

# The Effects of Data Governance in Theory and Practice



N.L. Martijn  
is a consultant at KPMG Advisory N.V.  
martijn.nick@kpmg.nl



R.A. Jonker  
is a partner at KPMG Advisory N.V.  
jonker.ronald@kpmg.nl

**Nick Martijn and Ronald Jonker**

Although we have considerable experience in implementing data governance as part of a data management organization and we have seen that the effects are consistently positive, it remains difficult to establish the exact effects of data governance. In this article, we pinpoint the relationship between data governance and an organization (theory). Subsequently we use client cases to examine the relationships between data governance and business performance (practice). These two together provide insight in the effects of data governance.

## Introduction

Enterprise data have become highly important. They were initially seen as a by-product of business processes, for example for financial recording. Nowadays, data are considered to be a valuable asset in and of themselves. This value is primarily provided through two important applications: business processes and reporting. First, increasingly complex and globalizing business processes require the support of reliable data. For example, the international container shipping industry requires timely and accurate data to feed its logistical planning; insufficient data quality leads to inefficiencies in the supply chain. Second, enterprise data are used for reporting to, for example, the regulator. Tax reporting, for instance, depends on reliable recordings of the company's purchases. Put differently, enterprise data are being used to steer and direct business performance, while simultaneously enabling compliance.

Notwithstanding its importance, the implementation of data management at companies is not always sufficient. Although many companies have well-managed IT systems, the responsibilities and accountabilities with regard to enterprise data are mostly unclear. We see that, when no clear policies, rules and controls are defined within the organization about who is responsible for which data, the

overall quality of the data is likely to deteriorate. Poor data quality leads to inefficiently organized business processes, and cause companies to run the risk of non-compliance with the relevant regulator(s).

KPMG has broad experience in advising clients about sustainable data management using data governance (such as [Waal12]). In most cases, we see that clients are satisfied with the solution provided and experience the added value of data governance. Nevertheless, it appears to be difficult to get sponsorship for a sustainable data management organization in advance. We consider an insufficient understanding of the effects and functioning of data governance in an organization to be the root cause of this. How does data governance affect an organization? And what is the quantitative effect on business performance? In other words: what is the business case for improving data governance?

In this article, we provide a view on data governance in an organization in order to elucidate the manner in which it affects enterprise data and business processes. By outlining cases in the food, retail and logistics sector, we set the first step towards gaining insight into the effects of data governance on an organization.

## What is data governance?

The starting point of this article should be the definition of data governance. A commonly used definition has been formulated by Thomas ([Thom06]): “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.” Essentially, data governance is an overarching concept that defines who is responsible and has which role for what data at what point in the process. Roles and responsibilities for enterprise data alone, however, are not sufficient to improve data quality in a sustainable way. Data governance needs to be complemented by other elements of data management that together form a comprehensive data management framework.

We defined a framework for data management (see Figure 1), based on our experience in the field of Enterprise Data Management and commonly used data management frameworks such as DAMA DMBOK ([Mos10]). The framework consists of nine building blocks, representing the approach for Enterprise Data Management. *Data Governance* is part of the approach, together with, for example, data *Definitions, Standards & Quality*, providing a single source of truth within an organization. The package of measures creates a sustainable data management

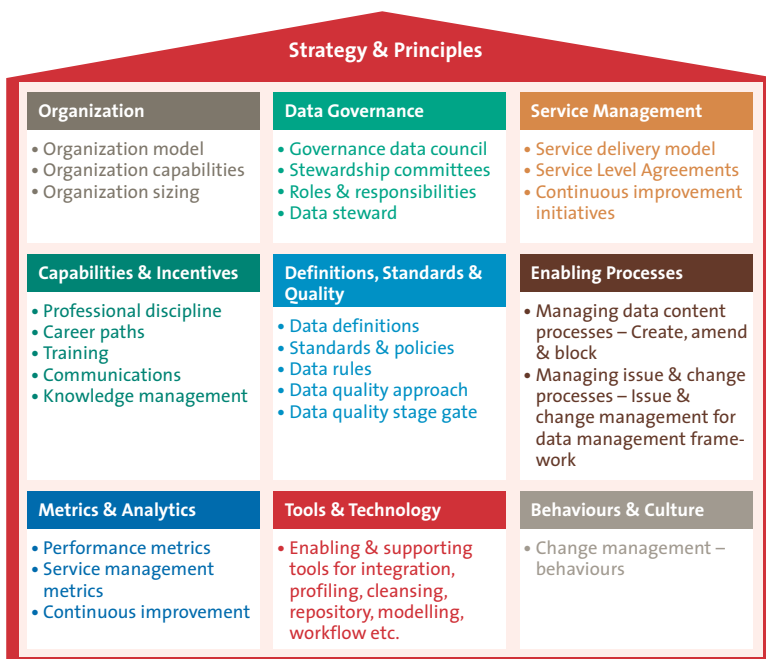


Figure 1. KPMG Enterprise Data Management Framework.

organization, given that they are properly embedded in the organization, guided by a strategy and principles for enterprise data.

