Organising Accountabilities for Data Quality Management
– A Data Governance Case Study

Kristin Weber*, Lai Kuan Cheong*, Boris Otto*, Vanessa Chang*

*Institute of Information Management
University of St. Gallen
Müller-Friedberg-Strasse 8
9000 St. Gallen, Switzerland
krystin.weber@unisg.ch
boris.otto@unisg.ch

*School of Information Systems, Curtin Business School
Curtin University of Technology
Kent Street, Bentley,
Western Australia 6102, Australia
lkcheong@yahoo.com
vanessa.chang@curtin.edu.au

Abstract: Enterprises need corporate data quality management (DQM) that combines business-driven and technical perspectives to respond to strategic and operational challenges that demand high-quality corporate data. Hitherto, companies have assigned accountabilities for DQM mostly to IT departments. They have thereby ignored the organisational issues that are critical to the success of DQM. With data governance, however, companies implement corporate-wide accountabilities for DQM that encompass professionals from business and IT. This study examines a large organisation that has adopted an ad-hoc data governance model to manage its data. It was found that its DQM efforts were hampered mainly by the lack of clear roles and responsibilities and the lack of mandate to carry out data quality improvement initiatives. In order to promote effective DQM, this research identifies a data governance structure with the emphasis on collaboration between business and IT to support organisations.
1 Introduction

Companies are forced to continuously adapt their business models. Global presence requires harmonised business processes across different continents, customers ask for individualised products, and service offerings must be industrialised [cf. KÖ06]. These factors certainly impact the business process architecture and the IT strategy of organisations. Ultimately, however, data of high quality is a prerequisite for fulfilling those changing business requirements and for achieving enterprise agility objectives [NL06a]. In addition to such strategic factors, some operational domains directly rely on high-quality corporate data, such as business networking [Ve00; Ma04; Ye04], customer management [ZG03; RC05; CM06], decision-making and business intelligence [SZW03; PS05], and regulatory compliance [Fr06].

Data quality management (DQM) focuses on the collection, organisation, storage, processing, and presentation of high-quality data. In addition, there are organisational issues that must be addressed, such as maintaining sponsorship, managing change and expectation, defining accountabilities, enforcing mandates, and handling political issues [Wa98a; En99; Ep96; Th06]. Despite these organisational aspects, responsibility for improving data quality and managing corporate data is often assigned to IT departments [Fr06]. Also, many companies try to cope with data quality (DQ) issues by simply implementing data management or data warehouse systems. Surveys on data warehousing failures reveal that organisational rather than technical issues are more critical to their success [WFA04].

Successful DQ programs identify the organisational processes behind DQ [BN07]. In order to address both organisational and IT perspectives an integrated approach to DQM is required. With data governance, companies implement corporate-wide accountabilities for DQM that encompass professionals from both business and IT. Data governance defines roles and assigns accountabilities for decision areas to these roles. It establishes organisation-wide guidelines and standards for DQM and assures compliance with corporate strategy and laws governing data.

There are only few academic resources about data governance. Apart from a few DQM approaches dealing with accountabilities [Re96; En99], an elaborate analysis of the interaction of roles and responsibilities, and the design of decision-making structures is missing. For our research, we therefore incorporate data governance sources from consultants, analysts and practitioners [e.g., Sw03; De06a; DL06a; MS06; NL06b; Ru06; Th06; BN07].
The conclusion from these sources is that data can be managed more effectively and successfully through the adoption of a data governance structure. This paper outlines a case study of an Australian utility company ("Water Inc.") to determine the justification for formal data governance. Our main contribution is to demonstrate how formal data governance structures contribute to successful DQM based on this case study. The paper exposes how the state of the art in data governance was adopted by the company in order to build a virtual, boundary-spanning organisation for managing data quality effectively.

The remainder of the paper is structured as follows: Section 2 introduces related work on data quality management and data governance. It provides an overview of formal data governance structures. Section 3 outlines the research methodology and the case study company "Water Inc.". Section 4 describes the initial situation at "Water Inc." and details its data governance approach. Section 5 discusses the insights of the case study. The last section, Section 6, summarises the paper and its contribution.

