Determining the Effects of Data Governance on the Performance and Compliance of Enterprises in the Logistics and Retail Sector

Nick Martijn, Joris Hulstijn, Mark Bruijne, Yao-Hua Tan

Determining the Effects of Data Governance on the Performance and Compliance of Enterprises in the Logistics and Retail Sector

Nick Martijn[1], Joris Hulstijn , Mark de Bruijne , Yao-Hua Tan

Delft University of Technology,
Faculty of Technology, Policy and Management
nickmartijn@gmail.com, j.hulstijn@tudelft.nl, m.l.c.deBruijne@tudelft.nl, y.tan@tudelft.nl

Abstract. In many of today's enterprises, data management and data quality are poor. Over the last few years, a new solution strategy has emerged, known as data governance: an overarching methodology that defines who is responsible for what data at what point in a business process. Although positive effects on the business performance and compliance of enterprises are seen in practice, a substantiated method for determining the effects of data governance has not yet been developed. This paper reports on explorative research to develop such a specification method. Through a conceptualization of data governance based on literature, case study analysis of clients of a large consultancy firm and interviews with representatives of companies that have recently implemented data governance, an effect specification framework was developed. Using the interviews, initial steps towards validation were performed.

Keywords: data governance, effect specification

1 Introduction

Enterprise data is becoming increasingly important. Data was initially seen as a by-product of business processes, used for example for financial recording (Lake & Crowther, 2013). Nowadays, data is considered to be a valuable asset in and of itself (Bughin et al., 2010). This value is primarily provided by two applications: measuring business performance and compliance reporting. Firstly, 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. Lack of data quality leads to huge losses (Steinfield et al., 2011). Secondly, enterprise data is used for financial reporting and for other kinds of compliance reporting. Companies have to comply to certain laws, such as Sarbanes-Oxley for companies listed on the US stock exchange, or Solvency II for insurance companies (Eling et al., 2007). These laws demand that companies demonstrate to the regulator that they are compliant, which requires evidence. Regulatory compliance creates additional data requirements. It is not sufficient to supply evidence; an audit trail is

[1] This paper summarizes (Martijn 2014). The research has been conducted with the support of Marinka Voorhout, specialist in Enterprise Data Management. We gratefully acknowledge her contribution.

also required (Jiang & Cao, 2011). Not meeting data requirements can lead to severe financial consequences. Consider the $3.75m fine Barclays bank received from US Financial Industry Regulatory Authority (BBC 2013). So enterprise data is used to gain insight in business performance, while also enabling compliance (Cheong & Chang, 2007; Golfarelli et al., 2004; Loshin, 2012).

Notwithstanding its importance, the standard of data management is often poor (Haug et al., 2011). Although most companies have well-managed IT systems, the responsibilities for maintaining specific kinds of data are mostly not incorporated (Redman, 2001). It is has been shown that when no clear policies, rules and controls are defined within the organization about who is responsible for what data, overall data quality will deteriorate (Batini et al., 2009). Poorly governed data may generate losses, as incomplete or erroneous information can mean a serious strategic disadvantage or lead to inefficiently organized business processes (Steinfield, et al., 2011), in addition to posing the risk of being deemed non-compliant.

Over the last few years a new solution strategy has emerged: *"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"* (Thomas, 2006). Essentially, data governance is an overarching methodology that defines who is responsible for what data at which point in the process. There is more to it though, such as internal controls, information systems architecture, standardization of data formats, corporate culture and use of technology, such as monitoring tools. Taken together, data governance measures can assure that enterprise data will be of sufficient quality. Data quality is seen as the most important aspect influencing usability of data for business processes and reporting (Friedman & Smith, 2011).

It turns out to be relatively hard to specify the effects of data governance projects and interventions. How do the application of various tools and techniques affect data quality? And subsequently, how does improved data quality affect business performance and compliance reporting? These are fundamental questions that have received relatively little attention. The effectiveness of data governance projects is only known from practical experience (De Waal & De Jonge, 2012). Data governance frameworks, such as DAMA DMBOK (Mosley et al., 2010), claim that they can determine these effects, but these are not fully scientifically substantiated.