## Theory: Conceptualizing data governance

In our data management projects, we have experienced that data governance and the accompanying data management measures from the EDM Framework have a positive effect on the performance and compliance of businesses. The measures elevate the data quality to the level of fitness-for-purpose for the organization and its regulator. Put differently, data quality should be improved up to the point at which data are a sufficiently reliable source of information to be fully supportive to the enterprises’ business processes and reporting.

Notwithstanding our broad experience in Enterprise Data Management projects, helping many organizations to create a sustainable data management organization, it is difficult to determine the exact impact on the business performance. For example, what is the effect of data governance on the Operation Expenditure (OPEX) of a company in the logistics sector? 10% decrease? 30% decrease? As one can imagine, it is difficult to find a straightforward answer to this question. One of the main reasons is that data governance, including all accompanying measures in a data management organization, has an enterprise-wide effect. As is essentially the case with all types of governance projects, data governance is initiated top-down. Consequently, there are many factors that ultimately determine the effect on the performance of the business. Consider incorrect product data in the ERP system of a food retailer. The supply chain planning and processes experience inefficiencies due to incorrect product data entries. Erroneous dimensions of products can, for example, lead to incorrect truck-loading schedules, requiring the expensive ad-hoc deployment of additional trucks. Furthermore, data governance should be a part of an Enterprise Data Management approach in order to function properly. It is therefore difficult to abstract the specific effects of data governance from the effects of the EDM framework implementation.

Due to this complexity, it is crucial to start with the conceptual relations between data governance, enterprise data quality and business performance. The context is added by mapping the relations on a typical organization, creating a generalized high-level overview of the effects of data governance in an organization. See Figure 2.

We start at the bottom layer. The *Technology* layer includes computers, databases and so forth, but also other devices that contain data used in the enterprise. Other elements include, for example, data stored in the cloud, on a server somewhere in the world that is (securely) connected to the internet.

The data in the *Technology* layer forms the basis of the Create, Read, Update, Delete (CRUD) processes and other data processes in the overlying *Data* layer. CRUD processes are the four basic processes performed with data in databases. Other important data processes are the acquisition, validation, exchange and archiving of data. For example, when a piece of data, such as information about the delivery address, size, weight and content of products, is used in the (supply chain) process, the data is “Read” for use in the business. From the viewpoint of causal relations, this means that if the *Technology* is of a lower quality, the quality of the processes that can be performed with the data will also be lower. If, for example, the data in the database are poorly accessible, the data processes also suffer from this.

The processes in the *Data* layer create a certain data quality. When, for example, a sales order is created and processed, the data quality of this order depends on the functioning of these processes. This originates from the fact that data quality is a multifaceted concept, entailing elements such as timeliness. A low data quality – the degree to which the data are fit for purpose for the enterprise – is pervasive, costly and can cause high inefficiencies.

The primary business processes and reporting are situated in the *Business Processes* layer. These are both affected by the data quality, as a sufficient data quality is required to enable efficient business processes and reporting. For example, in the production process of a certain product, the amount and composition of the planned products must be known. If the data quality is too low, for example due to failing CRUD processes, it could induce errors in the production process, as well as in the reporting of the enterprise processes, such as error logs or financial data.

Business performance and regulatory compliance are at the end of the causal relations chain. These are altered by the business processes and the reporting in the *Business Process* layer. At first, the efficiency of business processes largely reflects the business performance of the enterprise. This relation is relatively straightforward, as a lower throughput in the production process will cause the