2 Background

2.1 Data Quality Management and Data Governance

We refer to data quality management as quality-oriented data management, i.e., data management focusing on collecting, organising, storing, processing, and presenting high-quality data. The importance of organisational issues, such as maintaining sponsorship, managing change and expectation, defining accountabilities, enforcing mandates, and handling political issues, result from conflicts between different departments, lines of business, and legal entities in which DQM operates. DQM is boundary-spanning and provides many stakeholders (e.g., CxOs, sales, controlling, procurement, IT, business units, customers, public authorities) with high-quality corporate data by balancing their different interests (e.g., company-wide requirements, laws and regulations vs. local and regional differences). Because of these particularities of DQM, large multi-business companies are likely to have difficulties with institutionalising DQM, i.e., defining accountabilities, assigning people accountable for DQM within the organisational structure, and enforcing DQM mandates throughout the company.

Data governance — as part of DQM — addresses these particular issues within corporate structures. Data governance specifies the framework for decision rights and accountabilities to encourage desirable behaviour in the use of data. Hence, data governance encompasses two aspects: first, establishing accountabilities for DQM; and second, defining corporate-wide guidelines and standards for DQM. This paper focuses on the first aspect, because this aspect of data governance causes difficulties in most

---

1 Name of the company changed by the authors.
2 The term data is often distinguished from information by referring to data as "raw" or simple facts and to information as data put in a context or data that has been processed [HLW99; FS03]. In line with most data or information quality publications, we use the terms data and information interchangeably throughout the paper.
3 In the absence of academic definitions of data governance, this definition was adapted from the IT governance definition of [Weo04].
companies. Little practical guidance and few case studies of successful implementations exist. Furthermore, established accountabilities are the pre-requisite for the second data governance aspect—defining and implementing guidelines and standards.

Academic research on data governance is in its infancy. DQM approaches, such as Total Data Quality Management (TDQM) [Wa98b; Wa98a; HLW99], insufficiently address the allocation of accountabilities. They mainly outline DQM activities and decision areas. TDQM defines only one accountable role, the information product manager, which ensures that relevant, high-quality information products are delivered to information consumers. Few DQM approaches, such as Total Quality data Management (TQdM) by [En99] and DQ for the information age by Redman [Re96] deal with more than one accountable organisational position or role, roles related to several organisational levels, and their tasks and responsibilities.

However, an elaborate analysis of the interaction of roles and responsibilities, and the design of decision-making structures is missing. Hence, companies might find it difficult to establish and maintain organisational structures designed to assure and sustain high-quality data throughout the enterprise. Findings of a recent survey among data management professionals indicate that data governance is rarely adopted [Ru06]. Only 8% of respondents had deployed a data governance initiative, 17% were in the design or implementation phase.

2.2 Organising Data Quality Management

Hierarchical organisational structures fail to support corporate DQ in situations where data is used across organisational boundaries [Th06, 25], such as ensuring regulatory compliance or supporting global process harmonisation. Therefore, data governance establishes a “virtual organisation”. It defines roles4 and their responsibilities for DQM across organisational boundaries. It establishes committees5 to make important unbiased DQM-related decisions to achieve the best result for the whole organisation.

For identifying the roles and organisational bodies involved in DQM, Thomas [Th06, 81ff] distinguishes three kinds of data governance approaches. In Governance via Management managers execute decision-rights and set data-related rules. The governance organisation is the existing organisational structure. Responsibilities for DQM are not formalised. Governance via Stewardship is a more formal approach that provides separate roles and responsibilities. The governance organisation is made up or a hierarchy of data stewards and data owners. Data stewards typically define, create, or use data and may formulate rules. The Governance via Governance approach clearly distinguishes between governors (make rules and resolve issues) and data stewards (work with data, ensure compliance with rules, and raise issues). Governance bodies complement hierarchical management structures to address boundary-spanning DQM issues.

4 A role bundles different business tasks that are carried out by a single person (employee) or an organisational unit as well as their area of responsibility and competencies.
5 A committee is a politically-neutral organisational body that unites stakeholders from different parts of the organisation.
Data governance surveys [Ru06], case studies [De06b; DL06b], and reports and books from analysts and consultants [Sw05; DL06a; MS06; NL06b] follow that governance via governance approach. These sources distinguish between three and five organisational roles building the data governance organisation. Dycké and Levy [DL06a] and English [En99] describe more specialised roles – they distinguish twelve and nineteen roles respectively.

The analysis of these sources results in a set of four typical roles and one committee – the data quality board. Table 1 describes the roles and the committee. It provides a short description, the level and part of the organisation to which the roles typically belong, and the alternative names found in the sources. Names in brackets only partly match with either the description or organisational assignment.