In this paper, we therefore develop a framework to make it possible to specify the effects of data governance interventions. The paper is a summary of the graduation research reported in (Martijn 2014). We make use of case studies of firms, which have recently undertaken data governance projects. These cases were collected with the help of a large consultancy firm with extensive experience in helping clients improve their data quality through a data governance framework, in particular in the financial, logistics and retail sector (De Waal & De Jonge, 2012).

Concerning the choice of research method, note that organizational factors may affect the effectiveness of data governance interventions, but cannot be unambiguously operationalized. Thus, the case study approach is most appropriate, as the boundaries between the phenomenon (data governance) and the context (organizational effects) is relatively unclear (Boschi, 1982; Xiao et al., 2009).

The research proceeds as follows. Based on literature, we develop a conceptualization of data governance and its drivers (Section 2). We then develop a method to specify the expected effects of data governance interventions (Section 3). The conceptualization is relatively generic: it must be further specified for each case. Making use of client dossiers of companies that have recently adopted data governance measures, we show how the concepts can be further operationalized. Based on interviews with representatives of companies from the retail and logistics sector that are currently implementing data governance measures, we take initial steps towards validation of (Section 4). Full validation would require more cases, and would require comparison of the outcomes with other, independent, specification techniques.

2 Conceptualizing Data Governance

There is a lot of research on data quality and the effect it has on the use of information systems (Strong et al 1997). However, not much scientific research is specifically dedicated to the reverse question. How does data governance improve data quality, and consequently increase business performance and compliance? There are several data governance frameworks, of which the DAMA Data Management Body of Knowledge is most commonly used (Mosley et al., 2010). Such frameworks provide an overview of data governance measures to increase the data quality at an organization. The DAMA approach summarizes the following best practices (Mosley 2010): data architecture management, data development, database operations management, data security management, reference and master data management, data warehousing and business intelligence management, document and content management, meta data management, data quality management, all centered around the data governance. A problem with such frameworks is that there are generic and professionals have to adapt them to their own situation. The framework lists different kinds of activities, both technical and strategic. How can we structure their dependencies?

We have made a conceptualization of data governance, shown in Figure 1. As argued in the introduction, data quality is an essential property driving business performance and compliance. This will therefore be used as the guiding notion. To structure the diagram and locate the various activities, we use an enterprise architecture, based on the layers of Winter and Fischer (2006). From the bottom up: technology, software and integration architecture (merged here), process architecture, and business architecture. On the right we added organizational architecture, as we focus on governance aspects, involving roles, responsibilities and institutional arrangements.

Reviewing the structure from the bottom up, the first part consists of the technology, software and integration architecture. Here we find physical devices (gates; RFID readers etc.), computer systems (databases; networks etc.) and software applications (ERP systems; workflow management etc.) to store, retrieve and process information. In addition, we also find protocols and procedures for exchanging information and for integrating different modules. In this layer the actual data is situated.

Above this infrastructure layer, the process architecture is located, including the processes that are carried out with the data. Data processes are composed of four basic operations: Create, Read, Update and Delete (Martin, 1983; Polo et al., 2001). These CRUD operations determine the status of data elements at any point in the process, so we could say that this layer also contains the data architecture, which determines how data is being handled. When a certain piece of data, such as delivery address or price of a product is used, the data is mostly pulled from the infrastructure through an Enterprise Resource Planning (ERP) system.

The business architecture layer includes the value-adding processes, such as purchasing, sales, manufacturing or transport. Also internal control and risk management, compliance management, and reporting (financial statements, tax reports) are located here. To enable successful business processes, an effective data and process architecture is required. Consider for example the process of acquiring resources from a supplier, manufacturing products, and selling them on to a customer, consisting of steps like: receive, pay, manufacture, store, sell, dispatch. These steps in the primary process give rise to data operations. For example, receiving resources in a warehouse means the creation of new data objects representing the type of resources, storing the stock levels for these resources, updating inventory in the general ledger, changing the status of the corresponding purchase order, etc. The way in which basic 'CRUD' operations are implemented, largely determines data quality (Wand & Wang, 1996).