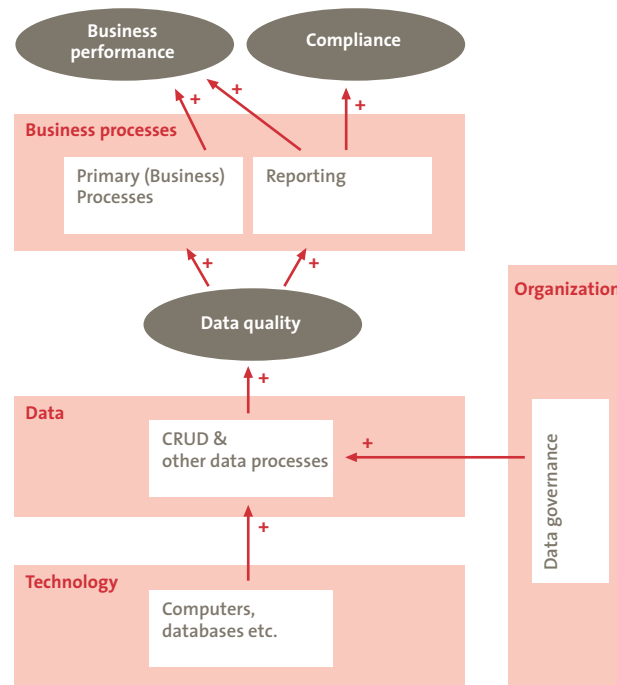


Figure 2. Data Governance Conceptualization.

production to decline, and therefore the business performance to decline (considering a constant product quality and market price). Second, the reporting influences the business performance and compliance of an enterprise. On the one hand, reporting is used for forecasting in business processes and performance overviews, on the basis of which the processes can be managed (monitoring). On the other hand, reporting (audit trails) is used to prove compliance with several laws and rules, depending on the country and sector.

As an overarching factor, the *Organization* layer includes data governance. It affects the CRUD and other data processes in the *Data* layer, as the organization will be more in control of its data processes as a consequence of data governance. It entails, for instance, roles and responsibilities regarding enterprise data. When no responsibilities are set, errors in the data might not be rectified, or only recognized much later in the process. If a responsible employee has been appointed, it can be assumed that fewer errors will actually enter the system, or that errors will be rectified earlier, leading to more efficient processes.

Although the conceptualization of data governance may seem straightforward, especially to those who are frequently involved in data management projects, it is an important starting point in measuring the effects of data governance in an organization. As the diagram shows, it can be concluded that data governance is part of the organizational structure, meaning that it is an overarching layer in the organization. Data governance influences the entire organization. This enterprise-wide effect drastically increases the difficulty of measuring the effects.

# *The enterprise-wide effect of data governance drastically increases the difficulty of measuring the effects*

## **Practice: Cases of data governance effects**

The conclusion from the conceptualization of data governance is clear: measuring the generalized effects of data governance is neither an easy nor an unambiguous exercise. To tackle this problem, we used three client cases to determine the effects *ex post*. Which effects do the companies experience? We interviewed three company representatives in the food, retail and logistics sectors who were closely involved in a data management project, being either data or IT managers. One additional representative from a company in the financial sector was interviewed as a comparing case.

The questions asked were twofold. First, do clients agree with our conceptual view on data governance? The answer to this question was positive in all interviews. More importantly, which effects of data governance did the clients experience in their businesses?

As expected, we gathered information on the effects of implementing data governance as part of a data management project. In line with earlier conclusions, the data governance effects as such are difficult to separate fully from the effects of other measures in the Enterprise Data Management Framework. Data governance alone simply does not suffice to sustainably increase data quality. The effects determined can therefore be attributed to the implementation of the data management organization, of which data governance is a component.

### **Case A: Dutch company in the food retail sector**

The first company was a large Netherlands-based company operating mainly in the national food and retail sector. The company processes products for resale at over 400 local sales points, but does not manufacture any products itself. The supply chain of this company is highly complex and crucial to the efficiency of the company's business processes. This case was characterized by insufficient data quality in the product database, primarily caused by absent formalized responsibilities with regard to the enterprise data.

Before data governance and accompanying data management measures were introduced at this client, the data quality throughout the supply chain was insufficient to ensure efficient business processes. The company's product database is the backbone of daily operations and con-

tains tens of thousands of products from various suppliers. At that time, it contained hundreds or even thousands of erroneous entries – and counting. This put a large burden on the efficiency of operations, as firefighting the consequences of errors in the product database had become a daytime job. Even critical errors occurred from time to time, which completely stalled the supply of products to the local sales points. The unavailability of products at the local sales points was not only a potential risk to the business profit, but also to the company's reputation. Especially in a highly competitive market, this was very concerning.

An important root cause for the deteriorated data quality in the product database was the overall low attention paid to the product data. Few employees felt responsible for the database which was used by many in the company. Triggered by this conclusion, we advised the implementation of a data management organization, including data governance.