<table>
<thead>
<tr>
<th>Role</th>
<th>Description</th>
<th>Organisational Assignment</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Sponsor</td>
<td>Provides sponsorship, strategic direction, funding, advocacy and oversight for DQ</td>
<td>Executive or senior manager (e.g., CEO, CFO, CIO)</td>
<td>Strategic information steward [En99], executive level [NL06b], executive sponsor [MS06], (executive council) [De06b]</td>
</tr>
<tr>
<td>Data Quality Board</td>
<td>Decides for corporate-wide standards and controls its implementation</td>
<td>Committee, chaired by chief steward, members are business unit and IT leaders as well as data stewards</td>
<td>Business information stewardship team [En99], data governance council [DL06a; MS06], data governance committee [Ru06b]. GRC Board [De06b]. Trustee Council [DL06b]. (Legislative level) [NL06b]</td>
</tr>
<tr>
<td>Chief Steward</td>
<td>Puts the board’s decisions into practice, enforces the adoption of standards, helps establish DQ metrics and targets</td>
<td>Senior manager with data management background</td>
<td>Master data coordinator [Sw05]. Director of data management [DL06a]. Chief steward [MS06]. Corporate steward [Ru06]. Lead stewards [De06b]. (Data czar) [DL06a]</td>
</tr>
<tr>
<td>Business Data Steward</td>
<td>Details the corporate-wide DQ standards and policies for his or her area of responsibility from a business perspective</td>
<td>Professional from business unit or functional department</td>
<td>Information professionals [Re96]. Business information steward [En99]. Business data steward [DL06a]. Subject area steward [NL06b]. Master data lead [Sw05]. Domain steward [Ru06b]. Business steward [MS06]. Subject matter expert [DL06b]</td>
</tr>
<tr>
<td>Technical Data Steward</td>
<td>Provides standardised data element definitions and formats, profiles source system details and data flows between systems</td>
<td>Professional from IT department</td>
<td>Database steward &amp; information architecture steward [En99]. Technical steward [MS06]. Source system data steward [DL06a]</td>
</tr>
</tbody>
</table>

Table 1: Set of Data Quality Roles

The actual number of roles may vary from company to company. However, the roles presented build a balanced and useful set when focusing on the strategic notion of
corporate DQM. The data governance organisation is virtual, i.e., it does not create a new organisational structure; it aligns and coordinates existing roles and cooperates with established committees and roles, such as IT governance boards, central purchasing committees, data architects or data security officers.

Against this background, this paper describes a case study of an Australian utility company. Several incidents that took place related to the lack of maintenance on network asset were the main business driver for starting a DQM initiative. Data for network assets need to be of high quality for the maintenance program to be planned efficiently. This paper outlines the process of implementing a data governance organisation within the company and how it benefited from a formal data governance (via governance) approach in its DQM initiative.

3 Research Methodology

3.1 Case Study Research

We adopted a case study research method which is particularly suitable for understanding phenomena within their organisational context [BGM87; Bi89; Yi02]. Case study research allows researchers to carry out their studies in a natural setting, learn about the actual process of managing data quality and generate theories from practice. This also allows researchers to answer the ‘how’ and the ‘what’ questions in order to understand the nature and complexity of the processes taking place. This is an appropriate research method for an area that had few previous researches [BGM87]. As outlined above, this is particularly true for data governance. Furthermore, data governance is an emerging topic that needs further development and research [WO07].

The chosen utility company ("Water Inc.") manages data from various disparate systems or sources, and the data involves different line of business units from multiple users and stakeholders. The current user base is more than 100 users and data quality is the priority concern for the organisation. The data collected from interviews with the IT Managers and Data Managers were primarily qualitative in nature. The interviews were recorded and later transcribed and analysed. The interview questions were designed to ascertain the IT view of data-related issues and problems and the business view of data-related issues and problems within the organisation. In addition, the study also investigated the resolution and mediation methods that the organisation employed; and the deployment of such methods. Additional information was also gathered from reviewing data management documentation of the organisation. The documentation review provided insight into the style of data quality management, methodology used for developing and enhancing application systems, and the interaction between IT and business.

3.2 Water Inc.

Water Inc. is a large utility organisation in Australia with complex data integration issues. This organisation provides essential services to approximately 890,000 industrial,
commercial, and residential customers. Its main mission is to provide reliable supply by maintaining the network and to restore the supply in the event when the supply was interrupted. It is also responsible for building new network to meet the demands of existing and future customers.

It employs more than 1,850 core staff with an asset base of nearly AU$3 billion. It operates in a regulated market to ensure quality of service and the accessibility of its network to all service retailer and supplier. Recently, Water Inc. was split into four different business entities (regional and metropolitan retail business, generation business and network/infrastructure (asset) management). Through this restructure, it had been charged with AU$2.23 million of investment to increase the network reliability by 25% over the next four years.