Kahn et al. (2002) see data quality as the degree to which data is fit for purpose, i.e. meeting company-specific requirements, see also Juran et al. (1999) and Wang et al. (1996). Data quality in turn largely determines the effectiveness of the business processes, and also influences the reporting quality (Eppler & Helfert, 2004), thereby affecting the drivers: business performance and regulatory compliance. Low data quality is pervasive, costly and can cause high inefficiencies (Eppler & Helfert, 2004; Fisher & Kingma, 2001; Wang et al., 2001).

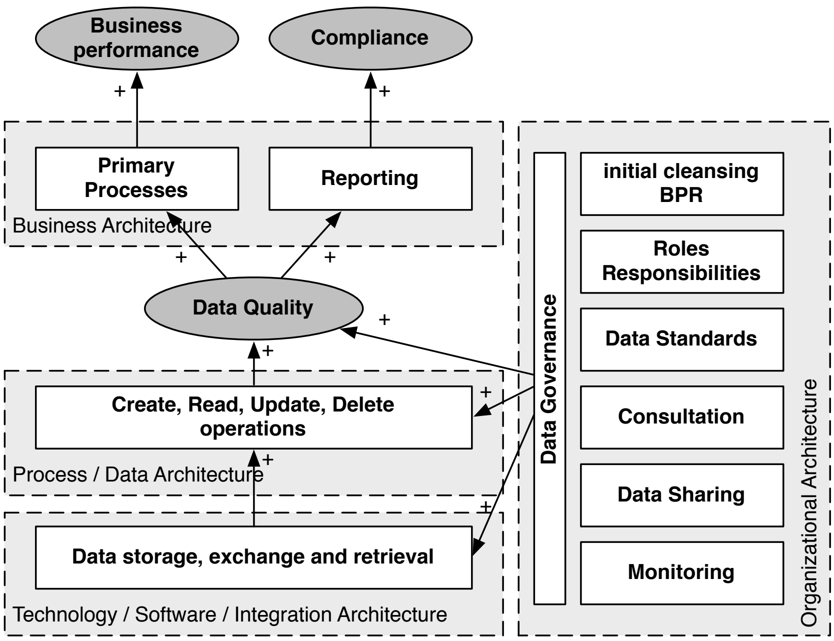

Figure 1. Conceptualization of data governance and its drivers

The organizational architecture contains institutional bodies like the board of directors, managing staff, change advisory board (CAB) and so forth. As data governance is a top-down methodology, we also locate most of the data governance efforts here. In particular, it entails roles for making people responsible for enterprise data. Data governance can influence all other layers. There is also a direct influence link between data governance and data quality, because of initial cleansing activities that are often part of a data governance project (see Section 2.1).

Business performance and regulatory compliance are at the end of the causal chain. First, the data that is used in business processes, directly affects business performance (Neely et al., 2002). Errors lead to missed deliveries, dissatisfied customers, etc. Also internal reports are used for forecasting, budgeting etc. Second, the reporting function influences regulatory compliance, as evidence needs to be produced of being compliant with laws and regulations (Jiang & Cao, 2011) .

2.1 Elements of Data Governance

Given this conceptualization of data governance, what does it actually involve? Data governance consists of various interrelated elements:

- *Initial cleansing* and *business process redesign* interventions are required before data governance can function properly within an organization. This sets a basic data quality level at the start of a data governance project. Often it also involves redesign of CRUD processes, compare process redesign (Hammer, 1990).
- *Roles and responsibilities* are essential to prevent lack of clear ownership for data management. When someone is made responsible, it can be assumed that less errors will enter the system, and that errors are detected and solved earlier, leading to more efficient processes (Mosley et al., 2010).
- *Data standards* describe how to represent, process, use and handle enterprise data. Implementing data standards, in combination with a governance structure that enforces the standards, leads to higher data and process quality. Use of data standards is a prerequisite for other data governance interventions. Standards also make it possible to measure data quality levels to indicate progress.
- C*onsultation* is meant to improve communication between departments (horizontal), and between management levels (vertical). Besides communication enhancements in primary processes and workflows, so called consultation platforms are recommended to improve the adaptability of data governance measures themselves. Errors should be identified and traced by feedback from users (Orr, 1988), so the company can learn from experience.
- *Data sharing* with supply chain partners is important for efficient alignment within supply chains, both internally, as well as externally (Steinfield et al 2011). This includes monitoring of data provided by the supplier to ensure sufficient quality. Data protocols can be part of the contract provisions.
- *Monitoring* should provide continuous insight in the current quality of the data to facilitate manageability of data by responsible employees, for instance, implementing tools that produce real-time data quality overviews on a dashboard. Again, this requires the ability to measure data quality level.

This list shows that data governance intervention requires a form of governance: it cuts across all layers and departments, which requires management support. Business must be involved, as they should define the information needs. Standards must be enforced. Budget to make the required changes to the IT infrastructure must be secured. Furthermore, even if we narrowly define data governance as the implementation of roles and responsibilities over enterprise data (Thomas, 2006), it cannot be abstracted from other data management aspects such as the use of standards and tools. After all, the roles and responsibilities are meaningless without technical and organizational means to support employees in executing these responsibilities. Therefore data governance is seen as a 'package deal': these elements strengthen each other.

3 Deriving a causal model

Using the diagram in Figure 1, we developed a causal model to specify the effects of data governance on an organization, shown in Figure 2. As such models are typically domain specific, the contribution lies in the method to derive the model. The model is developed on the basis of scientific literature and case study research (Section 4). Insights are based on dossier reviews at a large consultancy firm and interviews with clients who have recently been advised on data governance. Versions of the model were validated and adjusted on the basis of interviews with clients.

To scope the research, we decided to focus on cases from the logistics and retail sector. Data quality within supply chains is highly important (Li & Lin, 2006). Business performance can be operationalized using key performance indicators from the Supply Chain Operation Reference (SCOR) model. This model provides a standard method to review the performance of a supply chain, see Lockamy and Mccormack (2004), Xiao, et al. (2009) and Hwang et al. (2008). The KPIs include seven elements to determine customer service level: right product, right customer, right time, right place, right condition, right quantity and right costs (Fawcett & Fawcett, 2014).

The literature research, project dossiers and interviews with experts produced hypotheses for relationships between data governance measures, data quality, and ultimately business performance and compliance. The relations are shown as arrows in Figure 2. We use the following semantics: A –[+]– > B means a positive influence: when A increases, B should also increase. Conversely, A –[-]– > B means a negative influence. When A increases, B should decrease.

Starting at the left part of Figure 2, the various data governance elements will lead to better CRUD operations in business processes and subsequently, to better data quality. Initial cleansing will improve data quality directly. Consultation will improve communication between departments, which may lead to better IT responsiveness: the ability of the IT department to meet business demands. The enforcement of standards will reduce introduction of mistakes; in addition, it will make it easier to measure data quality. After all data quality is defined as fitness for purpose, where the purpose is reflected in company policies and data requirements, such as for instance those suggested by the SCOR model. Improved data sharing between partners in the supply chain, will improve data quality from suppliers. So also external factors play a role.

In the middle, better data quality helps to improve supply chain forecasting quality (Shankaranarayanan & Cai, 2006). Inherently, if forecasts are not reliable, the primary processes will be run less efficiently (Gunasekaran et al., 2004).

Improved data quality also decreases administrative costs. According to a worldwide investigation by Gs1 (2011), low data quality causes significant administrative costs referred to as shrinkage, the difference between what is shipped by the supplier and what is finally sold to the customer. Furthermore, efficiency in the primary process reduces operating expenditure (OPEX). Operating expenditure consists of all costs associated with operating a supply chain, such as transport and transaction costs. In particular, administrative costs have a large impact on OPEX.