Nowadays, the company has fully implemented roles and responsibilities concerning its enterprise data, as well as most of the other aspects related to the data management organization, such as a meeting structure to continuously improve the product data.

Most importantly, has data quality improved? Is the company more in control of its product data? And, are business processes more efficient now that data governance and accompanying data management measures have been introduced into the company? The overall conclusion is positive. Of the thousands of errors in the system, practically none remained, the rest having been structurally solved. The company has indicated that about 50% of all data-related working hours have been saved. Weekly product data mutations, such as price changes at the local selling points, are an example of this improvement. Due to errors in the product database, mutations had to be performed throughout the week in a far more ad-hoc manner. For instance, altered prices turned out to be incorrect, requiring extra manual adjustments by the local sales point employee. After implementation, almost no ad-hoc changes were required. In conclusion, by structurally improving the data quality through data governance, the company was able to structurally improve its supply chain efficiency and reduce operational expenditure. This validates the relations in the Data Governance Conceptualization.

## Case B: Dutch company in the food producing sector

The second case is a Netherlands-based company operating internationally in the food producing sector. The company manufactures products in various countries all over the world. The efficiency of the supply chain is therefore of great importance to the company's business performance. The unique feature of this case is the strong focus on the centralization of data management.

Similar to the situation in Case A, the data quality was insufficient for efficient supply chain processes. The subject of the matter was the data quality at local production facilities all over the world. The facilities all had separate databases, containing data in different formats and quality. For the global operator, it was therefore practically impossible to gain insight into the process efficiency at the production facilities, let alone control the production costs. This for example led to situations in which two production facilities in the same country had different vendors for the same product, missing out on larger-scale benefits and wasting valuable transportation costs. Incorrect invoicing and critical errors in the production process also frequently occurred. These inefficiencies were a serious problem for the profitability of the business.

As we see more often, the production facilities had such a strong focus on the daily routine of producing and transporting products to sales points that few employees could oversee the impact and importance of data quality in the system. A great deal of workarounds were in place to take care of duplicate vendor entries, and some fields in the system were used for other purposes than intended, creating a maze of data entries in the system. Again, the responsibility for the company's data had not been formalized, facilitating the deterioration of data quality.

The effects of the implementation of data governance and accompanying data management measures were also positive in this case. Data management was centralized to the global head office, taking responsibility for the data in the local production facilities. This included a large data-cleansing operation, removing about 40% of all data from the local systems. In conjunction with the centrally implemented data management organization, introducing, for example, enterprise-wide data standards, structured communication between head office and local production facilities, and data quality tooling, the data quality increased significantly. By deploying the new insights in the production processes for business strategy purposes, the costs of production could be significantly lowered. This conclusion is consistent with the relations in the Data Governance Conceptualization.

## Case C: Dutch company in the logistics sector

The third case is an international company operating in the logistics sector. The company ships products all over the world. As the process of shipping products is the core business of this company and the market in which it operates is highly competitive, the company relies heavily on the quality of data within its supply chain. The drive to increase data quality is the particularly interesting aspect of this case.

Data quality is essential to the efficiency of business processes at this company. On one hand, the company has to comply with increasingly complex regulatory requirements in many countries. On the other hand, decreasing demand in the market in which this company operates requires a highly efficient supply chain. In the past, the company encountered problems with customs and with an overall reduction in its competitiveness, caused by insufficient data quality. For example, some shipments could not be handled by customs in specific countries or were delayed due to insufficient or incorrect data on the shipments. Ultimately, this led to decreasing customer satisfaction, which the company simply could not afford.

Case	Type of market	Main characteristic	Impact on the business
Case A	Food retail	Insufficient data quality in the product database, primarily caused by absent formalized responsibilities concerning enterprise data.	The company was able to structurally improve supply chain efficiency and decrease operational expenditure. For example, 50% of all data-related working hours were saved.
Case B	Food production	Strong focus on the centralization of data management.	Costs of production could be significantly lowered and business strategy improved due to improved data quality and better insight into the enterprise data.
Case C	Logistics	Strong focus on data quality due to a supply chain in a competitive market and compliance issues.	The decision making and efficiency within the supply chain improved significantly, leading to fewer compliance issues in the global logistics and higher customer satisfaction.
Case D	Financial sector	Customer information stored in many systems throughout the company, leading to many inconsistent, incomplete, incorrect and double entries in several systems.	The improved data quality increased client satisfaction and improved compliance with the regulator.

Table 1. Comparison of cases.

