responsibles and IT.

In our example, the SmartBank loan account manager has named a data steward dedicated to the loan application management.

- **IT**: part of the data integration team, they support the implementation and development of DEIs and KVIs.

In our example, IT is a central team in the headquarters of SmartBank (Figure 21.14).

21.5.5.2 Description

This chart is a summary of the continuous data excellence process described just before. It demonstrates explicitly how in the DEF, data and information are governed and driven from the business requirements, by the business through data stewards into data management. The data accountable will define the context, BERs and KVIs. By setting target for the KVIs, he or she will give mandate to the data steward to meet it by fixing data issues that impact the value.

21.5.5.3 Example

In our example, the loan account manager has given instruction to the data steward to have a KVI ratio above 94% (i.e., 81,780€). The data steward notes that the current ratio is 64.37%: by fixing records #1 and #7,
he will add 26,000€ to the KVI (56,000€) and reach a 94.25% ratio. Record #1 is managed by the data responsible of BankA in the East, and record #7 is managed by the data responsible of BankA in the West. The data steward will send a requirement to both data managers of BankA to fix those records. Record #1 will be fixed by asking the account manager of BankA in the East region to request the payslip for Jean Dupont and update the loan application system. Record #7 will be fixed by requesting the default record of the customer.

If the root cause analysis shows that there is a defect in a process in BankA, some action might be decided to fix it in order to avoid further defects and value impact.

### 21.5.5.4 Conclusion

This example shows that the DEF is not a traditional scientific total data quality approach driven by KPIs, but a pragmatic contextual data quality approach driven by value (KVI).

### 21.5.6 Governance Model

This new approach to cultivate the data assets requires a totally different mind-set as shown in Figure 21.15.
Shifting from

- Data as a personal power asset to data as an enterprise asset: moving from traditional hierarchal organization to collaborative networks organization
- Compliance data management to a BER- and KVI-driven governance: moving from top-down standardization and MDM approach to “fit for purpose” value-driven DG
- Knowledge within experts to knowledge embedded in the business culture: moving the paradigm from “having the knowledge is the power” to “sharing the knowledge is the power”
21.5.7 Maturity Model

Obviously, a KPI-driven organization will not move overnight to the value-driven management. This is a progressive step-by-step evolution illustrated by the data excellence maturity model (Figure 21.16), which tracks organization’s progress from the early stages of DG (which is described as “chaotic”) to the most mature stages where data are utilized as a core enterprise asset (described as “predictive”; Figure 21.16).

This can be implemented smoothly with short-term returns and small investments for the first phases into a virtuous circle that will lead the enterprise into the data excellence culture (Figure 21.17).

You can achieve DG and high data quality without earthquake in the organization and without costly initiatives. Be progressively incremental with clear communicated business objectives and value of each business rule to be implemented (Figure 21.17). For instance, DG initiative at Nestlé was started with one (1) business rule (El Abed, 2007).

21.6 Conclusion and Future Research on Data Governance

The industrial age was focused on “how” to improve and optimize processes: this approach has proven successful for increasing volumes and lowering costs. However, in the information age characterized by a deluge of data and rules, processing data in a scientific way proves to be time consuming and very costly. In the twenty-first century, business is driven by data and better governance of data means better business. Thus, more organizations start to recognize that they need better data management processes capable of combining business-driven and technical perspectives in order to respond to strategic and operational challenges demanding high-quality corporate data. Both researchers and practitioners consider DG as a promising approach for companies to improve and maintain the quality of corporate data, which is seen as critical for being able to meet strategic business requirements, such as compliance or integrated customer management.

DG specifies the framework that encompasses professionals from both business and IT for decision rights and accountabilities to control and to protect the investments made in data in order to transform them in a real asset for the company. It helps an organization define itself and its strategy and then embed those definitions into core business data ensuring that all of its business processes, systems, projects, and strategies are aligned to create trust and enable consistent executive decision making. DG is to the organization what breathing is to the body. When it is working right, you hardly notice, but when it is not, everything quickly becomes more difficult and labored.

We have been discussing in this chapter the implications of DG for both academics and practitioners and presented the state of the art in the domain. Both sides have proposed a number of DG models and frameworks, which can assist organizations to take off with a DG program. Even though slight differences exist between these models, they all share the basic principle of what a DG program should do: “Defining how to deal with data-related issues.”

However, a number of limitations need to be considered. Many models are based on organizational structures that should fit all companies alike. They have, thereby, neglected the fact that each company requires a specific DG configuration that fits a set of contingencies. Furthermore, knowledge has been transferred from IT governance research to DG, and the research on contingencies influencing IT governance models is used as starting point for the contingency research on DG. So far, the proposed contingencies and their impact lack validation in the context of DG. DG model might be characterized by additional parameters, such as a time dimension, which respects the fact that the configuration might evolve over time. In the literature, these issues are investigating and the future direction of DG model is to look toward a contingency approach for DG. Also, one key is to shift the approach from “data-driven” to “value-driven” mind-set meaning that the “business transaction” must be put at the center of DG rather than “data” as advocated by data management trends. This approach for defining and implementing the DG model would help companies to align with shareholders’ accountability, to assure a sustainable just-in-time data quality enabling the realization of the ultimate “business transaction”
between the shareholders and the customers. The purpose of the “enterprise” is to realize the “business transaction” between the shareholders and the customers and exchange the expected value for each. Therefore, DG must be a mean to enable flawless execution of business strategy maximizing the enterprise potential and revealing the business value of the ultimate transaction.

References


Delez T. Data governance driven master data management, developing and optimizing the MDM function within business operations, *Firmenich Business Case for the Proceedings of the Master Data Management Conference*, Berlin, Germany, September 2010.


El Abed W. Data excellence framework from vision to execution and value generation in global environment, *Proceedings for the Information and Data Quality Conference*, San Antonio, TX, 2008b.


O’Kane B., Radcliffe, J., White, A., Friedman, T., and Casonato, R. Cool vendors in information governance and master data management, Gartner Research, ID Number: G00226098, April 2012.