When primary processes are more effective and efficient, for example when timeliness of deliveries is increased, the level of customer service will increase (Stevenson & Hojati, 2007). Customer service crucially affects sales. In addition, customer responsiveness is defined as the manner in which the business can meet demands of customers (Friedman & Smith, 2011). This property is related to the infrastructure: can it adapt. If a number of basic data elements are collected and processed reliably, new combinations can be engineered relatively easily. This improves the ability to forecast and report but also the ability to construct new customer services. These factors will help to increase sales and thus business performance (Neely, et al., 2002).

In the lower part of the diagram, higher data quality leads to improved internal controls and reporting quality, which by definition increases regulatory compliance. In most cases, an audit trail of enterprise processes and cash flow is required. Reporting quality is lower, when material (i.e. crucial) errors are not detected, or when relevant aspects of behavior are not reported, or not even recorded. Data quality is closely related to the notion of reliability, which involves accuracy (data correspond to reality) and completeness (all relevant aspects of reality are recorded) (Strong et al 1997).

In addition, reporting quality is affected by regulatory responsiveness, the ability to deal with compliance demands (Friedman & Smith, 2011). Data governance also affects this responsiveness variable: when basic figures are recorded and processed reliably, with an audit trial, new combinations of reports can be constructed reliably.

The relation between compliance reporting and data quality becomes even more crucial, when we apply innovative ideas of regulatory supervision, in which data is pulled from the source. For example, a 'data pipeline' infrastructure could facilitate reliable exchange of information in a trade lane, with access for customs authorities, but also for authorized traders (Klievink et al., 2012). Also in the XBRL-GL vision, data items are recorded and 'tagged' close to the source (Cohen, 2009). This makes it possible to record an audit trail with meta-data about provenance of data items. Given such basic elements with their provenance, new reports can be constructed reliably.

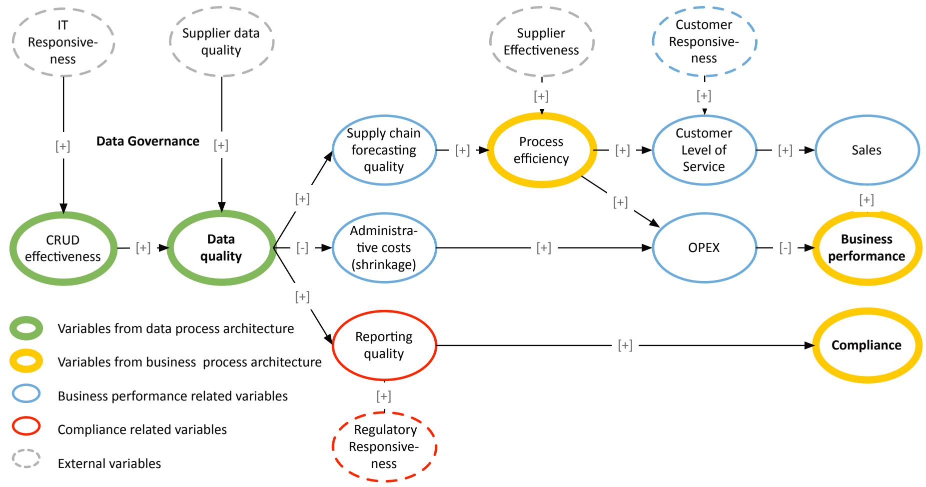

Figure 2. Causal diagram: expected effects of Data Governance

The discussion above shows, that it is in fact possible to operationalize the effects of data governance. Initially this will be qualitative, but once the organization has gained some experience also quantitative measures can be used. The resulting model always depends on the specific case; what matters is the line of reasoning. We identify variables that can be used to monitor the effects of data governance interventions.