*The company has indicated that about 50% of all data-related working hours have been saved*

As a solution, data governance and data management measures were implemented, including, for example, roles and responsibilities for enterprise data and data quality tooling. The entire approach was focused on decreasing the costs of the supply chain, to enable the company to maintain its position in the global market. The approach turned out to have a tremendous impact on the efficiency of the supply chain. The company estimates that the return on investment of creating and implementing business rules as standards for enterprise data alone is more than 3.5 million euros. The reason for this success is that the company is more in control of the quality of its enterprise data. It has the ability to not only indicate whether or not data are available, but also to check for other aspects of data quality, such as correctness. All brought together in a data quality dashboard, the decision making within the supply chain has improved significantly. This led to a more efficient supply chain with fewer compliance issues, leading to a higher customer satisfaction. The improved business performance and compliance due to data governance validates the Data Governance Conceptualization.

#### Case D: Dutch company in the financial sector

The last case we review involves a large international company in the financial sector. In contrast to supply-chain-driven companies, business performance is mostly based on client relations. The company does not produce, ship or sell any physical products. Customer satisfaction is essential to the existence of this company.

An interesting feature of this case is that, before the introduction of data governance and accompanying data management measures, client information was stored in many systems throughout the company. This led to many inconsistent, incomplete, incorrect and double entries in several systems. Of course, with its strong reliance on customer relations and satisfaction, this put a heavy burden on the business performance. Data cleansing had to be performed constantly, as the data quality continued to deteriorate.

Centralized data management, including data governance, was implemented in order to increase the overall data quality at the company. The governance of data in different systems was centralized, creating one single source of truth. As we have also seen at supply-chain-driven companies, data governance and accompanying data management measures significantly increased the data quality and the overall control over enterprise data. Continuous data cleansing is now no longer needed and the data quality remains fit for purpose, which has significantly increased customer satisfaction. In an international customer satisfaction index, the company was the first in the business to reach positive numbers, largely supported by the data management approach. Other examples of effects are the saving of FTEs in the manual processing of customer data, which has now become unnecessary due to structurally improved data quality, as well as increased compliance with various regulators, such as a high-quality customer data supply to the regulator in the US, which demands data on US citizens. Again, this case validates the relations in the Data Governance Conceptualization.

#### Comparing the cases

The conclusion of the four cases seems clear: data governance and accompanying data management measures are an effective concept to structurally increase enterprise data quality towards fitness-for-purpose within the company. The application differs slightly from case to case, but the effects are similar. The company in Case A focuses on product data, while the company in Case D

aims to increase customer data quality. Regardless of their exact application, data must always be supportive to the primary processes of a company, similar to what we concluded in the conceptualization of data governance. Whether the core enterprise data are customer, product or other data does not matter. Especially now that data are becoming more and more important in business processes, are becoming more of an asset to the company, the relevance of data management is increasing. Differences in effects between the financial, the retail and logistics sector are minimal, which will also be the case in other sectors that rely on enterprise data in their primary processes.

## Conclusion

In this article, we have conceptualized data governance and reviewed information on the effects of data governance at four companies. The most important conclusion is that data governance, as part of an enterprise data management approach, has a positive effect on the quality of enterprise data. We see that companies are better in control of their data, structurally increasing the quality. The improved data quality has a considerable effect on the efficiency of business processes and the business performance, but also positively increases compliance with the regulator. The exact determination of the effects, however, remains a specific and complex task, as many company-specific factors that determine business performance and compliance are influenced by data governance. Thereby, it turns out to be difficult to abstract data governance from other measures in the Enterprise Data Management Framework. For now, the fact that the experiences with Enterprise Data Management are consistently positive motivates us to continue delivering this message, helping future clients in creating a sustainable data management organization.

## References

- [Mos10] M. Mosley, D. Henderson, M.H. Brackett and S. Earley, *DAMA guide to the data management body of knowledge (DAMA-DMBOK guide)*, Technics Publications, 2010.
- [Thom06] G. Thomas, *Alpha males and data disasters: the case for data governance*, Brass Cannon Press, 2006.
- [Waal12] A. de Waal and A.W.A. de Jonge, *Data Governance bij een grote verzekeraar*, Compact 2012/2.

## About the authors

**N.L. Martijn** is a consultant at KPMG Advisory N.V. and advises companies in various sectors in the development and implementation of data governance and other data management measures.

**R.A. Jonker** is a partner at KPMG Advisory N.V. and Service Leader for Enterprise Data Management. He advises companies on the development and implementation of enterprise-wide data management.

*In an international customer satisfaction index, the company was the first in the business to reach positive numbers, largely supported by the data management approach*