Water Inc. is obliged to regularly report to a government regulatory authority. The government regulatory authority requires maintenance plans and proof that the company had performed its duties in maintaining public safety, ensuring reliable supply, and efficient management of its network infrastructure. The building and maintenance of the network infrastructure requires the cooperation of several divisions: the asset management division (setting the strategy for managing network assets), field services (packaging of inspection and maintenance program), and works delivery (planning and construction of new infrastructure).

A few years ago, several incidents related to the lack of maintenance on network asset affected public safety and reliability of supply. This resulted in an inspection of the conditions of the network assets. Data for these assets need to be current and accurate so that a maintenance program can be planned efficiently. With this maintenance program, the condition of the assets can be accurately determined and aging assets can be targeted for replacement ensuring the reliability of the service network.

Water Inc. is also in the process of migrating its asset related data from its legacy asset management systems to Ellipse, an Enterprise Resources Planning (ERP) system and is also planning to replace its custom-built Geographical Information System (GIS) with an off-the-shelf GIS. The successful migration of these legacy systems depends largely upon the quality of the asset and spatial data. For example, in order to locate assets the XY coordinates and GPS references must be accurate; in order to identify the number of customers affected by an asset malfunction the connectivity of the network must be accurate. Recently, a dedicated Data Management department was created to manage asset data and its related information systems for the Asset Management division. This department is responsible for ensuring data quality and providing strategic direction for the asset related data initiatives. It consists of three teams: First, the Data Management (DM) team is responsible for providing strategic and tactical direction on data related issues, conducting quality assurance and data cleansing activities; second, the Data Services team enters data for asset related application systems (distribution network) and provides underground asset information to external parties conducting trenching activities; and third, the GIS Strategist team is responsible for the strategic direction of any spatial data related issues, such as positional accuracy and graphical representation of data.
4 Data Governance at Water Inc.

This research focuses on the activities within the DM team at Water Inc. This team acts as data stewards. It has knowledge of the business processes and an in-depth knowledge of the asset-related information systems. Currently, it is facing difficulties with managing asset-related data as an enterprise asset. The DM team feels it is managing data in a reactive and ad hoc manner, it has no direct access to its source database, it has difficulties obtaining consensus on data-related issues, its data improvement projects were mostly overrun and over budget, and any attempt to set data-related standards had not been taken seriously by other groups of the business and IT. In light of this, the DM team started a data governance initiative and the following remediations were instigated to address these issues:

- **Reactive data management.** The DM team establishes a data quality strategy to improve and maintain data quality. Based on the strategy, a work plan for IT related project and resource management is set-up.

- **Asset data not easily accessible.** The asset data is stored in several disparate databases and maintained by different application systems.

- **Data improvement project mostly overrun and over budget.** A dedicated project manager is appointed to liaise between the business and IT and to ensure that the data stakeholders’ expectation and communication are managed effectively. There is still a lack of IT tools for data profiling and data cleansing.

- **Difficulty in setting data standards.** Data standards are important for the migration of data from legacy systems to the new ERP system [CS01]. The DM team plans to purchase a metadata repository.

- **Difficulty in obtaining consensus on data related issues.** Data users are spread across the organisation, spanning different divisions. A data custodianship policy was drawn up but was difficult to put into action as it did not have a governance structure to enforce this policy.

This data governance model corresponds to the governance via stewardship approach [Th06]. However, this form of governance is not working for the DM team due to the following reasons: first, a lack of mandate from senior executives results in no power to act; second, sometimes DQM projects have no priority as senior executives do not understand their importance; and third, a lack of clear roles and responsibilities. Obviously, there is no connection between those at the operational level who knows the problem and those who have the power to make decisions but are removed from the problem. In order to address the issue, the DM team put together a plan for the governance via governance approach of data governance with executive support of the Asset Management General Manager. It resulted in a data governance organisation based on the following roles and responsibilities:

- **Data Governance Council.** Membership of this council consists of executives from various divisions who have an interest in the management of asset data. They are responsible for endorsing policies, resolving cross
divisional issues, engaging the IT council at the strategic level, strategically aligning business and IT initiatives, and reviewing budget submission for IT and non IT related projects.

- **Data Custodian.** Asset data is managed by the data custodian on behalf of Water Inc. He or she is responsible and accountable for the quality of asset data. The data custodian resolves issues raised in user group meetings. If issues become political and impacts stakeholders from other divisions, they are escalated to the DG council level. The data custodian is also responsible for endorsing data management plan, endorsing data cleansing plan, ensuring data is fit for purpose, converting strategic plans into tactical plans, change management, and stakeholder management.