Some of these variables are well defined and are measured on a routine basis: OPEX, sales, customer service level. Other variables can be defined, once a good operationalization is found. CRUD effectiveness can be determined by a process review, as is typically done by an operational auditor. For instance, consider the property that all information needs should be covered, or that no data should be requested that is not used later. Such properties can be determined using a create-use matrix, in which all data elements are plotted against the activities and roles by which they are created, and subsequently used. Data quality can be specified as the converse of the number of deviations per volume of data from a set of agreed data requirements. So we have *DQ = volume of data / number of deviations*. In the supply chain domain, it makes sense to use the SCOR indicators as a starting point for those requirements. The point is to trace supply chain delivery errors back to the information systems that cause them. Supply chain forecasting quality is determined by comparing forecasts with actual performance. Reporting quality is typically determined as a by-product of a financial audit: the number of deviations is reported as well. The variables IT responsiveness, customer responsiveness and regulatory responsiveness should be seen as intermediate variables, to factor in the ability of the organization to adapt. Typically, these can't be measured, which is why they are indicated as dashed ellipses.

That leaves only process efficiency, business performance, and compliance. Those variables can be seen as outcomes of data governance interventions, not inputs that can be used to control and adjust. These outcomes can in fact be measured, but that topic is out of scope. Consider for instance the Business Balanced Scorecard to measure business performance. Here we focus on the effects of data governance.

	Market	Main characteristic	Impact on the business
A	Food retail	Insufficient data quality in the product database, primarily caused by lack of formalized responsibilities.	The company was able to improve supply chain efficiency and decrease operational expenditure. For example, 50% of working hours on data related activities was saved
B	Food production	Strong focus on the centralization and standardization of data management	Costs of production could be lowered. Business strategy improved due to improved data quality and better insight in enterprise data
C	Logistics	Strong focus on data quality due to a supply chain in a competitive market and compliance issues	Decision making and efficiency within the supply chain improved significantly, leading to fewer compliance issues and higher customer satisfaction
D	Financial sector	Customer data stored in many different legacy systems, leading to inconsistent, incomplete, incorrect and double entries	Improved data quality increased client satisfaction and improved regulatory compliance

Table 1. Overview of cases and main impact on the business

4 Towards validation

To assess the adequacy and usefulness of the causal model, representatives of four companies that have recently implemented data governance were interviewed. The interviews were conducted at companies from the retail and logistics sector, which aligns with the research scope. For comparison, one case from the financial sector was included. We found no structural differences in responses from the domains.

The interviewees are mostly (IT) managers that were closely involved in the implementation of data governance measures, based on advice from the consultancy firm. Generally, their overall opinion on data governance is highly positive. They see the added value of data governance and acknowledge the positive effect on data quality and therefore on compliance and business performance. Issues identified are mostly on an operational level (Table 1). An example is erroneous product information leading to inefficiencies due to data errors (case A).

Firstly, representatives of all four companies confirmed the presence of the main causal relations in the framework. They too experienced that data governance measures have an effect on the implementation of CRUD processes, on the data quality level, and on the primary processes. This can be seen as a first step towards validation of the methodology that was developed in this research. Secondly, based on these interview outcomes, the early findings from the project dossier research and literature research, summarized in Figure 1, can also be validated. Thirdly, the interviews led to the observation that the further ‘downstream’ a factor is in the framework, the more difficult it is to specify precisely. It could be that this originates from the fact that most interviewees had a strong 'data-view' on the business, in which CRUD processes and data quality play an important role. For most of the interviewees, business issues are out of their scope. Fourthly, it is confirmed that data governance is a top-down methodology, as was found in the literature research. Measures taken to implement data governance, such as standards and tooling (monitoring) are forced onto the company by management responsible for data governance, backed by general management. This result supports the organizational theory behind the conceptualization.

5 Conclusions

Business performance measurement and compliance reporting are driving an increasing demand for improved data quality in enterprise systems. Data governance has emerged as a solution concept. It is hard to specify the intended effects of a data governance project before the start of the project: what is the business case? It is even harder to specify the actual effects afterwards, both qualitatively, i.e., has data quality improved; has client market responsiveness improved, and quantitatively, i.e., how much has business performance improved; how much costs have been saved?

In this research we have studied literature about data governance, both from theory and practice. We have reviewed dossiers of recent data governance projects conducted at a consultancy firm, and have held interviews with experts and with clients of this firm. This material has led to two outcomes:

(1) a conceptualization of data governance, positioning it within the organizational part of an enterprise architecture, and indicating the effects on its drivers, namely business performance and regulatory compliance.

(2) a causal model, with hypotheses about the influence of data governance on variables representing data quality and other intermediate notions, and ultimately on business performance and compliance.

Precisely specifying the effects of data governance interventions is exceptionally complex. Data governance is an overarching and top-down methodology. Many sector-specific factors are involved. Moreover, many of the factors cannot be operationalized. For instance, every measurement of data quality – correspondence to data requirements that represent fitness for purpose – depends on the definition that is applied by a specific enterprise. The data requirements, standards or data rules are always changing, because businesses are dynamic and respond to market conditions. Furthermore, in terms of effect determination, data governance cannot be abstracted from other data management practices. Data governance involves a package deal, of measures that mutually strengthen each other. All this makes a generic specification of the effects of data governance impossible.

We can however provide a method of how data governance effects can be specified within a sector (using e.g. the SCOR model in supply chain management), or within a specific enterprise. The causal model in Figure 2 can serve as a 'back bone' for such a method. It follows the logic of increasing positive effects (customer satisfaction; sales) and reducing negative effects (operational expenditure), by reducing data errors and improving process efficiency, as well as reporting quality.

The set-up of the research certainly has limitations. First, the research is focused on the logistics and retail sector. When another sector is considered, only the main relations in Figure 2 can be used, not the choice of measures. Second, because data governance is an overarching methodology, in which measures strengthen each other, it is impossible to study the individual effectiveness of interventions. Such evaluations would be valuable in practice to improve efficiency of projects. Thirdly, the research is scoped towards the expected benefits of data governance. Especially for business purposes, in which benefits are often weighed against costs, it would be useful to gain insight in the costs in order to build a proper business case for data governance projects. Consider IT investments, costs of additional personnel, costs of maintenance of tooling, and increased controls. These investments should be balanced against the costs of poor data governance. This is a useful topic for future research.

References

1. Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM Computing Surveys (CSUR), 41*(3), 16.
2. Boschi, R. (1982). Modelling exploratory research. *European Journal of Operational Research, 10*(3), 250-259.
3. Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly, 56*(1), 75-86.
4. Cheong, L. K., & Chang, V. (2007). The need for data governance: a case study.
5. Cohen, E. (2009). XBRL's Global Ledger Framework. *International Journal of Disclosure and Governance, 6*(2), 188–206.
6. De Waal, A., & De Jonge, A. W. A. (2012). Data Goverance bij een grote verzekeraar. *Compact, 39*(2), 11-17.
7. Eling, M., Schmeiser, H., & Schmit, J. T. (2007). The Solvency II process: Overview and critical analysis. *Risk Management and Insurance Review, 10*(1), 69-85.
8. Eppler, M., & Helfert, M. (2004). *A classification and analysis of data quality costs.* Paper presented at the International Conference on Information Quality.
9. Fisher, C. W., & Kingma, B. R. (2001). Criticality of data quality as exemplified in two disasters. *Information & Management, 39*(2), 109-116.
10. Friedman, T., & Smith, M. (2011). Measuring the Business Value of Data Quality. Stamford, CT, USA: Gartner.
11. Golfarelli, M., Rizzi, S., & Cella, I. (2004). *Beyond data warehousing: what's next in business intelligence?* Paper presented at the Proceedings of the 7th ACM international workshop on Data warehousing and OLAP.
12. GS1. (2011). Australia Data Crunch Report. Sydney, Australia: GS1.
13. Gunasekaran, A., Patel, C., & McGaughey, R. E. (2004). A framework for supply chain performance measurement. *International journal of production economics, 87*(3), 333-347.
14. Hammer, M. (1990). Reengineering Work: Don't automate, obliterate. *Harvard Business Review, Jul/Aug*, 104-112.
15. Haug, A., Zachariassen, F., & Van Liempd, D. (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management, 4*(2), 168-193.
16. Hwang, Y.-D., Lin, Y.-C., & Lyu Jr, J. (2008). The performance evaluation of SCOR sourcing process—The case study of Taiwan's TFT-LCD industry. *International Journal of Production Economics, 115*(2), 411-423.
17. Jiang, K., & Cao, X. (2011). Design and implementation of an audit trail in compliance with US regulations. *Clinical Trials, 8*(5), 624-633.