- **Data Steward.** Data stewards have detail knowledge of the business process and data requirements. At the same time they also have good IT knowledge to be able to translate business requirements into technical requirements. They are led by the data custodian and are responsible for carrying out the tactical plans. They act on behalf of the data custodian in stakeholder management, change management, asset related information systems management and project management. They manage user group meetings, train and educate data users.

- **User Groups.** Data stakeholders from various divisions are invited to the user group meetings. These key data stakeholders consist of people who collect the data, process and report off the data. Technical IT staff is also invited to these meetings so that their technical expertise is available during the meeting. This is also a venue where urgent operational data issues can be tabled. The data users are responsible for reporting any data related issues, requesting functionality that would help them collect data more efficiently, and specifying reporting requirements.

The data governance structure in Figure 1 shows the business engagement with IT at the strategic, tactical and operational levels. This level of engagement ensures that IT and business are kept informed and IT initiatives align with the business data governance objectives.
5 Discussion

The data governance structure at Water Inc. provides a structured framework for mitigating the risks of DQM. It is scalable to include other divisions as the data governance efforts within the DM team mature. The data custodianship and user groups structure can be adopted by other division with the data governance council acting as the 'organisational glue'. Water Inc. had shown that business and IT need to work together in order to manage corporate data effectively.

The roles and committee used in the data governance structure correspond on the whole to the set described in Section 2.2. The data governance council corresponds to the data quality board, the data custodian to the chief steward and the data stewards correspond to the business data stewards. The Asset Management General Manager is the executive sponsor of the data governance initiative. IT technical staff fills the role of the technical data steward. Because of the operational nature of the user groups they have no equivalent in a typical governance structure.

The DM team had provided insights into the difficulties in managing data as corporate asset without proper authority. The following summarises the findings of this research:

- *The justification for formal data governance.* This study had shown that managing the data quality of enterprise data is not effective without a formal data governance model. The reason for this is because of the lack of clear roles and responsibilities among data stakeholders. Data governance also
assists business in engaging IT (vice versa) to manage corporate data collaboratively.

- **The process of setting up a formal data governance program.** The first step to setting up a formal data governance program is to determine a data governance structure. The structure provides escalation authority and a transparent decision-making process. Roles and responsibilities are defined so that members within the data governance structure are held accountable for their actions. Water Inc. had largely achieved this, however as this structure was recently introduced the success of its implementation cannot be determined in this study.

- **Ability to carry out actions as a result of a formal data governance structure.** Given the clear structure, the DM team is able to purchase a data profiling and metadata repository tool. The data profiling tool will allow the DM team to discover anomalies more efficiently. The metadata repository tool captures information about data so that it can be accessed by the whole organisation. Metrics measuring the quality of data had also been developed. The publication of these metrics will help management to determine the success of data improvement initiatives.

- **Simple data governance structure and framework.** The DM team did not want a cumbersome framework that could cause bottlenecks and delays in existing and future projects. With a proper structure and framework, the DM team is able to steer strategic projects to conform to and maintain good data governance. Nevertheless, Water Inc. moved from the governance via stewardship to the governance via governance approach.

## 6 Conclusion and Future Work

Companies need data quality management that combines business-driven and technical perspectives to respond to strategic and operational challenges demanding high-quality corporate data. Data governance specifies the framework for decision rights and accountabilities as part of corporate-wide DQM.

This paper investigates whether effective data quality management can be achieved without formal data governance. Some insights into data governance initiatives in a large utility organisation were obtained from this study. This paper underscores the importance of a data governance structure together with policies and procedures for managing data effectively. A data governance framework also enables collaboration from various levels of the organisation and it also provides the ability to align various data related programs with corporate objectives. This paper highlights that data governance provides a structured framework for mitigating the risks of data management.

This research has thrown up many questions in need of further investigation. Given that Water Inc. had only recently introduced a data governance structure, a longitudinal study would give a better indication of the benefits and success of the data governance
structure implementation. This case study was conducted in isolation. Future research could investigate multiple case studies and compare the implementation of data governance programs between organisations of different sizes.

With this paper we encourage IT and business to work together to achieve high quality data. The case study had shown that IT has the technical know-how but business sponsorship is important to give data quality programs visibility and direction.

References


[De06b] Dember, M.: Case Study: Gaining Efficiencies as a Result of Implementing a Data Governance Program. Presentation 2006.


[Ru06]Russom, P.: Taking Data Quality to the Enterprise through Data Governance. The Data Warehousing Institute, Seattle 2006.