18. Juran, J. M., Godfrey, A. B., Hoogstoel, R. E., & Schilling, E. G. (1999). *Juran's quality handbook* (Vol. 2). New York, NY, USA: McGraw Hill.
19. Kahn, B. K., Strong, D. M., & Wang, R. Y. (2002). Information quality benchmarks: product and service performance. *Communications of the ACM, 45*(4), 184-192.
20. Klievink, B., Van Stijn, E., Hesketh, D., Aldewereld, H., Overbeek, S., Heijmann, F., & Tan, Y.-H. (2012). Enhancing Visibility in International Supply Chains: The Data Pipeline Concept. . *International Journal of Electronic Government Research, 8*(4), 14-33.
21. Lake, P., & Crowther, P. (2013). *Concise Guide to Databases:* London, UK: Springer.
22. Li, S., & Lin, B. (2006). Accessing information sharing and information quality in supply chain management. *Decision Support Systems, 42*(3), 1641-1656.
23. Lockamy, A., & McCormack, K. (2004). Linking SCOR planning practices to supply chain performance: An exploratory study. *International Journal of Operations & Production Management, 24*(12), 1192-1218.
24. Loshin, D. (2012). Evaluating the Business Impacts of Poor Data Quality. Silver Spring, MD, USA: Knowledge Integrity Inc.
25. Martijn, N. (2014) Exploring the Effects of Data Governance, Msc Thesis, Delft University of Technology, Faculty of Technology, Policy and Management.
26. Martin, J. (1983). *Managing the data base environment*. Prentice Hall.
27. Mosley, M., Henderson, D., Brackett, M. H., & Earley, S. (2010). *DAMA guide to the data management body of knowledge (DAMA-DMBOK guide)*: Technics Publications.
28. Neely, A. D., Adams, C., & Kennerley, M. (2002). *The performance prism: The scorecard for measuring and managing business success*: Prentice Hall Financial Times London.
29. Orr, K. (1988). Data Quality and Systems Theory. *Communications of the ACM, 41*(2), 66-71.
30. Polo, M., Piattini, M., & Ruiz, F. (2001). *Reflective Persistence (Reflective CRUD: Reflective Create, Read, Update and Delete).* Paper presented at the Sixth European Conference on Pattern Languages of Programs (EuroPLOP).
31. Redman, T. C. (2001). *Data Quality: The Field Guide*. Wobum: Butterworth-Heinemann.
32. Shankaranarayanan, G., & Cai, Y. (2006). Supporting data quality management in decision-making. *Decision Support Systems, 42*(1), 302-317.
33. Steinfield, C., Markus, M. L., & Wigand, R. T. (2011). Through a glass clearly: standards, architecture, and process transparency in global supply chains. *Journal of Management Information Systems, 28*(2), 75-108.
34. Stevenson, W. J., & Hojati, M. (2007). *Operations management* (Vol. 8): McGraw-Hill/Irwin Boston.
35. Strong, D. M., Y. W. Lee and R. Y. Wang (1997). Data Quality in Context. Communications of the ACM 40(5): 103-110.
36. Thomas, G. (2006). *Alpha males and data disasters: the case for data governance*: Brass Cannon Press.
37. Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM, 39*(11), 86-95.
38. Wang, R. Y., Mostapha, Z., & Yang, W. L. (2001). *Data Quality*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
39. Wang, R. Y., Strong, D. M., & Guarascio, L. M. (1996). Beyond accuracy: What data quality means to data consumers. *J. of Management Information Systems, 12*(4), 5-33.
40. Winter, R., & Fischer, R. (2006). *Essential layers, artifacts, and dependencies of enterprise architecture.* EDOCW'06.
41. Xiao, R., Cai, Z., & Zhang, X. (2009). An optimization approach to cycle quality network chain based on improved SCOR model. *Progress in Natural Science, 19*(7), 881-890.